Terminological systems and prognostic models as instruments for quality assessment in intensive care

Raiez, F.

Citation for published version (APA):
Development and Application of a Framework for Maintenance of Medical Terminological Systems

Ferishta Bakhshi-Raiez, Ronald Cornet, Nicolette F. de Keizer

Journal of the American Medical Informatics Association 2008; 15: 687-700
ABSTRACT

Objective: Terminological Systems (TSs) need to be maintained in order to sustain their utility. This paper describes a study aiming at the standardization of the maintenance processes of medical TSs by capturing the criteria for the management of the maintenance processes into a framework. Furthermore, this paper describes application of the framework, which sheds light on the current practice of TS maintenance.

Design: Observational study.

Measurements: By means of a literature study, criteria for the maintenance of TSs were obtained and categorized into a framework. The current practice of TS maintenance was explored by a survey among organizations that maintain a TS. Results were stratified by the size of the TS being maintained.

Results: From sixty-three relevant articles, criteria for the maintenance processes of TSs were extracted and organized into four components. The primary component “Execution” concerns the core activities of the maintenance process. The other three components “Process management”, “Change specifications” and, “Editing tools” support the core activities of the component “Execution”. The survey had a response rate of 40% (37 of 93). The answers reflect the large variation in the number of criteria that are satisfied for the participating organizations. Overall, maintenance of larger TSs seems to satisfy more criteria.

Conclusions: The framework is an important step towards standardization of the maintenance of medical TSs and can be used to eliminate shortcomings in this process. Surveying the current practice showed that there is ample room to improve the maintenance processes of medical TSs, especially for the smaller TSs.
5.1 Introduction

Representing electronically stored medical data in a structured and standardized way is important for its use and re-use. To this end, numerous terminological systems (TSs) have been developed such as the International Classification of Disease Ninth Edition (ICD-9) (1), the Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT) (2), MedDRA (3), LOINC (4), Gene Ontology (5), and the Foundational Model of Anatomy (FMA) (6). A TS interrelates concepts of a particular domain and provides their terms and possibly their definitions and codes (7). The increase in number of TSs is demonstrated by the growth of the UMLS Metathesaurus that now integrates 143 (UMLS 2007AB release) TSs (8). Throughout different generations, these TSs have developed from single-purpose, inextensible systems to extensible multi-purpose systems (9).

Maintainers of medical TSs recognize that TSs are not static. Errors may be corrected and new concepts are added with the emergence of new diseases (e.g. the Severe Acute Respiratory Syndrome (SARS)). Furthermore, terms denoting concepts and their usage change over time (e.g. for AIDS and HIV). Finally, some concepts (e.g. hysteria) go out of fashion or become obsolete as domain knowledge changes.

There is a need to maintain TSs in order to sustain their utility (10-13). In line with the definition of (software) system maintenance, we define maintenance of a TS as “all activities that are performed on the TS after the first release” (14;15).

Maintenance of medical TSs becomes a challenging problem as their size increases. The SNOMED CT, for instance, contains 376,046 medical concepts (active and retired), associated with 1,060,424 terms describing these concepts, and related to each other by a hierarchy consisting of 1,359,435 relationships (July 2007 release) (2). Furthermore, continuous changes in TSs may lead to inconsistencies when they are not properly handled (12). Besides, frequent changes in TSs can lead to difficulties in the processing and tracking of historical data (12). Finally, processing changes can be labour-intensive and time-consuming. For TSs such as SNOMED CT (2), DICOM (16), MedDRA(3), RxNorm (17), and LOINC (4) teams of 5 to10 full-time equivalents are responsible for the maintenance processes (18).

The need for standardization of the maintenance of medical TSs has been discussed by several authors (12;13;19-26). Without a well-structured maintenance process, TSs cannot provide the utility required by today’s complex electronic health record applications. However, while the present medical literature has dealt with the semantic and ontological issues of maintenance, few papers concern the management aspects of the maintenance process. The objective of this study is twofold. The first goal is to develop a framework that summarizes the features that a TS’s maintenance process should cover as described in the literature. This framework can be used as a reference to design, evaluate and improve the
maintenance process of (new and existing) TSs. The second goal is to apply the developed framework to gain insight into the current maintenance practice of existing TSs and to evaluate whether the introduction of the framework triggers the wish for maintenance process redesign.

5.2 Materials and Methods

5.2.1 Literature search

In order to gain insight into the most important aspects of TS maintenance, we searched the literature for features and procedures related to the management of the maintenance of medical TSs.

A search in the medical literature from 1966 until 2006 was performed using Medline. The following MeSH headings (indicated by an asterisk) and terms were used: “Terminology*”, “Vocabulary*”, “Terminological System” and “Record coding system”, combined with “Guideline*”, “Maintenance*”, and “Upgrading”.

In addition to the Medline search, a supplementary search was performed including papers referenced by other papers and the Internet. To explore the Internet, from October 2006 to January 2007, Google searches were performed using the same key words as described above. Papers found with this supplementary search were considered relevant when they described the maintenance process of a (terminological) system. The supplementary search was not restricted to the medical domain.

5.2.2 Methods for developing the framework

In 2004 a preliminary literature search resulted in twenty-nine criteria related to the management of the maintenance of medical TSs. Discussion meetings with several experts (i.e. two medical information scientists, one terminology expert/ontologist and one software engineer) were organized to analyze and categorize these criteria. Consensus results were captured in a draft version of the framework. This draft version was evaluated via a first questionnaire by the maintainers of 27 TSs and the results were published in (27).

The extensive literature search described in the previous section was performed to identify additional criteria for the maintenance process. Again, discussion meetings with the experts were held to identify and organize the final set of criteria and to develop the final version of the framework. The framework categorizes the components of the maintenance process and provides criteria and requirements for designing the maintenance process of medical TSs.
5.2.3 Exploration of Current Practice

In May 2006, a web-based questionnaire was made available to maintainers of 75 TSs included in the UMLS (28), to maintainers of 17 TSs (not included in the UMLS) who responded to the National Committee on Vital and Health Statistics (NCVHS) ‘Terminology Questionnaire’(18;27) and the maintainers of DICE (Diagnoses for Intensive Care Evaluation) terminological system, which is developed and maintained in our own department (29). The objective of our questionnaire was twofold. Firstly, to compare the current practices with respect to the maintenance processes of existing TSs with the proposed framework. Secondly, to assess whether the introduction of the framework triggered the wish for maintenance process redesign. For the remaining 68 TSs included in the UMLS no valid contact information, i.e. email addresses, was available. To increase the response rate reminders were sent two weeks, four weeks and six weeks after the first request.

The web-based questionnaire covered all topics described in the framework. The respondents could gain more information on each topic by clicking on a link that would show a pop-up window with a short description of the corresponding part of the framework. For each topic, the respondents were asked how their current maintenance activities were organized and what the desired situation would be. Most of the questions were multiple-choice with an “Other, namely” option to enter complementary free text remarks. A copy of the questionnaire is available from our website (30).

The analysis of the questionnaire responses mainly focused on the extent to which the maintenance process of different TSs corresponded to the framework and if the framework triggered the wish for maintenance process redesign. Furthermore, this research studied the differences in the maintenance processes of the various TSs and whether the size of the TS is related to its maintenance process. To this end, the results are summarized per quartile, based on the size (i.e. number of concepts) of the included TSs.

5.3 Framework for Terminological System Maintenance

Many publications considering the semantic and ontological aspects of the maintenance process of medical TSs were found (12;19;31-34). However, organizational aspects of the maintenance process have received relatively little attention within the medical domain. The Medline search from 1966 until 2006 resulted in twenty-eight relevant publications. The supplementary search (including papers referenced by other papers and the Internet) resulted in another thirty-five publications. These publications spanned not only the medical domain, but also domains such as software engineering and information engineering.
As described in the preliminary results (27), the criteria and procedures for the management of the maintenance process of medical TSs could be subdivided into four components. “Execution” is the primary component and covers the core activities of the maintenance process such as collection of proposals for changes, validation of proposals for changes, implementation of changes, verification of changes, documentation of proposals and the implemented changes, and version management. The other three components support carrying out the core activities of primary component “Execution”. These components encompass: 1) “Process management”, describing the coordination and management of the maintenance process and the disciplines involved, 2) “Change specifications”, describing the possible changes that occur in the TSs and how to deal with them and 3) “Editing tools”, in most cases software applications that are used for various activities.

In total, 31 criteria for the maintenance process of TSs were identified in the present study. The earlier list of 29 (27) criteria was rearranged, expanded and refined with additional specifications and operational definitions. Table 5.1 presents the criteria (numbers 1 to 11) and the related sub-elements of the primary component “Execution”. For instance, criterion number 1 concerns the collection of proposals and the sub-elements describe what information (e.g. proposer ID, Version number etc.) should be included in the standardized forms. Table 5.2 presents the criteria (numbers 12 to 31) and sub-elements of the supporting components.

Figure 5.1 captures these criteria for the maintenance process into a framework that can be used as a guideline to manage the components of the maintenance process. Below, the four components are described for the situation in which there is a maximal adherence to the criteria. It should be noted that in some cases not all criteria are feasible and that an optimal management of the maintenance process does not necessarily satisfy all criteria.

5.3.1 Execution
The component “Execution” summarizes the core activities (i.e. sub-processes) of the maintenance process and describes how to carry them out. Figure 5.2 gives an overview of these sub-processes (3;13;35-38). Furthermore, on the left-hand side of the flow chart, the disciplines responsible for each step of the execution process are depicted.

Collecting proposals for changes
To optimize the process of change proposal, proposals are submitted using structured and standardized forms (3). For a large number of TSs in use, change suggestion forms are available from their websites (38-47). Mostly, in these forms the person who makes the suggestion, i.e. the proposer, provides information on: 1) the version number of the TS for which the change is required, 2) the corresponding terms and codes for the concept for
which the change is required, 3) change type and, 4) additional information on the concepts depending on change type (e.g. a definition of the concept, in case a new concept is to be added).

**Validating the proposals and verifying the changes**

To ensure the quality of the changes, the proposals for changes are validated. For each proposal, the maintenance team determines whether or not the change is desired (e.g. the change does not lead to redundancies) and whether it is possible to incorporate the proposed change into the TS (e.g. the proposed change does not violate the terminology domain or structure). Furthermore, the impact of the proposed change on the TS model is determined before implementing it (e.g. impact on hierarchical relations) (48-52). The International Classification for Nursing Practice (ICNP) for example, uses the following criteria to facilitate decision-making (38):

- Suggested change is appropriate for the domain of the TS,
- The suggested change does not lead to redundancy,
- If a suggested term is redundant, it will be reviewed for use as a synonymous term,
- If a suggested term is retained as a synonym, a "preferred term" would be identified,
- If a new concept is suggested, this is expressed in a clinically relevant way and the definition is consistent with scientific knowledge,
- If a new concept is suggested, the new concept does not violate the TS structure.

Other TSs use similar criteria to evaluate the proposals for changes (53-55).

Once a change has been incorporated into the TS, the maintenance team verifies the change on completeness, textual errors, consistency of the hierarchical relations, consistency of mappings to other TSs and consistency of mappings to other languages (48-52). After this in-house verification, feedback is given to the proposers to inform them on the incorporated changes and to ask them to verify the changes (56).

ABC Coding Solutions for example, works closely with individuals and institutions requesting changes to assure that the changes are optimally incorporated into the terminology (53). Once a change has been accepted for implementation in the TS, a draft version of the proposed ABC Terminology including the modification is developed and presented to the requesting party and other subject matter experts (i.e. a team of independent domain experts) for verification.

**Documenting proposals or changes**

To keep track of the proposals and the incorporated changes, it is important to have a well-structured and standardized documentation. The documentation can be managed on paper
or electronically (56). For each proposal, the following information is documented (33;56;57): proposal date, proposer ID, concept ID, change type, validation method, and the acceptance or rejection date.

For each change that is incorporated into the TS, additional information is documented. The amount and type of documented information depend on the change type and the number of concepts involved in the change operation (19;56). Next to the concept-specific information, the documentation includes the implementation date, change type, editor’s ID and finally the version number of the TS into which the changes are incorporated (33;50;57). Furthermore, if the functionality is supported by the editing tool in use, it is possible to link a “history note” to each concept, that automatically keeps track of the changes made to that concept (49;58). For instance, the concept-level history tracking mechanism that has been developed for the NCI Thesaurus contains the following information for each concept (59):

- History_ID, i.e. record number in the database.
- Concept_Code, i.e. identifier of the concept.
- Concept_Name, i.e. preferred name of the concept.
- Action, i.e. change type.
- Edit_Date, i.e. timestamp.
- Edit_Name, i.e. name of the edited NCI Thesaurus schema.
- Host_IP, i.e. address of editor's workstation.

**Version management**

Version management concerns the distribution of updated versions of the TS. It is important that the users are provided with updated versions of the system on a regular basis (50;57;60). Updates must be referable to unique consistent version identifiers. This identifier is for instance used in the communication with the users for collection of change proposals (61;62).

New releases of a TS can be distributed in different ways. One option is to provide a full version of the system each time an update of the system is available (non-incremental update). Another option is to provide a list with the accepted changes that can be imported into the previous version of the system (incremental update) (61). The choice of the update method depends on the volume and the complexity of the changes made. The update frequency is sufficiently low in order to quickly accommodate new concepts and repairs (63). On average, twice a year a new release of a TS is launched (36;57;60). However, this update frequency does depend on the content and the purpose of a TS. For instance, the Current Procedural Terminology (CPT®) containing three types of categories (I-III), has a different updating scheme for each category depending on its intended usage (54;64). Once approved by the Editorial Panel, the newly added Category I CPT concepts
are distributed annually whereas the category II CPT concepts are distributed every two years. Since Category III CPT concepts are used to report emerging technologies and must respond quickly to changes in treatment methods, Category III CPT concepts are updated twice a year. The update frequency also depends on how the TS is distributed. Printed revisions will necessarily be produced less frequently than revisions of a computerized system. Electronic updates have the advantage of being accessible more quickly or even instantaneously (49). Furthermore, distribution of updates may need to be coordinated with updates of related TSs.

5.3.2 Process management

Coordination
The medical TS’s maintenance process may be coordinated or not (23). Collaborative maintenance model allows the TS’s users to change the TS themselves; centralized maintenance model requires the intervention of a maintenance team for this activity. Although the collaborative maintenance model allows the users direct read and write access to the TS, it may lead to inconsistent or redundant content since not all users are qualified to change the TS. Therefore, we suggest to apply the centralized maintenance model that requires that the maintenance process is coordinated and carried out by a qualified maintenance team (13;23;60). This maintenance team is easily accessible and responds quickly to proposals and questions (65).

People involved
In order to carry out the maintenance process correctly, people from relevant disciplines take part in the maintenance team (20;60). Users and domain experts are involved because of their knowledge of the domain of the TS (60;61). Terminology experts and ontologists are involved because of their knowledge of the structure and the architecture of the TS (61).

Table 5.1 An overview of the criteria for the primary component Execution.

<table>
<thead>
<tr>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Proposals for changes in the TS are standardized in written forms, containing information on:</td>
</tr>
<tr>
<td>• Proposer ID, i.e. identification of the person who suggests a change.</td>
</tr>
<tr>
<td>• Version number, i.e. the unique version number of the TS for which the change is requested.</td>
</tr>
<tr>
<td>• Concept information, i.e. concept ID, concept code and concept name.</td>
</tr>
<tr>
<td>• Change type, i.e. change operation as described in the change model.</td>
</tr>
<tr>
<td>• Change definition, i.e. description of the desired change.</td>
</tr>
</tbody>
</table>
2. Proposals for changes in the TS are validated, on:
   - Necessity, i.e. change is relevant for the domain of TS and does not lead duplicate information.
   - Possibility to incorporate a change, i.e. change is consistent with scientific knowledge, does not lead to redundancies and does not violate the TS structure.

3. In case of uncertainty, the acceptance of proposals is based on group consensus.

4. In case a proposal is rejected, feedback is given to the proposer.

5. In case a proposal is accepted, changes made in the TS are verified by the maintenance team, on:
   - Completeness, i.e. all aspects of the change are taken into consideration.
   - Textual correctness, i.e. textual errors are corrected.
   - Consistency of hierarchic relations, the change does not lead to inconsistencies in the relations.
   - Consistency of mappings to other TSs, i.e. the change does not lead to inconsistencies in the mappings to other TSs.
   - Consistency of mappings to other languages, i.e. the change does not lead to inconsistencies in the mappings to other languages.

6. After incorporation of the changes into the TS, feedback is given to the proposer, to:
   - Inform the proposer, i.e. notify that the suggested change is implemented.
   - Ask them to verify the changes, i.e. validate whether the adopted change is in compliance with the suggestion.

7. Documentation is structured and standardized.

8. Proposals for changes are documented, with:
   - Proposal Date, i.e. the date on which the suggestion for a change was received.
   - Proposer ID, i.e. information on the proposer like name, address etc.
   - Concept information, i.e. Concept ID, Concept code and Concept name.
   - Change type, i.e. suggested change operation based on the change model.
   - Acceptance method, e.g. whether consensus rounds were used.
   - Status, i.e. is the suggestion accepted or rejected.
   - Acceptation/rejection date, i.e. the date on which it was decided to accept or reject the change.

9. Changes made in the TS are documented, with:
   - Concept information, i.e. concept ID, concept code and concept name.
   - Implementation date, i.e. the date on which the change is incorporated into the system.
   - Change type, i.e. change operation that was performed.
   - Editor ID, i.e. identification of the person who processed the change.
   - Version ID, i.e. version number of the TS in which the change was adopted.

10. New versions of the TS have a unique identification number, including the publication date and are distributed as:
    - Complete new version, i.e. the whole knowledge of the TS is provided.
    - Incremental update, i.e. list with changes that can be imported into the previous version.

11. On average, twice a year a new release of the system is launched (depending on the type of TS).
Table 5.2: An overview of the criteria for the three secondary components.

<table>
<thead>
<tr>
<th>Criteria</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Process Management</td>
<td></td>
</tr>
<tr>
<td>12. There is a team responsible for the management of the maintenance process.</td>
<td></td>
</tr>
<tr>
<td>13. The maintenance team is easily accessible.</td>
<td></td>
</tr>
<tr>
<td>14. The response time of the maintenance team for proposals and questions is short.</td>
<td></td>
</tr>
<tr>
<td>Coordination</td>
<td></td>
</tr>
<tr>
<td>15. Different disciplines participate in the maintenance team.</td>
<td></td>
</tr>
<tr>
<td>• Users: domain expert because of their knowledge of the domain of the TS.</td>
<td></td>
</tr>
<tr>
<td>• Terminology expert because of their knowledge of the structure and architecture of TS.</td>
<td></td>
</tr>
<tr>
<td>• Software Engineers because of their knowledge of technical possibilities and functions of the TS.</td>
<td></td>
</tr>
<tr>
<td>• Coordinators are responsible for the management of the maintenance process.</td>
<td></td>
</tr>
<tr>
<td>Persons involved</td>
<td></td>
</tr>
<tr>
<td>16. Only qualified people are able to make changes in the TS.</td>
<td></td>
</tr>
<tr>
<td>17. Access to the TS is secured, using:</td>
<td></td>
</tr>
<tr>
<td>• Identification, i.e. identity registration.</td>
<td></td>
</tr>
<tr>
<td>• Authentication, i.e. identity verification.</td>
<td></td>
</tr>
<tr>
<td>• Authorization, i.e. access rights verification.</td>
<td></td>
</tr>
<tr>
<td>Security</td>
<td></td>
</tr>
<tr>
<td>18. The codes that are assigned to concepts are context free and non-significant e.g. random.</td>
<td></td>
</tr>
<tr>
<td>19. Concepts that are no longer in use are not removed from the TS. Instead, these concepts are kept in the TS and are marked obsolete.</td>
<td></td>
</tr>
<tr>
<td>20. In case of removal of a concept (which is not in accordance with the criteria), the code assigned to that concept is not reused.</td>
<td></td>
</tr>
<tr>
<td>21. Within the TS there are no limitations regarding the number of concepts, hierarchic levels and terms that can be added.</td>
<td></td>
</tr>
<tr>
<td>Change specifications</td>
<td></td>
</tr>
<tr>
<td>22. There is a ‘change model’ that specifies all types of changes (e.g. delete, add, adjust, etc.) that can occur in the TS in concordance with the change policies.</td>
<td></td>
</tr>
<tr>
<td>Change policy</td>
<td></td>
</tr>
<tr>
<td>23. The maintenance team uses an editing tool for the maintenance process.</td>
<td></td>
</tr>
<tr>
<td>24. The editing tool is secured with login name and password.</td>
<td></td>
</tr>
<tr>
<td>25. The editing tool contains a module for collecting proposals.</td>
<td></td>
</tr>
<tr>
<td>26. The editing tool contains a module to support the consensus process.</td>
<td></td>
</tr>
<tr>
<td>27. The editing tool supports the input of changes into the TS.</td>
<td></td>
</tr>
<tr>
<td>28. The editing tool supports automatic validation controls.</td>
<td></td>
</tr>
<tr>
<td>29. The editing tool generates reports for documentation.</td>
<td></td>
</tr>
<tr>
<td>30. The editing tool supports managing different versions of the TS.</td>
<td></td>
</tr>
<tr>
<td>31. The editing tool supports distribution of new versions.</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.1 Framework for the maintenance of medical terminological systems. The numbers in brackets refer to the related criterion.
Figure 5.2 Flow chart of the sub-processes of the component “Execution”. At the left-hand side of the chart the roles responsible in each step of the execution phase are depicted.
Software engineers are involved because of their knowledge of the technical possibilities and functions of the system (61). Finally, one or more members of the maintenance team are also responsible for the organization of the maintenance process (61).

**Security**

An adequate user administration ensures that only qualified persons can make changes to the TS. This is realized by securing access to the TS by identification (i.e. identity registration), authentication (i.e. identity verification) and authorization (i.e. access rights verification) (23;56;58;66;67).

**5.3.3 Change specifications**

**Change policy**

Technical decisions during the development of a TS can impact its capacity to grow, change and remain usable over time (49;50;63). To this end, a change policy needs to be adopted when developing a TS.

First of all, medical knowledge is constantly changing and consequently the classification of the medical concepts is also changing. AIDS, for instance, is now known to be an infectious disease, but this was not always the case. In order to deal with this kind of changes in medical classification, the codes attached to concepts are independent of hierarchical position or other contexts, i.e. the codes are unique and non-significant (31;50).

Second, removing concepts from the TS is not permitted as this can disrupt analysis and interpretation of historical data (31;50). It is mandatory to retain these concepts in the terminology content and mark them obsolete for, for instance, retrieval purposes. In case of removal of a concept (which is not in accordance with the criteria), the code assigned to that concept is not reused so that consistency of data over time is not violated (63).

Finally, within the TS there are no limitations for the number of concepts, hierarchic levels and terms that can be added (31;48).

**Change operations**

The goal of making changes in a TS is to keep the TS up to date, while ascertaining the consistency of the concept model (13). The maintenance team uses a “change model” based on the change policies described above to define permitted “change operations” such as insertion of new concepts (19;68). Among other things, this change model can be used when submitting proposals for changes. Furthermore, the maintainers can use the change model to document the change operations.

In general, a distinction is made between the change operations that do and those that do not affect the hierarchy of the concept model (19;68). Change operations that do not affect
the hierarchy, e.g. changes in concept descriptions such as addition of synonyms or deletion of terms, are relatively easy to implement. Change operations that do affect the hierarchy, e.g. changes in concept definition or addition of concepts, are more complex because they must adhere to the constraints set by the change policy and often also lead to changes in existing concepts. Differences in the TS structure and content may lead to differences in allowable change operations within each TS. In general, for a TS that distinguishes concepts from terms, the change model includes addition, obsolete marking and amendments of the concepts, relationships and terms (12). Additional change operations may be applicable to terminological systems based on their typology and content architecture (for example, if they provide cross mapping or composition rules) (19). The National Cancer Institute (NCI) Thesaurus for example uses a change model, that satisfies the change policies described above, and contains the following change operations (59):

- Create concept, i.e. assign Unique ID number, Concept Code and Concept Name.
- Modify concept, i.e. additions, deletions or changes that do not change the concept’s definition.
- Split concept, i.e. a concept is redefined by partitioning its defining attributes over two concepts, one of which retains the original concept's code, and the second one is newly created during the split action. Ambiguities in the original concept's meaning are clarified by narrowing its definition.
- Merge concepts, i.e. multiple ambiguous concepts are combined into one concept.
- Retire concept, i.e. a concept is marked obsolete: previous taxonomic placement information is retained.

5.3.4 Editing tools

The management of TS maintenance is a complex, resource-intensive, and time-consuming task, but the right set of tools can improve the quality and efficiency of this process. Dedicated software tools are used to maintain a TS (13;24;32;57;69-75). Such tools can provide support in different phases of the maintenance process. In most of the cases these tools are used to create and edit the knowledge within the TS. Editing capabilities include standard word processing features, e.g. the ability to add, modify, and delete characters, words, and lines and to easily navigate between concepts and relations (48;49). Furthermore, the tools can be used for consistency control, e.g. to identify duplicates, inappropriate coding, errors in relationships, and inconsistencies after the modification of concepts (48;49). Finally tools might also provide functionalities to support the management of the maintenance process such as managing access rights (security) and collection of proposals (58;71;73).
5.4 Exploration of Current Practice

Thirty-seven of the 93 (40%) invited organizations filled in the questionnaire. The list of respondents is given in appendix 5.1.

The median number of concepts within these TSs is 15,500. The variation in number of concepts is large (200-1000,000 concepts). To make the results comparable and generalizable, the TSs are divided into four quartiles based on their number of concepts. Quartile I includes 10 TSs containing less than 3,950 concepts. Quartile II to quartile IV each include 9 TSs containing respectively 3,950 to 15,500 concepts, 15,500 to 46,155 concepts and more than 46,155 concepts. The number of concepts listed in here includes both active and retired concepts.

Table 5.3 and Table 5.4 present the results of the survey for each criterion. Figure 5.3 gives an overview of the number of criteria that are satisfied by the maintenance processes of the participating organizations. Overall, considerable differences exist between the TSs. Some organizations satisfy most of the criteria (e.g. number 3) while others fail almost all criteria (e.g. number 6). More specifically, there are considerable differences between the large TSs in quartile IV (satisfying on average 20 criteria) and the smaller ones in quartiles I (satisfying on average 14 criteria), II (satisfying on average 16 criteria) and III (satisfying on average 14 criteria).

Furthermore, the variation in the number of criteria satisfied by the TSs in quartile IV is less than the variation observed in the other quartiles. Analysis of the complementary free text remarks did not reveal noteworthy information.

Table 5.3 The results of the survey for the primary component “Execution”. For each criterion, the related question number in the questionnaire is provided. In the results part, the numbers provide the percentages of the participating organizations within each quartile satisfying the criterion.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Question number</th>
<th>Results of questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quartiles</td>
<td>I</td>
</tr>
<tr>
<td>Collecting proposals</td>
<td>3.2</td>
<td>50</td>
</tr>
<tr>
<td>1. Proposals for changes in the TS are standardized in written forms, containing information on:</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>• Proposer ID.</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>• Version number.</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>• Concept information.</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>• Change type.</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>• Change definition.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. Proposals for changes in the TS are validated on:
   - Necessity.
   - Possibility to adopt a change.

3. In case of uncertainty, the acceptance of proposals is based on group consensus.

4. In case a proposal is rejected, feedback is given to the proposer.

5. In case a proposal is accepted, changes made in the TS are verified by the maintenance team, on:
   - Completeness.
   - Textual correctness.
   - Consistency of hierarchic relations.
   - Consistency of mappings to other TSs.
   - Consistency of mappings to other languages.

6. After the implementation of the changes into the TS feedback is given to the proposer to:
   - Inform the proposer.
   - Ask them to verify the changes.

7. Documentation is structured and standardized.

8. Proposals for changes are documented, with:
   - Proposal Date.
   - Proposer ID.
   - Concept ID.
   - Change type.
   - Acceptance method.
   - Acceptation/rejection date.

9. Changes made in the TS are documented, with:
   - Concept information.
   - Implementation date.
   - Change type.
   - Editor ID.
   - Version ID.

10. New versions of the TS have a unique identification number, including the publication date and are distributed as:
   - Complete new version.
   - Incremental update.

11. On average, twice a year a new release of the system is launched (depending on the type of TS).

*: Consists of 10 TSs; #: Consists of 9 TSs
Figure 5.3 An overview of the number of criteria that are satisfied by the maintenance process of the participating organizations within each quartile. In total 31 criteria were investigated.

Figure 5.4 An overview of the number of additional criteria that are desired by the participating organizations within each quartile.
Table 5.4 The results of the survey for the three secondary components. For each criterion, the related question number in the questionnaire is provided. Unless stated otherwise, in the results part, the numbers provide the percentages of the participating organizations within each quartile satisfying the criterion.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Question number</th>
<th>Results of questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quartiles</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I*</td>
<td>II#</td>
</tr>
<tr>
<td>Process Management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. There is a team responsible for the management of the maintenance process.</td>
<td>1.1</td>
<td>100</td>
</tr>
<tr>
<td>2. The maintenance team is easily accessible.</td>
<td>1.2</td>
<td>30</td>
</tr>
<tr>
<td>3. The response time of the maintenance team for proposals and questions is short (Mean days (± SD)).</td>
<td>1.3</td>
<td>26.8 (52.1)</td>
</tr>
<tr>
<td>Coordination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Different disciplines participate in the maintenance team:</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>- Users/domain expert.</td>
<td>40.0</td>
<td>40</td>
</tr>
<tr>
<td>- Terminology expert.</td>
<td>30</td>
<td>56</td>
</tr>
<tr>
<td>- Software of Engineers.</td>
<td>30</td>
<td>33</td>
</tr>
<tr>
<td>- Coordinators.</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>5. Only qualified people are able to make changes.</td>
<td>1.4</td>
<td>60</td>
</tr>
<tr>
<td>6. Access to the TS is secured, using:</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>- Identification.</td>
<td>30</td>
<td>67</td>
</tr>
<tr>
<td>- Authentication.</td>
<td>30</td>
<td>67</td>
</tr>
<tr>
<td>- Authorization.</td>
<td>30</td>
<td>67</td>
</tr>
<tr>
<td>Persons Involved</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. The codes that are assigned to concepts are context free:</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>- Random.</td>
<td>30</td>
<td>67</td>
</tr>
<tr>
<td>- Hierarchy related.</td>
<td>30</td>
<td>67</td>
</tr>
<tr>
<td>8. Concepts that are no longer in use, are not removed from the TS. Instead, these are marked obsolete.</td>
<td>2.3</td>
<td>30</td>
</tr>
<tr>
<td>9. In case of removal of a concept (which is not in accordance with the criteria), the code assigned to that concept is not reused.</td>
<td>2.4</td>
<td>30</td>
</tr>
<tr>
<td>10. Within the TS there are no limitations regarding the number of concepts, hierarchic levels and terms that can be added.</td>
<td>2.5</td>
<td>60</td>
</tr>
<tr>
<td>Change specifications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. There is a ‘change model’ that specifies all types of changes (e.g. delete, add, adjust, etc.) that can occur in the TS in concordance with the change policies.</td>
<td>2.2</td>
<td>50</td>
</tr>
</tbody>
</table>

Framework for Maintenance of Medical Terminological Systems
Chapter 5

<table>
<thead>
<tr>
<th>Editing tool</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. The maintenance team uses an editing tool for the maintenance process.</td>
<td>4.1</td>
</tr>
<tr>
<td>13. The editing tool is secured with login name and password.</td>
<td>4.3</td>
</tr>
<tr>
<td>14. The editing tool contains a module for collecting the proposals.</td>
<td>4.2</td>
</tr>
<tr>
<td>15. The editing tool contains a module to enable the consensus process.</td>
<td>4.2</td>
</tr>
<tr>
<td>16. The editing tool supports the input of changes into the TS.</td>
<td>4.2</td>
</tr>
<tr>
<td>17. The editing tool supports automatic validation controls.</td>
<td>4.3</td>
</tr>
<tr>
<td>18. The editing tool generates reports for documentation.</td>
<td>4.2</td>
</tr>
<tr>
<td>19. The editing tool supports managing different versions of TS.</td>
<td>4.2</td>
</tr>
<tr>
<td>20. The editing tool supports distribution of new versions.</td>
<td>4.2</td>
</tr>
</tbody>
</table>

*: Consists of 10 TSs; #: Consists of 9 TSs

5.5 Exploration of the desired situation

Due to space limitations, the desired situation as indicated by the respondents is not extensively described in the tables. Figure 5.4 gives an overview of the number of additional criteria that the participating organizations wish to fulfill for their maintenance processes after being confronted with the framework. As shown in Figure 5.4, the respondents do not wish to make large changes in their maintenance process to meet more criteria. In most of the cases, the desired changes concerned the functions of the editing tools. This is mainly the case for TSs in the fourth quartile. Furthermore, the respondents of almost all TSs wish to improve the efficiency of collecting proposals through the use of standard forms and by involving more disciplines in this process. Moreover, particularly for TSs in the first quartile, the respondents wish to reduce their response time for processing the proposals.

5.6 Discussion

Without a well-structured and standardized maintenance process, TSs cannot provide the quality required by today’s medical applications. However, although the need for standardization of the TS maintenance process is recognized in literature, few publications
are available on this topic. In this study we contribute to this issue by enumerating the necessary features for the management of the maintenance process and by putting them as criteria into a framework. This framework can be used as a reference or guideline to design, evaluate and improve the maintenance process of a TS. Also the current TS maintenance processes were studied by means of a survey based on this framework.

5.6.1 Literature review
Available literature on the maintenance of medical TSs mainly focuses on technical aspects of the maintenance process and to our knowledge no other attempts have been made to standardize the organizational issues around the maintenance process. As far as we know, this study is the only work that provides an extensive survey of the maintenance process of medical TSs.

Information that was considered suitable and relevant for our study mainly came from publications outside the medical domain. Although many publications, guidelines and standards on maintenance of software systems are available, these mainly focus on the technical aspects of the maintenance process.

In general, the maintenance processes as used for information systems are also applicable to medical TSs (14, 63; 76-78). However, medical TSs distinguish themselves from these information systems by their dynamic character and their complex domain (12). Hence, when considering the maintenance of the medical TSs, timeliness is an important issue. This is reflected by the criterion of a short response time (i.e. time needed to handle a proposal or a question) of the maintenance team with respect to proposals and questions. Still, in the existing literature, no concrete indication is provided for what the response time should be. In our study, the mean response time of the maintenance teams varies between 26.8 days for quartile I, and 6.7 days for quartile III. Almost all respondents wish to reduce their response time for processing the proposals. This is particularly the case for TSs in the first quartile.

5.6.2 Current Practice
The survey revealed the diversity in and the limitations of the maintenance processes and thereby emphasized the importance of the development and application of the suggested framework. Especially for small TSs with a limited number of concepts, many criteria are not met. This can be attributed to the fact that the maintainers of larger TSs are better equipped to extensively organize their maintenance processes. Furthermore, large TSs are generally more complex and thus necessitate extensive management of the maintenance processes. Accordingly, the results show that a majority of the large TSs (i.e. TSs in quartile IV) fulfil most of the criteria.
Chapter 5

The criterion of using a “change model” forms an exception since the maintainers of smaller TSs more often use a “change model” than the maintainers of the large TSs. However, a closer look at the component “change specifications” shows that the larger TSs satisfy more of the remaining criteria in this component and have extensive change policies. So we believe that these large TSs implicitly use a change model to describe the change operations based on the change policies, but may not recognize it as such.

5.6.3 Desired situation

It is notable that many of the respondents of the survey, even when their maintenance process was incomplete and was not compliant with most of the criteria, did not wish to redesign their maintenance process after being confronted with the framework. Most of the wanted changes regard the functions of the editing tools and are especially mentioned by the respondents of the larger TSs (in the fourth quartile). Improving functions of an editing tool requires a one-time effort and will increase the efficiency of the maintenance process and decrease the work pressure on the maintainers.

TSs differ and the maintainers do not always have the need or resources to organize the maintenance process as completely as possible. For a large TS with a large number of users it is advisable to design a maintenance process that satisfies most of the criteria. However, for a small TS that is used by a limited number of users, some of the criteria may be superfluous and even overexpensive. For instance, while it is essential that the core activities of the component “Execution” maximally adhere to the criteria, an extensive Editing tool to support collection of proposals is not always necessary. The framework presented here can be applied to a TS in conformity with the needs and the possibilities of its maintainers. The maintainers can use the framework as a guideline to prioritize the processes that have to be implemented and to decide which criteria should be fulfilled given their own possibilities and needs.

5.6.4 Limitation of the study

To gain insight into the current maintenance practice of existing TSs and to evaluate whether the introduction of the framework triggers the wish for maintenance process redesign, we used self-reporting instead of more objective (observational) measurements. The criteria cover rather specific topics that are in general only answerable by the maintainers of a TS and therefore only the maintainers could provide complete answers to related set of questions. For our study, especially because we aimed to include a large number of TSs, it was not doable to use objective observation methods to get the answers. Furthermore, we did not validate the questionnaire in terms of reproducibility through its application to a given TS by several independent judges as it was not possible find several
independent judges who were all sufficiently familiar with the same TS to answer the specific questions.

Finally, the framework and the questionnaire do not deal with emerging practices for collaborative maintenance such as semantic wikis. In semantic wikis, wiki technology is used as a TS maintenance environment, reducing entry barriers for its maintainers (79). The idea of a semantic wiki is that qualified users can directly add new concepts to the TS, or refine or modify existing ones while taking the semantic structure of the TS into account (80). The model of such a maintenance process is characterized as centralized or semi-centralized (23). Some organizations have started to adopt these techniques for the maintenance of their terminological systems (81;82). Although semantic wiki systems are becoming more and more popular as tools for content and knowledge management, the use of semantic wikis for TS maintenance is not well described in the literature on TS maintenance. The lack of specific literature and the fact that the framework prescribes the use of centralized maintenance model are the reasons that the use of semantic wikis as potential editing tools for TS maintenance was not specifically included in the framework or in the questionnaire. Future work should take a closer look at these techniques and their usefulness for TS maintenance.

Conclusions

The framework developed within this study summarizes the principal notions that are important for the management of the maintenance process of medical TSs. It is applicable to all kinds of medical TSs and provides their maintainers with the criteria for a well-organized maintenance process.

The survey showed that larger TSs fulfil most of the criteria mentioned in the framework whereas smaller TSs fail more criteria, probably as this incurs costs and overhead that are too high and/or unnecessary.

Acknowledgements

The authors thank the NCVHS for providing the results of their questionnaire. We express our gratitude to the respondents of our questionnaire.

Bibliography

32. Oliver DD. Change Management and Synchronization of Local and Shared Versions of a Controlled Medical Vocabulary Stanford; 2000.


74. Elhanan G, Cimino JJ. Controlled vocabulary and design of laboratory results displays. Proc AMIA Annu Fall Symp 1997;719-23.