

Published in final edited form as:

Artif Intell Med. 2009 May ; 46(1): 55–66. doi:10.1016/j.artmed.2008.08.009.

Versioning computer-interpretable guidelines: Semi-automatic modeling of ‘Living Guidelines’ using an information extraction method★

Katharina Kaiser^{a,*} and Silvia Miksch^b

^aVienna University of Technology, Institute of Software Technology & Interactive Systems, Favoritenstrasse 9-11/188, A-1040 Vienna, Austria

^bDanube University Krems, Department of Information & Knowledge Engineering, Dr. Karl-Dorrek-Strasse 30, A-3500 Krems, Austria

Summary

Objective—Clinical practice guidelines (CPGs) are means to provide evidence-based medical knowledge. In order to make up-to-date “best” scientific evidence available these documents need to be updated on an ongoing basis. An effective method to accomplish this aim is offered by the so-called “*living guidelines*”: *Living guidelines* are documents presenting up-to-date and state-of-the-art knowledge to practitioners. To have guidelines implemented by computer-support they have to be formalized in a computer-interpretable form in a first step. Due to the complexity of such formats the formalization process is burdensome and time-consuming. Automating parts of the modeling process and, consequently, modeling updates of these guideline documents are demanded.

Methods and material—The LASSIE methodology supports this task by formalizing guidelines in several steps from the textual form to the guideline representation language Asbru using a document-centric approach. LASSIE uses information extraction techniques to semi-automatically accomplish these steps. We apply LASSIE to support the implementation of living guidelines.

Results—Based on a living guideline published by the Scottish Intercollegiate Guidelines Network (SIGN) we show that adaptations of previously formalized guidelines can be accomplished easily and fast. Thereby, the different versions of guideline documents are compared and updates are identified. Due to the traceable formalization method of linking text parts and their corresponding formal models, we are able to inherit unchanged models from previously formalized versions. Thus, we only need to formalize updated text parts using the semi-automatic formalization method LASSIE.

Conclusion—We propose a simple, time-saving, but effective method called LASSIE to formalize new guideline versions of previously formalized CPGs. Furthermore, models that have been added or modified by knowledge engineers in previous versions can also be transferred easily. This will result in a faster implementation of new guideline versions also known as *living guidelines* to provide up-to-date knowledge necessary for accomplishing the daily work of health care professionals.

★This is an extended and revised version of K. Kaiser and S. Miksch. Formalizing ‘Living Guidelines’ using LASSIE: A Multi-step Information Extraction Method, Proceedings of the 11th Conference on Artificial Intelligence in Medicine (AIME), 2007.

© 2008 Elsevier B.V. All rights reserved.

*Corresponding author. Tel.: +43 1 58801 18839; fax: +43 1 58801 18899. .
kaiser@ifs.tuwien.ac.at (K. Kaiser), miksch@ifs.tuwien.ac.at (S. Miksch).

Keywords

Information extraction; Clinical practice guidelines; Living guidelines; Computer-interpretable guidelines; Guideline representation; Treatment processes; Asthma

1. Introduction

Clinical practice guidelines (CPGs) are important instruments for standardization and uniform improvement of quality in medical care and are defined as “systematically developed statements to assist practitioners and patient decisions about appropriate healthcare for specific circumstances” [1].

According to [2] the five initial steps of the development cycle of a CPG are (1) identifying and refining the topic of the CPG, (2) convening and running guideline development groups, (3) assessing the evidence about clinical questions or conditions on the basis of systematic reviews, (4) translating the evidence into a recommendation within the CPG, and (5) external review of the guideline. This development process takes at least two years. Thus, CPGs can be out of date as soon as they are produced, since new research findings are continuously published. To overcome this problem sometimes the shelf life of a guideline is identified; either by a date (e.g., this guideline will be reviewed in 2 years) or by a statement that the review date will be determined by the availability of new evidence (e.g., this guideline will be considered for review as new evidence becomes available). Alternatively, we can consider a new option—the *living guideline*. A living guideline is one that remains under review on an ongoing basis, with updates published at set intervals (e.g., annually).

The review of the guideline (i.e., a new article in the specified field is available) may have various characteristics. On the one hand it can add additional evidence and thus alter the evidence level of a recommendation. On the other hand it can lead to a new recommendation or it may change an existing one. However, in the majority of cases only small text parts are changed; often only the reference to the new article is added or the reference to an obsolete article is removed.

Modeling CPGs in a computer-interpretable form is a prerequisite for various computer applications to support their application. However, transforming guidelines into a formal guideline representation is a difficult task. In [3,4] we have proposed a semi-automatic methodology called LASSIE to model treatment processes in multiple steps using information extraction (IE).

We will now show that we can use LASSIE to support the formalization of living guidelines. Applying this method, which traces both the general formalization steps and the changes to new versions has the potential to reduce the modeling effort. The Scottish Intercollegiate Guidelines Network (SIGN)¹ has already published a living guideline [5]. Based on the documents provided we will show that adaptations of formalized guidelines can be accomplished easily and fast.

In the next section we will discuss some work on guideline formalization tools and guideline versioning methods. Afterwards, we will give a short introduction in LASSIE. In Section 4 we describe the adaptation of LASSIE for supporting living guidelines followed by a case study illustrating our methodology. In Section 6 we discuss the applicability and usability of our methodology. Section 7 summarizes our work and represents our conclusions.

¹<http://www.sign.ac.uk> (Accessed on February 15, 2008).

2. Related work

In this section, we present relevant work describing guideline formalization tools and approaches for guideline versioning.

For formalizing clinical guidelines into a guideline representation language (see [6,7] for an overview and comparison) various tools exist. We can classify such tools in document-centric and model-centric tools.

2.1. Document-centric approaches

Markup-based tools utilize a document-centric approach. Thereby, the original guideline document is systematically marked-up by the user in order to generate a semi-formal model of the marked text part.

The *GEM Cutter* [8] was one of the first exponents of this approach transforming guideline information into the GEM format [9]. *Stepper* [10] is a tool that formalizes the initial text in multiple user-definable steps corresponding to interactive XML transformations. The *Document Exploration and Linking Tool/Addons (DELT/A)* [11] supports the translation of HTML documents into any XML language. It uses links between the text part in the original document and its corresponding XML model (see Fig. 1). To generate a specific model user-definable *macros* can be used. *Uruz*, part of the *Digital Electronic Guideline Library (Degel)* framework [12], is a web-based markup tool that supports indexing and markup using any hierarchical guideline-representation format. It enables the user to embed in the guideline document terms originating from standard vocabularies.

2.2. Model-centric approaches

In model-centric approaches a conceptual model is formulated by domain experts. The relationship between the model and the original document is only indirect.

AsbruView [13] uses graphical metaphors to represent Asbru plans [14]. *AREZZO* and *TALLIS* [15] support the translation into PROforma [16] using graphical symbols representing the different types of tasks of the language. *Protégé* [17] is a knowledge acquisition tool that supports the translation into guideline representation languages EON [18], GLIF [19], and PROforma [16]. It uses specific ontologies for these languages, whereas parts of the formalization process can be accomplished with predefined graphical symbols. *AREZZO*, *TALLIS*, and *Protégé* offer a flowchart-based representation of the processes.

2.3. Guideline versioning

Unfortunately, guideline versioning has not been adequately addressed by now. There are two approaches dealing with versioning.

Peleg and Kantor [20] propose a model-centric approach for GLIF [19]. Thereby, the underlying GLIF ontology is extended by version information. A versioning tool was developed that supports the creation of a new CPG model or the modification of an existing one as well as the displaying of versions of a CPG model, highlighting the differences.

Seyfang et al. [21] describe the formalization of ‘living guidelines’ using a document-centric approach. They start with an HTML version of the guideline and use different intermediate representations to derive a formal model of the guideline. The first intermediate representation is the *Many-Headed Bridge (MHB)* [22] that is used to bridge the gap between informal representations, such as free text and tables, and more formal representations, such as Asbru, GLIF, or PROforma. The DELT/A tool [11] is used to mark-

up text chunks and generate the corresponding model in the intermediate representation (see also Fig. 1). The original marked-up guideline document is then manually updated to the new version by highlighting both newly added and removed text fragments. Using the DELT/A tool the highlighted text fragments are selected to visualize the corresponding MHB chunk in order to make the necessary changes.

Besides dealing with changes during the development process of guidelines, the following approaches focus on changes during the adaptation process.

Peleg et al. [23] developed a model-based approach for adapting guidelines to local circumstances and encoding them using the GLIF3 formalism [19]. Thereby, at first, the original guideline is encoded in GLIF3 using the Protégé knowledge acquisition tool. In a second step, definitions of concepts are clarified, data items of the initial encoding are matched to the data entities in the local electronic medical record (EMR), and local practice is analyzed. After this, in a third step, the initial encoding is revised that it fits with local practice and it includes formal definitions of data queries and decision criteria that are mapped to the EMR's schema and available data. Finally, the encoding is manually checked by clinicians and knowledge engineers as well as test cases are executed using the GLIF Execution Engine (GLEE) [24]. If problems are encountered during validation, steps 2 and 3 are repeated until validation results are correct.

Terenziani et al. [25] proposed a context-adaptable approach to clinical guidelines. They describe how to extend GLARE² [26] in order to cope with context adaptation. To accomplish a resource-based adaptation, they developed an automatic pre-compilation tool which, taking into account a general clinical guideline expressed in the GLARE formalism and a list of resources (e.g., resources of a given hospital), provides a context-based clinical guideline, in which all non-executable actions (paths) are automatically pruned. In [27] they describe an extension of their work to deal with upgrade adaptation. Thereby, the underlying data model of GLARE is augmented to support versioning.

But still, using the mentioned tools the modeling process is complex and labor intensive. Methods are needed to automate parts of the modeling task.

3. LASSIE—modeling treatment processes using information extraction

In the course of time most of the guideline representation languages have produced a very extensive syntax. They can present a multitude of different information and data. Due to this complexity, the modeling process of a CPG into such a language is a very challenging and burdensome task. As a one-step or even a two-step modeling process was shown to be not sufficient to the modeler [28,22], we apply a multi-step transformation process that facilitates the formalization process by various intermediate representations (IRs) obtained in stepwise procedures.

Our multi-step transformation methodology (cf. Fig. 2), called LASSIE³ (see [3,4]), supports the document-centric approach by marking the original guideline document and generating the particular models for each marked text part. It is intended to be a semi-automatic approach. This enables the user not only to correct the transformations, but also to augment them by implicit knowledge necessary for a subsequent execution. After each step the user is able to view the results using the DELT/A tool [11] (see also Fig. 1).

²Guideline Acquisition, Representation and Execution.

³modeLing treAtment proceSSes using Information Extraction.

We can identify the following benefits of the multi-step approach and in the following of the IRs:

1. IRs support a concise formalization process.
2. IRs provide different formats as well as separate views and procedures for various kinds of information.
3. We can apply specific heuristics for each particular kind of information.
4. A simpler and more concise evaluation and tracing of each process step is accomplishable.

IRs are specific templates used by IE methods to present the desired information. The IE methods use a terminology based on the Medical Subject Headings (MeSH)⁴ [29] and manually generated extraction patterns (see Listing 1 for example patterns).

CPGs present effective treatment processes. One challenge when authoring CPGs is the detection of individual processes and their relations and dependencies. We can generate simple representations of treatment instructions (i.e., actions), which are independent from the final guideline representation language. Based on this independent representation we can transform the information in further steps into the guideline languages. To extract processes from CPGs we proceed in several steps which serve to filter segments of text containing treatment instructions from the documents and to generate processes. Fig. 2 represents our approach whereas the first two steps result in representations independent of the subsequent guideline representation language.

The first step is to extract the relevant sentences containing treatment instructions by marking-up the original guideline document (see Fig. 3(a)). The subsequent step is to combine several sentences to one action and to structure the actions and detect relations among them (see Fig. 3(b)). These two steps provide a basis for the subsequent transformation of the process information into several guideline representation languages. In [3,4] we have demonstrated that it is possible to formalize processes using IE for modeling guidelines in the Asbru format.

4. Adaptation of LASSIE for 'Living Guidelines'

Using LASSIE, a unique identifier (i.e., the DELT/A link) marks information transformed from one step to the next. We now apply LASSIE to support the formalization of living guidelines. The documents provide us the information that has changed: Adaptations (i.e., modifications in a sentence and newly added sentences) of every new revision are marked by arrows and highlighted in terms of color (or in different gray scales) (see Fig. 4). Sentences that were deleted in the new version are not represented or marked in the new guideline document.

The coloring results from the tagging of the guideline document both on a semantic level (e.g., by using `< span class= " update2005" >` to indicate an update of the 2005 version of the guideline) and on a representation level (i.e., the definition of the 'update2005' style using different background coloring, etc.). Both tagging and definition are accomplished by the guideline developing organization. Thus, the number of guideline versions is not restricted. Each version identifier has to be unique and the particular coloring should be visually distinguishable.

⁴<http://www.nlm.nih.gov/mesh/> (Accessed on February 15, 2008).

We now propose a new method utilizing this information. Thereby, the new guideline is not going to be modeled from scratch, but already modeled parts from previous versions are inherited. Thus, only new text parts have to be modeled (see Fig. 5).

As LASSIE is a multi-step methodology, we have to satisfy each step for the living guideline's formalization.

4.1. Pre-processing

As the input of LASSIE's first step is the XHTML-conform guideline document, we have to preprocess the document to receive a unified document format. We accomplish this by XSLT scripts, HTML Tidy,⁵ and manual post-processing in order to obtain not only a well-formed but also a hierarchically well-structured XHTML document.

We have built an extensible model for representing text documents containing sections, subsections, and so on, their headings, paragraphs, lists, tables, and figures or rather their links. The various taggings for marking text of different revisions are also preserved in this model.

4.2. Marking-up the new guideline version

LASSIE's first step is to detect relevant sentences and text parts in the guideline document. Text parts are thereby list entries that may not be complete sentences, but are referred to as sentences in the remaining article.

The output of LASSIE's first step are two files (see also Fig. 3(a)):

1. The marked-up guideline document, where relevant sentences are marked and tagged by a DELT/A link, and
2. A file containing all relevant sentences and their corresponding DELT/A links.

We use these files of the previous guideline version to detect unchanged relevant sentences in the new guideline version. We proceed by parsing the new guideline document and search for each sentence marked-up in the previous version. Thereby, we have to consider not only equal sentences but also equal contexts of them. This is necessary as a marked-up sentence can appear repeatedly in the document and we have to assign the correct DELT/A link in the new document. Thus, a sentence is defined by its section and paragraph. For sentences that remain unchanged in the new version and that were marked-up in the previous version the markup information (i.e., the DELT/A link id) is transferred. For each sentence in the new guideline that is marked as updated as a part or whole we apply the first step of the LASSIE methodology (see Fig. 3(a) and [4] for details) in order to detect relevant sentences for further processing. Relevant sentences of the old guideline version that are not found in the new document can be seen as removed.

For each sentence of the new guideline that has been marked as relevant by the LASSIE methodology we assign also a version id. Furthermore, we have to be aware to not assign an obsolete DELT/A link to a new sentence.

Thus, we obtain the new marked-up guideline version and are now able to extract the processes in order to gain a representation independent of the final guideline language. After this step, the user is also able to view the resulting files with the DELT/A tool and make corrections.

⁵<http://tidy.sourceforge.net> (Accessed on February 15, 2008).

4.3. Further transformation of the extracted information

After obtaining the new marked-up guideline document we can proceed with the subsequent steps coming up with LASSIE. That means, we can inherit models of subsequent representations that correspond with text parts that were not changed in the new guideline version. For new or changed models the particular processing step of LASSIE is applied. For instance, to detect processes we proceed as following.

Within the next step of the LASSIE methodology relevant sentences are structured and relationships between sentences are found (see Fig. 3(b)). The output of this step is a representation (*ActionIR*) containing actions, relations controlling the process flow between these actions, and the structure illustrating the hierarchy and nesting of groups of actions.

An *action* contains the action sentence, possible assigned annotation sentences, treatment instruments, information about the dosage, duration or iteration of a drug administration, and conditions. If the action is part of a selection (i.e., at least one action has to be chosen among several), it is given a selection id. DELT/A links are inherited from the *SentenceIR* representation in order to provide the traceability of the modeling process.

In order to obtain actions from our new version of the marked-up guideline, we can inherit action and annotation sentences from the previous *ActionIR* version. Furthermore, new relevant sentences of the current guideline version are classified in action and annotation sentences. When sentences are classified as annotation they must be assigned an *action* sentence. If an action and its assigned annotation sentences were not changed in the new version, the complete action node is inherited to the new *ActionIR* representation. Otherwise, the action node and its additional information has to be generated by LASSIE. Additionally, a version id is assigned for these new nodes. Likewise, we are able to inherit relations between action nodes if neither of the two action nodes has changed. Otherwise, we have to detect new relations using LASSIE. The third part of the *ActionIR* representation, the structure of the actions, is then generated by LASSIE.

The output of this step (see Fig. 6) is then a new version of the *ActionIR* representation, which can be viewed with the DELT/A tool. Changed information is identifiable by the version id. The user may then make corrections or add new information to the representation.

5. Case study

We tested the applicability of our method to a real *living guideline*. Based on the *British guideline on the management of asthma* [5] developed by SIGN in its version of 2005, we generated the previous guideline versions (i.e., from 2004) due to the non-availability of the old documents.⁶ This was possible because SIGN offers a document which clearly describes every adaptation (i.e., modification, adding, removal) of the text. For evaluating the method, we only used Section 4 (*Pharmacological Management*) of the guideline. It describes an important part of the asthma treatment and contains also updated text parts.

5.1. Formalizing the original guideline version

We preprocessed the old guideline document to comply our unified document format. Starting with the old guideline document, we used LASSIE to generate the particular models necessary for formalization in Asbru. We automatically generated the intermediate

⁶We were not able to receive the older guideline versions from SIGN.

representations and adapted them according to our needs. The document consists of 509 sentences. 139 of them were classified as relevant for further processing (see Fig. 7).

5.2. Formalizing the new guideline version

The next step was to model the new guideline version using our new method. Therefore, we prototypically implemented our method to automate this task and adapted our implementation of LASSIE to enable the processing of *living guidelines*.

5.2.1. Preprocessing—We preprocessed the new guideline document in order to gain a unified document format complying the XHTML format.

5.2.2. Markup of new guideline version—Afterwards, we automatically searched for unchanged sentences that were marked in the previous guideline version and added the corresponding DELT/A links into the current document. Now, we were able to have LASSIE check the adapted sentences for relevancy. The new version of Section 4 consists of 515 sentences. We were able to inherit 133 sentences of the old version, which means that six relevant sentences were either changed or removed in the new version. 13 updated or new sentences were found and checked with LASSIE, which classified 10 as relevant (see also Table 1). The new relevant sentences were marked and assigned a new DELT/A link as well as a version id.

5.2.3. Action generation and further transformations—Before proceeding with the next step, we had to cross-check the new added or modified sentences. We were able to verify the relevance of 10 new sentences classified by LASSIE. Furthermore, we also classified the remaining three sentences as not relevant for our tasks.

The new sentences were then classified in action or annotation sentences. The latter were then assigned to action sentences. We received five action sentences and five annotation sentences. Four of the annotation sentences were assigned to two previously available action sentences (see Fig. 8 for an example); one to a new action sentence. Thus, the remaining unchanged action models were inherited from the previous version.

Again, this step is cross-checked by a human. The inspections can be accomplished fast, as, on the one side, added and modified sentences (or clauses) are highlighted in the original guideline document and, on the other side, models in intermediate representations of deleted sentences can be easily detected by DELT/A's functionality.

The same procedure is done for all subsequent steps in an analogous manner.

6. Discussion

Our study shows that using a document-centric approach — the LASSIE methodology in combination with the DELT/A tool — offers distinct benefits in modeling *living guidelines*. By using automatic methods for comparing and matching the different guideline versions as well as computerizing to CIGs, a fast adaptation of the new document is possible. As in *living guidelines* there will not be radical changes from one version to the succeeding version, inheriting of previous models is a simple, time-saving, but effective method for modeling computer-supported guidelines. Models that have been added or modified by knowledge engineers in previous versions can also be transferred easily. Furthermore, in the intermediate representations the new models are marked by their version ids to enable a prompt identification. Thus, the user is able to perform adaptations quickly and conveniently.

Unfortunately, we are not able to provide detailed data about the time a typical adaptation would cost. Guideline modeling is a very complex and cumbersome task and the modeling effort extremely depends on the skills as well as on the training of the guideline modeler. Nevertheless, inheriting unchanged models and only modeling adaptations should cause a decrease of modeling effort, especially as there are only minor modifications from one version of a living guideline to the next.

These fast adaptations would also support the integration of computer-interpretable guidelines (CIGs) in patient data management systems (PDMSs). Having the possibility of relatively fast adaptations of CIGs, on the one hand, the dissemination of CIGs can be fostered and, on the other hand, the acceptance of CPGs as well as of CIGs can be enhanced, as health care is geared to up-to-date medical evidence consequently.

A limitation of our methodology is that minor changes (e.g., adding or removing a new reference, or changes like “Nedocromil sodium is of some benefit ...” to “Nedocromil sodium is *also* of some benefit ...”) in the text may result in applying a new relevance check, sentence classification, action generation, and so on, which will require an evaluation by a human afterwards. In methods described in Section 2.3 such minor changes may be checked and accomplished by a human user more efficiently and therefore saves a lot of time and cumbersome work.

Furthermore, we have to mention that the IRs do not contain the models of all versions, only the actual ones. Thus, it is not possible to have one file for all versions, but one file for each version of a representation.

In future versions we plan to integrate terminology support, in order to control changes of expressions (e.g., the name of a disease was changed) and — in case of the same underlying concept — to consider this rather non-change in our methodology (cf. ‘Previous Indexing’ in the MeSH thesaurus). This will support the integration of CIGs into the PDMS as well as the compatibility and comparability between different CPGs’ versions. Therefore, the analysis of past CPGs, their usefulness, and further comparative medical studies can be easily performed. This could have an impact on the evaluation of the CPGs’ versions and ease the development of improved CPGs in the future.

7. Conclusion

Living guidelines are documents presenting up-to-date and state-of-the-art knowledge to practitioners. To support their application they have to be brought in a computer-interpretable form, which is a difficult task.

We proposed a method applicable on documents previously being formalized using a document-centric approach. Thereby, the guideline document is marked-up and corresponding formal models are generated. Our method utilizes these links between the textual document and the formal models. It inherits formalized models of the previous guideline version by re-linking them to their corresponding text parts in the new guideline version. Only changed or added texts have to be analyzed and modeled. The formalization task is thereby accomplished using the LASSIE methodology. It is a semi-automatic approach using IE and various intermediate representations to model different kinds of information in various granularities. Our case study showed that the modeling effort can be reduced considerably by applying our LASSIE methodology.

By re-using previously formalized models of guidelines we are able to quickly and effectively formalize new guideline versions.

Acknowledgments

This work is supported by "Fonds zur Förderung der wissenschaftlichen Forschung FWF" (Austrian Science Fund), grant L290-N04.

References

- [1]. Field, MJ.; Lohr, KN., editors. Clinical practice guidelines: directions for a new program. National Academies Press, Institute of Medicine; Washington, DC: 1990.
- [2]. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. *BMJ*. 1999; 318:593–6. [PubMed: 10037645]
- [3]. Kaiser K, Akkaya C, Miksch S. How can information extraction ease formalizing treatment processes in clinical practice guidelines? A method and its evaluation. *Artificial Intelligence in Medicine*. 2007; 39(2):151–63. [PubMed: 16962747]
- [4]. Kaiser, K.; Miksch, S. Modeling treatment processes using information extraction. In: Yoshida, H.; Jain, A.; Ichalkaranje, A.; Jain, LC.; Ichalkaranje, N., editors. *Advanced computational intelligence paradigms in healthcare*. 1. Vol. vol. 48 of *Studies in Computational Intelligence (SCI)*. Springer Verlag; Berlin, Heidelberg: 2007. p. 189-224.
- [5]. Scottish Intercollegiate Guidelines Network (SIGN). Scottish Intercollegiate Guidelines Network (SIGN). British Thoracic Society; Nov. 2005 British guideline on the management of asthma. A clinical national guideline.
- [6]. Peleg M, Tu SW, Bury J, Ciccarese P, Fox J, Greenes RA. Comparing computer-interpretable guideline models: a case-study approach. *Journal of the American Medical Informatics Association (JAMIA)*. 2003; 10(1):52–68. [PubMed: 12509357]
- [7]. de Clercq PA, Blom JA, Korsten HHM, Hasman A. Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artificial Intelligence in Medicine*. 2004; 31(1):1–27. [PubMed: 15182844]
- [8]. Polvani, K-A.; Agrawal, A.; Karras, B.; Deshpande, A.; Shiffman, R. GEM cutter manual. Yale Center for Medical Informatics; New Haven, CT: 2000.
- [9]. Shiffman RN, Karras BT, Agrawal A, Chen R, Marengo L, Nath S. GEM: a proposal for a more comprehensive guideline document model using XML. *Journal of the American Medical Informatics Association (JAMIA)*. 2000; 7(5):488–98. [PubMed: 10984468]
- [10]. Ruži ka, M.; Svátek, V. Mark-up based analysis of narrative guidelines with the Stepper tool. In: Kaiser, K.; Miksch, S.; Tu, SW., editors. *Computer-based support for clinical guidelines and protocols; Proceedings of the symposium on computerized guidelines and protocols (CGP 2004); 2004; Prague, Czech Republic: IOS Press; p. 132-6. vol. 101 of studies in health technology and informatics*
- [11]. Votruba, P.; Miksch, S.; Kosara, R. Facilitating knowledge maintenance of clinical guidelines and protocols. In: Fieschi, M.; Coiera, E.; Li, Y-CC., editors. *Proceedings from the Medinfo 2004 world congress on medical informatics, AMIA; 2004; San Francisco, CA: IOS Press; p. 57-61.*
- [12]. Shahar, Y.; Young, O.; Shalom, E.; Mayaffit, A.; Moskovitch, R.; Hessing, A. DEGEL: a hybrid, multiple-ontology framework for specification and retrieval of clinical guidelines. In: Dojat, M.; Keravnou, E.; Barahona, P., editors. *Proceedings of the 9th conference on Artificial Intelligence in Medicine in Europe, AIME 2003; 2003; Protaras, Cyprus: Springer Verlag; p. 122-31.*
- [13]. Kosara R, Miksch S. Metaphors of movement: a visualization and user interface for time-oriented, skeletal plans. *Artificial Intelligence in Medicine*. 2001; 22(2):111–31. [special issue: information visualization in medicine]. [PubMed: 11348843]
- [14]. Shahar Y, Miksch S, Johnson P. The Asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artificial Intelligence in Medicine*. 1998; 14:29–51. [PubMed: 9779882]
- [15]. Steele, R.; Fox, J. Tallis PROforma Primer—introduction to PROforma language and software with worked examples. Advanced Computation Laboratory, Cancer Research; London, UK: 2002. Technical report

- [16]. Fox J, Johns N, Rahmanzadeh A. Disseminating medical knowledge: the PROforma approach. *Artificial Intelligence in Medicine*. 1998; 14:157–81. [PubMed: 9779888]
- [17]. Gennari JH, Musen MA, Fergerson RW, Grosso WE, Crubézy M, Eriksson H. The evolution of Protégé: an environment for knowledge-based systems development. *International Journal of Human Computer Studies*. 2003; 58(1):89–123.
- [18]. Musen MA, Tu SW, Das AK, Shahar Y. EON: a component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Informatics Association (JAMIA)*. 1996; 3(6):367–88. [PubMed: 8930854]
- [19]. Boxwala AA, Peleg M, Tu SW, Ogunyemi O, Zeng Q, Wang D. GLIF3: a representation format for sharable computer-interpretable clinical practice guidelines. *Journal of Biomedical Informatics*. 2004; 37(3):147–61. [PubMed: 15196480]
- [20]. Peleg, M.; Kantor, R. Approaches for guideline versioning using GLIF. In: Musen, MA., editor. *Proceedings of the 2003 American Medical Informatics Association (AMIA) annual symposium*; 2003; Washington, DC: AMIA; p. 509-13.
- [21]. Seyfang, A.; Martinez-Salvador, B.; Serban, R.; Wittenberg, J.; Miksch, S.; Marcos, M. Maintaining formal models of living guidelines efficiently. In: Bellazzi, R.; Abu-Hanna, A.; Hunter, J., editors. *Proceedings of the 11th conference on Artificial Intelligence in Medicine (AIME'07)*, LNAI 4594; 2007; Amsterdam, NL: Springer Verlag; p. 441-5.
- [22]. Seyfang, A.; Miksch, S.; Marcos, M.; Wittenberg, J.; Polo-Conde, C.; Rosenbrand, K. Bridging the gap between informal and formal guideline representations. In: Brewka, G.; Coradeschi, S.; Perini, A.; Traverso, P., editors. *European conference on Artificial Intelligence (ECAI-2006)*; Riva del Garda, Italy: IOS Press; 2006. p. 447-51. vol. 141 of *frontiers in artificial intelligence and applications*
- [23]. Peleg, M.; Wang, D.; Fodor, A.; Keren, S.; Karnieli, E. Adaptation of practice guidelines for clinical decision support: a case study of diabetic foot care. In: ten Teije, A.; Lucas, P.; Miksch, S., editors. *Proceedings of the workshop AI techniques in healthcare: evidence-based guidelines and protocols (at ECAI'06)*; 2006;
- [24]. Wang D, Peleg M, Tu SW, Boxwala A, Ogunyemi O, Zeng Q. Design and implementation of the GLIF3 guideline execution engine. *Journal of Biomedical Informatics*. 2004; 37(5):305–18. [PubMed: 15488745]
- [25]. Terenziani, P.; Montani, S.; Bottrighi, A.; Torchio, M.; Molino, G.; Correndo, G. A context-adaptable approach to clinical guidelines. In: Fieschi, M.; Coiera, E.; Li, Y-CC., editors. *Proceedings from the Medinfo 2004 world congress on medical informatics, AMIA*; 2004; San Francisco, CA: IOS Press; p. 169-73.
- [26]. Terenziani P, Molino G, Torchio M. A modular approach for representing and executing clinical guidelines. *Artificial Intelligence in Medicine*. 2001; 23(3):249–76. [PubMed: 11704440]
- [27]. Terenziani, P.; Montani, S.; Bottrighi, A.; Molino, G.; Torchio, M. Clinical guidelines adaptation: managing authoring and versioning issues. In: Miksch, S.; Hunter, J.; Keravnou, E., editors. *Proceedings of the 10th conference on Artificial Intelligence in Medicine in Europe, AIME 2005*; 2005; Aberdeen, UK: Springer Verlag; p. 151-5.
- [28]. Balsler, M.; Coltell, O.; van Croonenborg, J.; Duelli, C.; van Harmelen, F.; Jovell, A. Procure: integrating formal methods in the development process of medical guidelines and protocols. In: Kaiser, K.; Miksch, S.; Tu, SW., editors. *Computer-based support for clinical guidelines and protocols; Proceedings of the symposium on Computerized Guidelines and Protocols (CGP 2004)*; 2004; Prague, Czech Republic: IOS Press; p. 103-7. vol. 101 of *studies in health technology and informatics*
- [29]. National Library of Medicine. Medical subject headings. the library (updated annually)

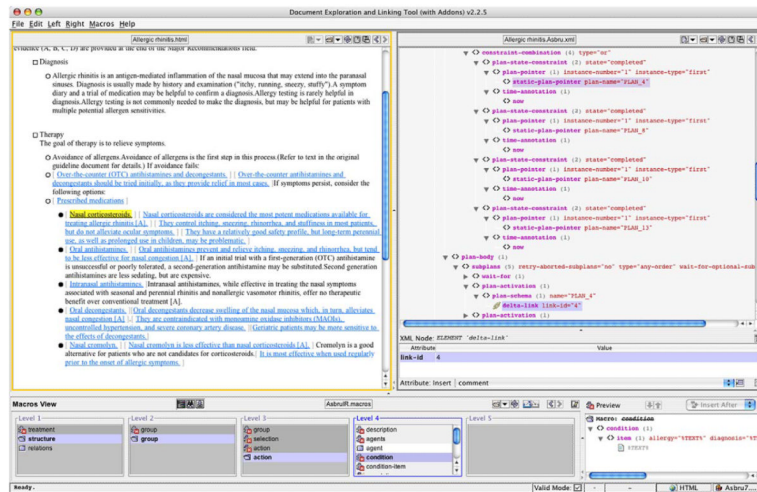


Figure 1. The DELT/A tool showing the textual guideline in the left pane, the corresponding Asbru format in the right pane, and the *macros* view in the bottom pane.

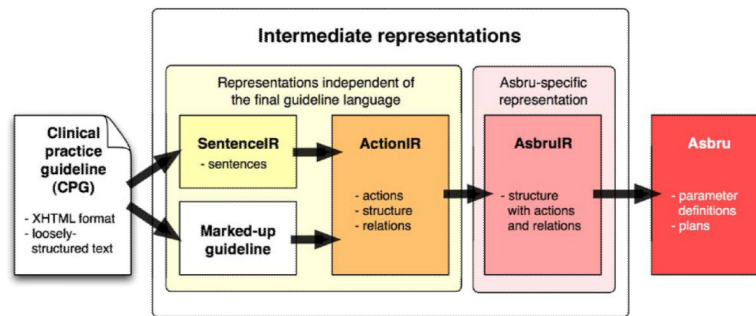


Figure 2. Steps to (semi-)automatically gain an Asbru representation of CPGs. To gain process information from a CPG the first two steps (gaining the *SentenceIR* representation and the marked-up guideline document in the first step, as well as the *ActionIR* representation in the second step) are accomplished in order to have a representation independent of the final guideline language.

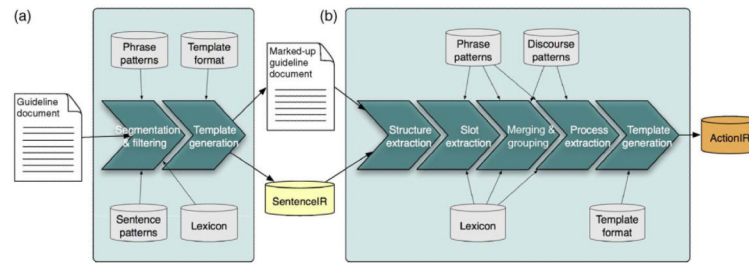


Figure 3.

Pipelines of the steps to gain *SentenceIR* and *ActionIR* representations. (a) Detecting relevant sentences. We split this task into two modules: (1) segmentation and filtering module and (2) template generation module. (b) Finding processes and extracting required information. We split this task into five modules: (1) the structure extraction module, (2) slot extraction module, (3) merging and grouping module, (4) process extraction module, and (5) template generation module.

Fluticasone provides equal clinical activity to BDP and budesonide at half the dosage. The evidence that it causes fewer side-effects at doses with equal clinical effect is limited.

2004 →

2005 → Mometasone is a new inhaled steroid and the relatively limited number of studies suggests it is equivalent to twice the dose of BDP-CFC. ¹⁶⁴ The relative safety of mometasone is not fully established. Ciclesonide is a new inhaled steroid. Its efficacy and safety relative to other inhaled steroids has not been fully established.

4.2.4 OTHER PREVENTER THERAPIES

2004 → Inhaled steroids are the first choice preventer drug. Long-acting inhaled beta2 agonists should not be used without inhaled corticosteroids. ¹⁶⁵ Alternative, less effective preventer therapies in patients taking short-acting beta2 agonists alone are:

- Chromones
 - 2004 →
 - 2005 → Sodium cromoglicate is of some benefit in adults ¹⁷⁰ and is effective in children aged 5-12 ¹⁷²
 - > 5-12 Evidence level 1+
 - Nedocromil sodium is also of some benefit in adults and children >5 ^{170, 519}
 - > 5-12 Evidence level 1++ | 5-12 Evidence level 1+
 - There is no clear evidence of benefit with sodium cromoglicate in children aged <5 ¹⁷²
- Leukotriene receptor antagonists have some beneficial clinical effect (and an effect on eosinophilic inflammation) ^{165, 666, 172}
- >12 Evidence level 1++ | 5-12 Evidence level 1++ | <5 Evidence level 1++

Figure 4. Excerpt of the 2005 version of the “living guideline” [5]. Adaptations of every new revision are marked by arrows and highlighted in terms of color (or in different gray scales).

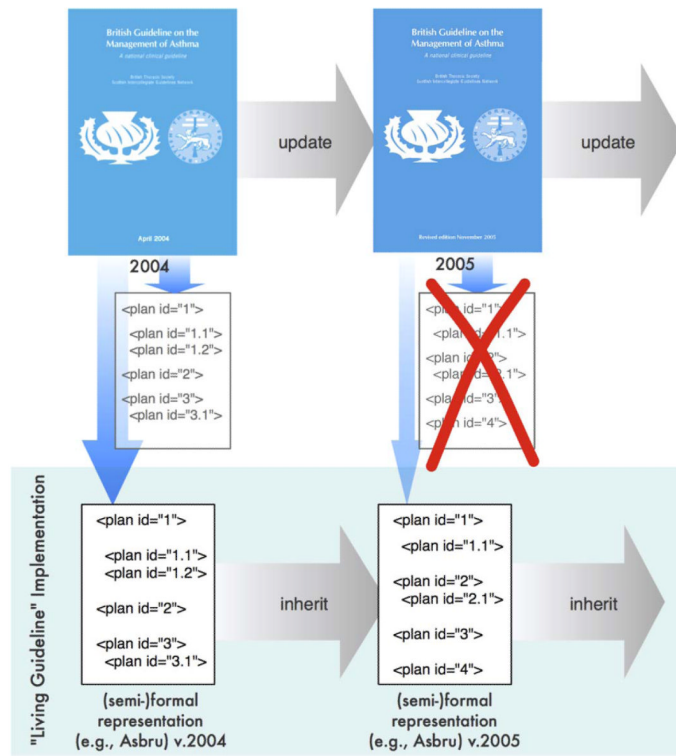


Figure 5. Formalizing a living guideline using LASSIE. The documents provide us the information that has changed. After comparing the new documents with the previous ones we are able to adapt the former formalized documents using LASSIE.

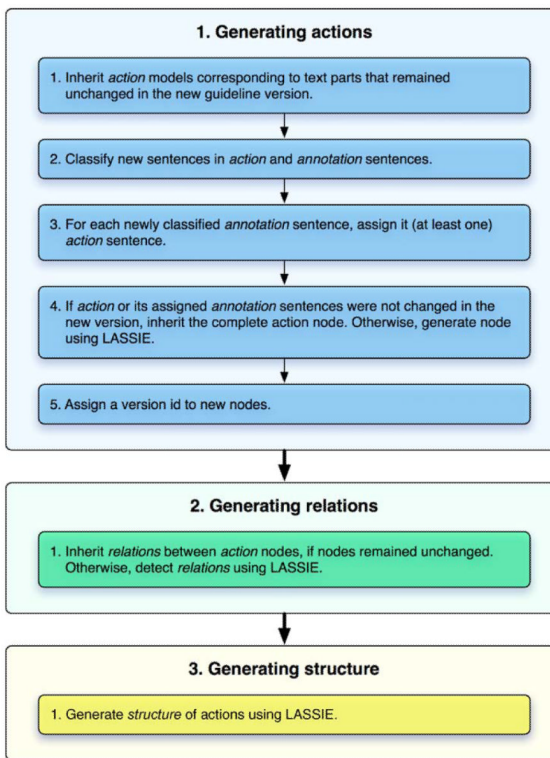


Figure 6. Steps to transform extracted information into actions.

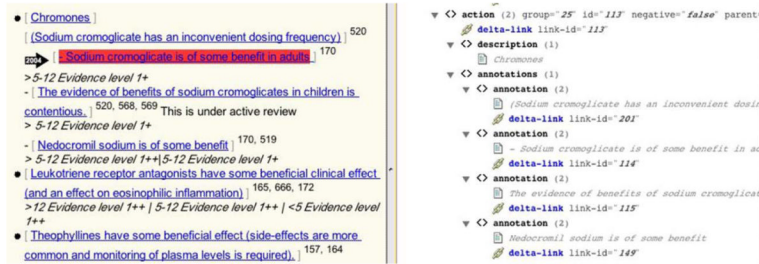


Figure 7.

Part of the annotated SIGN guideline of 2004 viewed with the DELT/A tool. On the left side the annotated HTML format and on the right side the corresponding *ActionIR* format are shown.

The figure displays two side-by-side views of a clinical guideline. The left view shows the annotated HTML format, where text is color-coded and includes evidence levels and action markers. The right view shows the corresponding ActionIR format, which is a structured JSON-like representation of the same content.

Left Panel (Annotated HTML):

- Chromones
- Sodium cromoglicate is of some benefit in adults and children aged 5-12. ¹⁷⁰ **and**
- is effective in children aged 5-12. ¹⁷²
- 5-12 Evidence level 1+
- Nedocromil sodium is also of some benefit in adults and children >5. ^{170, 519}
- 5-12 Evidence level 1++ | 5-12 Evidence level 1+
- There is no clear evidence of benefit with sodium cromoglicate in children aged <5. ¹⁷²
- Leukotriene receptor antagonists have some beneficial clinical effect (and an effect on eosinophilic inflammation). ^{165, 666, 172}
- >12 Evidence level 1++ | 5-12 Evidence level 1++ | <5 Evidence level 1++

Right Panel (ActionIR):

```

action (3) group="25" id="113" negative="false" parent="Chromones"
  delta-link link-id="113"
  description (1)
    Chromones
  annotations (2)
    annotation (2) version="2005"
      - Sodium cromoglicate is of some benefit in adults and children aged >5.
      delta-link link-id="223"
    annotation (2) version="2005"
      Nedocromil sodium is also of some benefit in adults and children aged >5.
      delta-link link-id="224"
    annotation (2) version="2005"
      There is no clear evidence of benefit with sodium cromoglicate in children aged <5.
      delta-link link-id="225"
  action (2) group="25" id="43" negative="false" parent="Leukotriene receptor antagonists"

```

Figure 8.

Part of the updated annotated SIGN guideline of 2005 viewed with the DELT/A tool. On the left side the annotated HTML format and on the right side the corresponding *ActionIR* format are shown. Updates of 2005 are highlighted in blue on the left side (compare Fig. 7). The sentences were classified as annotation sentences and thus were assigned an action sentence. Unchanged sentences and their corresponding models were inherited.

```

<number or range> ::= <number> ((' to '|' '.' ) <number>)?
<time> ::= <number or range> <time-unit>
<dosage> ::= <number or range> <dose-unit>
<iteration> ::= ('TID'|'BID'|'QD'|('Q'|'every')) <time> | <number or range> ('times'|'doses') ('per'|'a') <time-unit>
<duration> ::= <time>
...
<therapy instrument> ::= <pharmacological substance> | <drug> | <surgical procedure> | <remedy>
<trigger> ::= 'indicate' | 'activate' | 'perform' | 'prescribe' | 'treat' ...
<relevant sentence> ::= (<therapy instrument> <dosage> <duration> <iteration> <condition> <diagnosis>)*
<trigger>
(<therapy instrument> <dosage> <duration> <condition> <diagnosis>)*
<relevant clause> ::= (<dosage> <duration> <iteration> <condition> <therapy instrument>)*
...

```

Listing 1.

Extraction patterns used to populate IRs expressed in BNF. Above are definitions of some phrase patterns, followed by sentence patterns.

Table 1

Number of sentences in Section 4 of the “British guideline on the management of asthma” [5] changed from version 2004 to 2005

Version	2004	2005
Sentences (overall)	509	515
Sentences classified as relevant	139	143
Changed or added sentences (classified as relevant)		13 (10)

133 out of the 139 sentences were preserved from the 2004 version. 13 sentences have changed or were added in the 2005 version, whereas 10 sentences were classified as relevant.