Modeling Clinical Protocols using Semantic MediaWiki: the Case of the Oncocure Project

Claudio Eccher[1], **Antonella Ferro**[2], **Andreas Seyfang**[3], **Marco Rospocher**[4] and **Silvia Miksch**[3]

Abstract. A computerized Decision Support Systems (DSS) can improve the adherence of the clinicians to clinical guidelines and protocols. The building of a prescriptive DSS based on breast cancer treatment protocols and its integration with a legacy Electronic Patient Record is the aim of the Oncocure project. An important task of this project is the encoding of the protocols in computer-executable form — a task that requires the collaboration of physicians and computer scientists in a distributed environment. In this paper, we describe our project and how semantic wiki technology was used for the encoding task. Semantic wiki technology features great flexibility, allowing to mix unstructured information and semantic annotations and automatically generate the final model with minimal adaptation cost. These features render semantic wikis natural candidates for small to medium scale modeling tasks, where the adaptation and training effort of bigger systems cannot be justified. This approach is not constrained to a specific protocol modeling language, but can be used as a collaborative tool for other languages. When implemented, our DSS is expected to reduce the cost of care while improving the adherence to the guideline and the quality of the documentation.

1 INTRODUCTION

The unprecedented growth in the scientific understanding and management of diseases poses the serious problem of applying this knowledge in the clinical practice. Clinical protocols adapt the available knowledge in books, articles, and clinical guidelines to the local resources and conventions at a specific site. They are a means to improve the quality of and reduce the undesired variations in care by efficiently disseminating the existing knowledge about the state of art. While they are more concise than clinical guidelines, they still can easily exceed 50 pages and handling them in paper form in daily practice can be tedious. Practitioners compliance with clinical protocols and outcomes can be promoted and improved by computerized Decision Support Systems (DSS) supporting guideline-based or protocol-based care in an automated fashion at the time and location of decision making, especially when used in combination with an Electronic Patient Record (EPR) and integrated in the clinical workflow [15, 2]. To this end, the Oncocure project aims at designing and implementing a prescriptive guideline-based DSS integrated with a legacy Oncological EPR (OEPR) in use in the Medical Oncology Unit (MOU) of the S. Chiara Hospital of Trento (Northern Italy). The DSS is based on the Asbru [11] encoding of protocols of breast cancer medical therapies used in the unit.

[1] Fondazione Brunio Kessler, Trento Italy, email: cleccher@fbk.eu
[2] Medical Oncology, S. Chiara Hospital, Trento, Italy
[3] Danube University Krems, Austria
[4] Fondazione Brunio Kessler, Trento, Italy

The translation process when dealing with protocols shares important characteristics with the modeling of clinical guidelines. The difference is that the protocols are much more concise than guidelines and their structure is nearer to a formal representation. Therefore, the transformation effort is smaller for protocols. Nonetheless, it is far from trivial, since protocols, like guidelines, are combinations of written text and informal diagrams, and contain implicit knowledge and assumptions about the care process and the medical background, which need to be acquired during the modeling process. It requires the collaboration of physicians and computer scientists: only the former have the knowledge to grasp the deep clinical meaning integrated in the protocols, and only the latter have the training in creating formal models. Moreover, members of the encoding team may not be located in the same building, or even in the same town, and may not be able to physically participate in meetings.

In the case of the Oncocure project, the modeling phase required collaboration between oncologists and computer scientists located in Trento and in Vienna. Interaction consisted of a continuous exchange of text and pseudo-code based documents, which had to be maintained, updated and all changes had to be traced. Thus, we needed a web-based tool for intuitive collaborative editing to elaborate the model and exchange notes. In this paper, we present the Oncocure project and describe the use of Semantic MediaWiki (SMW) as a collaborative framework for the knowledge acquisition phase of project. The proposed approach allows remotely located people to actively participate to the encoding of the breast cancer protocols into the skeletal plan-representation language Asbru.

In Section 2 we describe the Oncocure project. In Section 3 and 4 we present the modeling phase and the process of protocols encoding using SMW, by showing a running example. Finally, in Section 5 we discuss the related work and draw some conclusions.

2 THE ONCOCURE PROJECT

Clinical Practice. Inside the MOU of the S. Chiara Hospital of Trento, each oncologist is specialist in one or more type of cancer (e.g., breast cancer, colon-rectal cancer, etc.). However, in the daily routine at the hospital, he/she is often required to treat patients with other types of cancer. Medical treatment to cancer patients is also provided in the Internal Medicine wards of the peripheral hospitals of the Province of Trento, which lack a specific oncology service, under the supervision of an oncologist of the MOU. In such cases, the DSS will be particularly useful, as it efficiently assists the physician in looking up the required knowledge.

Central to the care process are the periodical encounters with the patients, in which the physician visits the patient and decides the appropriate strategy on the basis of an objective examination and labo-

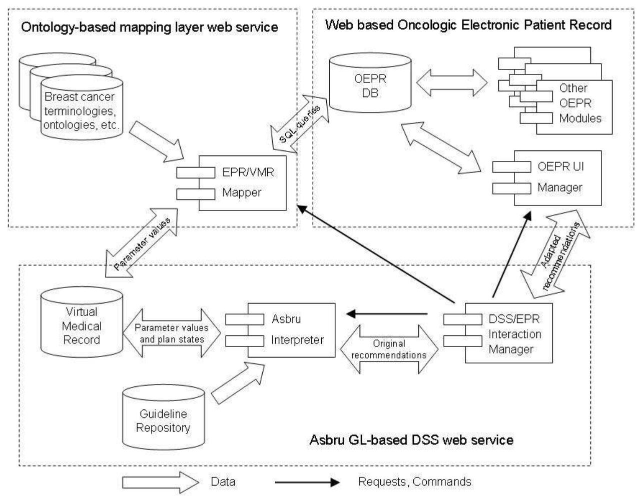


Figure 1. Block diagram of the DSS integrated with the legacy OEPR. The VMR and ontology-based mapping layer separates the DSS from the OEPR.

ratory and diagnostic imaging exams: whether to continue, suspend or interrupt an ongoing medical therapy or initiate a new one.

Infrastructure. Since 2000, a web-based OEPR, developed in our laboratory [4], has been in use in five hospitals of our Province. OEPR provides a first level of "passive" support to the shared management of cancer patients between the oncology unit of the S. Chiara Hospital and the peripheral hospitals, allowing to store patient and cancer data as well as past and ongoing therapies and outcomes. Currently, all the cancer patients are managed through the system, which by now stores more than 10 000 cases.

On this legacy infrastructure, we intend to add intelligent tools for actively supporting the physicians in their everyday clinical practice. To this end, the two year *Oncocure* project started in April 2007, with the main aim to design and develop a prescriptive guideline-based DSS for giving active support at important decisional steps of the oncological care process. The DSS is based on the Asbru language, used to encode the protocols of pharmacological therapies for breast cancer in use in the MOU of the S. Chiara Hospital, and integrates the Asbru interpreter [13]. The system will generate patient-specific reminders about the most appropriate therapeutic strategy recommended by the cancer treatment protocols in the presence of the specific disease and patient conditions. To this end, the DSS must be seamlessly integrated with the legacy OEPR and the local clinical workflow. At the same time, we wanted an architectural solution ensuring loose coupling between the DSS and the legacy OEPR, in order to build a easy maintenance system integrable with different clinical information systems. For this reason, the DSS will be built as a Web service invokable by the physician from the OEPR User Interface (see Figure 1). Moreover, we utilize the Virtual Medical Record (VMR) approach [5], which supports a well-defined structured data model for representing information related to individual patients. A breast cancer VMR will be defined, in which the parameters required by the DSS, extracted or abstracted from the OEPR database through an ontology-based mapping layer, are stored along with the states reached by the interpreter during guideline execution. Ontologies, in fact, are widely acknowledged as the essential glue to ensure semantic consistency to data and knowledge [10]. This approach allows to deploy our solution at other points of care.

Aim and Evaluation. An important part of the work is the test phase, planned for the last three months of the project, which will give indications about the real effectiveness and the impact on the clinical practice of a protocol-based DSS integrated into the care-flow. According to the oncologists, we expect several benefits from the deployment of the Oncocure system, especially for those oncologists of the MOU who do not specifically follow the breast cancer disease and for the clinicians of the peripheral hospitals who provide breast cancer medical treatments.

- Improved adherence to protocols, because all the decisions can be taken after consultation of the protocols. This, in turn, can improve the care quality by favoring the use of the best-evidence.
- Reduction of labor of the clinicians, because for each case only the relevant protocol fragments are displayed. This avoid the necessity for physician to leaf through protocols to find the right recommendation, especially in the presence of the patient.
- Improved documentation, because each decision of the physician can be automatically logged in the EPR. If a valid justification is required from the care provider who refuses a recommended treatment, this information can be used to verify *a posteriori* the actual applicability of the protocols and to improve their quality.

3 MODELLING BACKGROUND

In contrast to guidelines, which are mostly in textual form, the internal protocols are mainly constituted by informal "box and arrow" diagrams accompanied by short explanation text. In Figure 2 we show an example of the diagram for the adjuvant therapy decision of hormone responsive patients. Notwithstanding they are more formal and concise than guidelines, they still present ambiguities and incompleteness that must be resolved. Moreover, to smoothly integrate the system in the clinical workflow, it is necessary to explicit the care process tasks in which protocols are used and define the system requirements. Consequently, the knowledge acquisition phase was still considered the most important task of the Oncocure project.

Most of the first year of the project (10 months), in fact, has been devoted to the acquisition of the explicit knowledge written in the protocols and the implicit knowledge on the real care process carried out in the ward, and to the codification of the Asbru model. This knowledge acquisition phase was conducted through regular meetings between a computer scientists and an expert oncologists in Trento, with the intervention of a computer scientists in Vienna by phone or through the preparation of documents, email, etc., discussed during the encounters. Although the size of the protocol and the resulting Asbru model is not large, we still needed a multi-user, web-based tool to manage all the pieces of knowledge which accumulated over time, and to ensure that everyone is working on the same updated version.

Semantic MediaWiki [17] belongs to a new generation of tools supporting the integration of Web 2.0 and Semantic Web approaches [8], which have been developed to meet the needs of the semantic modeling community to easily create, share, and connect content and knowledge. While preserving the freedom in format of the overall page, it introduces knowledge annotations with a formal semantics for their use. This is an advantage over editors like Uruz [14] which enforce the structure of Asbru plans onto the data entry forms. Moreover, all the important functionalities (access control and permissions, tracing of the activity, semantic search, etc.) are already provided by the SMW framework, without needing to install specific client applications.

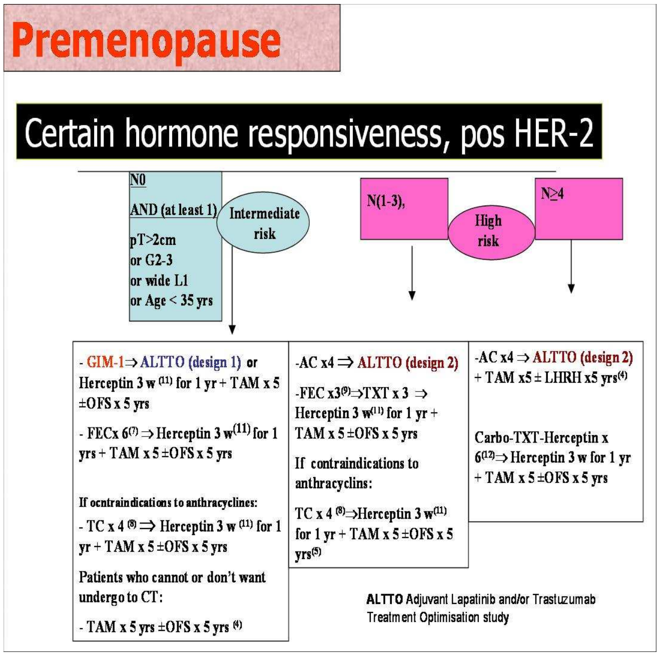


Figure 2. An example of the informal diagrams of the adjuvant treatment decision breast cancer internal protocols prepared by the oncologist. It is quite concise, but a certain amount of knowledge is not explicit (e.g., the preferences/conditions for the choice of one of several therapies.)

SMW employs binary relations and attributes to render machine-processable the shared knowledge of wiki. Key semantic annotations of SMW are the following:

- *Categories*, which classify the pages according to their content.
- *Typed links*, which express a relation between pages. New types can be created by the user on the fly by just employing them for annotations.
- *Attributes*, which specify simple datatype properties related to the content of a page.

The access to the wiki pages is protected by login and password. Different level of permissions can be defined for users.

The SMW framework provides RDF exports of the pages, which can be transformed into other representations. This means that, while the overall wiki page can be organized as the authors prefer, the collection of semi-formal fragments can be automatically transformed into a basis for the final Asbru model.

One of the disadvantages of SMW is the requirement to handle the special syntax. This is generally not perceived to be easy by non-IT people involved in the authoring process.

4 THE ENCODING PROCESS

The basic idea is that an Asbru guideline model is expressed as a collection of interrelated SMW pages connected by typed links. A SMW page corresponds to an Asbru building block of the guideline and may contain:

1. typed links connecting the page to other building block pages (e.g., a plan and its subplans),
2. attributes for expressing element data that must be present in the model (e.g., the title of a plan), and

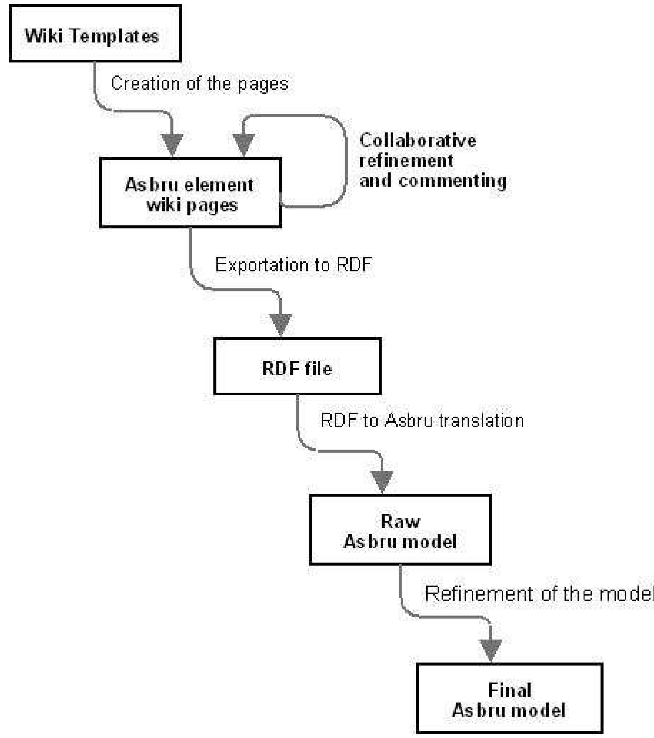
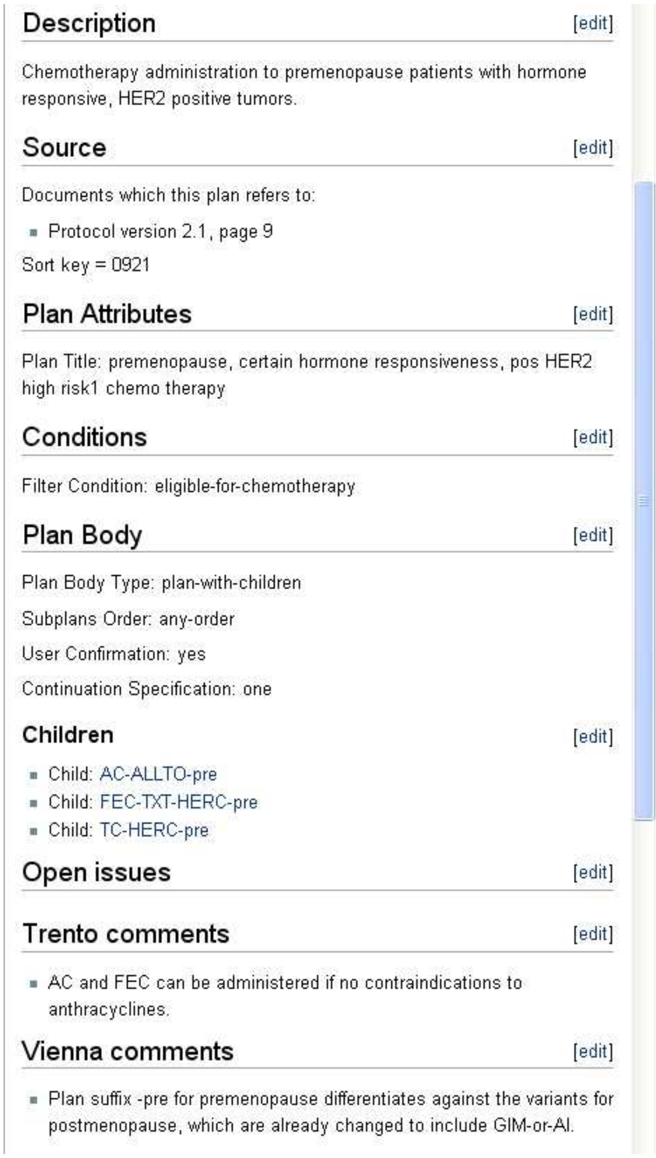


Figure 3. Diagram of the process to encode the protocols in Asbru using SMW. The activity in bold is the phase in which model pages are collaboratively revisioned and consolidated.

3. free text, organized in sections (e.g., Description, Open issues, etc.) to add comments and explanations for users not trained in the representation. The comments can be added by each user and are not translated in the final model.

Each page is assigned to one of a set of predefined SMW categories according to the type of Asbru element it models, by means of a specific semantic annotation. In Figure 3 we show the diagram of the process steps for encoding the protocols in Asbru using SMW.

Creation of the SMW Pages. To facilitate the creation of wiki pages, we have defined a set of template pages for the most frequently used Asbru building blocks. For this first version of the tool, we have defined a template for each plan body type (plan with subplans, plan activation, ask, user performed, cyclical plan, variable assignment, if-then-else), two templates for parameters (raw data definition and qualitative parameter definition), and two templates for abstraction (qualitative scale definition and secondary qualitative entry.) A new Asbru element can be defined simply by cutting and pasting the template into the editing area of the new wiki page, completing the fields and removing those not required (e.g., unnecessary conditions in plans). However, to overcome the drawback of handling SMW syntax of slots, we designed a Java page generator applet that presents fill-in forms representing a typical page layout. The set of fill-in forms is generated from the chosen template so that the page layout can be adapted to changing requirements during the project with minimal effort. Of course, this takes some freedom in the organization of the content from the users, but it relieves them from learning the syntax and it helps prevent editing errors. Also, the use of the page generator does not prevent later individual changes to pages, without any restraints. In Figure 4 we show a snapshot of the wiki page representing the plan that models the chemotherapy treat-



Description [edit]

Chemotherapy administration to premenopause patients with hormone responsive, HER2 positive tumors.

Source [edit]

Documents which this plan refers to:

- Protocol version 2.1, page 9

Sort key = 0921

Plan Attributes [edit]

Plan Title: premenopause, certain hormone responsiveness, pos HER2 high risk1 chemo therapy

Conditions [edit]

Filter Condition: eligible-for-chemotherapy

Plan Body [edit]

Plan Body Type: plan-with-children

Subplans Order: any-order

User Confirmation: yes

Continuation Specification: one

Children [edit]

- Child: AC-ALLTO-pre
- Child: FEC-TXT-HERC-pre
- Child: TC-HERC-pre

Open issues [edit]

Trento comments [edit]

- AC and FEC can be administered if no contraindications to anthracyclines.

Vienna comments [edit]

- Plan suffix -pre for premenopause differentiates against the variants for postmenopause, which are already changed to include GIM-or-AI.

Figure 4. Snapshot of the SMW page of a plan with subplans representing the chemotherapy treatment options in middle box of Figure 2. In the *Open issues*, *Vienna comments* and *Trento comments* sections each participant can add comments and explanations. Note the *Sort key* entry, which allows to relate the wiki page (and the generated Asbru plan) to the precise position in the source document.

ments recommended in the middle box of Figure 2. A fragment of the wiki code of this page, with the semantic annotations defining the plan type, is displayed in Figure 5.

An important knowledge slot is the reference to the exact location in the source document, on which the model fragment described by a particular wiki page is based. This was inspired by MHB and the DELT/A tool [12], where links are inserted automatically. In Oncocure, the source document is technically difficult to access (graphics in a non-editable data format). Therefore, we resorted to manually inserting a code with page number and a separate order key within the page (Sort key in Figure 4) in the wiki model. Using both, we can print the wiki model in the precise order of the source document, which allows the easy side-by-side comparison of source and model also by non-IT people.

```
== Plan Body ==

Plan Body Type:
       [[has Plan Body Type:=plan-with-children]]

Subplans Order: [[has Subplans Order:=any-order]]

User Confirmation: [[has User Confirmation:=yes]]

Continuation Specification:
          [[has Continuation Specification:=one]]

=== Children ===

* Child: [[has Child::AC-ALLTO-pre]]
* Child: [[has Child::FEC-TXT-HERC-pre]]
* Child: [[has Child::TC-HERC-pre]]
```

Figure 5. Fragment of the SMW code defining the plan in Figure 4. Only semantic annotations, included between double square brackets, are translated in RDF and then in Asbru. For example, the current plan is defined as a plan with subplans ([[has Plan Body Type:=plan-with-children]]) with three children (AC-ALLTO-pre, FEC-TXT-HERC-pre, and TC-HERC-pre).

```
<plan-body>
 <subplans retry-aborted-subplans="no"
  type="any-order" wait-for-optional-subplans="no">
   <wait-for>
     <one/>
   </wait-for>
   <plan-activation>
     <plan-schema name="AC-ALLTO-pre"/>
   </plan-activation>
     <plan-activation>
       <plan-schema name="FEC-TXT-HERC-pre"/>
    </plan-activation>
    <plan-activation>
      <plan-schema name="TC-HERC-pre"/>
    </plan-activation>
 </subplans>
</plan-body>
```

Figure 6. Fragment of the valid Asbru model generated from the SMW semantic annotations shown in Figure 5. The Asbru model is the translation of the RDF dump generated using the SMW built-in function.

Page Refinement and RDF Export. Pages can be refined and commented by each member of the modeling team. Once the team has agreed upon the general modeling issues, a built-in functionality allows to choose a set of pages and export them in a RDF file. Categories, pages and semantic annotations are exported as classes, instances, and properties, respectively.

Asbru Model Generation and Refinement. A second custom Java applet transforms the RDF export to an Asbru XML model according to the Asbru DTD (see Figure 6). This model has those fields set to their final value, for which sufficient and precise information is available. For those cases, where the SMW model contains only a sketch in free text (e.g., the conditions on plan state transitions in this first version), the available information is inserted as a comment and the refinement is left for the next modeling phase. The Asbru model is then further refined and completed (e.g., transforming the pseudo code conditions into Asbru propositions) through a dedicated Asbru modeling software like DELT/A.

5 RELATED WORK AND CONCLUSION

A number of Clinical Practice Guidelines (CPG) frameworks representing task, plan and decision structures have been proposed in recent years. Some of them have been complemented with active software tools for guideline execution [3]. See Peleg et al. [9] for a detailed comparison and Mulyar et al. [7] for a pattern-based analysis of CPGs. In both comparisons, Asbru [11] scored favorably.

Guideline Modeling Tools can be roughly classified into (1) model-centric and (2) document-centric approaches. In the model-centric approach, a conceptual model is formulated by domain experts. Thus, the relationship between the model and the original document of the clinical guideline is only indirect. In the document-centric approach markup-based tools are used to systematically mark up the original guideline in order to generate a semi-formal model of the marked text part (see Leong et al. [6] for an overview.) In our project, both approaches merged to a considerable extent, as the protocol is already in a structure similar to a formal model. One could say that we followed a model-centric view, without loosing contact with the document structure, by keeping precise references to the source document. In a previous project [16], an intermediate notation, MHB [12], was used to model a breast cancer guideline. The protocol modeled in our current project, however, does not contain as many dimensions of knowledge as the previously modeled guideline. Also, the smaller bridging effort and the tense time frame of the project suggested going from the original and already well-structured document to Asbru, without an intermediate representation. Still, we needed a multi-user, web-based tool that had to allow a significant amount of freedom of the form, which is one of the prime advantages of wikis.

In this paper, we present the Oncocure project and propose to use SMW as a collaborative tool for knowledge acquisition of treatment protocols. To our knowledge there is only a paper in biomedical informatics describing the use of an extended SMW (BOWiki) for the collaborative annotation of gene information [1]. Main advantages of the solution we propose are:

1. the provision of a lightweight infrastructure that allows each participant to easily follow and contribute to the encoding process,
2. the possibility to mix informal content with semi-formal fragments, and
3. the automatic generation of valid skeletal Asbru models.

The oncologist expressed a positive attitude toward this collaborative modeling tool, because she was able to easily understand and verify the computer scientists's modeling work, using both the page structure and the informal annotations and comments. SMW, however, still requires the knowledge of the Asbru language to correctly complete the semantic tags; it does not solve the problem of allowing a domain expert to work alone in encoding guidelines to maintain/update the knowledge base. However, this is a problem shared by all knowledge acquisition projects. Its solution requires a broader approach than just applying software, e.g., user training in formal modeling. Although our work is focused on using Asbru as encoding language, our approach can be generalized to encode protocols in other languages in a collaborative way, by defining templates with slots specific for the language used.

This first prototype is limited to the most frequently used Asbru elements. Further work is needed for the definition of templates for more complex blocks. We are also working to fully integrate the Java applets into the SMW framework so as to generate wiki pages and Asbru code from within the pages themselves.

ACKNOWLEDGEMENTS

This work has been carried out in the context of the project "*Oncocure - Information and Communication Technologies for clinical governance in oncology: Design and development of a computerized guideline-based system for supporting and evaluating evidence-based clinical practice*", funded by the Fondazione Caritro of Trento.

REFERENCES

[1] M. Backaus and J. Kelso. BOWiki - a collaborative annotation and ontology curation framework. WWW 2007, Banff, Canada. May 8-12, 2007.

[2] D.W. Bates, G.J. Kuperman, S. Wang, T. Gandhi, A. Kittler, L. Volk, C. Spurr, R. Khorasani, M. Tanasi-jevic, and B. Middleton, 'Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality', *J Am Med Inform Assoc*, **10**, 523–530, (2003).

[3] P. de Clercq, J. Blom, H. Korsten, and A. Hasman, 'Approaches for creating computer-interpretable guidelines that facilitate decision support', *Artif Intell Med*, **31**, 1–27, (2004).

[4] C. Eccher, F. Berloffa, E. Galligioni, B. Larcher, and S. Forti, 'Experience in designing and evaluating a teleconsultation system supporting shared care of oncological patients', in *Proceedings of AMIA Symposium 2003*, p. 835, (2003).

[5] P.D. Johnson, S.W. Tu, M.A. Musen, and I. Purves, 'A virtual medical record for guideline-based decision support', in *Proceedings of AMIA Symposium 2001*, pp. 294–298, (2001).

[6] T.Y. Leong, K. Kaiser, and S. Miksch, 'Free and open source enabling technologies for patient-centric, guideline-based clinical decision support: a survey', *2007 IMIA Yearbook of Medical Informatics, Methods Inf Med*, **46**(1), 74–86, (2007).

[7] N. Mulyar, W.M.P. van der Aalst, and M. Peleg. A pattern-based analysis of clinical computer-interpretable guideline modeling languages. J Am Med Inf Assoc, 14(6), (to appear).

[8] N.F. Noy, A. Chugh, and H. Alani, 'The CKC challenge: exploring tools for collaborative knowledge construction', *IEEE Intelligent Systems*, **23**(1), 64–68, (2008).

[9] M. Peleg, S. Tu, J. Bury, P. Ciccarese, J. Fox, R.A. Greenes, R. Hall, P.D. Johnson, N. Jones, A. Kumar, S. Miksch, S. Quaglini, A. Seyfang, E.H. Shortliffe, and M. Stefanelli, 'Comparing computer-interpretable guideline models: a case study ap-proach', *J Am Med Inf Assoc*, **10**(1), 52–68, (2003).

[10] D.M. Pisanelli, M. Battaglia, and C. De Lazzari, *ROME: a Reference Ontology in Medicine*, 485–93, New Trends in Software Methodologies, Tools and Techniques, IOS Press, Amsterdam, 2007.

[11] A. Seyfang, R. Kosara, and S. Miksch. Asbru 7.3 reference manual. Technical report, Vienna University of Technology, (2002).

[12] A. Seyfang, S. Miksch, M. Marcos, J. Wittenberg, C. Polo-Conde, and K. Rosenbrand, 'Bridging the gap between informal and formal guideline representations', in *17th European Conferencen on Artificial Intelligence*, eds., G. Brewka, S. Coradeschi, A. Perini, and P. Traverso, pp. 447–451. IOS Press, (2006).

[13] A. Seyfang, M. Paesold, P. Votruba, and S. Miksch, *Improving the Execution of Clinical Guidelines and Temporal Data Abstraction in High-Frequency Domains*, volume Computer-Based Medical Guidelines and Protocols: A Primer and Current Trends of *Health Technology and Informatics*, forthcoming, IOS Press, 2008.

[14] T. Shahar, O. Young, E. Shalom, M. Galperin, A. Mayaffit, R. Moskovitch, and A. Hessing, 'A framework for a distributed, hybrid, multiple-ontology clinical-guideline library, and automated guideline-support tools', *Journal of Biomedical Informatics*, **37**, 325–344, (2004).

[15] F.A. Sonnenberg and C.G. Hagerty, 'Computer interpretable guidelines: where are we and where are we going?', *2006 IMIA Yearbook of Medical Informatics, Methods Inf Med*, **45**(Suppl 1), S145–158, (2006).

[16] A. ten Teije, M. Marcos, M. Balser, J. van Croonenborg, C. Duelli, F. van Harmelen, P. Lucas, S. Miksch, W. Reif, K. Rosenbrand, and A. Seyfang, 'Improving medical protocols by formal methods', *Artif Intell Med*, **36**(3), 193–209, (2006).

[17] M. Völkel, M. Krötzsch, D. Vrandecic, H. Haller, and R. Studer. Semantic Wikipedia. WWW 2006, Edinburgh, Scotland. May 23-26, 2006.