SAGE Guideline Modeling: Motivation and Methodology

Samson W Tu[1], James Campbell[2], Mark A Musen[1]

[1]*Stanford Medical Informatics, Stanford University, Stanford, CA, USA*
[2]*University of Nebraska Medical Center, Nebraska Health Systems, Omaha, NE*

Abstract. The SAGE (**S**tandards-Based Sharable **A**ctive **G**uideline **E**nvironment) project is a collaboration among research groups at six institutions in the US. The ultimate goal of the project is to create an infrastructure that will allow execution of standards-based clinical practice guidelines across heterogeneous clinical information systems. This paper describes the design goals of the SAGE guideline model in the context of the technological infrastructure and guideline modeling methodology that the project is developing.

Introduction

The SAGE (**S**tandards-Based Sharable **A**ctive **G**uideline **E**nvironment) project, a collaboration among research groups at IDX Systems Corporation, the University of Nebraska Medical Center, Intermountain Health Care (IHC), Apelon, Inc., Stanford Medical Informatics, and the Mayo Clinic, seeks to create the technological infrastructure for integrating interoperable computer-based guidelines into enterprise clinical information systems (CISs). This paper describes the design goals of the SAGE guideline model in the context of the technological infrastructure and the guideline modeling methodology that the project is developing. We discuss considerations that led us to develop a new model, the suite of models that define the interface between a guideline knowledge base and a CIS, and a deployment-driven approach to guideline knowledge base development.

1. Design Goals

The literature on guideline models is full of methods for formalizing clinical guidelines and protocols [1-4]. What is the justification for starting yet another guideline model?

Three considerations led us to our decision to start a new model. Past efforts have gone into developing shared models for representing medical decisions and clinical guidelines [1, 5]. However, as an experiment to share Medical Logic Modules (MLMs) across two institutions [6] indicated, it takes more than a formalism for medical logic to accomplish sharing of computable medical knowledge. Lack of standards in terminologies and in data models for patient information required re-coding of significant parts of the MLMs. Similarly, work in United Kingdom to develop guideline-based decision support for primary care suggested that reuse of a guideline knowledge base is possible once an infrastructure that includes medical record query interface, terminology mediation, and act

interface is in place [7]. With the emergence of clinical standards such as Health Level Seven's Version 3 (HL7 v3) Reference Information Model (RIM) [8] and College of American Pathologists' SNOMED Clinical Terms [9], we believe that we have the opportunity to build a guideline model from ground up to take advantage of these infrastructural standards in a systematic way. As we will discuss in the paper, making use of standards for modeling guideline is not a straight forward process. Rarely do existing standards completely satisfy the requirements of guideline modeling. Thus the elucidation of the complex relationship between existing standards and requirements of guideline modeling and deployment is one of the themes of the SAGE project.

The second consideration is SAGE's approach to the integration of guideline-based decision support with the workflow of care process. That the success of clinical decision-support systems (DSSs) depends heavily on how the system is integrated into the care process is widely recognized. Interpretation of the integration problem, however, varies widely. For alert-and-reminder systems, integrating into the workflow can mean the timing, modality, and format of notification. In hospital environments, the protocol for managing a specific medical condition may drive the workflow that sequences care tasks and schedules resources [10]. The SAGE project takes the approach that, as a provider of decision-support services to CISs, SAGE will not be in control of host systems' workflow management. Thus, in the SAGE modeling approach, we are not required to model detailed workflow as, for example, University of Pavia's careflow methodology proposes. Instead, the SAGE system will respond to *opportunities for decision support* in the care process. We need to model enough of the workflow contexts to recognize appropriate events that should trigger decision-support services. Upon receipt of such triggering event, the SAGE DSS will deliver, through existing functions of the CIS, guideline-based recommendations appropriate for members of a care team. The implication of this approach for the guideline modeling is that guideline knowledge must support operations in an event-driven reactive system and it must take into account clinical and organization contexts such as care setting and provider roles. Instead of just creating an electronic version of a clinical practice guideline, guideline modeling in SAGE formalizes guideline knowledge being used in specific scenarios and settings.

The third consideration in our decision to start a new guideline line is that, in recent years, much interchange and cross-fertilization have taken place in the guideline modeling community. Starting with workshops such as the ones sponsored by InterMed in 1999, Open Clinical in 2000, and University of Leipzig in 2001, and continuing with a number of comparison papers (such as [11]), workers in the guideline modeling community have gained much better understanding of the commonalities and differences among different guideline modeling approaches and of the design choices made in them. The SAGE project has given us the opportunity to take advantage of the prior work, including the GLIF3 and other models, to create a synthesis.

In summary, the SAGE project seeks to create a guideline model that

- uses standardized components that allow interoperability of guideline execution elements with the standard services provided within vendor clinical information systems.
- includes organizational knowledge to capture workflow information and resources needed to provide decision-support in enterprise setting
- synthesizes prior guideline modeling work for encoding guideline knowledge needed to provide situation-specific decision support and to maintain linked explanatory resource information for the end-user

2. Design Decisions

In this section, we will describe how the members of the SAGE project work toward the first two design goals.

2.1 A Suite of Models and Services to Support Guideline Modeling and Execution

To achieve interoperability of guideline decision-support system (DSS) with vendor clinical information systems (CIS), we make explicit a suite of models and services that together define the interface between DSS and CIS.

An organizational model that defines available clinical and administrative events, roles, settings, and resources provides the vocabulary to describe the contexts in which GDSS provides decision-support services. Thus, a guideline (for example, a diabetes guideline in which a patient is enrolled) may be triggered by a patient *check-in event* generated at a *primary care outpatient clinic* where guideline-based alerts are generated for providers who play the *roles* of *clinic nurse* and *primary care physician*. A guideline is encoded using a simplified view of a patient's medical record data, called a Virtual Medical Record (VMR) [12] that is ultimately based on the HL7 RIM. The SAGE VMR, for example, models allergy information as instances of an 'AdverseReaction' class that has attributes such as 'code,' 'substance,' 'reaction,' and 'effective time' (time during which a patient is to be allergic to the substance). The VMR classes, by themselves, still allow several degrees of freedom in representing patient information (e.g. the code slot in AdverseReaction may be 'allergic drug reaction' (SNOMED CT 74069000) or more restrictive 'vaccine allergy' (SNOMED CT 294640001). Detailed clinical models, also called *Clinical Expression Models* (CEMs), spell out, by placing constraints on attributes of VMR classes, precisely how patient data would be represented. For example, a CEM for "Anaphylactic reaction to hepatitis B vaccine" may specify that such data will be modeled as instances of *AdverseReaction* class where the code slot has value 'vaccine allergy,' the reaction slot is constrained to be a concept subsumed by 'anaphylactic reaction,' and the substance slot is constrained to be a kind of 'hepatitis B vaccine.'

Terms from terminologies are the atomic units of meaning that we use to make assertions through information models such as VMR and CEMs. However, concepts used in clinical guidelines often do not match precisely the term hierarchies in standard medical terminologies. The concept of 'pulmonary problem excluding asthma' in for example, is unlikely to have an exact equivalent in any standard terminology. Thus, the SAGE project has developed several strategies to define guideline concepts from standard terminologies. The first technique is to use a reference terminology's own compositional method for defining new concepts. Using SNOMED CT, for example, we can define to terms such as 'severe wound' as a {'wound lesion' (SNOMED CT 239155007) associated severity 'severe' (SNOMED CT 24484000)}. The second technique is to using a notation, which we call *Concept Expression*, to define a term as Boolean combinations of other terms (e.g. 'pulmonary disease excluding asthma' as a {'disease of lung' (SNOMED CT 19829001) AND NOT 'asthma' (SNOMED CT 195967001)}).

2.2 Deployment-Driven Knowledge-Base Development Process

To ensure that a guideline formalized in a SAGE knowledge base is informed by the usage scenarios of the guidelines in the care process, SAGE project developed a seven-step deployment-driven guideline modeling methodology [13] (see Figure 1). Once the decision to implement a guideline has been made, the SAGE guideline knowledge base development

methodology requires that clinicians first create clinical scenarios that are detailed enough to support integration of recommendations from that guideline into clinical workflow. These usage scenarios identify opportunities for providing decision support, the roles and information needs of care providers, events that may activate the guideline system, and guideline knowledge relevant in these scenarios. In the second step, clinicians analyze the information content of the desired guideline recommendations and distill, from guideline texts, medical literature, and their clinical expertise, the knowledge and logic needed to generate these recommendations. This distillation process requires clinicians to select, interpret, augment, and operationalize guideline statements in terms of unambiguous concepts and of data that may be available. Concepts identified as part of the required guideline logic are instantiated as detailed clinical data models (the CEMs described before). The fifth step of the methodology calls for specifying guideline concepts in terms of standard terminologies. As we discussed earlier, the use of standard terminologies may require significant extensions and must be defined in the context of the detailed clinical data model. The sixth step is the translation of the clinical scenarios and guideline logic into a computer-interpretable form using the SAGE guideline model as the ontological structure. Finally, before a formalized guideline can be installed and used in a local institution, its medical content must be reviewed and revised (in what we call the *localization* process) and its data models, terminologies, and organization assumptions (roles, events, and resources) must be mapped to those of the local institution (in what we cal the *binding* process).

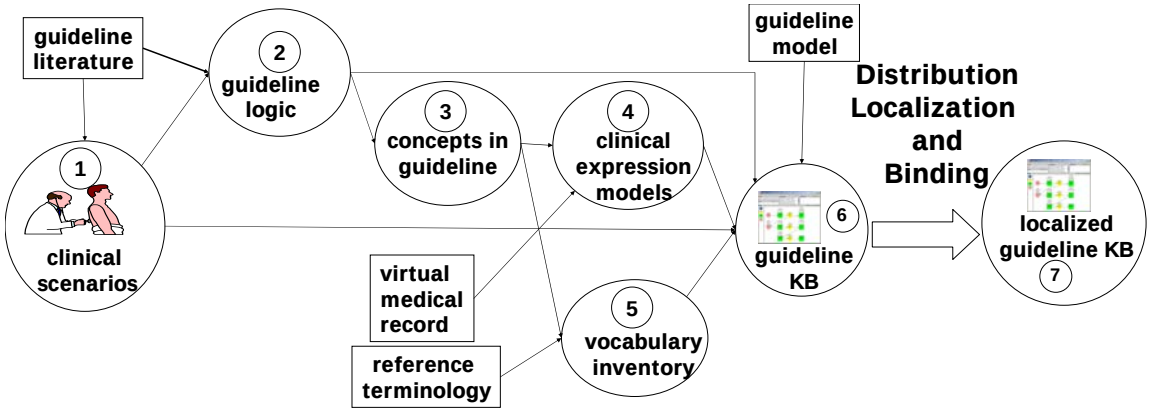

Figure 1. The SAGE guideline knowledge-base development process. The process is driven not only by the guideline literature, but by use cases for clinical decision support based on carefully defined clinical scenarios. The guideline knowledge base is support by a series of terminology, information, and organizational models.

3. Conclusion

Much of the recent literature on guideline mark-up and guideline modeling emphasize representation issues. The implementation and deployment of DSSs for guideline-based care, on the other hand, depend crucially on the infrastructure that allows such systems to query for data, to send recommendations and reminders, and to interact with users as an information source. The SAGE project attempts to define the infrastructure as a series of standard organization, terminology, and information (VMR and CEM) models. It works with standard organizations, such as Health Level Seven, to develop industry consensus for these standard models. A deployment-driven methodology makes use of these models in the guideline encoding process. The methodology helps to identify opportunities for guideline-based interventions at specific points in the care process, and it allows the identification and

distillation of guideline knowledge that is required in these scenarios. At the conclusion of the project, the SAGE project will have demonstrated that guideline knowledge bases developed using these model and this methodology can be deployed at the three clinical sites (Mayo, Nebraska, and IHC).

Acknowledgement

This work has been supported by the U.S. National Institute of Standards and Technology, Advanced Technology Program, Cooperative Agreement Number 70NANB1H3049.

References

[1] Peleg M, Boxwala A, Ogunyemi O, Zeng Q, Tu S, Lacson R, Bernstam E, Ash N, Mork P, Ohno-Machado L, Shortliffe EH, Greenes RA. GLIF3: The evolution of a guideline representation format. Proc AMIA Symp. 2000; pp. 645-649.

[2] Tu SW, Musen MA. A flexible approach to guideline modeling. Proc AMIA Symp. 1999; pp. 420-424.

[3] Johnson PD, Tu SW, Booth N, Sugden B, Purves IN. Using scenarios in chronic disease management guidelines for primary care. Proc AMIA Symp. 2000; pp. 389-393.

[4] Shiffman RN, Karras BT, Agrawal A, Chen R, Marenco L, Sujai Nath M, GEM: A proposal for a more comprehensive guideline document model using XML. *J Am Med Inform Assoc* 2000. 7: 488-498.

[5] Hripcsak G, Clayton PD, Pryor TA, Haug P, Wigertz OB, Van der lei J. The Arden Syntax for Medical Logic Modules. Proc Annu Symp Comput Appl Med Care. 1990; pp. 200-204.

[6] Pryor T, Hripcsak G. Sharing mlm's: An experiment between Columbia-Presbyterian and LDS Hospital. Proc Annu Symp Comput Appl Med Care. 1993; pp. 399-403.

[7] Johnson PD, Tu SW, Jones N. Achieving reuse of computable guideline systems. Medinfo 2001; pp. 99-1003.

[8] Health Level 7, *Hl 7 Reference Information Model*. 2003: http://www.hl7.org/library/data-model/RIM/modelpage_non.htm.

[9] Wang AY, Sable JH, Spackman KA. The SNOMED Clinical Terms development process: Refinement and analysis of content. Proc AMIA Symp. 2002; pp. 845-849.

[10] Quaglini S, Stefanelli M, Cavallini A, Micieli G, Fassino C, Mossa C, Guideline-based careflow systems. *Artif Intell Med* 2000. 5(22): 5-22.

[11] Peleg M, Tu SW, Bury J, Ciccarese P, Fox J, Greenes RA, Hall R, Johnson PD, Jones N, Kumar A, Miksch S, Quaglini S, Seyfan A, Shortliffe EH, Stefanelli M, Comparing computer-interpretable guideline models: A case-study approach. *J Am Med Inform Assoc* 2003. 10(1): 52-68.

[12] Johnson PD, Tu SW, Musen MA, Purves I. A Virtual Medical Record for guideline-based decision support. Proc AMIA Symp. 2001; pp. 294-298.

[13] Tu SW, Musen MA, Shankar R, Campbell J, Hrabak K, McClay J, Huff SM, McClure R, Parker C, Rocha R, Abarbanel R, Beard N, Glasgow J, Mansfield G, Ram P, Ye Q, Mays E, Weida T, Chute CG, McDonald K, Mohr D, Nyman MA, Scheital S, Solbrig H, Zill DA, Goldstein MK. Modeling guidelines for integration into clinical workflow. Medinfo 2004; submitted