Designation: E2210 – 06

An American National Standard

# Standard Specification for Guideline Elements Model version 2 (GEM II)—Document Model for Clinical Practice Guidelines[1]

This standard is issued under the fixed designation E2210; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

## 1. Scope*

1.1 This specification updates a standard representation for storing and organizing the heterogeneous information contained in clinical practice guidelines. This specification is intended to facilitate translation of natural-language guideline documents into a format that can be processed by computers. It can be used to represent document content throughout the entire guideline life cycle. Information at both high and low levels of abstraction can be accommodated. This specification is based on the guideline elements model (GEM) created at the Yale Center for Medical Informatics and designed to serve as a comprehensive XML-based guideline document representation.

1.2 This specification refers to and makes use of recommendations from the World Wide Web consortium, the W3C.[2]

1.3 *Standard Guideline Schema*—This specification defines a standard Schema for clinical practice guidelines. The Schema is included in Annex A1.

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory requirements prior to use.*

## 2. Referenced Documents

2.1 *W3C World Wide Web Consortium:*
XML 1.0 Recommendation [3]
XML Schema 1.0 [4]

## 3. Terminology

3.1 *Definitions:*

3.1.1 *document type definition (DTD)*—the formal definition of the elements, structures, and rules for enabling platform-independent data access via XML, or for marking up a given type of SGML document.

3.1.2 *extensible markup language (XML)*—standard from the World Wide Web Consortium (W3C) that provides for tagging of information content within documents, offering a means of representation of content in a format that is both human and machine readable. Through the use of customizable style sheets and schemas, information can be represented in a uniform way, allowing for interchange of both content (data) and format (metadata).

3.1.3 *health level 7 (HL7)*—a standards organization traditionally focused on standards for healthcare information interchange. HL7 messages are the dominant standard for peer-to-peer exchange of clinical text-based information. More recently, HL7 has developed a comprehensive object model of the healthcare enterprise and the first level of an XML clinical document architecture.

3.1.4 *HL7 clinical document architecture (CDA)*—a document markup standard for the structure and semantics of exchanged clinical documents. A clinical document is a documentation of observations and other services with the following characteristics: persistence, stewardship, potential for authentication, wholeness, and human readability. A CDA document is a defined and complete information object that can exist outside of a message and can include text, sounds, and other multimedia content.

3.1.5 *hypertext markup language (HTML)*—the language used in creating a web page. Its origin is an implementation of SGML DTD. It provides tags regarding the way a document should be displayed in the text of an HTML document, which act as commands that a browser interprets when downloading an HTML file.

3.1.6 *namespaces*—provide a simple method for qualifying element and attribute names used in XML documents. This is accomplished by associating a particular tag set by associating a prefix with a URI reference. XML namespaces provides a mechanism for authoring compound documents (documents consisting of elements and attributes from multiple DTDs or schemas) in such a way that will provide global identification without collisions of names that are the same but are used differently.

3.1.7 *parser*—a specialized software program that recognizes markup in a document and differentiates the content from

[1] This specification is under the jurisdiction of ASTM Committee E31 on Healthcare Informatics and is the direct responsibility of Subcommittee E31.35 on Healthcare Data Analysis.

Current edition approved Dec. 1, 2006. Published January 2007. Originally approved in 2002. Last previous edition approved in 2002 as E2210 – 02. DOI: 10.1520/E2210-06.

[2] http://www.w3.org

[3] http://www.w3.org/XML/

[4] http://www.w3.org/XML/Schema

***A Summary of Changes section appears at the end of this standard.**

the markup. A parser that reads a DTD and checks and reports on markup errors is a validating XML parser. A parser can be built into an XML editor to prevent incorrect tagging and to check whether a document contains all the required elements.

3.1.8 *XML Schema*—provides a means for defining the detailed structure, content and semantics of XML documents. XML Schema was approved as a W3C Recommendation approved on 2 May 2001 and with a second edition incorporating many errata was published on 28 October 2004 that provides a means for defining the detailed structure, content and semantics of XML documents. XML Schema defines the elements that can appear within the document and the attributes that can be associated with an element. It also defines the structure of the document: which elements are child elements of others, the sequence in which the child elements can appear, and the number of child elements. It defines whether an element is empty or can include text. The schema can also define default values for attributes.

3.1.9 *stylesheet*—the XSL transformations (XSLT) describes a vocabulary recognized by an XSLT processor to transform information from an organization in the source file into a different organization suitable for continued downstream processing. The extensible stylesheet language (XSL) describes a vocabulary recognized by a rendering agent to reify abstract expressions of format into a particular medium of presentation.

3.1.10 *valid XML document*—a document that is well-formed, with internal or DOCTYPE reference to element definition of tags within the document.

3.1.11 *well-formed XML document*—an XML document that conforms to the syntax as specified by the W3C XML 1.0 recommendation.

3.1.12 *World Wide Web Consortium (W3C)*—develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential as a forum for information, commerce, communication, and collective understanding.

3.1.13 *XHTML*—HTML documents that are well formed and can be processed by an XML parser.

3.1.14 *XLL/XLINK/XPOINTER*—XLL, the extensible linking language, is divided into two parts, XLinks and XPointers. XLink, the XML linking language, defines how one document links to another document. XPointer, the XML pointer language, defines how individual parts of a document are addressed. XLinks point to a URI (in practice, a URL) that specifies a particular resource. The URL may include an XPointer part that more specifically identifies the desired part or section of the targeted resource or document. XPointer, the XML pointer language, defines an addressing scheme for individual parts of an XML document. XLinks point to a URI (in practice, a URL) that specifies a particular resource. The URI may include an XPointer part that more specifically identifies the desired part or element of the targeted resource or document. XPointers use the same XPath syntax as XSL transformations to identify the parts of the document they point to, along with a few additional pieces.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *clinical practice guidelines*—systematically developed statements to assist practitioner and patient decisions about appropriate healthcare for specific clinical circumstances.[5]

3.2.2 *guideline elements model (GEM)*—an XML-based guideline document model that promotes translation of natural language guideline documents into a format that can be processed by computers. Developed at the Yale Center for Medical Informatics, GEM serves as the basis for this specification.[6]

3.2.3 *guidelines interchange format (GLIF)*—a proposed representation for guideline logic created by the INTERMED Collaboratory.[7]

3.2.4 *national guidelines clearinghouse (NGC)*—a website sponsored by the U.S. Agency for Healthcare Quality and Research that disseminates information about qualifying guidelines. It includes a structured vocabulary for describing several aspects of guidelines.[8]

3.3 *GEM Definitions*:

3.3.1 See Table A1.1 in Annex A1.

## 4. Significance and Use

4.1 *GEM Representation*—The guideline elements model (GEM) was created to unify representations created by health services researchers and by informatics specialists. Specification E2210 Schema is based on the GEM knowledge representation. It is intended to be:

4.1.1 *Comprehensive*, that is, capable of expressing all the knowledge contained in a guideline. Existing health services models of guidelines are inadequate for expressing the complexity of knowledge components in sufficient detail to facilitate electronic translations. On the other hand, existing informatics models are insufficient to model constructs that express and support guideline validity. Lack of confidence in the validity of guideline recommendations may ultimately limit end user adherence.

4.1.2 *Expressively adequate* to express the complexities and nuances of clinical medicine while remaining *informationally equivalent* to the original guideline. Most tagged elements in the Specification E2210 Schema store the actual language of the guideline, thereby remaining true to the original. Moreover, this Schema does not require recommendation knowledge to be structured in a temporal sequence, an often artificial transformation necessary for algorithmic representations.

4.1.3 *Flexible*, that is, a useful model must be able to deal with the variety and complexity of guidelines. The representation should permit modeling at high and low levels of granularity so that guidelines can be interpreted at different levels of abstraction. The Specification E2210 Schema allows markup using high-level tags or deeper analysis using elements from lower levels in the hierarchy. In addition, the open XML document model can be modified easily if necessary to accommodate missing semantic constructs.

[5] *Guidelines for Clinical Practice: From Development to Use*, Institute of Medicine, National Academy Press, Washington, DC, 1992.

[6] http://ycmi.med.yale.edu

[7] http://www.glif.org.

[8] http://www.guideline.gov

4.1.4 *Comprehensible*, that is, it should match the stakeholders' normal problem-solving language and allow domain experts to describe their knowledge with little effort. The Specification E2210 Schema markup does not require knowledge of programming. The markup process parallels physical highlighting of a document and should be learned easily by nonprogrammers.

4.1.5 *Shareable across institutions*. The use of XML for knowledge representation and markup provides unparalleled cross-platform compatibility.

4.1.6 *Reusable* across all phases of the guideline life cycle.

4.2 *Conformance*—A document is tested for conformance to this specification by a validating XML parser according to the W3C XML 1.0 recommendation.[9] A conformant document must validate without either well-formedness or validity errors, according to XML 1.0. A conformant document must also conform to constraints expressed within the prose of this specification; however, this specification does not express a formal means of testing conformance to such additional constraints. A document must be valid according to the Schema specified in this specification in order to conform to this specification.

4.3 *Use*—The Guideline Elements Model has been the subject of considerable interest and application and has become the leading exemplar of document-centered guideline knowledge representation. It has been applied by national specialty societies in the U.S. for guideline development. Shahar in Israel has employed GEM within the DeGeL architecture to create a digital guideline library.[10] In Paris, Georg and colleagues found the GEM representation to be superior to their then current guideline system (ASTI) for encoding therapeutic guidelines.GEM has been incorporated within the GUIDE architecture in Pavia, Italy; it has been used to teach informatics students about guidelines by Rector in the UK; it was incorporated within the CPGA architecture by Purves in the UK; and it is being used in New Zealand for referral guideline dissemination. In Canada, Jones has used GEM to generate tailored patient education materials and Kershaw has applied the system to create a web-enabled best-evidence retrieval system. GEM is featured and linked on the Open Clinical website in the UK.

4.3.1 Workers at Yale have found that parsing guideline recommendations into decision variables (and values), actions, and directives facilitates their encoding in controlled vocabularies such as SNOMED and LOINC and promotes the creation of rules based on the recommendations.

## 5. Procedure

5.1 *GEM Architecture*—As shown in Fig. 1, the root <schema> element contains two components: <GuidelineDocument> and <GemBasicType>. The next tier of the <GuidelineDocument> hierarchy defines a series of high-level elements that include <Identity>, <Developer>. <Purpose>,

[9] http://www.w3.org/TR/2000/REC-xml-20001006

[10] Shahar, Y., Shalom, E., Mayaffit, A., Young, O., Galperin, M., Martins, S., et al, "A Distributed, Collabaorative, Structuring Model for a Clinical-Guideline Digital-Library," Musen, M. A., editor, *AMIA 2003 Symposium*, Washington, DC, 2003, pp. 589-593.



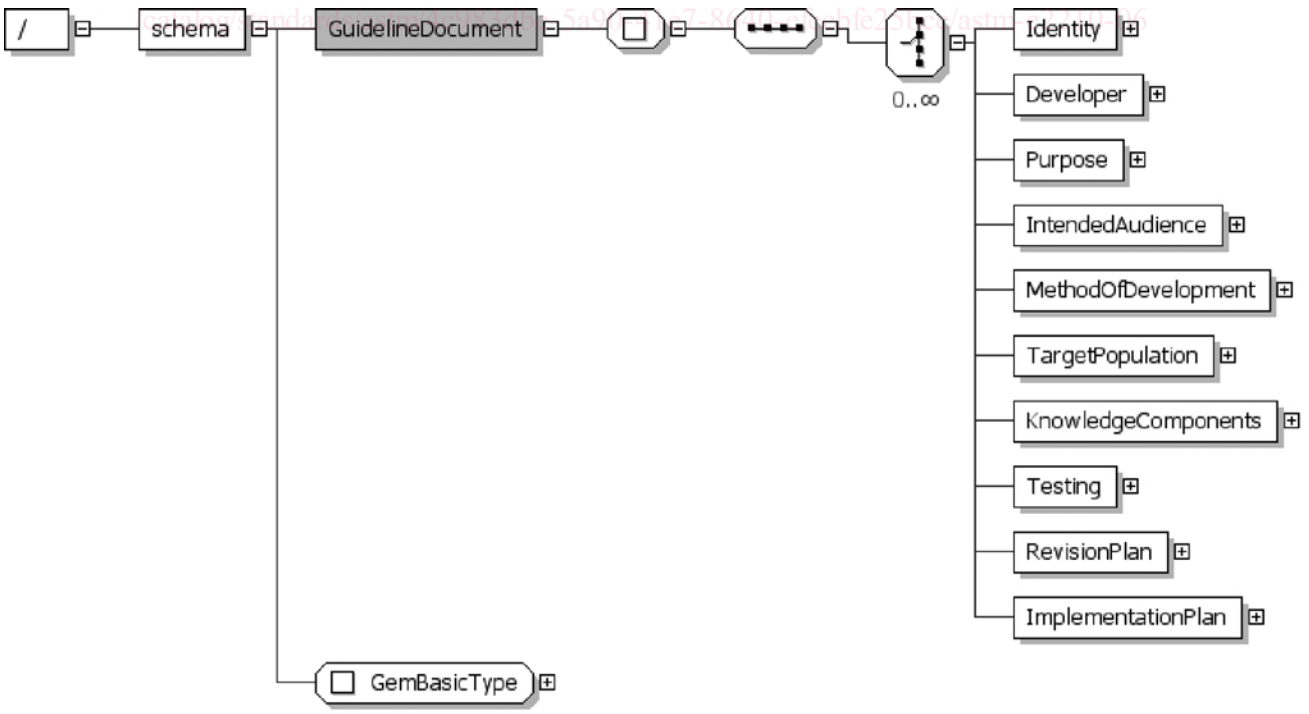


**FIG. 1 Top Level of the GEM II Schema**