MEDINFO 2007
HISA Ltd Australia
© 2007 The authors. All rights reserved.

Qualitative Characterization of Query Structure in the Arden Syntax

Robert A. Jenders

Cedars-Sinai Medical Center and the University of California, Los Angeles, USA

Abstract and objective

Context: Arden Syntax is an American National Standard Institute (ANSI) standard used to encode executable knowledge as Medical Logic Modules (MLMs). Consensus is elusive regarding a standard format for database queries in knowledge bases (KB). Objective: Analyze a sample of MLMs in order to characterize the syntax of database linkages. Method: A convenience sample of 331 MLMs drawn from 6 distinct clinical decision support systems (CDSS) was examined. MLM types included 42% lab, 23% clinical assessment and 14% medication. Results: Only half of the systems include the Structured Query Language (SQL) in queries. Only two used standard terminologies. Queries from all six systems contained at least some site-specific content, some of which was included only at run-time. Conclusion: Arden Syntax database linkages vary in format, and significant work remains to converge on a standard for Arden Syntax queries.

Keywords:

decision support systems, clinical; knowledge bases

Introduction

Arden Syntax is an HL7/ANSI standard for procedural knowledge representation, and the medical logic module (MLM) is the unit of representation. Nevertheless, some site-specific changes must occur in order for a KB in standard format to be transferred from one site to another—the "curly braces" challenge, named for the construct that encodes site-specific references in the Syntax. A database query is represented by a READ statement, which may contain site-specific references in the curly braces combined with standard operators such as aggregation and constraints. Key to minimizing the need for KB changes when sharing MLMs is the standardization of these database linkages. The present work qualitatively characterizes the structure of database queries from different institutions and CDSS. This will inform development of a standard for database linkages and estimation of the work needed to converge on the standard.

Methods

MLMs were obtained from six distinct source systems: sample knowledge bases of the 3 principal vendors of Arden-compliant CDSS and three academic medical centers with independently developed CDSS. The READ statements in the collected MLMs were isolated in order to characterize their structure. The analysis focused on these characteristics: the use of ANSI-standard SQL; whether query elements were added at run time that were not encoded in the MLM; the use of coded terminology; and the extent to which the query was represented entirely within the curly braces or partially in the standard component of the READ statement.

Results

A total of 331 MLMs were pooled from the 6 source CDSS, including 19 from 3 vendor KBs and 312 from the 3 academic medical centers. MLMs concerned mainly with lab tests were the most common (138/331 = 42%), followed by clinical assessment or classification (75/331 = 23%) and medication (45/331 = 14%). The remainder addressed administrative and miscellaneous topics. Each MLM contained at least one READ statement. Within the MLM cohort of each CDSS, query structure did not vary by MLM type. Half of the source systems used SQL to link to databases, while the others used proprietary expressions. Four of the six systems included the entire query in the READ statement, while the other two relied on partial run-time inclusion of query elements. Only two of the systems used coded vocabulary terms to refer to data elements. Each of the other four appeared to use a local, controlled terminology that was neither coded nor standard. All six source CDSS used the curly braces to represent at least part of the query; none used the standard parts of the READ statement exclusively. Three of the six included SQL queries in the curly braces, while two others reserved the curly braces for local terminology references.

Conclusion

A sample of MLMs drawn from a variety of Arden Syntax-compliant CDSS demonstrates the absence of a uniform query structure. Other standards such as SQL and controlled terminologies are used but do not predominate. Run-time inclusion of query elements complicates efforts to represent queries in a standard, shareable format. Considerable work remains to achieve convergence on a standard for representing database queries in the Arden Syntax.