Automated Extraction of Computable Clinical Decision Support Rules from an Electronic Health Record System

Branden C. Hickey, M.A.S., Marc L. Sainvil M.S., Adam M. Bartscher, Robert R. Freimuth, Ph.D., Jane L. Shellum, M.H.A., M.A.S., Davide Sottara, Ph.D.

Kern Center for the Science of Health Care Delivery

Mayo Clinic, Rochester, MN

Introduction

In 2018, Mayo Clinic replaced many of its clinical applications with a single electronic health record (EHR) system. In the process, several clinical decision support (CDS) rules were built in the new EHR.

Benefits to the Mayo Clinic practice

- Consolidated authoring environment and editing experience
- Centralized repository
- Better integration with the clinical workflow

Limitations

- The EHR does not provide robust knowledge management utilities for the CDS rules including a lack of standards-based representations
- Lack of visibility and traceability - "provenance and pedigree" of the logic
- Subject Matter Experts' (SME) inability to validate and endorse the clinical knowledge assets once they have become technical artifacts

Motivations

- As manifestations of clinical knowledge, each rule must be validated and endorsed by knowledge oversight groups for patient safety and regulatory purposes.
- Transparency of the rule logic is required to assess correctness, quality, impact, and overall outcomes.

Objectives

To address the gaps, this project aims to:

1. Develop and implement an automated pipeline (figure 2) that:
   a. Reverse engineers the native implementation of CDS event-condition-action (ECA) rules
   b. Stores the rule logic and content as discrete data elements
   c. Translates and transforms the data elements into a standard representation and presentation
2. Integrate the automated pipeline into the editorial workflow
3. Publish the logic into a formal content management system to support and facilitate the presentation (figure 3) in a human-readable, rather than technical rendering

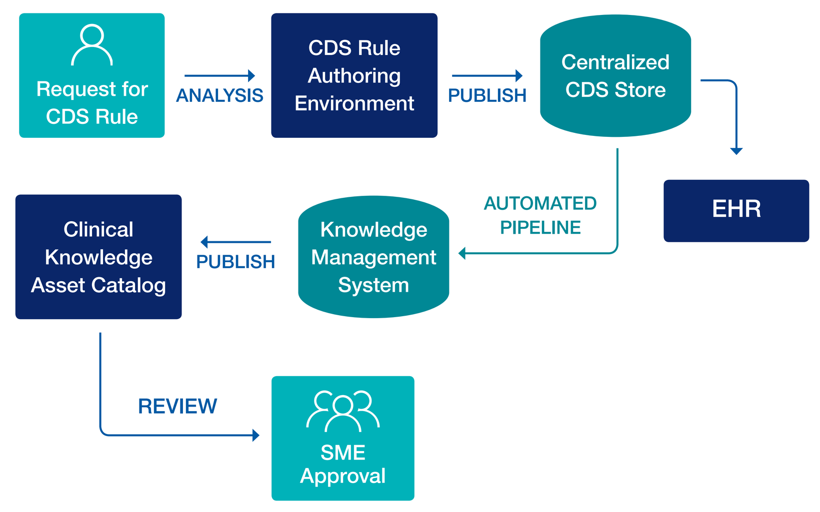

Figure 1:
CDS rule editorial workflow with integrated automated pipeline

Methods

- The most common patterns [1] used in the EHR's CDS logic were identified and then formalized using an ANTLR grammar [2].
- An in-depth analysis five rules was conducted to determine how the patterns were manifested in the EHR's configuration parameters.
- 1400 clinically relevant CDS rules were identified; 218 of these were prioritized by SMEs based on impact on the practice.
- For each rule, information was extracted and standardized using the HL7 KNART [3] metamodel. Logic was expressed using CQL [4] and validated [5].
- Coded concepts and value sets were linked and mapped in the process.
- A user-centric design study was conducted to determine the most effective way to present the logic.
- A further XSLT-based transformation rendered the KNART rules into properly styled XHTML documents.
- Metadata, including identification and versioning, was added so that both the KNART and HTML renditions can be published into a knowledge management system.

Automated Extraction Pipeline

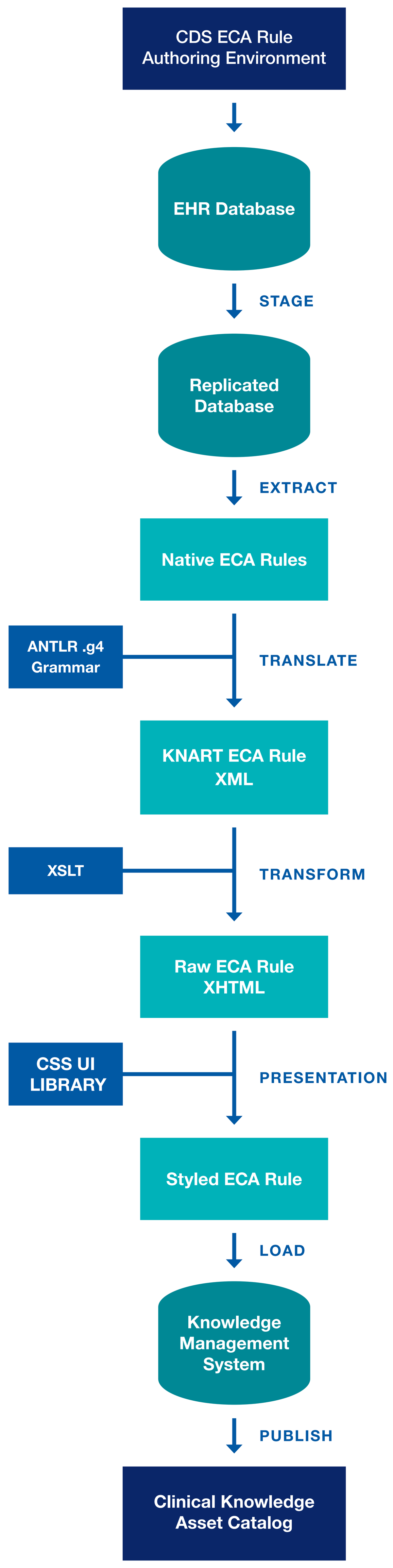

Figure 2:
The automated transformation pipeline

The first phase is authoring of the CDS rule. From there, a seven step process occurs; six steps are automated, one step is manual.

1. Stage data from CDS EHR store to a replicated DB
2. Extract relevant rule data to a separate cache
3. Translate the data into a XML standard representation KNART ECA specification
4. Transform XML using a XSLT stylesheet into XHTML web language
5. Apply final styling for presentation
6. Manually load into a knowledge management system
7. Publish into a knowledge asset catalog for subject matter expert review and validation

Clinical Decision Rule View

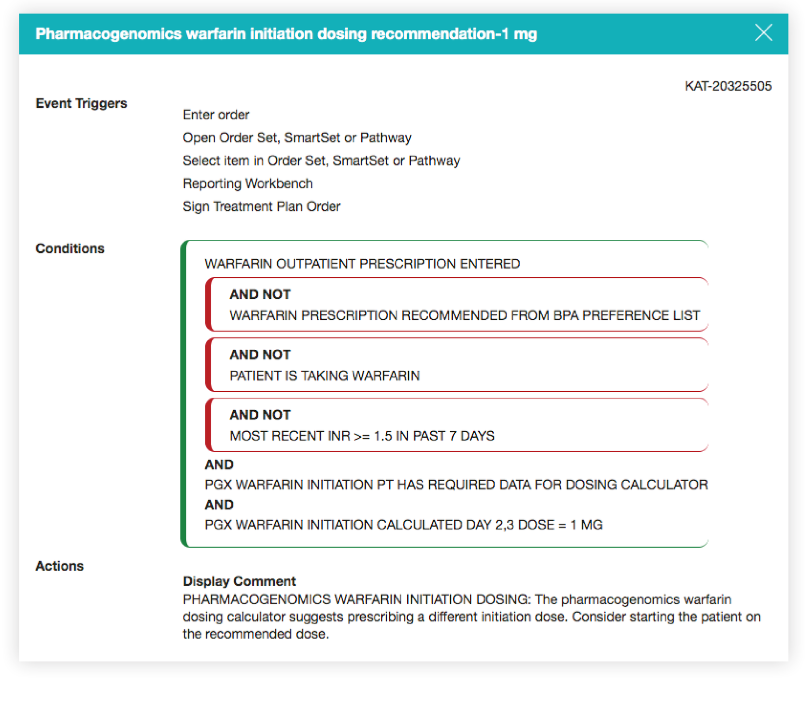

Figure 3:
The subject matter expert view of an extracted CDS ECA rule

The condition section coloring shows green for positive statements, and red for negated ones. In general, the color coding is aligned to the underlying logic formalism (CQL).

Results

Of the 218 CDS ECA rules selected for extraction, 211 (96.79%) have been loaded into our knowledge management system using the automated pipeline. The process takes, on average, 20 minutes (i.e. <10 seconds/rule)

Without automation, the analysis, extraction, normalization and publication into the KCMS requires ~4 hours/rule of a clinical knowledge engineer specialized in clinical Informatics.

The extraction process is repeatable for individual rules, or batches thereof, and supports versioning as new rules and/or versions thereof need to be published.

Discussion/Conclusion

The preliminary implementation of this automated pipeline proved the feasibility of the approach. Even if the initial set of published rules is relatively small, the process is repeatable as CDS rules tend to follow similar patterns [1]

The current combination of the machine and human readable representations (KNART and HTML, respectively) enabled several scenarios:

- Subject matter experts can verify that the actual implementation matches their requirements, and can easily propose updates.
- Providers impacted by the CDS rules have been able to understand why and how certain recommendations have been delivered.
- Conversely, reporting and analytics tools can be built to detect changes in the EHR implementation, and to identify commonalities.

However, there are several limitations and opportunities for enhancements:

- The CQL-based formalization of the logic is limited to Boolean operator and simple expressions; atomic expressions (e.g. most recent INR) are treated as black boxes.
- The ability to standardize rules is hampered because the EHR uses proprietary codes as opposed to standard code systems (e.g. SNOMED-CT).
- A few rules (7) were not published due to the complexity of the logic, which resulted in a presentation form unfit for review. The pipeline currently does not support strategies to simplify/optimize the logic and/or its presentation.
- The methodology is tailored to ECA rules with business priority. We expect gaps as we plan to scale to the other 1400 CDS rules and other clinical knowledge assets (order sets, health plans) implemented within the EHR system.
- The inability to push the documents into the knowledge management system limits the scalability of the approach.
- As an alternative to KNART, the methodology needs to be migrated to FHIR, using the Resource PlanDefinition [6] which derived from the same conceptual model [7] and other Clinical Reasoning resources as they become more mature.

This extraction pipeline gives the ability to take daily, automated snapshots of the CDS rule logic in the system, capturing any changes to the underlying records. This will, in turn, give the ability to better understand the impact of the CDS rules, the behavior of the system, and the people interacting with it. This will prove invaluable to the Mayo Clinic practice.

References

1. Greenes RA, Sordo M, Zaccagnini D, Meyer M, Kuperman GJ. Design of a standards-based external rules engine for decision support in a variety of application contexts: report of a feasibility study at Partners HealthCare System. Stud Health Technol Inform. 2004;107(Pt 1):611-5.
2. ANTLR. (n.d.). Retrieved from https://www.antlr.org/index.html.
3. Health Level Seven International: HL7 Standard: Clinical Decision Support Knowledge Artifact Specification, Release 1.3 (Jul 2015), available at: http://www.hl7.org/implement/standards/
4. Health Level Seven International: HL7 Standard: Clinical Quality Language Specification, Release 1 STU4 (CQL 1.4), available at: https://cql.hl7.org/
5. Mohammad Hekmatnejad, Andrew M. Simms, Georgios Fainekos Model Checking Clinical Decision Support Systems Using SMT
6. Health Level Seven International: HL7 Standard: FHIR, Release 4, available at: https://www.hl7.org/fhir/plandefinition.html
7. Kawamoto, Ken et al. Health eDecisions (HeD): a Public-Private Partnership to Develop and Validate Standards to Enable Clinical Decision Support at Scale.