Evaluating and classifying the readiness of technology specifications for national standardization

RECEIVED 17 March 2014
REVISED 8 May 2014
ACCEPTED 10 May 2014
PUBLISHED ONLINE FIRST 28 May 2014

Dixie B Baker[1], Jonathan B Perlin[2], John Halamka[3]

ABSTRACT

The American Recovery and Reinvestment Act (ARRA) of 2009 clearly articulated the central role that health information technology (HIT) standards would play in improving healthcare quality, safety, and efficiency through the meaningful use of certified, standards based, electronic health record (EHR) technology. In 2012, the Office of the National Coordinator (ONC) asked the Nationwide Health Information Network (NwHIN) Power Team of the Health Information Technology Standards Committee (HITSC) to develop comprehensive, objective, and, to the extent practical, quantitative criteria for evaluating technical standards and implementation specifications and classifying their readiness for national adoption. The Power Team defined criteria, attributes, and metrics for evaluating and classifying technical standards and specifications as 'emerging,' 'pilot,' or 'ready for national standardization' based on their maturity and adoptability. The ONC and the HITSC are now using these metrics for assessing the readiness of technical standards for national adoption.

Key words: Standards, Meaningful Use, Certification

PERSPECTIVE

BACKGROUND

The American Recovery and Reinvestment Act (ARRA)[1] of 2009 clearly articulated the central role that health information technology (HIT) standards would play in improving healthcare quality, safety, and efficiency. By offering financial incentives for adopting certified electronic health record (EHR) technology and using it meaningfully, ARRA launched the healthcare industry on a transformational path and aggressive timeline toward the realization of a nationwide HIT infrastructure. ARRA assigned responsibility for creating this infrastructure to the Office of the National Coordinator for Health Information Technology (ONC), with support from two new federal advisory committees—the HIT Policy Committee and the HIT Standards Committee.

Success rested firmly on the development of national standards against which EHR technology could be certified—standards for enabling technical, semantic, and operational interoperability so that healthcare information could be privately and securely exchanged among, and effectively used by, healthcare providers, researchers, public health, and consumers across the USA. The decision to adopt (or not to adopt) a technology standard or implementation specification as a national standard would have enormous impact on the industry. The choice of a standard could affect the usability and cost of the technology that is produced, the intelligibility and security of health information, the convenience of operational workflows, and the safety, quality, and availability of care. The 'Initial Set of Standards, Implementation Specifications, and Certification Criteria for Electronic Health Record Technology'[2] published on July 28, 2010, represented the first step in an incremental approach to adopting national standards for certifying EHR technology to be implemented and meaningfully used by eligible professionals and eligible hospitals seeking to qualify for Stage 1 incentive payments under the Medicare and Medicaid EHR Incentive Program. With the plethora of existing and emerging standards available, there was a need for an objective process and set of metrics for evaluating and classifying HIT standards and implementation specifications relative to their readiness for adoption as national standards.

Evaluation of Nationwide Health Information Network pilot standards

ARRA's 'meaningful-use' incentives sought to attract a broad population of providers ranging from small clinical practices to large integrated delivery networks, each expecting to use the Nationwide Health Information Network (NwHIN) to securely exchange clinical information with other providers and with state and federal partners. The country needed a well-defined set of transport, security, and content components from which a provider organization could select those most suitable for meeting its information-exchange needs. The two initial pilots for what would ultimately become the NwHIN had produced

Correspondence to Dr Dixie B Baker, 2034 Eisenhower Ave., Suite 270, Alexandria, Virginia 22314-4678, USA; dixie.baker@martin-blanck.com

For numbered affiliations see end of article.

a number of implementation specifications, and the ONC wanted to know which of these specifications were ready to be adopted as national standards, which specifications needed further development and piloting, and which specifications were probably inappropriate for broad adoption across multiple use cases.

So in early 2011, the ONC asked the Health Information Technology Standards Committee (HITSC) to assemble an NwHIN Power Team for the purpose of evaluating the implementation specifications developed through the Exchange[3] and Direct[4] pilots with respect to their usability and scalability to support nationwide health information exchange. The Power Team would recommend those specifications that could be integrated and deployed to support the secure transport and exchange of electronic health information at a national scale, and would identify where further work may be needed. Outputs from this work would help inform ONC decisions regarding future investments in additional pilots and further specification development. The focus of this work was at the national level—any use of the specifications within enterprises or among partners within a regional health information exchange was not addressed.

The evaluation included 10 technical specifications developed for the Exchange pilot and two specifications developed by the Direct Project. From this initial work, the Power Team defined five criteria for evaluating the readiness of each of the 14 specifications to be adopted as a national standard:

1. Need (low, moderate, high)
2. Maturity of Specification (low, moderate, high)
3. Maturity of Underlying Technology (emerging, maturing, mature, declining)
4. Deployment/Operational Complexity (low, moderate, high)
5. Market Adoption (low, moderate, high)

The Power Team's recommendations were used to inform both the development of the 2014 edition of EHR standards and certification criteria[5] and ONC's investment in the development of a transport specification based on a representational state transfer (REST) approach. In addition, the ONC recognized the potential value of building on the Power Team's initial work to more completely define objective measures for evaluating and classifying the readiness of standards and implementation specifications to be adopted at the national level.

METHODS

In February 2012, the ONC requested that the NwHIN Power Team be reconvened for the purpose of building upon this initial work to develop comprehensive, objective, and, to the extent practical, quantitative criteria for evaluating technical standards and implementation specifications and classifying their readiness for national adoption. The evaluation criteria and classification process would provide the ONC with a granular and robust methodology for evaluating and classifying standards and implementation specifications, and for identifying those standards that would comprise a portfolio of nationally recognized standardization activities. The criteria and process would need to be applicable to the evaluation and classification of standards for vocabulary, content, transport, services, and security.

The NwHIN Power Team defined an approach that included:

1. Review of relevant work
2. Refinement of criteria identified during NwHIN pilot standards evaluation
3. Definition of attributes that characterize each of the criteria
4. Identification of a set of objective and, to the extent practical, quantitative metrics for measuring each attribute
5. Definition of an evaluation and classification process for using the criteria and metrics to determine the suitability and readiness of technical standards and implementation specifications for adoption as national standards
6. Evaluation of the reliability and validity of the metrics and process

REVIEW OF RELEVANT WORK

With research assistance from the ONC support team, a search was conducted for relevant work on methods and criteria for assessing the maturity and experiential use of technology standards and implementation specifications. Results of the search were presented to the NwHIN Power Team members for discussion, assessment of applicability and utility, and suggestions on ways to refine the search. The search found criteria and methods for evaluating the readiness of technology and the maturity of processes, but very little work addressing the assessment of standards and implementation specifications. For many years, US government agencies and many private companies have used technology readiness level (TRL) as a metric for assessing the maturity of a particular technology. From the original TRL developed by the US National Aeronautics and Space Administration (NASA)[6] as a metric for rating the readiness of technology for use in space, many other US government agencies have adapted their own TRL scales to measure the maturity of a broad range of technologies. A complementary integration readiness level (IRL) also has been defined.[7] Both of these focus on the maturity of specific technologies and not the maturity of the standards and specifications that define these technologies. However, the Power Team recognized the value of the TRL as a useful metric for evaluating the maturity of technology components that may be incorporated in a standard or specification.

For decades, the capability maturity model integration (CMMI) model[8] has been used to measure the maturity of the processes an organization uses to develop software. However, the CMMI model does not assess the maturity of the standards and specifications an organization uses or their readiness for any particular purpose.

In 2007, the Health Information Technology Standards Panel (HITSP) defined a set of criteria for use in evaluating the readiness of standards to be selected for inclusion in interoperability specifications.[9] For each specific use case, the HITSP would assess the readiness of various standards for use together as

an interlocking (harmonized) set that would both meet the functional requirements of that use case and be compatible with standards HITSP had previously selected for other use cases. Based on an assessment of the readiness of various standards, the HITSP would select a group of standards judged most ready for use as an interlocking (harmonized) set to implement for a specific use case, while remaining compatible with existing HITSP standards selections across all use cases. The criteria were organized into six categories: suitability, compatibility with other HITSP harmonized standards, preferred standards characteristics, preferred standards developer organization and process, data element usage, and expected total costs of implementation and conformance criteria. Preferred characteristics of standards included approved, widely used, readily available, technology neutral, supporting uniformity, demonstrating flexibility, and international usage. Although the HITSP criteria focused on suitability for inclusion in a harmonized set of standards designed for a specific use case, some of the categories and characteristics identified are also applicable to evaluating and classifying the readiness of an individual standard or specification for adoption more generally, and therefore served to inform the Power Team's work.

Also useful is the Qualification and Selection of Open Source Software (QSOS) methodology[10] developed for use in assessing the functional fit and quality of open source software—essentially its readiness to be selected for a given purpose. The QSOS methodology is licensed under a free Gnu Free Documentation License (FDL) and considers factors such as maturity, adoption, development community and leadership, activity, training, support, documentation, platforms supported, modularity, licensing, maintainability, and future direction—all important factors in assessing the readiness of a specification to become a national standard. In addition, the QSOS methodology defines metrics for measuring these factors, which proved useful to the Power Team's effort.

REFINEMENT OF CRITERIA

Based on lessons learned from the pilot standard evaluation, and drawing from previous relevant work, the Power Team made several adjustments to its initial set of evaluation criteria. First, the team realized that 'Need' was not suitable as a criterion for evaluating the merits of a given standard or specification, but rather should be a consideration in deciding what standards are subjected to the evaluation process. Second, the team observed that deployment complexity and operational complexity are really two separate properties since a standard that may be difficult to implement may be quite simple to maintain during operations, and vice versa. Third, the team realized the need to consider intellectual property issues, which can be challenging during both deployment and operations. Fourth, recognizing the importance of maintaining symmetry in the rating scales, the team decided to uniformly adopt the 'low, moderate, high' rating scale, where 'low' was always least desirable and 'high' most desirable—and for semantic consistency with these ratings, changed the two 'complexity' criteria to measure 'ease' (ie, lack of complexity) instead.

DEFINITION OF ATTRIBUTES

Thus the refinement process produced six criteria—three of which represented maturity properties, and three how easily the specification or standard could be implemented and adopted. Each criterion was then characterized by a set of attributes, as identified below.

Maturity criteria

1. Maturity of Specification—the maturity of the specification or standard itself, considered apart from the technologies used to implement it

Attributes:

- Breadth of support
- Stability
- Adoption of specification

2. Maturity of Underlying Technology Components—the maturity of the technologies used in the specification or standard, including any 'nested' technologies upon which the primary technologies may depend, and the platforms that support these technologies

Attributes:

- Breadth of support
- Stability
- Adoption of technology
- Platform support
- Maturity of technology within its life cycle

3. Market Adoption—how widely the specification or standard has been adopted, both within and outside healthcare, and projections for its future adoption

Attributes:

- Installed healthcare user base
- Installed user base outside healthcare
- Interoperable implementations
- Future projections and anticipated support
- Investment in user training

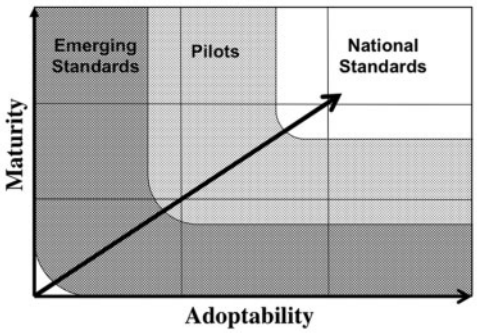
Adoptability criteria

1. Ease of Implementation and Deployment—how easily the technology specified can be implemented and deployed for use

Attributes:

- Availability of off-the-shelf infrastructure to support implementation
- Standard as success factor
- Conformance criteria and tests
- Availability of reference implementations
- Quality and clarity of specifications
- Specification modularity
- Separation of concerns
- Ease of use of specification
- Degree to which specification uses familiar terms to describe 'real-world' concepts

Figure 1: Standards and implementation specifications classification grid.



- Runtime decoupling
- Appropriate optionality

2. Ease of Operations—once the specified technology has been implemented, how easy it is to operate and maintain on a day-to-day basis

Attributes:

- Comparison of targeted scale of deployment to actual scale deployed
- Number of operational issues identified in deployment
- Degree of peer-coordination needed
- Operation scalability (ie, operational impact of adding a single node)
- Fit to purpose

3. Intellectual Property—how open and accessible the specification is for implementers

Attributes:

- Openness
- Affordability
- Licensing permissiveness
- Copyright centralization
- Freedom from patent impediments

IDENTIFICATION OF METRICS

Each of the attributes was characterized by a set of metrics representing 'low,' 'moderate,' and 'high' degrees of that attribute. These metrics were designed to be as objective and measurable as possible. The maturity and adoptability metrics are available in the online supplementary file.

The NwHIN Power Team conducted an evaluation exercise using these metrics to assess the readiness of the HL7 'Infobutton' specification[11] to become a national standard. Lessons learned from that exercise were factored into the final recommendations transmitted to the ONC for evaluating the readiness of technical standards and implementation specifications to become national standards.[12]

DEFINING AN EVALUATION AND CLASSIFICATION PROCESS

In May 2012, the Department of Health and Human Services published a Request for Information (RFI) entitled 'Nationwide Health Information Network: Conditions for Trusted Exchange'[13] that included a section that asked questions about a proposed process for classifying technical standards and implementation guides into three classes:

1. 'Emerging'—technical standards and implementation specifications that still require additional specification and vetting by the standards development community, have not been broadly tested, have no or low adoption, and have only been implemented with a local or controlled setting
2. 'Pilot'—technical standards and implementation specifications that have reached a level of specification maturity and adoption by different entities such that some entities are using them to exchange health information either in a test mode or in a limited production mode
3. 'National'—technical standards and implementation specifications that have reached a high level of specification maturity and adoption by different entities

The RFI included the grid shown in figure 1, depicting how a specification might be classified based on maturity and adoptability.

The RFI proposed an annual process to identify, review, and assess standards and implementation specifications using objective criteria and cited the NwHIN Power Team's on-going work to define such criteria.

The Power Team included these classification categories and grid in its final recommendations, which were presented to and endorsed by the HITSC in August 2012. The final recommendations stressed that the metrics should not be used to generate a 'score' as input to a numeric 'average' or to determine whether a minimum threshold has been reached, rather the evaluation process and metrics should be used to provide structure and discipline to what is essentially a qualitative evaluation and classification process. The team noted that the metrics are most effectively used to inform and justify a classification decision, rather than as a quantitative measure of goodness.

The Power Team recommended that the ONC select standards and specifications for evaluation based on industry needs for specific use cases, and that a description of the use cases be given to the evaluation team along with the standards and specifications. A lesson learned from the Infobutton evaluation exercise was that the specification being evaluated needs to be functional; if the specification itself is non-functional, then implementation guidance should be included as part of the evaluation. If alternative specifications exist for the same functionality, then a comparative evaluation may be considered.

EVALUATING THE RELIABILITY AND VALIDITY OF THE METRICS AND PROCESS

The Power Team recommended that the evaluation program incorporate a retrospective process for examining adopted

standards within the context of actual use in order to identify those that might need to be re-evaluated for potential retirement or replacement, and for refining the process and metrics as needed over time. Also, such retrospective evaluation should assess the validity and reliability of these metrics for use as a prediction rule to predict the future success of a new national standard.

CLASSIFICATION METRICS IN USE

The ARRA assigned to the ONC responsibility for reviewing and deciding whether to endorse each standard, implementation specification, and certification criterion for electronic exchange and use of health information recommended by the HITSC, and for reviewing federal HIT investments to ensure that they are meeting the objectives of the strategic plan.[1] Within the constraints of ever tightening federal budgets, it is essential that we derive the most value possible from our HIT investments. Further, a highly competitive marketplace demands that the standards we adopt and the investments we make in standards-development activities be selected based upon objective criteria and transparent methods that encourage innovation and are free of bias. The criteria and metrics developed by the HITSC NwHIN Power Team provide the ONC with a good, initial set of objective criteria and metrics, and a methodology for assessing existing and emerging standards to determine their readiness to become national standards, and the need for investments in standards-development efforts.

Since the HITSC transmitted the NwHIN Power Team's recommendations to the ONC, the HITSC has used these metrics to support its decision-making regarding what standards and implementation guides to recommend for adoption in regulations, and where ONC needs to consider investing in the development of new standards or the acceleration of the development of emerging standards. The ONC Standards and Interoperability Framework project teams also are using these metrics to guide their choices of standards and to measure the maturity of the reference implementations they produce. These metrics are not intended to produce a pass-or-fail 'grade,' but rather to provide an objective basis for substantiating our standards selections and for justifying our investment decisions. We expect to continue to refine these metrics over time as we continue to gain experience and insights through their use.

ACKNOWLEDGMENTS

We wish to thank the members of the Nationwide Health Information Network Power Team, the Health Information Technology Standards Committee, and the Office of the National Coordinator for Health Information Technology support team for their contributions to this work.

CONTRIBUTORS

DB led the Nationwide Health Information Network Power Team in defining these metrics and is the primary author of this paper. JH and JP chaired the Health Information Technology Standards Committee under which this work was performed, and led the review and endorsement of the recommendations reported herein.

COMPETING INTERESTS

None.

PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed.

SUPPLEMENTARY MATERIAL

Supplementary material is available online at http://jamia.oxfordjournals.org/.

REFERENCES

1. 111th Congress. American Recovery and Reinvestment Act. January 6, 2009. http://www.gpo.gov/fdsys/pkg/BILLS-111hr1enr/pdf/BILLS-111hr1enr.pdf (Accessed May 2014).
2. Department of Health and Human Services. Health Information Technology: Initial Set of Standards, Implementation Specifications, and Certification Criteria for Electronic Health Record Technology. *Fed Regist*. 2010; 75:144.
3. Office of the National Coordinator for Health Information Technology. Nationwide Health Information Network Exchange. Specifications linked to http://www.healthit.gov/policy-researchers-implementers/nationwide-health-information-network-exchange (Accessed May 2014).
4. The Direct Project. Specifications linked to http://directproject.org/content.php?key=getstarted (Accessed May 2014).
5. Department of Health and Human Services. Health Information Technology: Standards, Implementation Specifications, and Certification Criteria for Electronic Health Record Technology, 2014 Edition; Revisions to the Permanent Certification Program for Health Information Technology. *Fed Regist*. 2012;77:171. http://www.gpo.gov/fdsys/pkg/FR-2012-09-04/pdf/2012-20982.pdf (Accessed May 2014).
6. National Aeronautics and Space Administration (NASA). Technology Readiness Levels Demystified. August, 2012. http://www.nasa.gov/topics/aeronautics/features/trl_demystified.html (Accessed May 2014).
7. Sauser B, Gove R, Forbes E, *et al*. Integration maturity metrics: Development of an integration readiness level. *Info Knowledge Syst Manag*. 2010;9:17–46.
8. Software Engineering Institute, Carnegie-Mellon University. Capability Maturity Model Integration. http://www.sei.cmu.edu/cmmi/ (Accessed May 2014).
9. Healthcare Information Technology Standards Panel. HITSP Standards Readiness Criteria. HITSP 07 N 191. Version 1.9.6. April, 2007. http://publicaa.ansi.org/sites/apdl/Documents/Standards%20Activities/Healthcare%20Informatics%20Technology%20Standards%20Panel/HITSP%20Policy%20Documents/HITSP%2007%20N%20191%20Readiness%20Criteria%20Tier2%20v1.9.6.pdf (Accessed May 2014).

10. Atos Origin. Method for Qualification and Selection of Open Source Software (QSOS), Version 1.6. April 2006. http://www.qsos.org (Accessed May 2014).
11. Health Level Seven. Product Infobutton. http://wiki.hl7.org/index.php?title=Product_Infobutton#Product_Name_-_HL7_Version_3_Standard:_Context-aware_Information_Retrieval_.28Infobutton.29.2C_Release_1 (Accessed May 2014).
12. Health Information Technology Standards Committee. NwHIN Power Team: Final Recommendations for Standards Evaluation Criteria. 2012. http://www.healthit.gov/sites/default/files/archive/HIT%20Standards%20Committee/2012/2012-08-15/2012Aug15_HITSC_NwHINPT%28Baker%29.pptx (Accessed May 2014).
13. Department of Health and Human Services. Nationwide Health Information Network: Conditions for Trusted Exchange. Request for Information published May 5, 2012. https://www.federalregister.gov/articles/2012/05/15/2012-11775/nationwide-health-information-network-conditions-for-trusted-exchange (Accessed May 2014).

AUTHOR AFFILIATIONS

[1]Martin, Blanck & Associates, Alexandria, Virginia, USA

[2]HCA, Inc. (Hospital Corporation of America), Nashville, Tennessee, USA

[3]Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA