ASSESSMENT TECHNOLOGY INTEROPERABILITY REQUEST FOR INFORMATION

SUMMARY, ANALYSES, AND OBSERVATIONS

U.S. DEPARTMENT OF EDUCATION
WASHINGTON, DC 20202

APRIL 27, 2011

INTRODUCTION

On December 20, 2010, the U.S. Department of Education (Department) published in the *Federal Register* a Request for Information (RFI) on the subject of technology standards and interoperability for assessment (http://www2.ed.gov/legislation/FedRegister/other/2010-4/122010e.html). In that notice, the Department indicated that it anticipated using this information to help determine the appropriate interoperability standards for assessments and related work developed under the Race to the Top Assessment (RTTA) program. The Department also expects to use this information to help in the development of related standards-based programs.

The purpose of this document is to provide a summary of the responses to the RFI and to provide some overall observations from the Department regarding:

1) Interoperability and technology standards for assessment and related learning systems;
2) Areas where interoperability between assessment systems and learning systems may create a more effective and seamless digital educational system; and
3) A potential processes for undertaking the development of assessment technology standards.

In this summary document, the Department provides information for organizations interested in developing next-generation assessments, including observations as to where technology standards could be developed or selected in the assessment life cycle and on how those technology standards could affect broader work in the development of educational technology.

The intent of this document is to provide information and observations that may be useful to grantees under RTTA and other programs, members of the education community who are addressing difficult issues in the area of creating interoperable technology, and the broader public. This document does not impose any specific standards that must be used under any specific program, nor is it intended to endorse any specific approaches, methodologies, products, or organizations.

From reading the RFI responses and conducting internal research on interoperable technology, we conclude that there are many areas of assessment functionality that could be enhanced or enabled by effective assessment technology interoperability. Areas that show promise in advanced assessment functionality include:

- Measuring results of multiple, alternative learning progressions against the same curricular standards.
- Providing adaptive measurements for learning progressions.
- Enabling or enhancing simulations, games, and other advanced interactive environments as part of assessment and measurement solutions.
- Enhancing “artificial intelligence” technology for scoring of assessments.
- Enhancing the use of multiple, different input/output/interactive hardware and software devices for measuring student achievement, depending on student need or organizational infrastructure and capability.

- Enabling the broader use of "electronic learning records" that can securely share data between authorized assessment, learning, and administrative systems.

The development of this summary involved review, analysis, and consolidation of 22 responses to this RFI. The complete text of the responses to this RFI is available at http://www.ed.gov/oii-news/interoperable-assessment-technology-standards-public-responses.

AREAS OF ASSESSMENT INTEROPERABILITY

The Department's analysis of the RFI responses shows five main areas of potential assessment interoperability:

I. Assessment Instruments and Items: Format and Packaging
II. Initiation and Return of Assessment Administrations
III. Administration of Assessments
IV. Learning Outcomes Management
V. Learning Records Management

Below, we briefly define each of these five areas, provide a short summary of the responses received to the RFI, and highlight examples of how we might standardize assessment interoperability.

I. Assessment Instruments and Items: Format and Packaging

A. **Definition:** Format and packaging of assessment items refers to the content and material necessary to provide a stimulus for student response in a desired format. This includes the criteria to evaluate the responses. Multiple items aggregated together comprise the assessment "instrument."

Packaging also includes standard representations of metadata used for classification and discovery, such as what curricular standards are measured; what activities are used during the measurement; what format is used for measurement activity; pricing information; and copyright licensing information.

B. **Summary and Analysis:** Respondents indicated that there are many existing standard interoperable formats for assessment items and instruments, including: Question and Test Interoperability (QTI), Common Cartridge, and Sharable Content Object Reference Model (SCORM).

Several respondents also indicated that there were some limitations to the use of both QTI and SCORM, such as with standardization of how items are "tagged" in the system and the values permitted for tags. A number of respondents suggested that QTI was close to a complete solution for packaging assessment content. No respondent indicated that a complete solution was currently available.

Several respondents also contended that universal design for learning (UDL) provides a framework that may be useful for creating standard format and packaging of items that

are accessible to all students. In particular, several respondents cited the Accessible Portable Item Profile (APIP) project, which is based on QTI, as noteworthy for providing a way to standardize accessibility options in assessment items and instruments.

C. **Example 1:** In a print-based environment, assessment instruments and items could be a set of test questions grouped into a simple multiple-choice assessment, printed into a pre-defined layout, with associated information regarding what skills are being assessed.

Example 2: In a digital format, assessment instruments and items could be discretely stored and shared using formats, including QTI and SCORM.

Example 3: In a more advanced digital format, assessment instruments and items could be stored in a Flash or HTML5 interactive application. By itself, the use of Flash or HTML5 might be considered non-interoperable, but if integrated with standards for "Initiation and Return of Assessment Administrations" and "Learning Outcomes Management" (see below), such an activity could be interoperable. This example shows that selection of format and packaging standards must be considered in light of a full assessment life cycle (from initiation of assessment to the management of the results).

II. **Initiation and Return of Assessment Administrations**

A. **Definition:** Initiation and return of assessment administration is the process of starting an assessment activity and, once the activity is complete, returning the results from the assessment activity. (Note that this area does not include the format of the data sent during initiation or return; see Learning Outcomes Management below.)

B. **Summary and Analysis:** Respondents cited Learning Tools Interoperability **(**LTI) and variants as an example of an existing set of standards in the initiation and return of assessment administration that holds promise.

Respondents highlighted several of the complex issues involved in the standardization of initiation and return processes, including: 1) how to conduct secure and private communication between assessment systems, and 2) how to transfer information and the types and format of information that could be transferred between assessment systems.

C. **Example 1:** In a print-based environment, initiation and return of assessment administration could be as simple as a teacher verbally instructing students to take the test and then physically passing out the instruments.

Example 2: In digital environments, initiation and return of assessment administration could include the ability for a learning system to electronically transfer the student to a system that administers the assessment. This process could include both the handing off of the student to another system for assessment administration and for activities associated with the conclusion of the assessment activity, such as returning any results.

III. **Administration of Assessments**

A. **Definition:** Administration of assessment is the method of providing a stimulus (or "a prompt") to the student and providing a means to collect responses to the stimulus from the student.

B. **Summary and Analysis:** A number of respondents suggested that administration of assessments should not be standardized at this time. This is primarily because administration of assessments is an area of rapid innovation and often a stand-alone activity that does not require interoperability.

Some respondents also indicated that the content of specific assessment items and instruments should be standardized and that the items should be interoperable (i.e., format and packaging so that they may be used on multiple technology platforms). Respondents felt that standardizing the specifications for how these items and instruments are administered to the student (e.g., screen size, media type, and accessibility) was of less importance than the format and packaging because administration of assessment is an area of rapid innovation and not standardizing too specifically will enable continued innovation and market development. Multiple respondents cautioned, however, about the possible impact on validity of changes within how any single item is presented. Small changes in the administration of an assessment (e.g., font size, Internet connection speed) may impact whether the item is consistently measuring what it is intended to assess. Those individuals who are developing next-generation assessments will need to have a method or process to ensure the validity and reliability of the results.

C. **Example 1:** In a traditional print-based environment, the administration of the assessment could be as simple as the student reading the assessment booklet for questions and filling out the appropriate bubble on the response sheet.

Example 2: In digital environments, the administration of the assessment could be a student being given a complex interactive simulation and the assessment system observing the student's actions and determining his or her level of knowledge and ability without providing any traditional "prompted" stimuli or questions. The assessment system could internally convert this determination into a standard format for use by the "Learning Outcomes Management" system (see below).

IV. Learning Outcomes Management

A. **Definition:** Learning outcomes management is the collection and examination of assessment responses to determine the degree to which the student demonstrates the capabilities intended to be measured by the assessment.

B. **Summary and Analysis:** Respondents indicated that learning outcomes management data should include metadata about assessment metrics (e.g., the scale used for scoring and calculating the conditional standard error of measurement or item difficulty) and data about the student's specific assessment results (e.g., what score was achieved and on what scale; what standards were assessed; and what error bars are associated with the measurement).

A number of respondents indicated that there is a grey area between the collection of responses within an assessment and the scoring of the assessment for outcomes. Many respondents emphasized the importance of permitting flexibility in the administration of assessments while also standardizing the format and techniques for communicating assessment results. The general consensus of respondents was that if the outcome of an assessment is standardized in terms of format, content, and reporting results, then it is not significantly important that the method used for administering the assessment be standardized. Respondents also maintained that having standards for how items are characterized and how results are reported would create a sufficient level of interoperability while still permitting innovation in how test items are presented to the student.

There was also general consensus among respondents that standardization of reporting student results is important for the market and for the development of effective next-generation assessments because it allows technology providers to support each other while still providing market-differentiating innovation in the actual assessment administration and in the development of other student-focused learning solutions (e.g., tailored tutoring opportunities based on the student's assessment results).

C. **Example 1:** In a traditional print-based environment, learning outcomes management could involve collecting the response sheets from the students, feeding them into an automated scanning machine, and producing a report of the results for each student.

Example 2: In digital environments, a learning outcomes management system could include sending student data to an assessment system to determine what content areas should be tested. The assessment system would have the capacity to tailor an assessment to the student's learning needs and report results in a standard format that permits the receiving system to know what capabilities the student demonstrated, and to what degree, and also to generate assessment reports. This is distinguished from "Initiation and Return of Assessment Administrations" above in that the "sending" and "return" of data here defines the format or type for the data returned (rather than serving only as a transport mechanism for returning results data). It also may be necessary to develop multiple learning outcomes formats that are standardized for different purposes, such as educational levels (e.g., kindergarten vs. high school assessments).

V. **Learning Records Management**

A. **Definition:** Learning records management is the collection of outcome data along with other relevant student data for purposes of analyzing, reporting, aggregating, or processing these data within a larger learning management solution (in some cases involving more than one learning system).

B. **Summary and Analysis:** Some respondents indicated that the Schools Interoperability Framework Association (SIFA), a membership organization dedicated to creating rules and definitions for interoperability across education software programs, currently provides many capabilities to support learning records management, though it might not

currently provide a complete solution in highly heterogeneous environments with myriad learning management systems.

C. **Example 1:** In a traditional print-based environment, learning records management could include putting a copy of the assessment results into a grade book alongside other performance and achievement accomplishments of the student. It could also include aggregating a number of students' assessment results at multiple time points and reporting these findings to demonstrate academic growth or progress over time.

   Example 2: In digital environments, learning records management could include integrating a student's assessment results into a broader portfolio such as a digital grade book or student profile. It could also include digitally aggregating a number of students' results, removing personally identifiable information, and providing reporting or analysis output to those who do not have authority to access student-level data.

SUMMARY AND OBSERVATIONS REGARDING ASSESSMENT INTEROPERABILITY STANDARDIZATION

From the input provided, the Department has identified three areas of interoperability that should be considered for standardization to assist in the development of next-generation assessments; one area where standardization would not be productive; and one area where creating a standard set of rules and definitions may not need to occur concurrently with the first three areas but that should continue to receive attention.

- Areas where developers of next-generation assessments should consider creating common interoperability standards:
  I. Assessment Instrument and Item Format and Packaging
  II. Initiation and Return of Assessments
  IV. Learning Outcomes Management

- Areas where developers of next-generation assessments would probably accelerate their efforts, potentially reducing cost and improving functionality, by *not* standardizing:
  III. Administration of Assessments

- Areas where, while opportunities to standardize exist, it is a larger goal or undertaking than the assessment system (i.e., it involves creating standards for multiple student data record systems, not just the assessment system) and is most likely out of the scope for developers of next-generation assessments:
  V. Learning Records Management

AREAS WHERE INTEROPERABILITY IN LEARNING SYSTEMS MAY CREATE A MORE EFFECTIVE AND SEAMLESS DIGITAL EDUCATIONAL SYSTEM

Based on the input received from RFI respondents, the Department has identified some critical standardization and technology areas that are important for state education agencies, local education agencies, and software and assessment developers to consider as part of their effort to

transition from a traditional, print- and paper-based classroom to a digital learning environment where students' work is supported by diverse and effective technology infrastructure and systems. This standardization would supplement but not depend on any standardization work by the RTTA grantees or other developers of next-generation assessments.

There are two key areas of focus in our analysis:

A. Technology Landscape
B. Standards and Openness

A. **Technology Landscape**

From the RFI responses, five general areas of the "Technology Landscape" were identified: (1) the need for a learning ecosystem; (2) security; (3) data interchange and transport; (4) community; and (5) acquisition, purchasing, and fulfillment.

1. **Learning ecosystem**
   a) **Summary:** An effective digital assessment approach must be situated within a larger learning "ecosystem." The ecosystem consists of the entire learning system, including the assessments and the daily instruction the student receives in the classroom. The systems within the ecosystem must be able to coexist and interoperate. The development of a diverse learning ecosystem requires the creation and implementation of technical standards and common definitions that allow for interactions between the individual systems so that they may share information without requiring custom extensions or modifications.
   b) **Example:** An example of the operation of a diverse digital learning ecosystem could be the interplay between a district's student information system, a teacher's learning management system, and the formative, interim, and summative assessment systems used by a school or district. If these systems can communicate in a flexible, secure manner, it becomes possible to assemble complete learning systems from component parts, many of which may be able to be housed in the Internet "cloud," provided security, privacy, and legal issues can be accommodated.

2. **Security, privacy, and confidentiality**
   a) **Summary:** Security in the context of a learning ecosystem is broader than what is required just to ensure there is no security breach of individual test items before, during or after a test (which is itself extremely important). Security in this context involves the transfer of data across multiple entities and varying technology systems. It involves several factors:
      i) Trust/Policy
         (1) Organizations operating within a learning ecosystem would have legal operating agreements with each other to ensure that their exchanges are legitimate and lawful and not just technically secure.
         (2) This legal relationship would be identifiable within a standardized digital framework. That is, if a school district receives a digital request for assessment data to be transferred to a third party, the district's data systems should be able to compute whether this third party has been authorized (by

whatever legal authority) to exchange data, whether across the organization or for only a single student for a specific measurement activity.

(3) **Example:** Several organizations sign a common operating agreement, committing, for example, that they will only request the exchange data from each other when legally entitled to do so. This allows them to interchange data with confidence, knowing that each organization is legally bound to the others to abide by the terms of the agreement and that accountability for compliance rests within each organization. This would not necessarily mean that these organizations share all data with each other.

ii) Identity

(1) Organizations must be able to identify another organization within the digital learning ecosystem, so they can know with whom they are communicating. (e.g., if a school district receives a digital request for data from another school district, their systems must be able to compute that the requesting system is in fact who it claims to be.)

(2) Related to this, organizations could internally align their internal processes for identification to the larger learning ecosystem, creating opportunity for a single sign-on.

(3) **Example:** Organizations agree (via trust/policy) to employ a common public key infrastructure (PKI) so that all organizations can be uniquely identified with each other, and can encrypt and sign messages via PKI.

iii) Security

(1) Organizations must know that their confidential communications are secure. When a school district transfers digital data to another organization, no one else should be able to obtain an unauthorized copy or "listen in" on the transfer itself.

(2) **Example:** Organizations agree to use a common encryption and data-signing technique based on PKI, such as secure sockets layer (SSL). All communication can be secured from unauthorized access.

iv) Privacy

(1) Organizations must ensure adherence to all relevant legal requirements, such as the Family Educational Rights and Privacy Act (FERPA) and the Health Insurance Portability and Accountability Act (HIPAA), to ensure personally identifiable student information is not provided to unauthorized entities.

3. **Data Interchange and Transport**

a) **Summary:** Organizations must be able to interchange data of varying formats over the "transport mechanism" (e.g., a file transfer protocol). While standardized data formats are critical so that systems can recognize and read what has been transmitted, a digital learning ecosystem should be flexible enough to permit alternative formats to be transmitted over a common transport mechanism. It is critically important that these interchange formats are able to evolve and change over time without requiring changes to the underlying transport mechanism. The transport mechanism would generally include security features as defined above.

b) **Example:** Organizations can agree on a transport mechanism independent of any data interchange formats. Then, organizations can agree on specific data interchange

formats based on mutual requirements, which might include independent security features, depending on specific requirements of the format. An organization can participate in different interchange formats with various organizations while still relying on a single transport mechanism.

4. **Community**
   a) **Summary:** Organizations require a technical framework and a technical community to support implementation, development, innovation, troubleshooting, and standards setting.
   b) **Example:** In many standards development processes (education-related and otherwise), the technology standards organizations as well as technology trade associations often collaborate to serve as the technical community to support implementation, development, innovation, troubleshooting, and standards setting.

5. **Deployment, Payment, and Licensing**
   a) **Summary:** Organizations need technical mechanisms to acquire assessments and install them into learning management systems in a more efficient manner than is possible today.
   b) **Example**: A state wants to enable its school districts to select and deploy assessments for students in their districts on an individualized basis. School districts and their staff would have systems that allow them to identify appropriate assessments, perhaps from among a variety of providers, and install those assessments into their learning systems in a timely and efficient manner. Digital rights management, assessment interoperability, and a variety of technical standards detailed above are required to accomplish this task and should allow the state to acquire assessment systems and assessments from a variety of vendors and sources.

B. **Standards and Openness**

The Department required in the notice inviting applications for RTTA that the assessment technology standards developed by RTTA consortia be "industry-recognized open-licensed." There are a variety of categorical elements relating to the openness of the standards and the fact that a technology approach is standardized does not necessarily mean that the standard itself is "open."

In that notice inviting applications to the RTTA program, the Department did not specify in detail which categories of openness would be required. Possible categories of openness could include:

1. Openness of intellectual property. The intellectual property that makes up the standard is free for use by implementers for any purpose or requires only a nominal fee to access.

2. Openness for use. Types of usage of the standard are not restricted in terms of how the standard can be applied to solve any specific problem and would not be limited for use in a specific industry, technology, or strategy.

3. Openness for participation. Participation in the technical working group (TWG) that undertakes a standards-setting process is not a priori limited to a specific group of individuals or organizations. While membership dues or other fees are often required to cover the costs of standard-setting activities, any organization or individual who wishes to pay and participate is permitted to do so.

4. Openness of input. During the standards-setting processes, it is often common for the TWG to hold internal work groups, even when developing open standards, without public access for a period. An open standards process, however, involves the TWG reporting its progress and sharing materials regularly to provide the public opportunities to comment before the standards are finalized.

5. Openness of contributions. Intellectual property contributed during the standards-setting process has the same level of openness as the intellectual property of the standards themselves. This ensures that, in the future, contributors to the standards cannot restrict use of the standards by claiming rights or patent violations.

TECHNOLOGY STANDARDS DEVELOPMENT PROCESS

Based on input from the RFI respondents and additional research on current or best practices, this section identifies a potential process for defining interoperable technical standards. This process could be used by developers of next-generation assessments, as well as for purposes of standardization work more broadly, by interested entities as well as the Department. Existing standards organizations—such as the SIF Association, IMS Global Learning Consortium, the Postsecondary Education Standards Council (PESC), the Organization for the Advancement of Structured Information Standards (OASIS), and others—currently use a similar or identical process. In addition, other organizations, such as the National Institute of Standards and Technology (NIST) or the Department's National Center for Education Statistics' Common Educational Data Standards, possess expertise in the process of standardization and may provide useful advice in how to create successful assessment technology standards.

A general standardization process can be summarized as:

1. Develop organizational motivation to standardize
   - This process often includes multiple organizations and involves a willingness to commit staff, finances, intellectual property, and leadership focus to the effort.
2. Identify standards organizations to assist in the standardization process
   - Negotiate pricing, timelines and detailed process to complete.
3. Initiate work group, run by standards organization(s)
   - Set common Internet protocol policy and agreements.
   - Develop common data formats (such as XML), transaction sequences, and other details required to implement.
   - Undertake a common process to develop the standards.
4. Identify functional requirements
   - Systems and administrator functional requirements
   - Teacher, student, and other users' functional requirements

5. Test technology development, documentation, and tools creation
   - Provide for engineering, testing, and development of tools to support implementation of standards.
6. Provide certification environment
   - Provide certification methods for establishing compliance. This could be based on a fee or some other manner, depending on the market.