US Department of Education Assessment RFI Response from Advaiya, Inc.

Federal Register Doc. 2010-31881 Filed 12/17/2010.

Primary Contact: Patrick McDermott

Direct: (425) 739-3249 | **Main:** (425) 739-3202 | **Mobile:** (435) 459-2183
Advaiya, Inc. | patrick.mcdermott@advaiya.com | **www.advaiya.com**

January 17, 2011

3. Context for Responses

3.1 The primary intent of this RFI is to explore existing, in-process, or planned open technology standards, specifications, and technology products that support the management, delivery, and exchange of assessment content and the capture and exchange of assessment results. While the focus of this RFI is assessment technology standards, the Department recognizes that assessment generally occurs within the context of broader learning activities (whether online or offline) and, therefore, does not wish to restrict the range of responses to assessment-only approaches. The Department, therefore, also welcomes responses that address broader technology standards or approaches that are relevant to the handling of assessment management, delivery, or reporting. As mentioned earlier, the Department has required RTTA grantees to adopt a technical standard (or standards) that permit interoperability of the assessments and technology developed by that program. To help focus our consideration of the comments provided in the response to this RFI, we have developed several questions regarding the development of assessment technology standard(s) and their application to the RTTA program. Because these questions are only a guide to help us better understand the issues related to the development of interoperable technology standards for assessments, respondents do not have to respond to any specific question. Commenters responding to this RFI may provide comments in a format that is convenient to them.

Who we are and what we do:

Advaiya, Inc. is a software development company with a unique convergence of skills and experience with about 130 employees, headquartered in Seattle. Our focus is on technology as a platform for the enterprise within various industries, including education. We focus on emerging technologies and strive to be thought leaders in the market, developing new approaches and innovative solutions. We also advise large software firms on implementing technology strategies.

Our staff has extensive experience in education and working with the education market. In 2009, we launched a major new initiative for improving education data management and brought new staff on board with many years of experience in working with LEA and SEA data management, reporting, and systems integration, bringing in mature education data projects and intellectual property. These efforts have resulted in working with thought leaders in the education market to develop new strategies and new approaches that have significant potential for positively impacting education achievement in the United States. We are taking the results of this initiative into the Open Source community and putting it at the disposal of a governing body that can use it for these purposes.

We appreciate the opportunity to be able to respond to this RFI, not just with this document and the attached presentation, but also by ***contributing our intellectual property, that represents over 13 years of development efforts to the US Department of Education***, or whatever governing body the Department proposes.

We would like to introduce our response to this RFI, with our analysis of the main issues dealing with education data management in the United States. We will then provide responses to the individual questions raised in this RFI.

What are the underlying issues surrounding standards for assessments of learning?

Taken from a holistic, learner-centric view of assessment, there are three critical components in assessment technology standards:

1. **Capturing data** from a wide range of tools, vendors, and instruments;
2. **Moving data** from one application to another and keeping the data synchronized on various applications;
3. **Aggregating, storing, and reporting the data**, making it available to the stakeholders who can use it to further the educational goals of the organization.

A few real-life examples can help illustrate this:

Most vendors have concentrated on capturing data and providing value to their customers through their intellectual property (IP) that expresses their expertise in a particular assessment. The University of Oregon's DIBELS program is a great example of this. DIBELS is an effective tool for capturing oral reading fluency and other measures of early literacy, and the University of Oregon has invested significant resources in developing the IP for the program, which is why it is so widely used all over the United States.

Moving data from one application to another has been the mission of several of the standards that have been developed since the late 1990s. The Schools Interoperability Framework, or SIF, is the most common example for K-12 education in the US. This standard has grown incrementally since then, and is widely accepted in the industry, at least at the upper echelons.

SIF would theoretically allow a school district to send its student data from its Student Information System (SIS) into the DIBELS database. Adding new students to DIBELS could be automated through SIF as they enrolled and were entered in SIS. Assessment data could then be sent back from the DIBELS database and integrated into the district's longitudinal data system or data warehouse, so it could be integrated and analyzed and provided back to teachers within the context of other data.

This is more of a theoretical situation though, since the vast majority of school districts do not have the necessary Zone Integration Server or the know-how to hook it all together. In fact the average size of a school district in the US is just over 3,000 students and a district of this size would normally have very limited IT support of any kind, focusing its few resources on hardware and software support rather than Enterprise Application Integration (EAI), or Enterprise Resource Planning (ERP). Also, DIBELS does not support any method of moving student records into its data store, other than typing the data in by hand, or manually uploading .csv files. DIBELS does not currently support any kind of automated data feeds out of their system, other than a manual download of a text file with 312 fields, that most school district IT departments would have no idea what to do with.

Of course, if the typical school district only had a few sources of assessment data, this wouldn't be too huge of a problem. They can type in the student data online and view DIBELS reports on the DIBELS website and print them out as needed.

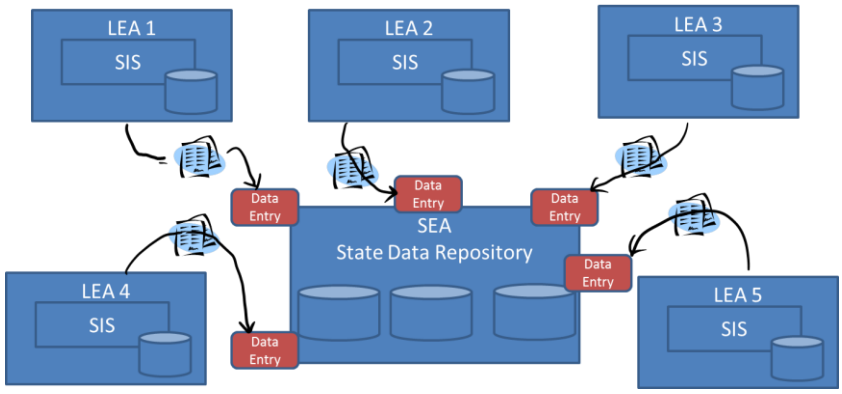
Unfortunately, this is rarely the case. Schools get data from their state in the form of end-of-level, and other high stakes assessments, SAT/ACT results, usually a great number of vendor-provided assessments, for reading, math, writing, English language learners, special needs students, gifted and talented programs, career and technology education programs, and a whole host of others. Districts that understand educational research are also conducting vast amounts of formative assessments, and virtually every teacher is constantly making classroom assessments that directly relate to the learning process. Each one of these assessments may provide reports, sometimes online, often in a proprietary format, but integrating the assessment data from multiple sources and tying it in to other relevant data, such as who was this particular student's teacher from last year, is virtually impossible for the vast majority of school districts in the US.

Having more (and better) assessments developed with Race to the Top funding won't really change the fact that most districts will still not be able to effectively integrate, report, and use this data to impact instruction in the classroom.

SIF has not yet realized its potential in helping to solve this problem. Where SIF has been most effective, in the relatively few districts that currently use it, is in passing data from one discrete data application to another, such as moving student data from an SIS system into a library application. Assessment data is much trickier: It comes from many sources and most districts don't know where it's supposed to be moved to. ***The problem is that the vast majority of districts do not have any longitudinal data system for aggregating assessment data over time from multiple sources, and hooking this data to its natural context, such as who were the teachers, which students did they teach, and what standards did the assessments relate to***.

As we have seen in the short examples above, the vendors can capture the data; SIF can move it; but where should it get moved to so that teachers can actually get to it and integrate it in their teaching? The third component of the assessment technology standards (aggregate, store, and report) is missing. Vendors have tried to fill in the gaps with more proprietary analysis tools and reports, but these still won't be able to get to other vendor's data, or the state's data, or, most importantly of all, the district's own formative assessment data, unless it was captured through their proprietary tool. **Because districts don't really have EAI/ERP capabilities (except for the very few), they are stuck with lots of data, but little actionable information that can impact learning in the classroom**.

The result of this situation can safely be described as data chaos. Assessment data is kept in a combination of proprietary vendor tools, locked up and often inaccessible at the LEAs, and stored in mainframe databases at the SEAs, resulting in inconsistent and redundant data in the LEAs and SEAs with very little of it finding its way into the classroom in a format that educators can actually use.



What is the most effective solution to this problem?

Many organizations are promoting greater use of standards to help improve educational data (Data Quality Campaign, CCSSO, the Common Data Standards Initiative, the National Education Data Model, the SIF Association, the Bill and Melinda Gates Foundation, the Dell Foundation, just to name a few). The most typical understanding of "standards" however is this: "Data standards are documented agreements on representations, formats, and definitions of common data elements intended to improve the quality and share-ability of education data" (National Forum on Education Statistics). **We believe that these elements of standards are necessary components for better educational data, but that they are not in themselves sufficient for improving the *access* to better educational assessment data for most school districts.**

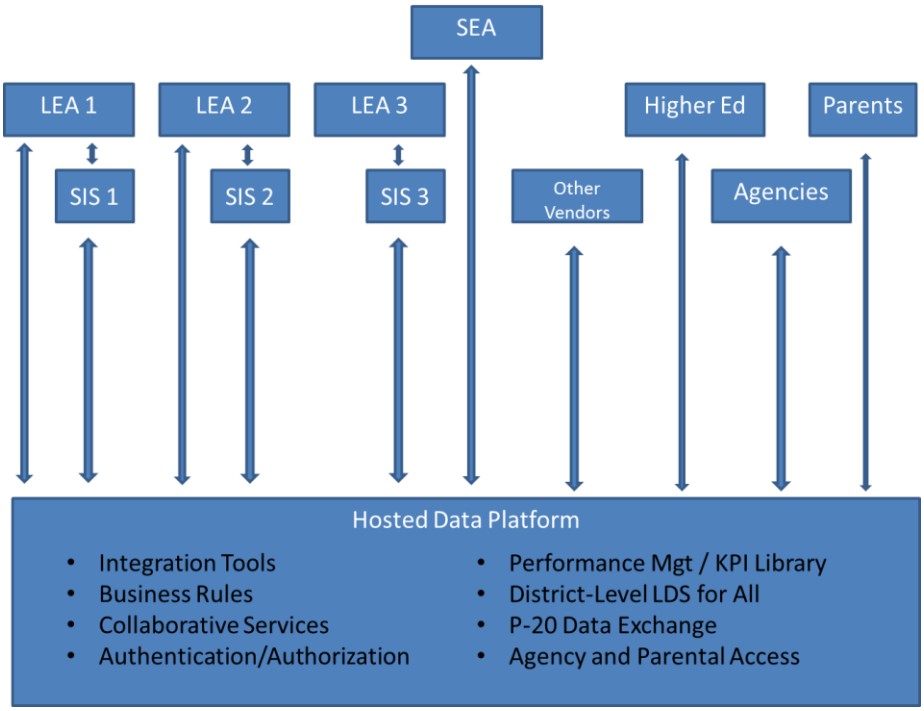
We believe there is one more critical element that needs to be added to our understanding of education data standards, and that is the actual implementation of standard, state-level databases in a secured cloud environment that can be used by any LEA or SEA that chooses to do so. That is to say that it is not just the *definition* of the data elements, but the actual hosted storage in a standardized relational database format for the data itself. While SEAs and LEAs could choose to utilize this web-accessible data storage as they saw fit, we propose that the vendors who receive RTTA funding would be obligated to provide data that their assessments generate back into these state-level standard databases, using SIF, or any other available techniques, based on the defined web services associated with the standard databases.

Data element definitions would be used to build integration tools from all the major SIS systems to keep secured student, class enrollment, academic standards, and teacher data synchronized with the cloud databases. This becomes the "context data" for the assessments.

The vendors would be authorized by either the SEAs or LEAs to access their context data on the cloud. The vendors would then retrieve the necessary identifiers for the SEA's/LEA's schools, classes, students, and teachers. Using their proprietary tools, the vendors would capture the assessment data and other value added data that pertains to their area of expertise and their own intellectual property. This is the "value-added data" for the assessments. FERPA guidelines would be followed in the same way as they are today when LEAs upload student data to vendor tools.

Once the value-added assessment data was captured by the vendors, it would be transported back to the cloud database within the standard data schema and its implemented relational table structure for assessment data. Standardized web-based and secured reports would be available for the SEAs/LEAs who wanted to use them. These

reports would be able to display both the context data and the value-added assessment data from any assessment and any context that the user had rights to see. These security rights would be controlled by the SEAs/LEAs. Vendors could create additional tools that helped analyze and add more value to this data directly on top of the cloud databases, bringing the benefits of innovative new ways of looking at multiple assessments from multiple vendors over time. Researchers could be granted special access to the data that did not disclose personally identifiable information.



How can this solution be developed within realistic parameters?

This proposed solution may seem daunting at first, but new technologies, new ideas, and new approaches make this solution realistic. Any implementation has to take into consideration the following:

1. The **cost** of developing the database schema, structure, the web services and documentation that allow vendors to get data into and out of the system, the standard web-accessible reports and tools, and the Key Performance Indicator (KPI) definitions that help define the taxonomy in a standard way (such as defining the student-teacher linkages in the data system);
2. The **required IT support** from the SEAs/LEAs to implement any system;
3. The **time** to develop the integration tools for the major SIS systems and for vendors to create ways to integrate data into and out of their proprietary systems, using SIF, or other applicable standards;
4. The **security and privacy** requirements of such a system, and ensuring compliance with FERPA and other applicable laws, regulations, and best practices;
5. Creating a **governance structure** to manage the standard database schema, web services, documentation, and other aspects for the benefit of all the stakeholders.

We believe that an effective implementation of this proposed solution that addresses each of these 5 considerations is achievable within a short amount of time and within tight financial constraints, because ***much of the core work has already been accomplished.***

In1997, a school district that was frustrated by the inability of commercially available tools to manage longitudinal education data for the entire enterprise began developing a data system to solve this problem. In 2009, we acquired this IP to generalize it and make it available to a wider community. This data system has been in constant use and improvement for many years now, reacting to new requirements and new circumstances.

From that initial effort, we have grown this system tremendously and have deployed it to the cloud in 51 separate state-level databases (for all 50 US states and the District of Columbia). Each of these 51 state databases currently has all of the school and LEA information for the entire state, as well as the state's academic standards and the new Common Core State Standards. The databases share essential functionality, yet allow each state and even each LEA to modify many items according to their needs. It also references third party identifiers to support better integration, such as SIF codes, NCES codes, Achievement Standards Network IDs for state and CCSS standards, provides fields for SEA and LEA codes, and many other hooks for integration. ***This system is a complete cloud-based education data platform, including persisted data storage, functions for retrieving, updating, and inserting new data, and the web services that can allow any application that has the proper authorization to interact with the data in the data system***.

These databases currently have schemas (including the supporting functions and web services) that address in a very detailed, fine-grained, yet flexible manner the following contextual areas:

- Informal Classroom Assessments (including observations and other informal methods)
- Formative Assessment (including assessment plans, scheduling, scoring, and reports)
- Summative Assessment and other High Stakes Assessments
- Education Plan Information (standards based)
- Interventions
- Schedules (student enrollments in classes and interventions and their associated teachers)
- Standards (local, state, and national—state and national preloaded; local can be added)
- Reporting (data can roll up easily to district and state levels)
- Course (including local, state, and national course definitions)
- Curriculum (subject areas, etc.)
- Contact Information
- Security (role-based: student, teacher, interventionist, administrator, parent, etc.)
- Directory Information
- Employment (staff, positions, assignments)
- Enrollment
- Groups (staff membership in PLC groups, etc.)
- Language (ELL information, etc.)
- Licensure (educator certification and licensing)
- Program Participation (special ed, gifted and talented, ELL, etc.)
- Person Information (staff, students, parents, demographics, etc.)
- Household Information (including parent/guardian information)
- System Tools (error logs, debugging, backups, recovery, data administration, etc.)
- And many others

We have taken a very broad approach in developing this data standard. **Our constant focus has been to support the learning process in a holistic way, rather than just figuring out how to move bits of text from one system to another**. Our approach therefore goes far beyond merely defining the transport mechanism, or the field data types, or the required elements in a block of xml. Instead, we have developed an entire data platform for education. We call this data platform the **Learning Management Infrastructure (LMI)**.

We have decided to share this system with the wider education community as an Open Source/Open Standards tool to promote better education data collection, management, reporting, and integration—directly supporting the learning process at all levels: From the classroom to the SEA and beyond.

What are the benefits of this proposed solution?

This proposal goes beyond what most education data organizations have contemplated in the past. It is a radical, but achievable solution. Yet any innovative solution will have the risks associated with trying something new. These risks have been mitigated through our experiences with actually creating and deploying these systems. We feel that the potential benefits of such a solution greatly outweigh any risks associated with this proposal.

From the educators' perspectives, easily accessible assessment data that is integrated both with the context of the learning environment, as well as the value-added data from many vendors' assessments, that have all been tied to academic standards will provide them useful information that can have **a major impact on classroom instruction**.

From the SEA/LEA perspective, this solution solves many problems. It provides a longitudinal data system to the vast majority of school districts that currently don't have such tools. It provides SEAs and LEAs with accountability and systemic improvement tools that takes a much more holistic view of educational data rather than just focusing on limited summative test data. All of this can be accomplished with very limited impact on the SEA/LEA IT support structures, **because the infrastructure already exists and is located on the cloud**.

From the vendor perspective, this solution encourages vendors to continue to develop IP in capturing data, allows them an easier path to get that data back to the users where they can integrate it with their other longitudinal data, and eases their burdens in figuring out how to work with the context of the data that they capture: the students, standards, and teachers. **This solution also opens up new market opportunities for vendors who want to develop innovative tools quickly and easily that can leverage an existing (and extensive) data structure**.

For example, LMI contains all the information necessary to securely connect parent information to individual students. So a vendor who wanted to create an application that helped parents understand and interact with assessments more effectively could easily leverage the LMI tools to connect the assessments, individual student scores, parent information, to their own proprietary information and IP that would help parents understand these assessments better. Imagine a parent logging into this vendor's application, which then retrieved their children's DIBELS scores, compared them with scores from specific CCSS standards derived from their children's state end-of-level results from last year to made specific recommendations for books and lessons that the parents could use to further their kids' reading abilities, while providing those children's teachers the opportunities to make comments and share insights directly with the parents in a secure manner.

From other perspectives, researchers and higher ed would be able to query the data with special accounts that protected privacy, yet offered unprecedented insights into what is working in promoting achievement. For example, because the databases use the same master data elements and identifying codes, it will be possible to ask such questions as which interventions are most effective for phonemic awareness of kindergarten ELL students? Or, which mathematics standards from the Common Core State Standards are most indicative of overall mathematical achievement? **The possibilities for major breakthroughs in education research are significant**.

Finally, this solution addresses other concerns raised by the Race to the Top program, as well as presumably the future reauthorization of the Elementary and Secondary Education Act. Many elements of the Race to the Top program involve data structures that are not currently widely available. For example, how can we promote data exchange between K-12 and higher ed? What is the exact nature of the teacher-student data linkage? What are the best practices in promoting data usage to improve instruction? All of these areas, and many others could be significantly facilitated with the data in the state-level cloud databases.

We are available at any time to discuss these ideas with any of the stakeholders who are interested in this proposal. Please see the contact information above. Please also see the attached presentation that addresses many of these same issues.

3.2 Questions About Assessment Technology Standards

General and Market Questions

3.2.1 Current Landscape. What are the dominant or significant assessment technology standards and platforms (including technologies and approaches for assessment management, delivery, reporting, or other assessment interoperability capabilities)? What is the approximate market penetration of the major, widely adopted solutions? To what degree is there significant regional, educational sub-sector, or international diversity or commonality regarding the adoption of various technology standards and capabilities, if any?

There are lots of standards available today, including the Schools Interoperability Framework (SIF), the Common Cartridge (CC) from IMS/GLC, the PESC standards for higher education, the National Education Data Model, and the Common Data Standards Initiative. However, none of these have yet been widely employed in school districts for assessment management.

These standards are currently most commonly used to transport packages of data from one application to another. The primary reason why these standards are not employed more for assessment data is that the vast majority of school districts do not have the IT resources to set up and manage a longitudinal data system. So, while the standards are capable of transporting data, in most cases, there is no one repository of assessment data available to transport the assessment data to.

3.2.2 Timelines. Approximately how long would it take for technology standards setting and adoption processes to obtain a technology standard that meets many or all of the features or requirements described in this RFI? What are the significant factors that would affect the length of that timeline, and how can the impact of those factors be mitigated? More specifically, would the acquisition of existing intellectual property (IP), reduction or simplification of specific requirements, or other strategies reduce the time required to develop these technology standards and processes?

If the recommendations proposed here were created from scratch, the timelines to complete these projects would run into several years. Taking an existing and functional data platform specifically designed for comprehensive education assessment data management would significantly reduce the required efforts and timelines. Advaiya is willing to provide much of the IP we have developed and acquired, that represents over 13 years of development efforts. The LMI data platform currently exists in 51 cloud-based, state-level databases. These data platforms allow for the contextualization and integration of summative, formative, and classroom assessments, covering the entire learning process. We would need probably about 9 months to improve documentation, help develop governance structures, formulate and draft Key Performance Indicators (KPIs), draft privacy and security guidelines, test the system with multiple applications and security issues, and create a website with resources for vendors who wanted to start developing against the LMI educational data platform.

After that, it would take several months for most vendors to work with the existing standards, such as SIF, to develop the mechanisms to move the SEA's/LEA's context data (students, standards, teachers, etc.) into their proprietary systems, and also to develop the mechanisms that would allow those vendors to load their value-added assessment data back into the state-level databases.

One such application that we have built, utilizes LMI for its data storage and retrieval. This application is called Metria (www.MetriaLearning.org). Metria is a free tool for teachers to manage formative assessment data, standards-based education plans, and interventions for their students. Metria is a functioning proof-of-concept application built on top of the LMI education data platform.

3.2.3 Process. What process or processes are appropriate for the adoption, modification, or design of the most effective technology standard in a manner that would answer many or all of the questions in this RFI? We are interested in learning the extent to which the uses of one or another process would affect the timeline required to develop the technology standards.

We recommend convening a meeting consisting of some technical and educational thought leaders, some organizations such as CCSSO, Data Quality Campaign, NCES, National Governors Association, the Dell Foundation, the Bill and Melinda Gates Foundation, the PARCC/SBAC leadership and a few SEA, LEA, and some vendor representatives under auspices of the US Department of Education to create a committee to develop a governance structure and outline the most appropriate process to achieve the US Department of Education's goals and begin work on implementation.

3.2.4 Intellectual Property. What are the potential benefits and costs to the Federal Government, States, and other end-users of different IP restrictions or permissions that could be applied to technology standards and specifications? Which types of licensed or open IP (e.g., all rights reserved, MIT Open License, or Gnu Public License) should be considered as a government technology standard? How should openness relating to the IP of technology standards be defined and categorized (e.g., Open Source Initiative-compatible license, free to use but not modify, non-commercial use only, or proprietary)

We recommend that the entire definition of the data schema, including scripts to create the database, tables, functions, stored procedures, as well as the code for the web services, be made freely available under an open source initiative-compatible license, such as the Shared Source Permissive License. Furthermore, the KPIs that define exactly how certain measurements are to be defined, both in descriptive human language, as well as in Structured Query Language (SQL) should also be posted for free distribution for anyone to use. The use of the actual databases on the cloud should be subject to Acceptable Use Agreements that each user must acknowledge before they are permitted to access the data system. No modifications should be allowed to the source code unless it is approved by the governance committee and unless it can guarantee backwards compatibility for applications already using the system.

Advaiya would be willing to consider turning over our database IP to a governance body under any licensing terms that this governance body felt was most appropriate.

3.2.4.1 Existing Intellectual Property. What are the IP licenses and policies of existing assessment technology standards, specifications, and development and maintenance policies? Are the documents, processes, and procedures related to these IP licenses and policies publicly available, and how could the Department obtain them?

These licenses are freely available on the Internet. We would be happy to provide copies of them to the Department.

3.2.5 Customizing. Can assessment tools developed under existing technology standards be customized, adapted, or enhanced for the use of specific communities of learning without conflicting with the technology standard under which a particular assessment tool was developed? Which technology standards provide the greatest flexibility in permitting adaption or other enhancement to meet the needs of different educational communities? What specific provisions in existing technology standards would tend to limit flexibility to adapt or enhance assessment tools? How easy would it be to amend existing technology standards to offer more flexibility to adapt and enhance assessment tools to meet the needs of various communities? Do final technology standards publications include flexible IP rights that enable and permit such customizations? What are the risks and the benefits of permitting such customization within technology standards? When would it make sense to prevent or to enable customization?

Good database design can provide structures for protecting master data (also called validation or lookup data), and ensuring that client data reflects this master data, while allowing for SEA/LEA differences in areas that don't impact the master data, and ensuring maximum flexibility for future growth and future changes. Inflexible database design is the biggest threat to any system for its long-term viability. We have tried to utilize best practices in our database design to ensure that this is not a problem now, or in the future.

As with any complex system, there will need to be a governance body to determine how to react to new circumstances and respond to suggestions and requests for changes in ways that benefit all the stakeholders. Changes should be allowed under two circumstances: if the governing body approves of the changes, and if the changes do not break backward compatibility of existing applications, possibly allowing this second condition if minor application changes can be implemented easily to remedy the situation.

3.2.6 Conformance and Testing. Do existing technology standards or technologies include specifications or testing procedures that can be used to verify that a new product, such as an assessment tool, meets the technology standards under which it was developed? What specifications or testing procedures exist for this purpose, e.g., software testing suites, detailed specification descriptions, or other verification methods? Are these verification procedures included in the costs of the technology standards, or provided on a free or fee-basis, or provided on some combination of bases?

The SIF standard provides compliance certification, although this seems to be more for marketing than anything else. A particular application may be SIF-compliant for a particular SIS system, but not for a library system, for example. However, SIFA does provide some very beneficial models of testing applications and this is definitely something that should be explored.

The governance body should have very specific categories of compliance for standards developed here, in order to ensure that conformance testing conveys a very crisp understanding of what the application can and cannot do.

3.2.7 Best Practices. What are best practices related to the design and use of assessment interoperability technology standards? Where have these best practices been adopted, and what are the general lessons learned from those adoptions? How might such best practices be effectively used in the future?

Best practices really depend upon the type of standard one is looking at. In terms of simply defining the fields, messaging protocols, and transport mechanisms, the SIF standard could be said to represent a best practice, although its early handling of assessment data was rather weak (it has been improved greatly since then).

In terms of data specification, the National Education Data Model and the Common Data Standards Initiative both do a good job of addressing the data specifications, but neither goes far enough to effectively capture all of the necessary context data surrounding assessment, while providing for maximum flexibility of design for customization and future changes. We feel that our database specification is more detailed in many areas and incorporates a more flexible and customizable overall design.

From the perspective of a cloud-based specification that represents a secured data store that can be used by many organizations, we are not aware of any such open source implementation in the education market other than our own. However two notable examples exist in other markets: Namely, SalesForce.com and Apache's Ofbiz framework.

SalesForce.com (www.salesforce.com) is a cloud-based customer relation management data platform that is used securely by thousands of companies all over the world. It is essentially "Platform as a Service" allowing its customers to conduct critical operations securely online without software, other than a browser, and it serves as a hub for a huge vendor community that builds and sells other innovative applications around SalesForce's cloud data platform.

Apache's Ofbiz framework (http://ofbiz.apache.org/) is an open-source e-commerce and ERP data platform. It has many of the design goals of SalesForce.com, but additionally, it is run as an open source project. It is not however hosted on the cloud.

Our proposal to the Department of Education is essentially to provide SalesForce-like functionality on the cloud within an Ofbiz-like open source framework for the education market in the United States. This would provide assessment management services and reports, coupled with ERP/EAI for education with shared storage and common master data on the cloud.

Additionally, the cloud itself should be viewed as a best practice. The cloud can provide all the functionality of locally-run servers, without the infrastructure, know-how, or risk of running these servers, and the cloud can provide this functionality at a fraction of the cost that it would take to run these services locally. The cloud facilitates collaboration around data, allowing organizations to easily share, rate, and comment upon various approaches, curricular plans, and assessments. The cloud can be as secure as local installations and has been proven as such by thousands of companies over the last several years who run applications such as SalesForce. The cloud scales well and has been proven to be able to handle the most intensive transactional loads. ***Most importantly, having the assessment data securely available in one location for an entire state, makes querying, reporting, and analysis of the data from multiple sources dramatically easier than any other solution***.

Technological Questions Regarding Assessment Technology Standards

3.2.8 Interoperable Assessment Instruments. What techniques, such as educational markup or assessment markup languages (see also http://en.wikipedia.org/wiki/Markup_language), exist to describe, package, exchange, and deliver interoperable assessments? How do technology standards include assessments in packaged or structured formats? How can technology standards enable interoperable use with resources for learning content? How can technology standards permit assessment instruments and items to be exchanged between and used by different assessment technology systems?

The SIF specification is definitely the most widely accepted markup language for education data in the US. The SIF Association (www.sifa.org) provides detailed information on packaged and structured formats, as well as transportation protocols and message formats. Because this specification has been developed over a long period of time and is widely accepted by the community, it should be used for transport and messaging.

However, as noted above, the transport of data is only one component of an overall education technology specification. The persistent storage, including aggregation, longitudinal analysis, and reporting is also a necessary standard.

The Postsecondary Electronic Standards Council (PESC http://www.pesc.org/) also defines markup language for higher education. As we begin to develop linkages between K-12 and higher ed, we will need to accommodate these standards for communicating with higher ed.

3.2.9 Assessment Protection. For this RFI, "Assessment Protection" means keeping assessment instruments and items sufficiently controlled to ensure that their application yields valid results. (See also paragraph below, "Results Validity.") When assessment instruments or content are re-used or shared across organizations or publicly, are there capabilities or strategies in the technology standards to assist in item or instrument protection? What mechanisms or processes exist to ensure that assessment results are accurate and free from tampering? Do examples exist of public or semi-public assessment repositories that can provide valid tests or assessments while still sharing assessment items broadly?

There are many ways that we can protect assessment data. Controlling authentication and authorization effectively is the first line of defense. Logging and recovery strategies can help if the first line of defense fails. Triggers are currently set up on the data platform to record any attempted tampering and these allow for logging such activity and rolling back the data to its original state if the activity was not authorized. We also record the date and time of modifications and who the logged on user was that made such modifications.

3.2.10 Security and Access. In what ways do technology standards provide for core security issues, such as access logging, encryption, access levels, and inter-system single-sign-on capabilities (i.e., one login for systems managed by different organizations)?

There are two stages to most security systems: authentication and authorization. Authentication validates that a person is who they say they are. There are widely accepted third party authentication tools that manage user identification and passwords. LMI currently supports this method of authentication.

Authorization is the process of determining which rights an authenticated user has over which data. The LMI data platform has a complete role-based security system for authorizing which users can see and/or edit which data. Because the LMI data platform has been designed from the ground up to interact with multiple applications, this process of authorization occurs in two phases. First, the calling application itself must be authorized (i.e. the SEA or LEA have given the calling application necessary rights to interact with the data). After the calling application is authorized, then the individual user who is using that application must be authorized through role-based security to determine if this particular user has access to see this data and if so, in what level of detail.

FERPA is one of the key regulations that helps determine what level of authorization is appropriate for a particular user.

LMI fully supports a role-based and FERPA-compliant system of authorization for both applications and individuals.

3.2.11 Results Validity. For this RFI, "Results Validity" means protecting the statistical validity and reliability of assessment instruments and items. How can interoperable instruments be managed to ensure they are administered in a way that ensures valid results? Are solutions regarding assurance or management of validity appropriate for inclusion in technology standards, or should they be addressed by the communities that would use the technology standards to develop specific assessments?

The validity of assessment instruments and items, when understood in terms of how those assessments are administered to students, should be addressed by the communities that would be using this technology. However, we should consider "results validity" in a broader context: Namely, do these results accurately demonstrate the learning that is occurring in a particular classroom, and can these results be used to improve the learning process?

This is one of the biggest issues currently facing educational assessment data. Any system of standards that does not address the wider issues of validity that are associated with the learning process will have very limited usefulness.

Taking a wider view of validity means focusing on the educational process, rather than just on bits of data out of context. It is the context in fact that is most critical. The context for educational assessment data are the elements of the wider learning process: Students, teachers, standards, assessments, results, etc. Assessment data that is removed from its context, or that is not fine-grained enough, has virtually no ability to impact classroom instruction.

Assessment data needs to be coupled to the context data that it is associated with. See next question.

3.2.12 Results Capture. How can technology standards accurately link individual learners, their assessment results, the systems where they take their assessments, and the systems where they view their results? How do technology standards accurately make these linkages when assessments, content, and other data reside across numerous, distinct learning and curriculum management systems, sometimes maintained by different organizations?

The best option for linking individual learners and other context information to their assessment results is to create a system where the external assessment system obtains its student information from the SEA/LEA data warehouse, including the identifiers that help link that assessment back to the student data, once the assessment is completed. Virtually every state currently has state student unique identifiers that can help make this link. Our data platform provides persisted storage for state unique student identifiers, LEA-issued local student numbers, and universally unique GUIDs (global unique identifiers) that can all be utilized to correctly identify students.

However, for assessment data to be useful for educators, it must also be coupled to all of the other context data that surrounds the entire learning process: students, standards, teachers, etc., including that students' prior classes and teachers, and prior assessments, as well as the academic standards those assessments relate to. The external assessment systems cannot be expected to relate their own assessments back to all of this context data. The assessment systems generally do not keep extensive records of which teachers taught what to which students, and almost never contain any references to prior student performance. Therefore, the most meaningful solution is the one proposed here, where the value-added assessment data is linked to the entire context data on an independent database that aggregates assessments from many sources over time.

One other critical link that must be made is that every assessment needs to be related to the academic standards that it addresses. This can be accomplished by using accepted academic standards identifiers, such as the Achievement Standards Network (ASN) SID identifier (see www.achievementstandards.org). This will allow multiple assessments to be compared and reliably related to each other, through the academic standards that the assessments measure.

We strongly recommend that RTTA require all vendors who receive RTTA funding to reference the ASN SID identifier for all standards. This is the only uniform way to identify state and Common Core State Standards across the United States that we know of.

We currently have a database of over 85,000 individuals state and CCSS academic standards, including the ASN SID identifiers and we would be happy to help address these issues.

3.2.13 Results Privacy. How do technology standards enable assessment results for individual learners to be kept private, especially as assessments results are transferred across numerous, distinct learning systems? How can such results best be shared securely over a distributed set of systems managed by independent organizations that are authorized to receive the data, while still maintaining privacy from unauthorized access?

Security is best accomplished in depth. That is to say, it is important that all processes have multiple levels of security, no one of which can be compromised without others still preventing unauthorized access. We protect privacy through many layers. We recommend the following measures should constitute the minimum security requirements for any assessment technology standards: Secured transport over https; Encryption of names, so that personally identifiable information (PII) can be decoupled from the rest of the context data, as well as from the assessment data; Measures that prevent packet tampering during transport; and database security, such as SQL injection attack prevention.

Privacy of results also requires more fine-grained security measures, including role-based security. In order to be compliant with the Family Educational Rights and Privacy Act (FERPA), including the recently amended regulations and case law, the role-based security system must take many factors into account. For example, displaying data that links student names to Free or Reduced Lunch program data is not even permissible for principals, let alone their classroom teachers. Too many assessment programs available today ask the SEAs/LEAs to upload student data, including special education status and Free and Reduced Lunch data directly to the vendors in an unsecured and unprotected manner. This is a clear violation of FERPA that happens all the time.

The solution here, is to keep this kind of context data secured on the state-level cloud databases and then report all of the assessment data from those databases, rather than from the individual vendors. Anonymization should be employed where necessary to comply with FERPA. See next question.

3.2.14 Anonymization. Do technology standards or technologies permit or enable anonymization of assessment results for research or data exchange and reporting? How do various technology standards accomplish these tasks? For example, where a number of students take a test, can their answers be anonymized (through aggregation or other techniques) and shared with researchers to examine factors related to the assessment (e.g., instructional inputs, curriculum, materials, validity of the instrument itself) without revealing the identity of the learners? Is this an area where technology standards can help?

As mentioned above, we recommend the encryption of student names, and only permit the decryptions of names when it is permissible under FERPA and other security regulations. We make extensive use of globally unique identifiers (GUIDs) that can uniquely identify a student, without permitting that GUID to be tied back to their unencrypted name. This allows us to employ many anonymization tools.

Another measure that must be required of any RTTA vendors under FERPA is known as “n-size protection.” In its simplest form, n-size protection means, for example, that if you display anonymized data for Mrs. Smith’s third grade math class, but you show scores disaggregated by race, you may not display some data if assessment data is displayed that shows scores for Asian students and there are only two of them in Mrs. Smith’s class. This is not generally permitted for groups of ten or fewer students (though as far as we know, this number is not an absolute). Any RTTA vendor must be required to be able to protect n-size data on their own reporting systems. Our state-level cloud databases can protect a variable value anonymized n-sizes.

3.2.15 Scoring and Analysis of Results. How can technology standards be used for the scoring, capture, recording, analysis or evaluation of assessment results?

It is our recommendation that scoring, capture, and recording be handled as value-added assessment data and therefore should be accomplished by the vendors as part of their intellectual property. Therefore, it should not be covered by assessment technology standards.

We recommend however that analysis and evaluation should be conducted where it can be effectively integrated with multiple assessments from multiple sources and with all of its context data. That is, this should occur on the state-level databases on the cloud. As such, these measures would be covered under assessment technology standards.

We recommend that Key Performance Indicators (KPIs) should be part of the assessment technology standards. The KPI standards should include two descriptions: One that explains the particular KPI in detail in English descriptive terms; and one that defines exactly how this KPI will be determined on the state-level database in Structured Query Language (SQL).

We are currently beginning this process on our state-level databases and would be happy to help with this work.

The exact nature of the student-teacher link should also be defined in the same way. We have developed the state-level databases with this requirement in mind.

3.2.15.1 Results Aggregation and Reporting. How can technology standards enable assessment results to be aggregated into statistical or other groupings? How can technology standards provide capabilities for results (aggregated or raw) to be reported across multiple technology systems? For example, if a learner takes an assessment in one system, but the results are to be displayed in another, how do technology standards address transferring results across those systems? How do technology standards address aggregation of results for a number of learners who are assessed in one system and whose results are displayed in yet another technology system? Can anonymization controls be included with aggregation and reporting solutions to ensure individual data privacy and protection (see also 3.2.14 above).

This is the core of our proposal. We recommend that this is accomplished through the creation of state-level databases on the cloud for the use of the SEAs/LEAs. The vendors would provide the value-added assessment data to these state-level databases and all of the reporting, aggregation, and analysis would occur there. Anonymized accounts would be available to these databases for researchers, higher education, and the US Department of Education. See the introduction to this RFI response for more details, as well as the attached presentation.

3.2.16 Sequencing. How do technology standards enable assessment items stored within an assessment instrument to be sequenced for appropriate administration, when the assessment consists of more than a single linear sequence of items? For example, how do technology standards address computer-adaptive assessments? How are the logic rules that define such sequencing embedded within a technology standard?

Unless the SEA defines sequencing for particular academic standards, we recommend that sequencing remain a value-added assessment data element that is the IP of the vendors. It should not be addressed as assessment technology standards.

3.2.17 Computer-Driven scoring. How do technology standards permit, enable, or limit the ability to integrate computer-driven scoring systems, in particular those using “artificial intelligence,” Bayesian analysis, or other techniques beyond traditional bubble-fill scoring?

This is value-added assessment data and should remain the intellectual property of the vendors. As such, it should not be covered under assessment technology standards.

3.2.18 Formative, Interim, and Summative Assessments. What technology and technology standards exist that support formative, interim, and summative assessments? What technology standards support non-traditional assessment methods, such as evidence, competency, and observation-based models?

Our existing state-level databases provide persistent storage for (as well as the web services to retrieve and load) summative assessment, formative assessment, and informal classroom assessments, such as observations, competency, and evidence. Furthermore, all assessment data, whether summative, formative, or classroom assessments, should be able to be related back to academic standards using a universal identifier, such as the Achievement Standards Network Standard Identifiers (SIDs).

Assessment technology standards should specify standard reports that combine data from all of these types of assessments and from multiple sources. This is a primary requirement for classroom educators to use the data to impact instruction.

We would be happy to share the technical details of how this is accomplished.

3.2.19 Learning and Training. What applications or technology standards exist that can apply assessment results to support learning and training? Are there technology standards or applications that support more than one of the following: Early learning, elementary/secondary education, postsecondary education, job training, corporate training, and military training?

Learning and training should be integrated into the assessment technology standards, especially for professional development purposes for educators. The SCORM standards (http://www.adlnet.gov/Pages/Default.aspx) address this kind of training and provide standards for dealing with it. We feel that our state-level databases could easily accommodate professional development training for educators, because the context data already exists for this within our database. Corporate and military training would be a bit more difficult to integrate because the context data for those is so different from a K-12 environment.

Our state-level databases can accommodate an unlimited number of standards systems, including state and CCSS academic standards for K-12, state academic standards for early childhood development, behavioral and citizenship standards, 21st Century Skills, and others. We currently have over 85,000 rows of standards data from various sources in our databases that are available as part of this Open Source project.

3.2.20 Repositories. What technology standards-based assessment instruments, questions, or item banks (or repositories and learning management systems) are used to manage and deliver assessments?

Each SEA and each vendor would normally have to have their own. This is probably going to remain the intellectual property of the vendors. As such, it should not be covered under assessment technology standards.

However, we would recommend that SEAs develop the content for their item banks and store that data in the state-level cloud databases where they can be provided to the vendors, as well as the LEAs, as web services for them to consume.

3.2.21 Content Lifecycle. How can technology standards be employed to support an assessment content lifecycle (creation, storage, edit, deletion, versioning, etc.)?

This is value-added assessment data and should remain the intellectual property of the vendors. As such, it should not be covered under assessment technology standards.

3.2.22 Interfaces and Services. What interoperability specifications for application program interfaces (APIs) or Web services interfaces to assessment management, delivery and tracking systems have been developed? How are they organized? What are the best practices related to their design and usage? How broadly have they been adopted, and what are the lessons learned from those who have designed or implemented them?

Web services are the way to make data most effectively available to vendors and other consumers of the data in a distributed environment. The biggest single criticism of SIF was that it did not use the web service standard. SIF has its own solution called Zone Integration Servers. Now in SIF 3.0, that should be released sometime this year, SIF will also support web services.

While web services are still not ubiquitous in the education market, they certainly are almost everywhere else. There are hundreds of books available about best practices in web service design and use and it would be difficult to even begin to address this here, but suffice it to say that we recommend using web services, rather than anything else, such as SIF's old Zone Integration Servers.

3.2.23 Internal Transparency and Ease of Use. Are there technology standards and communication protocol implementations that are "human readable?" What are the benefits and risks of "human readable" technology standards? Some technology standards are not comprehensible without tools to unpack, decode, or otherwise interpret the implementation data resulting from use of the technology standard. Other technology standards, such as HTML, RTF and XML, are largely readable by a reasonably sophisticated technical user. RESTful-designed Web services are often specifically intended to be readable by, and even intuitive to, such users as well. We ask commenters to consider the extent to which various technology standards possess native "human readability" and comprehensibility.

As long as the documentation for using the web services is clear enough, in our opinion it doesn't really matter whether the data that is transported is human readable or not. All access should come through the web services so the only reason one would ever need to go beyond that is for debugging and troubleshooting. Both of which can still be accomplished whether the data is human readable or not. Whether or not the data itself is human readable should have minimal impact on security, since the transport must be protected through secure means anyway. Transporting binary, but unsecured, data is still essentially no protection at all, in our opinion.

3.2.24 Discovery and Search. How is the discovery of items or instruments (or other elements) handled within a technology standard or technology? For example, are there search APIs that are provided to permit a search? How are metadata exposed for discovery by search engines or others?

There are many approaches to discovery and search that should be included in assessment technology standards. Metadata, including key words, as well as known identifiers, such as ASN SIDs for academic standards, NCES codes for secondary courses, and certain SIF codes, could all improve discovery and searchability. Full text searches should also be available for the content of academic standards.

3.2.25 Metadata. What kinds of metadata about assessments (i.e., information describing assessments) are permitted to be stored within technology standards or technologies? How do technology standards accommodate structured data (such as new State curriculum standards) that were not anticipated when the technology standard was designed? How are metadata describing unstructured (such as free-text input) and semi-structured data incorporated within assessment technology standards?

Metadata is especially useful for academic standards. This metadata should include key words, ASN SIDs, grade levels, and subject areas for every state or CCSS standard. We are currently building additional structures to support metadata for academic standards in our state-level databases. These structures consist of defining key words by running analysis on our 85,000 rows of state and CCSS academic standards and then pulling out the statistically significant terms and relating those back to the appropriate standards.

3.2.26 Recommendation, Rating, and Review. Do technology standards or technologies permit rating, review, or recommendations to be incorporated within an item, instrument, or other element? If so, in what ways? How are conflicting ratings handled? Do technology standards or technologies permit "reviews of reviews" (e.g., "thumbs up/down"

or “Rate this review 1-5”)? Is the rating or review system centralized, or are multiple analyses of the rating data permitted by distributed participants?

Recommendations, rating, and reviewing should definitely be encouraged and the structures for this data should be included in the assessment technology standard. We are currently building this functionality into our state-level databases. While this functionality is not yet fully available, it will be soon.

3.2.27 Content and Media Diversity. What types of diverse content types and forms of assessment content exist that extend beyond traditional paper-based assessments translated to an electronic delivery medium? We are interested in learning more about electronic delivery and interaction media, such as performance-based assessments, games, virtual worlds, mobile devices, and simulations.

This is value-added assessment data and should remain the intellectual property of the vendors. As such, it should not be covered under assessment technology standards. This is an area where vendors should be encouraged to compete to find the most effective ways to integrate new media into their assessments.

3.2.28 Accessibility. How do technology standards ensure that the platforms are accessible to all persons with disabilities? How can technology standards ensure the availability of accommodations based on the individual needs of persons with disabilities? What factors are important to consider so that accessibility capabilities can be included within an interoperable technology standard, both for end-users, as well as operators, teachers, and other administrators? How are issues related to Universal Design for Learning (UDL) relevant to standards for accessible use? How can technology standards provide for, improve, or enhance Section 504 and 508 of the Rehabilitation Act compliance for assessment technology?

Compliance with §504 and §508 should be treated just like FERPA compliance. However, unlike FERPA compliance, accessibility relates much more to the user interface, rather than the underlying data system. Therefore, it should be a requirement that vendors who receive RTTA funding should comply with these regulations, but it does not need to become part of the assessment technology standards that relate to the data platform.

3.2.29 English Learners. How do technology standards ensure that assessment platforms support the assessment, reporting of results, and other capabilities related to the assessment of English learners?

Our state-level databases include persistent storage for assessments of English language learners as well as other language-related fields (language or languages spoken at home, years in the US, country of birth, etc.). Language data and language assessment should definitely be included in the assessment technology standards.

Questions about process and IP for technology standards development include:

3.2.30 Transparency. How do the organizations that develop assessment technology standards approach development and maintenance activities? Is it common for such work to be performed in an unrestricted or open public forum? Are there examples of organizations conducting technology standards development through private (e.g., membership-driven) activities? Are the final work products produced through standards-development activities made publicly available in a timely manner? If not, when or for how long is it necessary to keep these products private? What circumstances require, justify, or benefit from protecting trade secrets or intellectual property?

We recommend that the intellectual property for the assessment technology standards be available as an open source project, with the source code available, but not modifiable on the cloud without approval by the governing board and assurances of backwards compatibility. However, the governing board, in order to be effective, needs to have a limited number of members. Any work products of the governing board should be public though, in our opinion.

This assumes that the assessment technology standards would relate only to the technology platform and not the content that is developed as test items, etc. That should remain the IP of the vendors and SEAs, or whoever else generates that content.

3.2.31 Participation. Does the development of assessment technology standards depend on membership fees from individuals and organizations who wish to contribute to development and maintenance activities? Are there requirements for "balance" within membership across different constituencies? What are the cost and structure of such memberships? Are there viable alternative methods for generating revenue necessary to conduct the work? What are the most realistic and useful ways to generate participation, fund work, and ensure public access to a technology standards-setting process?

Development of these kinds of standards requires huge amounts of work. In our opinion, a funding source should pay someone to do any remaining necessary development, including documentation, while reporting progress and issues to a governing board. The governing board should be chosen from members of a wider association, and membership in that association could charge minimal membership dues. The SIF Association (SIFA), we believe, is a good model to look at. However, SIFA is heavily influenced by the vendors. Many of the proposals here may not win immediate favor from the vendors, especially from some of the larger ones, who may feel that their business model is threatened. Any association that is organized to develop this concept further needs to ensure that non-commercial, educational interests are paramount. This could be accomplished by guaranteeing minimal board membership of governmental organizations and foundations, such as SEAs, CCSSO, DQC, the Bill and Melinda Gates Foundation, and the Dell Foundation, among others.

3.2.32 Availability. What are the costs associated with final publication of technology standards, and with all supporting materials for those standards, and can these assessment products be made available at nominal or no cost to users? Do technology standards require restrictions for use or application, including limitations on derivation, resale, or other restrictions? Is it appropriate to obtain patent, copyright, or trademark protections for assessment technology standards? Are the publications for technology standards and materials provided in a machine-readable, well-defined form? Are there restrictions or limitations on any future application of the publications and materials after initial release? Are developer-assistance materials (e.g., Document Type Definitions, test harnesses, code libraries, reference implementations) also made available free under an open-license? In what circumstances should technology standards-setting organizations retain rights or control, or impose restrictions on the use of publications, derivations, and resale or developer-assistance technologies, as opposed to open-licensing everything? When should materials be made freely available (that is, at no cost to the consumer) while still retaining most or all copyright license rights?

We recommend that supporting materials as well as the assessment technology standards themselves should be provided for free, though there will be costs associated with developing these materials. A funding source will need to be found to cover these development costs.

3.2.33 Derivation. For technology standards, do copyright licenses for publications and all supporting materials and software licenses for software artifacts permit the unrestricted creation and dissemination of derivative works (a.k.a. "open licensed")? Do such open licenses contain restrictions that require publication and dissemination of such works in a manner consistent with the openness criteria described by, for example, a GNU Public License (a.k.a. "viral licensed") or an MIT Public License (a.k.a. "academic licensed")? Are there policies or license restrictions on derivative works intended to prevent re-packaging, re-sale, or modifications without re-publication for assessment technology standards?

If we assume that encouraging innovation in the education technology market is a primary goal of developing assessment technology standards, then there are many possible ways to accomplish this. Some people argue that keeping derivative works under open licenses encourages more development efforts by many smaller companies. Others point out that protecting derivative works encourages more investment by companies who wish to add value to an open standards project. There is currently a lot of debate on this topic right now. We do not believe that a definitive answer to this very important question is yet available. There are some very creative ways of both sharing and

protecting IP that can lead to much higher levels of innovation. In some ways this is a core feature of our proposal as outlined in this response. We are willing to participate in any future discussions concerning this and other aspects of licensing and IP. We feel that these kinds of decisions will need to be made by a larger group, such as the eventual governing body that is created to govern these assessment technology standards.

3.2.34 Licensing Descriptions (for materials contained within the standard, not for the standard's licensing itself). How do technology standards address licensing terms for assessment resources described within the technology standard? Are there successful technology standards or approaches for describing a wide variety of license types, including traditional per-use licensing, Web-fulfillment, free (but licensed), open (but licensed, including commercial or non-commercial use permitted), and public domain status. Are there other resource licensing issues that should be addressed within a technology standard as a best practice?

We recommend that a committee investigate the various options for licensing, with the goals of encouraging innovation, while keeping the fundamental data structures open source. See also our response to the question immediately above.

As noted above, there is also a presentation that accompanies this response and is attached to this email.

We would very much like to make a presentation in person with a demonstration of the existing data system that we are willing to contribute to these efforts, at any time that is convenient. Please let us know if you would like any additional information.

Primary Contact: Patrick McDermott

Direct: (425) 739-3249 | **Main:** (425) 739-3202 | **Mobile:** (435) 459-2183
Advaiya, Inc. **|** patrick.mcdermott@advaiya.com **|** **www.advaiya.com**