# CHAPTER 2 – Lowering the Barrier to Adoption of Intelligent Tutoring Systems through Standardization

**Robby Robson[1] and Avron Barr[2]**
[1]Eduworks Corporation; [2]Aldo Ventures

## Introduction

To have an impact, a learning system must be effective and must be used. In the case of ITSs, studies have repeatedly and consistently shown significant learning gains (Dodds & Fletcher, 2004; Durlach & Ray, 2011; Kulik & Kulik, 1991; VanLehn, 2011). Despite their demonstrated value and over thirty-year history (Barr, Beard & Atkinson, 1975; Sleeman & Brown, 1982), the use of ITSs remains restricted to research projects and a few commercial applications. There are success stories, but in general, such systems are not being adopted at the rates that their effectiveness would justify. As stated by Blessing, Gilbert, Ourada & Ritter (2009) in a paper on authoring model-tracing cognitive tutors, "*ITSs, including model-tracing tutors, have not been widely adopted in educational or other settings, such as corporate training. Perhaps the most successful deployment of model-tracing tutors is Carnegie Learning's Cognitive Tutors for math, which are in use by over 1300 school districts and by hundreds of thousands of students each year. After this notable success, however, most educational and training software is not of the ITS variety*."

Multiple factors could be impeding adoption, but two factors in particular stand out. The first is that ITSs are designed to be standalone systems that do not communicate or interoperate with other systems used to support learning, education, and training. The second is the sheer complexity of ITSs and the concomitant effort it takes to develop them (Murray, 2003).

With this as motivation, this chapter explores how standardization might help ITSs fit into learning ecosystems and simplify their design. Here we suggest focusing on standards for exchanging learner information among systems, *not* on standards for internal components. We also point out that the requirement to exchange learner information emphasizes the problem of determining which adaptations are responsible for the positive learning effect sizes observed when using ITSs. We believe that the suggested approach will enrich learner models, encourage developers to separate their innovative and proprietary adaptation engines from the portions of ITSs that interoperate with other learning systems, and ultimately, transform ITS architecture in ways that will make them easier to implement and adopt.

## Standards

As Christensen & Raynor (2003) point out, interoperability and standardization enable competition and the growth of supply chains. In learning technology, for example, the emergence of learning management systems (LMSs) in the 1990s disrupted the print-based supply chain from authors to publishers to schools to learners. Standards such as IMS Content Packaging (IMS Global Learning Consortium, 2004), AICC Computer Managed Instruction (AICC, 2004), Sharable Content Object Reference Model (SCORM) 1.2 (Dodds, 2001), SCORM 2004 (ADL, 2006), and IMS Common Cartridge (IMS Global Learning Consortium, 2011) reestablished much of the same chain by allowing courseware to be produced independently and used by any compliant LMS. Arguably, the eLearning industry, which is a multi-billion dollar industry today (Adkins, 2011; Bersin, 2012; Global Industry Analysts, 2013), would not exist without these standards, and it is reasonable to assume that some standardization is needed to spur the adoption of ITSs.

At the same time, many proposed learning technology standards have failed to be completed or failed to achieve significant adoption. These proposals include standards for learner information and learner models, architectures, repository interoperability, intelligent agent communication, competency definitions, student identifiers, and many others. For example, the IEEE Learning Technology Standards Committee web site from December 2000 (IEEE, 2000a) shows 17 standards projects, almost all of which are still relevant today (including a standard for learner models). Less than half of these were completed in any form, and almost none have achieved any significant adoption. In most cases, failed efforts were driven more by an example and a vision rather than by a market/adopter pain point and real products. These failed standards, as described in Robson (2006), were "innovation and research-driven standards" rather than "market-driven standards." If standards are to be developed to spur the adoption of intelligent tutoring systems, they must be defined by market needs, or else they will have little effect.

## Market Needs

What are those market needs? If we accept that ITSs are effective, the immediate market needs of potential customers are to (1) implement them with as little integration effort and work disruption as possible; and (2) obtain the data that are required for learning management and talent management and that can be used to demonstrate pedagogical impact and monitor costs. Customers may tolerate some inconvenience to get a more effective learning solution, but they are probably not willing to reconfigure their entire learning infrastructure, retrain all of their users, radically alter their established workflows, or give up on tracking grades and course completion status.

Unfortunately, ITSs make little or no attempt to exchange even basic results data with other systems. This isolationism is typical of new learning technologies. Early LMS and assessment engine products were the same, and so are massive open online courses (MOOCS) and the Khan Academy. The following exchange from the Google Khan Academy Developers Group (Azevedo & Ojeda, 11/19/12) typifies the reluctance of new product developers to address interoperability and the frustration of early adopters who typically run their daily operations via institutional learning management systems. (Typos and spelling corrected).

> DA: *Hello developers. Is Khan Academy / Khan Exercises SCORM Compliant?*
>
> MO: *hi D. Sorry to say, we do not. Is there some definite advantage to supporting it other than this graphic from scorm.com?" (Graphic shows reduced costs from using SCORM)*
>
> DA: *Hello and thanks for the quick response. I find one big use for SCORM. If a student uses multiple learning platforms it would be nice for grades and progress to be shared across the platforms. One example would be: I'm attending a class in Khan Academy, like Algebra II, but I find another site/course and I want to make Algebra III in that new site/course. If I could export/import my certified data across platforms that would be very flexible for students. In the long run people will like that the time spent on Khan or another site/course is certified and flexible, like in real universities there is equivalence in subjects and grades/progress.*[1]

The immediate market needs of potential customers are to track results and hold down implementation costs. These are conditions for diffusion of the technology, but once diffusion starts, other requirements will appear. For example, since ITSs provide individualized learning, it is likely that students will frequently switch among different systems. If the one system has gathered data about a learner's cognitive or affective characteristics, other systems can make use of it. Existing standards such as IMS GLC

[1] We note that one of the motivations for moving from SCORM to the Experience API is that SCORM does not address learners working on multiple platforms.

Learner Information Package (IMS Global Learning Consortium, 2005) and the Europass (Cedefop, 2012) can be used for transcript data, curriculum vitae (CV) data, and high level competencies (e.g., language skills), but new standards will be needed for expressing and comparing learner information.

In addition, adoption will bring renewed demands for avoiding lock-in. Even if the same content cannot be plugged and played in multiple intelligent tutoring systems, customers will want to leverage content they have already acquired and will want to outsource ITS development to multiple sources. To achieve this requirement, there must be some reasonable separation between the "system" software, "content" such as text and multimedia presented to the learner, and "interfaces," including the user interface and those that exchange data with other learning systems. This is diagrammatically shown in Figure 2-1.

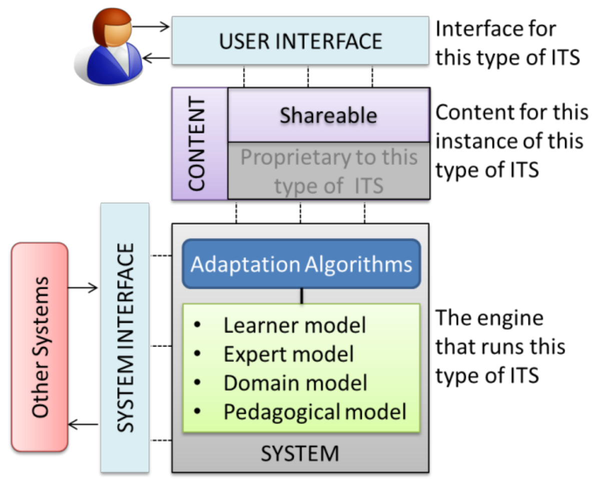


**Figure 2-1. System, Content and Interfaces**

## Standards for Learner Information

The first requirement for adoption is for ITSs to "play well" in current learning environments. This can be done by requiring conformance to SCORM and the Experience API or Tin Can API (Advanced Distributed Learning, 2013) or by adopting the IMS approach based on Learning Tools Interoperability (IMS Global Learning Consortium, 2012). The important point is that ITS developers should not ignore these standards if they want their systems to be adopted.

Requirements to exchange even very coarse data such as completion status and quiz results will naturally lead developers to reexamine their system architectures and, hopefully, lead to a separation of components such as that illustrated in Figure 2-1. For example, the Experience API, planned for the next generation of SCORM, uses an "actor – verb – activity" paradigm and is designed for compatibility with semantic inference engines (Poltrack, Hruska, Johnson & Haag, 2012). Using this application programming interface (API), a score on a quiz can be reported as a series of statements such as "Student completed quiz" and "Student scored 95." An inference engine embedded in an LMS or other learning system might additionally know that "Quiz assesses Quadratic Formula" and conclude that "Student demonstrates competency in Quadratic Formula." The requirement to generate such triples and support the API will suggest using a similar structure to store and exchange other data, including learner information. This may or may not be the optimal design choice for a particular ITS, but experience shows that developers of new systems often use standards as guidelines for functionality and design (Devedzic, Jovanovic & Gasevic, 2007).

Independent of how it is represented, the key question for standardization is what information should be exchanged. In other words, what should an ITS be telling other systems, including other ITSs, other enterprise learning systems, and other applications used by instructors, students, managers, and researchers? Since ITSs are valuable because of their positive effects on learning, this information should consist of the data responsible for attaining this effect. In other words, the question of what information should be exchanged is the question of what student data are required to achieve near optimal adaptation. This is a special case of the more general question, posed by Durlach (2012) and Ray Perez (personal communication), of what functionality in ITSs has the most effect on learning outcomes and how much of this functionality is needed in practice. We do not know the answer, but we can nonetheless make some reasonable conjectures concerning the types of learner information might be included:

- *Educational records and high level competencies such as language skills and flight certifications*. Standards exist for these data. It is not clear what inferences an intelligent tutor can make directly from them, but since human tutors find them valuable, there is an argument for including them.
- *Competencies (including Skills, Knowledge, Abilities, Outcomes, Objectives) and level of competence*. These have also been standardized and represent data that are crucial for talent management, job placement, and many other applications. They are also used for AICC/SCORM-type sequencing, and at a fine level of granularity, can be equated with domain topics.
- *Data in affective, motivational, and social dimensions*. Cognitive models are more common and better understood, but many adaptations rely on data in these dimensions (Dimitrova, 2009). For example, tutors may observe how students react to different types of stimuli and discover beliefs and attitudes that other systems could use to select instructional content and strategies.
- *Goals, including learning goals and mission/task goals*. These goals are also data that a human tutor would want to know and that might be important for adaptation.
- *Physical adaptations, such as location, device capabilities, ambient light, and accessibility data*. Accessibility data have been standardized, and these other data can clearly be used in many ways.

As an historical note, we believe that the task of standardizing learner data was first taken up by the IEEE Learning Technology Standards Committee. In 2000, the scope statement for the Learner Model working group (IEEE, 2000b) read: "*This standard will specify the syntax and semantics of a 'Learner Model,' which will characterize a learner (student or knowledge worker) and his or her knowledge/abilities. This will include elements such as knowledge (from coarse- to fine-grained), skills, abilities, learning styles, records, and personal information. This standard will allow these elements to be represented in multiple levels of granularity, from a coarse overview, down to the smallest conceivable sub-element. The standard will allow different views of the Learner Model (learner, teacher, parent, school, employer, etc.) and will substantially address issues of privacy and security*"

Its purpose consisted of five items:

1. *To enable learners (students or knowledge workers) of any age, background, location, means, or school/work situation to create and build a personal Learner Model, based on a national standard, which they can utilize throughout their education and work life.*
2. *To enable courseware developers to develop materials that will provide more personalized and effective instruction.*
3. *To provide educational researchers with a standardized and growing source of data.*
4. *To provide a foundation for the development of additional educational standards, and to do so from a student-centered learning focus.*
5. *To provide architectural guidance to education system designers.*

This project was known as "Personal and Private Information" (PAPI) and never turned into a standard for reasons beyond the scope of this chapter. It did, however, have some important attributes. It considered highly granular information, a consideration abandoned by the IMS Learner Information Package and Europass standards. It had personalized learning and architectural guidance as important use cases. Architectural guidance is important because the complexity of ITSs is a barrier to adoption. Interchange standards do not dictate how data are stored or processed internally, but they tend to influence how systems are designed. Devedzic et al. (2007), for example, give an extensive overview of eLearning standards aimed at developers of web-based eLearning systems because "*they establish high-level principles for organizing learning resources and developing web-based education (WBE) applications*." As things stand, an ITS constructs its internal model of a specific learner solely from its interactions with that learner. The modeling framework is constructed from cognitive and learning science, but the only data used to instantiate the models come from interactions with the ITS. The ability to retrieve learner information from other systems will change that and encourage developers to separate their innovative and proprietary adaptation engines from the portions of their systems that interoperate with other learning technologies.

### Models Should Not Be Standardized

We have suggested that learner information should be standardized but have avoided suggesting the same for *learner models*. In fact, we do not believe that standardization is appropriate for any of the models associated with the inner workings of ITSs.

In the abstract, ITSs adjust their instructional strategy based on a learner model, expert model domain model, and pedagogical model (Durlach & Ray, 2011), but very few real-world ITSs use all these components and the nature of each of these components can vary wildly. For example, the model-tracing cognitive tutors built by Carnegie Learning (Anderson, Corbett, Koedinger & Pelletier, 1995) encode expert knowledge in logical production rules and give hints to students when their input violates one of the rules, whereas ALEKS (2011) tracks which topics a student has mastered and, using data gathered from a large population of previous learners, infers which new topics the student is likely to be able to master next. In contrast to cognitive and model-tracing tutors, the AutoTutor family (Graesser et al., 2004; Hu et al., 2009) has less explicit encoding of domain knowledge. These tutors use semantic analysis to determine how relevant student input is to content that has been presented, and in some cases, adjust their strategies based on affective states determined by sensor data (D'Mello & Graesser, 2010; D'Mello & Graesser, 2012). An examination of examples cited in sources (Graesser, Jeon & Dufty, 2008; Murray, 2003; Ohlsson & Mitrović, 2006; Woolf, 2009) reveals even further diversity in the way that ITSs operate.

This diversity exists for a good reason. Motivated by the two-sigma effect size observed with one-on-one human tutoring, ITSs attempt to replicate this experience with technology (Kulik & Kulik, 1991; VanLehn, 2011). Each system's technical approach to building and using models to achieve this goal is its principal locus of innovation and a chief source of differentiation in the marketplace. There is a significant difference between macro-adaptations, which persist over time and can be used by multiple systems, and micro-adaptations, which are ephemeral and lie fully in the domain of a single system. The candidate standardization categories in the previous section are all macro-adaptations. Since ITSs derive their competitive advantage from micro-adaptations and the models that enable them, standardization of these models would not be accepted and, if accepted, may hinder innovation.

## Recommendations for GIFT and Future Research

GIFT (Sottilare, 2012) provides a testbed in which multiple ITSs can function. Its architecture contemplates that ITSs will exchange data with an LMS. Adoption of GIFT would encourage ITS developers to include LMS-compatible reporting mechanisms, which we argue is a necessary step for market diffusion of ITS technology.

GIFT also includes a learner module, which we would view as learner information exchange service. There are existing standards that can serve as the basis for constructing such a service, and any reasonable candidate standard for learner information exchange will be extensible enough to evolve over time.

The key question, however, is what data ITSs actually need to exchange to enhance learning gains and support the other intelligent systems and apps used by teachers and students. This is a difficult research question, and the history of learning technology standardization teaches that abstract and theoretical answers to questions of this nature do not lead to practical and adoptable standards. The best approach is likely repeated cycles of experimentation and observation. That approach may require creating tutors that implement well-controlled and limited functionality, measuring their effect on learning, and iteratively refining standards for learner information exchange based on the results. A framework such as GIFT is ideally suited as the "breadboard" for such experimentation.

## References


Adkins, S. (2011). The U.S. Market for Self-paced eLearning Products and Services: 2010-2015 Forecast and Aanalysis.

ADL. (2006). Sharable Content Object Reference Model (SCORM) 2004. (3rd ed.). Alexandria, VA: Advanced Distributed Learning initiative.

Advanced Distributed Learning. (2013). Experience API, from http://www.adlnet.gov/capabilities/tla/experience-api

AICC. (2004). CMI Guidelines for Interoperability (Rev 4.0 ed.): Aviation Industry CBT Committee.

ALEKS. (2011). ALEKS Home Page Retrieved October 20, 2011, from http://www.aleks.com

Anderson, J. R., Corbett, A. T., Koedinger, K. R. & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The journal of the learning sciences, 4*(2), 167-207.

Azevedo, D. & Ojeda, M. (11/19/12). SCORM, from https://groups.google.com/forum/?fromgroups=#!msg/khanacademy-developers/pEts3twbk5g/3Y9SL4sHgZoJ

Barr, A., Beard, M. & Atkinson, R. C. (1975). A rationale and description of a CAI program to teach the BASIC programming language. *Instructional Science, 4*(1), 1-31.

Bersin, J. (2012). LMS 2013: The $1.9 Billion Market for Learning Management Systems: Deloitte.

Blessing, S. B., Gilbert, S. B., Ourada, S. & Ritter, S. (2009). Authoring model-tracing cognitive tutors. *International Journal of Artificial Intelligence in Education, 19*(2), 189-210.

Cedefop. (2012). Europass Retrieved February, 2013, from http://europass.cedefop.europa.eu/en/home

Christensen, C. M. & Raynor, M. E. (2003). *The innovator's solution: Creating and sustaining successful growth*: Harvard Business Press.

Devedzic, V., Jovanovic, J. & Gasevic, D. (2007). The pragmatics of current e-learning standards. *Internet Computing, IEEE, 11*(3), 19-27.

Dimitrova, V. (2009). *Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling* (Vol. 200): Ios PressInc.

Dodds, P. (2001). Sharable Content Object Reference Model v 1.2: Advanced Distributed Learning Initiative.

Dodds, P. & Fletcher, J. D. (2004). Opportunities for New "Smart" Learning Environments Enabled by Next-Generation Web Capabilities. *Journal of Educational Multimedia and Hypermedia, 13*(4), 391-404.

Durlach, P. J. (2012). Vanilla, Chocolate, or Chunky Monkey: Flavors of Adaptation in Instructional Technology. *iFest 2012*, from http://www.adlnet.gov/wp-content/uploads/2012/08/Durlach_Adaption_in_IT_iFest-2012.pdf

Durlach, P. J. & Ray, J. M. (2011). Designing adaptive instructional environments: Insights from empirical evidence *Army Research Institute Report*. Arlington, VA.

Global Industry Analysts. (2013). Global eLearning Market to Reach $107.3 Billion by 2015. Reported on PRWeb., from http://www.prweb.com/releases/elearning/corporate_elearning/prweb4531974.htm

Graesser, A. C., Jeon, M. & Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes, 45*(4-5), 298-322.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A. & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods, 36*(2), 180-192.

Hu, X., Cai, Z., Han, L., Craig, S. D., Wang, T. & Graesser, A. C. (2009). *AutoTutor Lite*.

IEEE. (2000a). IEEE Learning Technology Standards Committee (LTSC) Retrieved February, 2013, from http://web.archive.org/web/20010217054051/http://grouper.ieee.org/LTSC/

IEEE. (2000b). IEEE P1484.2 Learner Model Working Group Retrieved February, 2013, from http://web.archive.org/web/20010221110739/http://ltsc.ieee.org/wg2/index.html

IMS Global Learning Consortium. (2004). IMS content packaging specification v 1.1.4.

IMS Global Learning Consortium. (2005). IMS Learner Information Package.

IMS Global Learning Consortium. (2011). IMS Common Cartridge Specification.

IMS Global Learning Consortium. (2012). Learning Tools Interoperability, from http://www.imsglobal.org/toolsinteroperability2.cfm

Kulik, C.-L. C. & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in human behavior, 7*(1), 75-94.

Murray, T. (2003). An Overview of Intelligent Tutoring System Authoring Tools: Updated analysis of the state of the art. Chapter 17 in Murray, T., Blessing, S. & Ainsworth, S. *Artificial Intelligence*.

Ohlsson, S. & Mitrović, A. (2006). Constraint-based knowledge representation for individualized instruction *Computer Science and Information Systems, 3*(1), 1-22. doi: 10.2298/CSIS0601001S

Poltrack, J., Hruska, N., Johnson, A. & Haag, J. (2012). *The Next Generation of SCORM: Innovation for the Global Force*. Paper presented at the The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC).

Robson, R. (2006). Globalization and the future of standardization. *Computer, 39*(7), 82-84. doi: 10.1109/mc.2006.231

Sleeman, D. & Brown, J. S. (1982). Intelligent tutoring systems.

Sottilare, R.A., Brawner, K.W., Goldberg, B.S. & Holden, H.K. (2012). *The Generalized Intelligent Framework for Tutoring (GIFT)*. Orlando, FL: U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED).

VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist, 46*(4), 197-221.

Woolf, B. P. (2009). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Burlington, MA: Morgan Kaufmann.