Missing Pieces: Infrastructure Requirements for Adaptive Instructional Systems

Avron Barr[1(✉)] and Robby Robson[2]

[1] Institute for Defense Analyses and IEEE Learning Technology Standards Committee, Aptos, USA
avron@ieee.org
[2] Eduworks Corporation and IEEE Standards Association, Corvallis, USA
robby.robson@eduworks.com

Abstract. As the market embraces adaptive instructional systems (AISs) and related data-intensive learning technologies, effective and economical deployment of products from multiple vendors will require critical infrastructure components. Four important types of data must be managed and securely shared across applications: the learner's background and objectives; a profile of the learner's current state of mastery; live data recording the learner's current activities; and metadata describing the available learning activities. Additionally, software tools to manage multi-product learning environments will be needed. This paper explains the importance of these infrastructure components, reviews their state of readiness, and anticipates the benefits they will offer to various stakeholders.

Keywords: Adaptive instructional systems · Artificial intelligence · Competencies · Data analytics · eLearning infrastructure · Experience Application Program Interface (xAPI) · Intelligent Tutoring Systems (ITSs) · Learner modeling

1 Evolving Digital Learning Architectures

Digital learning, i.e., the use of computer- and network-based learning applications, has been adopted in all major education and training markets: schools, colleges, enterprise training, and professional certification. ELearning systems have been used to save costs, broaden markets, improve learning, and make learning more convenient for the learners (anytime, anyplace). Since they were introduced in the 1990's, the focal point of digital learning has been the institutional Learning Management System (LMS). The LMS handles many administrative issues:

- Student rosters and assignments;
- Learning materials, contracts with publishers, and usage statistics;
- Tracking and reporting learner activity and performance;
- Class management and communications; and
- Room and equipment scheduling.

R. A. Sottilare and J. Schwarz (Eds.): HCII 2019, LNCS 11597, pp. 169–178, 2019.
https://doi.org/10.1007/978-3-030-22341-0_14

Since the advent of software as a service (SaaS), many organizations have adopted cloud-based LMS offerings, but the basic configuration remains the same: the multi-purpose LMS is the main learning-specific infrastructure component needed to offer and manage digital learning environments. However, both technological and market developments are changing the way digital learning will be deployed.

1.1 Market Changes

Modern learners often explore multiple cloud-based offerings in their pursuit of knowledge, in addition to those that have been assigned via an LMS. In fact, learners are likely to be affiliated with multiple learning institutions and on-line activity vendors simultaneously. Changes in career paths and reduced corporate investments in in-house employee training have created increased demand for recording and managing an individual's learning over their entire lifetime, from schooling to employment and beyond.

As a result, "learner portability" has replaced "content portability" as the central interoperability issue in digital learning. While we once focused on assuring that all content could run on all LMSs, our biggest problem now is securely sharing data about learner preferences, backgrounds, traits, and achievements across learning systems and across institutional and content provider boundaries. Only through such sharing can we assure that a learner's interactions with all of these systems is informed by the learner's past activity and current state of understanding. The demand for learner portability will only grow over time, with increasing global availability of online opportunities, the "gig economy," and labor market pressures to facilitate frequent re-training and career changes [1].

1.2 Cloud-Based Learning Activities

Simulations, gaming, virtual reality, augmented reality, and other newer technologies are being used more frequently to create immersive, exploratory, and collaborative learning and rehearsal environments. Increasingly, these more complex offerings are delivered via the cloud. The institutional LMS is not designed to manage a learner's activities outside of a single institution or to track learner data across these systems.

1.3 The Data Storm

Learning platforms now include mobile phones, tablets, and virtual/augmented reality devices as well as the instrumentation of learners with devices that measure their location, movements, biometrics, affective state, and instant-by-instant actions. Digital learning activities can generate detailed data about the learner's actions, choices, responses, hesitations, biometrics, affective state, and much more. As learners spend more and more time using these digital platforms, we anticipate a dramatic increase in the variety, velocity, and volume of data they produce.

1.4 Artificial Intelligence

Learning activities are getting smarter. Adaptive Instructional Systems (AIS), for example, use data about a learner's history, preferences, recent performance, and state of mind to offer personalized instruction. In an AIS, the path that a student takes through the material, the modality that instruction is delivered, the style of remediation, and the feedback about progress can all be tuned to the individual learner's needs and preferences. There are several other emerging product categories that use data-intensive algorithms to improve learning and help teachers, students, and administrators including Intelligent Tutoring Systems, learning analytics engines, coaches, recommenders, robo-graders, and pedagogical agents.

In general, the requirements for learner data are largely driven by AI technologies, which are increasingly being applied to education and training in many forms, including machine learning, learning analytics, language and speech processing, computer vision, and affect recognition. The use of AI technologies improves the way learning activities model and interact with learners, e.g. by personalizing their experience, processing written and spoken language, or diagnosing their misunderstandings. Some learning activity publishers, having access to data from tens of thousands or even millions of customers, use machine-learning algorithms to glean insights into learner behavior, effective remediation, and product improvement. As with all AI algorithms, the more data a system has access to, the better the results.

1.5 Architectural Variations

Different kinds of organizations will evolve towards different infrastructure architectures. Public schools, for instance, may rely in part on infrastructure components supplied by their school district or state Department of Education. Corporations, associations, and government organizations may build the entire ecosystem in-house, or rely, partially or completely, on SaaS offerings. Publishers will track how learners use their products in aggregate, but they may not be allowed to track individual leaner progress.

It is not possible at this time to offer concrete guidelines about how to construct an organization's future elearning ecosystem for several reasons:

- Installations in different market segments have different goals, constraints, and already-installed systems;
- The functionality of some infrastructure components is still evolving;
- There is no way to predict how the needed functionality will be combined and packaged in new products; and
- Some important interoperability standards are still in development.

Today, every instance is an experiment. By describing the key infrastructure components, we hope to inform the many architectural decisions that lie ahead for all providers of digital learning.

2 Essential Architectural Components

Modern learning ecosystems must extend the traditional, LMS-based architecture to effectively manage the dramatic increase in learner mobility across suppliers, the variety of learning activities, and the amount of learning data. Specifically, operators will need tools to help them manage several types of data as well as the overall ecosystem configuration:

- **Learner Background**: It is essential that all learning activities have access to data about the learner's history, current objectives, preferences, and current activities. The things any teacher or tutor would want to know about a new student.
- **Competency**: Characterization of the learner's level of mastery relative to the level required to achieve their objectives, as defined by a certifying institution or firm in terms of a competency framework.
- **Activity stream data**: Detailed data about what a learner is doing in real time, required by AI-enhanced learning activities and supplemental products like analytics engines and data visualization tools, which respond to data about what the learner is doing second by second.
- **Learning activity metadata**: Learners, as well as intelligent tutors and recommender systems, benefit from knowing specifics about the available learning activities and their relevance to the learner's current state of understanding and their learning goals.
- **Component registration and configuration management:** Typically, a modern learning organization will deploy multiple systems locally and integrate with an array of cloud-based data providers. Tools to support the easy (automatic) registration and management of those systems will eventually be a necessity.

Some of these needed infrastructure components are available commercially now. Some we expect will be available soon, because prototypes have been used in real applications. And some key components are still on the drawing board. Often commercial products combine functionality, e.g. by adding features to the LMS, so there isn't necessarily a one-to-one correspondence between the functionality needed and the emerging products categories.

2.1 The Learner's Background

Current commercial and research systems that model the learner's background, interests, preferences, objectives, and current state of understanding start off with a blank slate. This practice will eventually become problematic for teachers and learners who use multiple systems simultaneously. It will be necessary to decouple the learner model from the learning activity and to create a means that enables all products from all vendors to effectively share this important information about the learner. Moreover, since over the course of their lives, learners work across institutions and activity providers, the learner model will need to be persistent and independent of any single institution or provider.

To the best of our knowledge, there are no stand-alone products on the market now that maintain a learner model across activities and vendors. In fact, there is no agreed-upon standard format for the data contained in such a learner model. While there has been research in learner modeling for a decade (e.g., Bull and Kay [2]), the demand for products that support multiple learning activities has been more theoretical than market-driven. Now, however, there is an increased real-world need for multiple learning activities to share learner model data.

2.2 Competency Frameworks and Learner Profiles

In research systems, keeping track of the learner's developing mastery of the subject at hand is usually considered part of the Learner Model, along with the background information described above. Nonetheless, it makes sense to architecturally separate competency management functionality from Learner Models for several reasons:

- Background information is typically used and updated less frequently and by a different set of applications than competency information.
- The competencies of interest, the learner's objectives, may vary as the learner moves between learning activities, e.g. when studying Algebra and Geometry in the same semester.
- Descriptions of competencies and rules for managing them are idiosyncratic. Schools, employers, publishers, and government agencies all have their own way of thinking about competencies and ascribing mastery.
- Security considerations about the learner's progress and status may be very different from those for background data.

In addition to describing the learner's current state and learning objectives, competency information is used to describe course requirements; degree requirements; certificates and credentials; job requirements; and the intended use of learning materials and activities. Every educational institution, corporate HR department, and learning activity vendor has its own framework for describing relevant competencies.

Managing competency data requires new tools. Their functionality may differ a bit in different markets, and again, future products may implement different sets of features, but digital learning organizations will need tools to work with the following aspects of competency management:

- **Competency framework**: A structured representation of the knowledge, skills, and abilities of interest. Frameworks capture the titles and descriptions of competencies, relationships among competencies, and data such as levels and assessment methods.
- **Competency evidence**: Each organization defines the nature and format of the evidence it accepts in evaluating progress towards mastery.
- **Learner competency profile**: Profiles track each learner's level of mastery on each competency relative to their learning goals. Some organizations may incorporate a model of "decay over time" into the way they model the learner's state of knowledge.
- **Evidence evaluation and rollup rules**: Different organization have different rules for accepting evidence of learning and different algorithms for aggregating that evidence to draw conclusions about learners' progress. Rules for concluding the

learner's state of understanding of a higher-level competency, based on their mastery of sub-competencies, will also vary across organizations.

Some LMSs and enterprise talent management systems offer tools for building and maintaining competency frameworks. These tools, however, are focused on a single institutional context and do not conform to standards that have been developed to compare or share competency frameworks. More recently, projects such as the Competency and Skills System project (CaSS, www.cassproject.org) and OpenSALT (www.opensalt.org) have been launched with the capability of sharing competency frameworks across multiple systems. CaSS, in particular, enables frameworks to be imported and exported in standardized formats, conforms to web standards such as linked data, and enables users to:

- Create, store, and share competency frameworks, specifying local terminology, rollup rules, and associated assessments;
- Compare and exchange competency definitions among organizations, recognizing the likely differences in terminology and semantics [3];
- Specify relationships among competences and rules for concluding mastery from evidence and from mastered sub-competencies.

In addition, products are emerging that enable learner profiles to be stored, updated, and consumed by multiple learning, training, and staffing systems. These products include CaSS, MARI (www.mari.com), Viridis (www.viridislearning.com), Degreed (www.degreed.com), and many others, each with its own focus and target market. There are even more products that store learner profiles internally for their own use. Standardizing a learner profile component could enhance the value and lower the production cost of all of these systems [4].

2.3 Runtime Activity Stream Data

Perhaps the most advanced new infrastructure components, in terms of commercial product availability, are tools to collect and manage the runtime data generated during learners' actual learning activity. Many types of applications use this activity stream data [5], including:

- AI-enhanced tutors, recommenders, and adaptive instructional systems that track what a learner is doing in real time to make diagnoses and to offer useful remediation and recommendations;
- A variety of learning analytics and data visualization products that use real-time data to give feedback to students, apprise teachers about the status of their students, and issue early warning about learners who need attention;
- Publisher applications that monitor runtime data to analyze how learners use their materials, looking for insight into problematic content.

The Learning Record Store (LRS) is an emerging database product category for collecting and managing learner activity stream data. Most LMSs now incorporate an LRS, and most independent LRS products on the market are stand-alone "analytics" systems that typically offer more robust data visualization functionality than the LMSs that produce the data. The US Advanced Distributed Learning Initiative sponsored development of the protocols for sending activity stream data to an LRS and for querying an LRS [6]. LRSs can exchange data with any conformant learning activity and with each other.

The xAPI protocol is relatively mature and we expect will soon be published as an IEEE standard. Many installations now have multiple LRSs, each monitoring some or all learner data streams for specific purposes. Security and privacy considerations will influence the architecture of LRSs and the implementations of xAPI-based data sharing.

2.4 Learning Activity Metadata

Related to the problem of sharing competencies across institutions, where each has its own way of describing learning objectives and learner mastery, there are problems with current approaches to describing learning activities and materials. Current standards for describing learning material are largely bibliographic, with a few nods to pedagogical considerations, such as grade level, reading level, and links to national school curricula standards or textbook chapters [7]. In a world where AI-enhanced tutors and recommenders are trying to identify exactly the right next step for learners, based on knowledge about their past activity and current state of knowledge, the current schemes fall short.

Unfortunately, research doesn't bode well for a general solution. Pilot studies using the Advanced Distributed Learning Initiative's Total Learning Architecture [8, 9], a framework for integrating advanced learning activities and applications, showed that different recommenders and intelligent tutors used different schemes for tagging learning activities with the information they use to make decisions. For example, some recommender engines want to see content tagged as "intro vs. easy vs. hard," while others might base their recommendations on a formal pedagogical framework such as Bloom's taxonomy [10]. The precise parameters on which each recommender bases its decisions seem to be part of the recommender's "secret sauce" and hence a poor candidate for standardization at this time.

Of course, each learning activity needs to be tagged with the appropriate metadata. Publishers are able to adequately describe their own learning activities, for their own instructional environment and recommendation engines, but tagging content for general use by recommenders and other AI-enhanced products can involve a great deal of manual labor. AI language processing techniques have been explored for many years as a way to automate metadata tagging [11, 12].

2.5 Component Registration and Configuration Management

Finally, broad adoption of advanced systems and newly enabled pedagogical initiatives such as competency-based training or mastery-based education will require new tools. Administrators will need to add, remove, and monitor all of these infrastructure components and control communications with the tools that students and teachers are using and with the learning activities themselves, whether they are located locally or in the cloud. All stakeholders, not just administrators will benefit from these tools. For publishers, for instance, one might imagine automatic registration of a new cloud-based learning activity, including alignment of competency descriptions [13].

3 Infrastructure and the Economics of Adaptive Instruction

Privacy concerns will shape, and possibly impede, market acceptance of intelligent learning technologies [14]. The full commercial development of all of the infrastructure components needed to support the economic deployment of Adaptive Instructional Systems and related AI-enhanced products may be years off. To deal with natural privacy concerns, the industry needs to present a strong argument about the value of storing and sharing learner data, now. In our opinion, AI technologies will continue to drive the demand for more learner data, while privacy concerns will increasingly throttle the sharing of that data in many markets. Infrastructure tools that allow individuals and institutions to define the rules about the procurement and deployment of data are key to assuring the community and the public that their data is need and that it is managed professionally.

In addition to addressing data privacy, another benefit of the supportive infrastructure described earlier is facilitating the sharing of data from multiple vendors among AI-enabled products. In some markets, publishers with adaptive learning products in multiple subject areas have begun to share internal data elements, such as background and competency info, across their product lines. Sharing data across apps benefits learners by improving personalization. For example, a Geometry tutor might be able to use information about the learner's knowledge of Algebra.

As AIS products mature, and as customers acquire and deploy multiple systems from different vendors, the pedagogical, economic, and technical inefficiencies of the monolithic AIS model will become increasingly evident. Education and training organizations that are running multiple adaptive courses will begin to build out infrastructure to make it easier for them to install, deploy, use, monitor, evaluate, and change out AIS products. The resulting technical infrastructure will in turn simplify requirements and processes for product developers, teachers, publishers, and students, as well as the IT managers.

Market adoption of key standards is essential to realizing economic gains over the lifetime of AI-enhanced products. Every infrastructure element depends on the implementation of standard data protocols that are implemented by all the vendors. Several research projects, including the US Advanced Distributed Learning Initiative's Total Learning Architecture and the US Army Research Lab's Generalized Intelligent Framework for Tutoring [15, 16] have explored the issues involved in deploying complex adaptive systems. Both of these projects are simultaneously contributing to IEEE standards working groups that will define the software interfaces needed for these infrastructure components [17]. Current relevant projects at the IEEE Learning Technology Standards Committee include:

- xAPI: base standard, best practices, and soon, xAPI profiles
- Adaptive Instructional Systems: (concept and definitions, component interoperability, and best practices for evaluation of an AIS)

- Competency definitions and frameworks
- Child and Student Data Governance: definitions and best practices
- Federated Machine Learning: how to you use big data techniques when the data as distributed across many institutions that are not allowed to share?

At the present time, some of the key products needed to build the infrastructure described in this paper do not exist or are in a pre-product (custom-built) state of evolution. Also, we expect to see the functionality and configuration of these infrastructure products evolve differently across market segments, including: K12, higher education; enterprise and military training; and professional certification. Customers are just beginning to use AI-enhanced learning products, but they will soon find that infrastructure components are need to realize the potential benefits of AI and to manage these inherently more complex learning ecosystems. Operators will also need to establish policy in new areas, e.g., requiring that publishers generate certain data during learning and store it in the right place. Today, every installation is an experiment. Eventually, we will have the products and experience needed to bring the full power of AI to education and enterprise training.

References

1. Robson, R., Barr, A.: Learning technology standards - the new awakening. In: Sottilare, R., Brawner, K., Sinatra, A., Goldberg, B. (eds.) Proceedings of the Sixth Annual GIFT Users Symposium: US Army Research Laboratory (2018). https://www.gifttutoring.org
2. Bull, S., Kay, J.: Open learner models. In: Nkambou, R., Bourdeau, J., Mizoguchi, R. (eds.) Advances in Intelligent Tutoring Systems. SCI, vol. 308, pp. 301–322. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14363-2_15
3. Credential Engine: The credential engine - moving credentialing forward (2018). http://www.credentialengine.org/. Accessed June 2018
4. Robson, R., Barr, A., Fletcher, J.D.: Universal learner profiles. To appear as an Institute for Defense Analyses Report (in preparation)
5. Downes, A.: Learning analytics dimensions: learning experience analysis (2019). https://www.watershedlrs.com/blog/learning-experience-analysis
6. ADL Initiative: The xAPI overview (2018). https://www.adlnet.gov/research/performance-tracking-analysis/experience-api/
7. LRMI: LRMI - a project of DCMI (2018). From Dublin Core Metadata Initiative: http://lrmi.dublincore.org/. Accessed June 2018
8. ADL: Total Learning Architecture (TLA) (2017). From Advanced Distributed Learning initiative: https://www.adlnet.gov/tla/. Accessed 20 Jan 2017
9. Smith, B., Gallagher, P.S., Shatz, S., Vogel-Walcutt, J.: Total learning architecture: moving into the future. In: Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) (2018)
10. DeFalco, J.: Proposed standard for metadata tagging with pedagogical identifiers. In: Sottilare, R., et al. (eds.) Proceedings of the Workshop on Standards for Adaptive Instructional Systems at the Intelligent Tutoring Systems Conference, Montreal (2018)
11. Cardinaels, K., Meire, M., Duval, E.: Automating metadata generation: the simple indexing interface. In: International World Wide Web Conference Committee (IW3C2). ACM (2005). 1-59593-046-9/05/0005

12. Simko, M., Bielikova, M.: Automated educational course metadata generation based on semantics discovery. Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technology, Slovak University of Technology (2009)
13. Schema.org.: Alignment object (2018). From Schema.org.: https://schema.org/AlignmentObject. Accessed June 2018
14. Herold, B.: inBloom to shut down amid growing data-privacy concerns, 24 April 2014. From Education Week: http://blogs.edweek.org/edweek/DigitalEducation/2014/04/inbloom_to_shut_down_amid_growing_data_privacy_concerns.html. Accessed June 2018
15. Sottilare, R.A., Brawner, K.W., Goldberg, B.S., Holden, H.K.: The Generalized Intelligent Framework for Tutoring (GIFT). Concept paper released as part of GIFT software documentation. U.S. Army Research Laboratory—Human Research & Engineering Directorate (ARL-HRED), Orlando, FL, USA (2012)
16. Sottilare, R., Brawner, K., Sinatra, A., Johnston, J.: An updated concept for a Generalized Intelligent Framework for Tutoring (GIFT). US Army Research Laboratory, Orlando, FL, USA (2017)
17. Robson, R., Barr, A., Sottilare, R.: Overcoming barriers to the adoption of IEEE standards. In: AI in Education (AIED), London (2018)