Total Learning Architecture Data Model for Analytics and Adaptation

Brent Smith
ADL Initiative (Contractor)
Orlando, FL
brent.smith.ctr@adlnet.gov

Sae Schatz, Ph.D.
ADL Initiative
Washington, DC
sae.schatz@adlnet.gov

John Turner
Office of the Chief Data Officer
Washington, DC
john.d.turner120.civ@mail.mil

ABSTRACT

The Advanced Distributed Learning (ADL) Initiative started the Total Learning Architecture (TLA) project in 2015 with the goal of establishing a technical framework across education and training for data-driven lifelong learning. The TLA data strategy is built around commercial, open-source standards that organize learning-related data, syntactically and semantically, to support interoperability across diverse organizations and products.

As a policy-driven architecture, the TLA does not require any mandatory components; although, it does define required functions, organized into microservices. Additionally, the TLA defines common software interfaces, data standards, and design patterns for communicating and storing data. For example, learner performance data uses the IEEE Experience Application Programming Interface (xAPI) standard, and competency frameworks are encoded in the IEEE Reusable Competency Definitions standard. Learning experiences (such as information about educational courses) are described using IEEE P2881, the foundation for Learning Experience Metadata. Finally, ledgers of learners' performance are captured via the IEEE P2997 Standard for Enterprise Learner Records. These four data standards, coupled with an active governance strategy, form the core of the TLA.

In recent years, progress has been made toward providing adaptivity and personalization in technology-enhanced learning environments. However, the breadth of data made available through the TLA promises to greatly enhance those benefits—not only within a single system or learning activity, but longitudinally and across organizations, echelons, and functional areas.

This paper summarizes the TLA data strategy and key components of its architectural design. The paper also discusses how a coherent data framework can improve DoD's ability to analyze, visualize, and tailor learning experiences through robust, multifaceted data, and it shows how this work establishes a foundation for personalization and adaptation across the human capital supply chain.

ABOUT THE AUTHORS

Brent Smith is a software systems architect with over 25 years of experience in designing and developing distributed training and education systems for DoD. As the Advanced Distributed Learning (ADL) Initiative R&D Principal, Mr. Smith informs the vision for future distributed learning modernization and helps ensure the ADL Initiative research aligns with its overall strategy and Defense priorities.

Sae Schatz, Ph.D., is the Director of the ADL Initiative, a Department of Defense program focused on modernization, policy, and coordination for distributed learning. Dr. Schatz is an applied human-systems scientist, with an emphasis on human cognition and learning, instructional technologies, adaptive systems, human performance assessment, and modeling and simulation. She holds a doctorate in Modeling and Simulation from the University of Central Florida.

John Turner is the senior advisor to the DoD Chief Data Officer (CDO). In this role, Mr. Turner collaborates with DoD stakeholders to develop strategy on data governance, architecture, standards, and culture. Mr. Turner also represents the Defense Intelligence Enterprise within the Intelligence Community's (IC) CDO Council.

Total Learning Architecture Data Model for Analytics and Adaptation

Brent Smith
ADL Initiative (Contractor)
Orlando, FL
brent.smith.ctr@adlnet.gov

Sae Schatz, Ph.D.
ADL Initiative
Washington, DC
sae.schatz@adlnet.gov

John Turner
Office of the Chief Data Officer
Washington, DC
john.d.turner120.civ@mail.mil

INTRODUCTION

"Data is a strategic asset," according the memo from Deputy Secretary of Defense Kathleen Hicks, issued on May 5, 2021. The *DoD Data Strategy* similarly lauds the power of data, saying: "Improving data management will enhance the Department's ability to fight and win wars in an era of great power competition, and it will enable operators and military decision-makers to harness data to capitalize on strategic and tactical opportunities that are currently unavailable" (2020, p. i). Finally, DoD's *Personnel and Readiness Strategy for 2030* echoes these themes within the context of training, education, and talent management, saying:

> "…the Office of the Under Secretary of Defense for Personnel and Readiness (P&R) is setting its sights on the year 2030—driving to cultivate a technologically dominant force that is strategically ready, globally relevant, and flexibly sustainable. The cornerstone to this vision is achieving data dominance through digital modernization, seamlessly connecting all our data in real-time, and harnessing the skills of a generation of digital natives" (2020, p. 2).

As the saying goes, *data is the new oil*—among the most prized commodities of modern times. To unlock its promise, however, DoD must transform itself by embracing new approaches, advanced data engineering, enterprise-wide systems, and thoughtful governance mechanisms. Within the training and education functional area, the Total Learning Architecture (TLA) project is addressing these goals.

The DoD Advanced Distributed Learning (ADL) Initiative first conceived of the TLA in 2015 (e.g., Gallagher et al., 2017; Barr, Fletcher, & Morrison, 2020). The TLA is a set of common software interfaces, data standards, and design patterns for communicating and storing data about learning and development. The TLA is *not* some standalone system or piece of software; rather, it's a framework or a technical blueprint, akin to the technical designs for the internet. The TLA is meant to define how heterogeneous organizations and technologies can plug together into an integrated system-of-systems (a "learning ecosystem," as the jargon goes).

Technically speaking, the TLA framework defines an *enterprise architecture*, which is a "…a coherent family of parent and subsidiary architectures, to help modernize its nonintegrated and duplicative business operations and the systems that support them" (Senate Committee on Armed Services, 2012). Further, in an enterprise architecture the "member architectures (e.g., Air Force, Army, and Navy) conform to an overarching corporate or parent architecture and utilize a common vocabulary…[and] governance across all business systems, functions, and activities" as facilitated by "data standards, policies, procedures, and performance measures that are to be applied throughout the Department" (GAO, 2013). And per best practices, the TLA uses the Modular Open Systems Approach (MOSA) for its architecture, as directed by Defense guidance (e.g., DoD Instructions 8320.07 and 5000.88) and law (e.g., National Defense Authorization Act for Fiscal Year 2015, H.R. 3979 § 801), to enable enterprise-wide interoperability.

This paper summarizes the TLA data model and the approaches we're taking to implement that data strategy within Defense training and education systems, organizational structures, and policies.

OVERVIEW OF THE TLA DATA MODEL

Each individual training or education platform can benefit from a data-centric approach. However, a larger payoff comes from leveraging data at the enterprise-level, i.e., DoD-wide collection, sharing, dissemination, and analysis of data. The TLA data strategy (Gordon, Hayden, Johnson, & Smith, 2020) provides a common set of goals, data formats, technical interfaces, and business processes across learning and development functions to ensure data are usable across large-scale systems of systems—horizontally (office to office), vertically (up and down echelons), and longitudinally (over time). Key to managing this abundance of lifelong learning data are interoperable technical standards, linked vocabularies, and a federated catalog that provides pointers to authoritative data sources.

First, consider the data standards. For the TLA, these standards are being formalized via the Institute of Electrical and Electronics Engineers (IEEE), a preeminent voluntary consensus standards organization that facilitates the development, publication, and governance of technical standards. The TLA relies on four main data standards:

(1) **IEEE P2997**, Standard for Enterprise Learner Records for lifelong learner profiles
(2) **IEEE P2881**, Learning Experience Metadata for defining learning experiences (e.g., courses, scenarios)
(3) **IEEE 1484.20.1**, Reusable Competency Definitions for common descriptions of subjects and their levels
(4) **IEEE P9274.1**, Experience Application Programming Interface (xAPI) 2.0 for runtime learner performance

Figure 1 provides an overview of three of the four IEEE standards that comprise the TLA data strategy.

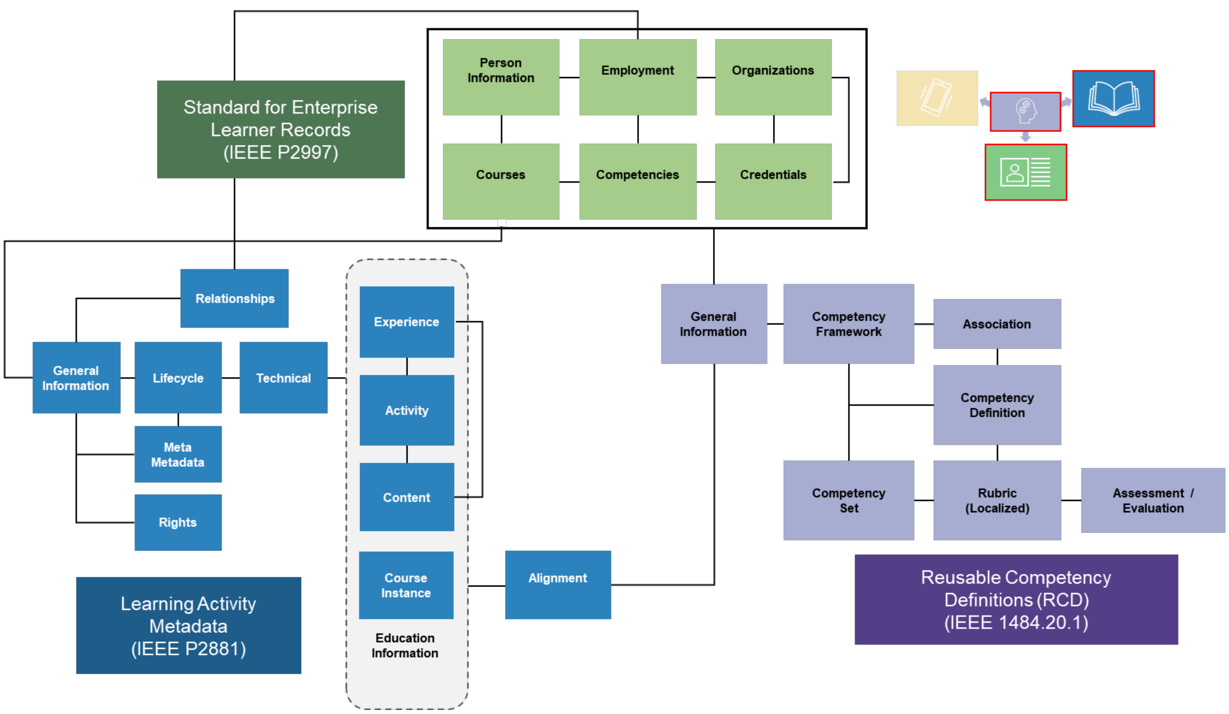

Figure 1. Total Learning Architecture's Logical Data Model (Partial View). *This image shows the major subcomponents of three of the four TLA data pillars. It omits xAPI, the data standard for collecting runtime performance, which is discussed later in the paper.*

TLA Data Pillars

Every device or software service in a TLA-defined learning ecosystem is a Learning Record Provider (meaning it pushes data out), and/or it's a Learning Record Consumer (meaning it ingests external data). The various components plug together, like LEGO® bricks, to form the comprehensive system. The specific composition of learning technologies within any particular organization will differ, and the arrangement of these systems can change over time. However, the overarching enterprise architecture will remain—similar to how the connection points on LEGO® blocks enable interconnectivity, even when the assembled castle or spaceship is modified.

Stated more technically, the TLA's constituent software services, devices, and data are loosely coupled, and they interact through specified data contracts. The TLA's data contracts don't depend on the nature of the upstream source or downstream use of any given message. This means that there's no single system responsible for coordinating the execution between components. (This statelessness is essential for the loose coupling required to be a true ecosystem.) Rather, the TLA framework assumes that there are enterprise software services and associated infrastructure (e.g., to enable semantic interoperability and maintain digital identity for users), but the framework merely defines the functions and interfaces without requiring specific technologies, configurations, or organizational owners. The resulting system of systems enabled by the TLA is asynchronous and event driven. This makes it perfectly adapted to using modern high-performance messaging systems and microservices to satisfy functional requirements.

The following sections describe each of the TLA data pillars and how they interoperate.

(1) Standard for Enterprise Learner Records

The IEEE P2997 Standard for Enterprise Learner Records facilitates the aggregation, management, and sharing of learner data generated from diverse, connected systems. Each individual within an organization will have an Enterprise Learner Record, which includes information about completed learning experiences, competencies, credentials, and employment history as well as administrative information (e.g., identification of the organization inputting data into the record). Additionally, each record includes local and global attributes about the person applicable to learning contexts, and these can be shared across connected systems using the Learner API.

The Standard for Enterprise Learner Records also defines an underlying data model for Learner Profiles (at the local level) and an API to communicate learner-record data among connected systems. Each Learner Profile also includes linkages to the evidentiary chains of learner data. These evidentiary chains (in contrast to disconnected outcome summaries, such as a paper certificate that is no longer linked to the original course grades) enable increased reliability and facilitate more diverse uses of the data. The Standard for Enterprise Learner Records data model also includes guidance on data quality, to further guide the suitability of data to ensure its effective use. (For more information, see Reardon & Gordon, 2020.)

(2) Learning Experience Metadata

The IEEE P2881 Learning Experience Metadata standard defines a framework for describing and sharing descriptive information about formal and informal learning activities, such as academic courses, training exercises, instructional simulation scenarios, or instructional videos. Within a large-scale organization such as DoD, each of these learning resources is assigned a unique identifier, and then data about them are stored in a local Experience Index. The metadata are maintained locally so that training and education owners (e.g., schoolhouses, training centers) can manage how they define and share information about the content they own.

Some of the Learning Experience Metadata attributes, such as a course's length or its objectives, may be populated during its development. Other metadata elements may be derived from other connected systems, such as post-course survey systems that provide students' ratings of a course (i.e., the `AggregateRating` paradata value calculated from accumulated survey scores). The TLA framework defines Activity and Resource Management functions to support the creation, review, update, and deletion of Learning Activity Metadata as well as the publishing of those experiences to other connected systems.

By encoding and exposing metadata about learning experiences across the enterprise, an organization can create a single catalog of training and education offerings—such as the Enterprise Course Catalog currently in development for DoD (see Reardon et al. 2020). Beyond that, these metadata elements enable other TLA-enabled systems to link learning experiences to learner performance and longitudinal learner records. In other words, these metadata elements can be used to variously support various functions, from personalized learning to acquisition lifecycle planning, organizational effectiveness evaluations, and task-skill alignment analyses.

(3) Reusable Competency Definitions

Competence is a set of demonstrable behaviors, characteristics, and skills that enable the efficient performance of a job (White, 1959). Competency-based learning, in turn, is an instructional technique that focuses on developing and assessing the mastery of competencies. There is much debate about the best way to define competencies as well as numerous tools and formats for expressing them. However, for the sake of the TLA, we're only concerned with how competency-related data are accessed, interpreted, and shared. From that data perspective, each competency is subdivided into its specific knowledge, skills, abilities, and other behaviors (KSAOs) as well as different levels of proficiency (e.g., novice, advanced beginner, competent, proficient, and expert). To demonstrate competence at a given level, an individual or team must show evidence of performance of certain behaviors or skills at that level of proficiency.

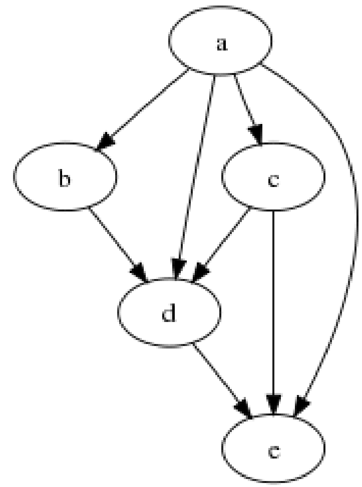

Figure 2. Portion of a Competency Framework Shown as a Directed Acyclic Graph

The IEEE 1484.20.1 Reusable Competency Definitions data standard defines a model for describing the content, required KSAOs, contexts, mastery levels, and credentials associated with competencies. The standard also defines Competency Frameworks, which articulate the relationships among competencies. The frameworks are hierarchical in nature, but a single competency may be used across numerous frameworks, creating a many-to-many relationship among competency elements.

Because of this complex relationship, the TLA framework recommends the use of graph databases for encoding competency elements. A graph database meets the requirements for semantic queries using nodes, edges, and properties to represent and store Reusable Competency Definitions and their associated frameworks. As shown in Figure 2, Directed Acyclic Graphs provide a natural way to express competencies, their relationships with each other, and their proficiency levels.

The TLA supports a network of federated competency listings where competencies can be tailored and aligned to the local context. This facilitates local control where appropriate, such as for adjudicating learner performance against local tasks, conditions, and standards. At an organizational level, the TLA requires a Competency Registry, the authoritative source of competency definitions and descriptions readable by both humans and machines. DoD will likely create an authoritative competency registry (or set of registries) in the future. Beyond DoD, other organizations publish competency sets (e.g., Department of Labor's O*NET), and some even provide access to interoperable networks of trusted competency registries (e.g., U.S. Chamber of Commerce's T3 Innovation Network).

(4) xAPI and the Master Object Model

The fourth data pillar focuses on runtime learner performance, and it uses the IEEE P9274.1 xAPI version 2.0 standard. The foundational xAPI standard defines its general structures and processes, and the corresponding xAPI Profile specification provides rules for associated controlled vocabularies applicable to different contexts (e.g., medical training) or interaction types (e.g., watching a video). xAPI-encoded data are stored in a Learning Record Store (LRS), which is formally part of the standard. At a local level, such as within a given training simulation, xAPI data are stored in a "noisy" LRS. These local data stores capture extensive and granular data, which may be inapplicable outside of the immediate situation. However, the TLA framework defines ways to federate LRSs together. This is particularly applicable for pushing local (noisy) data up, to be filtered and aggregated at the organizational and enterprise levels.

The TLA vocabulary for filtering and linking data—across data pillars and from local to enterprise levels—is called the Master Object Model (MOM). The MOM defines the data elements needed to link data about people with all of their learning experiences, contexts, competencies, credentials, and other key performance metrics. In other words, the MOM defines the object lifecycle of a single "thread of learning" that culminates in the reporting and evaluation of a learning event. These data are captured in MOM-conformant xAPI statements. As shown in Figure 3, MOM statements generated by each learning experience include the same general sequence, although the nature of the activities, and whether they are explicit or implicit, may change with each event.

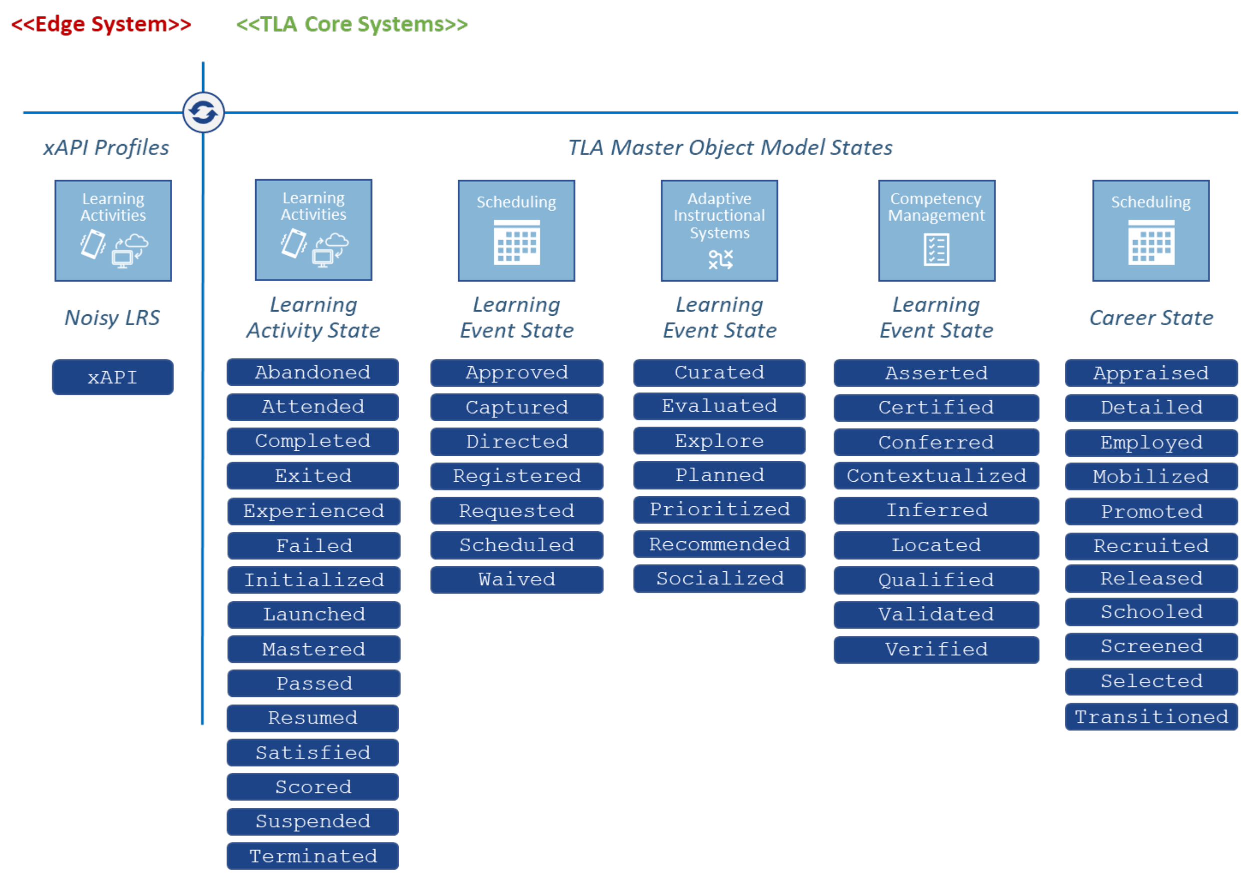

Figure 3. TLA Master Object Model (MOM). *TLA MOM statements are used to track learner progress against a set of learner goals and assigned courses (e.g., learning activities) that are established by other connected systems.*

TLA MOM statements are stored in a transactional LRS, and each MOM statement contains linkages to the noisy LRS to maintain the chain of evidentiary learner performance data. MOM statements are later aggregated using the IEEE P2997 Standard for Enterprise Learner Records. Each learner record in a profile also includes linkages to the different LRSs where the raw learner performance data were generated, different Experience Indices that store information about the learning experiences that produced the data, and the competencies that describe the subjects and criteria for how learners were evaluated. These linkages inform a "trust chain" of supporting evidence.

Supporting TLA Data and Components

The four-pillar TLA data strategy elegantly defines the major components of an enterprise learning ecosystem. However, any large data-centric system will necessarily have other data types and services. For example, Identity, Credentialing, and Access Management (ICAM) data are needed to accurately assign an identifier to each person within the ecosystem. To support generalizable functions like ICAM, the TLA framework defers to the DoD-wide guidance. In this case identity management is informed, in part, by DoD practices involving DoD ID credentials. These are directed by broader DoD policies (e.g., DoD Instruction 1000.30, DoD ICAM Reference Design) and aren't unique to the TLA. In this way, the TLA can be considered a learning-and-development overlay across the Departments larger enterprise architecture.

TLA Data Model Analogy: Intelligent Tutor

The TLA can support interoperability across diverse learning activities, technical systems, and business functions (e.g., manpower, personnel, acquisition, and readiness). It was designed for enterprise-level impacts while also allowing each organization to maintain its own data equities relative to educating, training, qualifying, or employing people within their purview. However, these big-picture concepts can be difficult to visualize, so for the sake of clarity let's consider a local (not enterprise) analogy: intelligent tutors.

Intelligent tutoring systems use data and artificial intelligence to provide adaptive learning experiences. For example, a mathematics tutor may include a series of algebra questions that adjust in difficulty based on a student's ability, performance, attributes, or emotional state. To accomplish this, the intelligent tutor monitors a student's activities against a model of expected behaviors, evaluates observed performance, and intervenes to optimize the experience.

The core components of an intelligent tutoring system are often generalized as a Domain Model (that formally describes the subject area, such as algebra), a Student Model (that describes the learner), and a Tutoring Model (that defines recommended pedagogical interventions based on a student's actions). Conceptually, these components align with the TLA framework, but the core difference is that an intelligent tutor is a single, fully encapsulated system while a TLA-enabled ecosystem is the emergent outcome of a diverse system of systems.

Consider this: Within a TLA-enabled ecosystem, someone's performance on a learning experience will be captured via the xAPI data standard and stored in a local (noisy) LRS. Similar to an intelligent tutor Student Model, this noisy LRS records key learner interactions and performance outcomes that can be used to immediately support adaptation. However, many learning experiences generate additional, useful data. For instance, simulation exercises often use After-Action Review systems that record video, or they incorporate third-party biometric sensors to assess learners' states. Other data sources might include location sensors, exercise management tools, or any systems that communicate additional context about the experience. Within a TLA-enabled ecosystem, all of these different Learning Record Producers can push data to an interconnected LRS.

As shown in Figure 4, the TLA's Experience Index acts as a federating function that describes connected systems in detail and provides linkages to their data. Each xAPI statement is timestamped so that key learner interactions can be correlated with other available data associated with that experience. The increased granularity afforded by the TLA's metadata can inform a much wider variety of interventions. In this way, the totality of data from each learning experience can be used to better inform instructor support, automated feedback, real-time analytics, and automated adaptation.

Another distinction is that intelligent tutors typically focus on micro-adaptation, that is, interventions that are made inside of a learning activity such as within a given math problem. Although these micro-adaptive interventions can be complex (e.g., modifying behaviors in a simulated scenario), they're still constrained to a given period of time and subject. In contrast, as different adaptive systems connect to a TLA-enabled infrastructure, they can recommend adaptive interventions across an organization's entire inventory of learning resources and over much longer time scales.

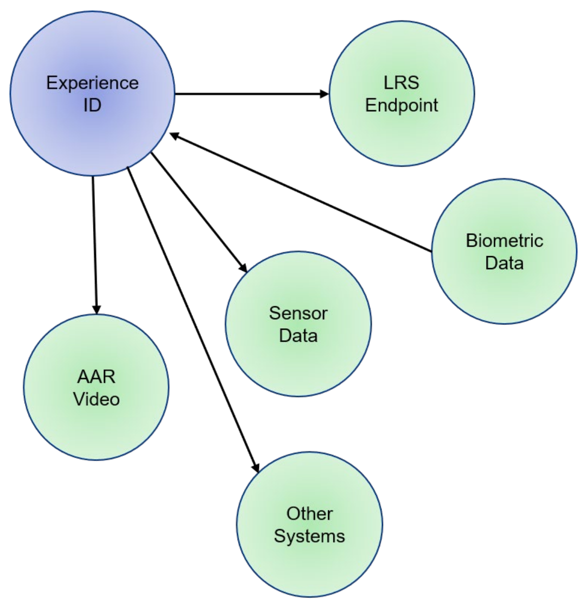

Figure 4. Federated Data Catalog. TLA metadata includes linkages to other sources of learner data created by a learning experience.

Eventually, a TLA-enabled ecosystem could facilitate adaptation across different subjects, platforms, and functions. Systems within a TLA-enabled infrastructure could use the rich data within it to adaptively sequence learning experiences across someone's entire career or to inform workplace decisions, such as promotions or career planning. In these sorts of applications, individual or workplace goals would be aligned to the competencies encapsulated in Reusable Competency Definition frameworks, and those competencies, in turn, would connect to the available inventory of all training and education resources as well as other human resources systems. For example, competency frameworks might include Rich Skill Descriptors, which are machine-readable strings of data that express information about who the skill applies to, the nature of the skill itself, and the context in which it is applied (Open Skills Network, 2021). Such data could inform team selection, upskilling, hiring, and other talent management decisions.

Looking Forward: From Learning and Development to Comprehensive Talent Management

Talent development is one portion of the larger talent management cycle. Unquestionably, the future of talent management involves the increasing use of artificial intelligence (driven by data) to better support employee, leadership, and organizational goals from hire to retire. Although the TLA framework is focused on enabling learning, it's being designed to eventually integrate with other human capital functions. For example, in 2020, the HR Open Standards (HROS) organization began incorporating xAPI into existing HROS API standards as an xAPI Profile extension for standardizing assessments. This work informed the TLA MOM's career state verbs, which were subsequently designed to capture a learner's career trajectory from different manpower and personnel systems.

As another example, Reusable Competency Definitions works in concert with the Credential Transparency Description Language (CTDL) developed by Credential Engine. CTDL enables rich descriptions of credential-related resources including credentialing organizations and specific credential subclasses such as degrees, certificates, certifications, and digital badges (Kitchens, Sutton, & Barker, 2021). Using CTDL, credentials can link competencies to other data categories, such as occupational specialty codes, position descriptions, or career pathways. Learner records can then include the competencies as well as the credentials conferred from different training experiences, assessments, or jobs. Relatedly, the U.S. Chamber of Commerce's Job Data Exchange (JDX) standard works with CTDL to link job position descriptions to credentials and Reusable Competency Definitions.

IMPLEMENTATION APPROACH

In May 2021, the Deputy Secretary of Defense signed a memo outlining five data decrees meant to help DoD achieve the vision laid out in its enterprise data strategy. These data decrees are:

1. Maximize data sharing and rights for data use: All DoD data is an enterprise resource.
2. Publish data assets in the DoD federated data catalog along with common interface specifications.
3. Use automated data interfaces that are externally accessible and machine-readable; ensure interfaces use industry-standard, non-proprietary, preferably open-source, technologies, protocols, and payloads.
4. Store data in a manner that is platform and environment-agnostic, uncoupled from hardware or software dependencies.
5. Implement industry best practices for secure authentication, access management, encryption, monitoring, and protection of data at rest, in transit, and in use.

The memo also empowers the DoD Chief Data Officer (CDO) to issue guidance regarding the DoD's data ecosystem, which includes creating a culture of data sharing by building a data ready workforce and implementing a modular open system architecture that uses technology to manage the lifecycle of data. The TLA is designed with these characteristics in mind, as well as the other "Guiding Principles" and "VAULTIS" goals defined in the *DoD Data Strategy*. (VAULTIS stands for visible, accessible, understandable, linked, trustworthy, interoperable, and secure.) The sections below briefly describe some of the TLA implementation designs that address these overarching DoD data directives.

Maximizing Data Sharing

The TLA approach inherently emphasizes data sharing, from its focus on open-source data standards, to its integrated data management policies and business processes. The TLA framework also relies heavily on Linked Data. Linked Data is a methodology for defining and exposing data vocabularies via published, structured metadata that can be interpreted by humans and machines to enable semantic interoperability. This ensures that different systems use specific terms in the same way. It also helps clarify the relationship among data elements, data formats, and pre-defined assemblages of terms. Linked Data is essential to preserving the meaning and context of data communicated between systems, without requiring the transmittal of the entire data definition with each data set. It helps abstract the definitions of data elements away from the data sources themselves, which improves data integrity, overall system resiliency, efficiency, and semantic interoperability.

As shown in Figure 5, a Linked Data and Schema Server is a recommended TLA core element. It provides a central service by connecting to the enterprise data fabric. The DoD Linked Data and Schema Server will provide a single source of truth (i.e., authoritative data source) for those data definitions and will establish immutable Internationalized Resource Identifiers (IRIs) for each term and schema that all DoD technologies can reference. It will also include DoD-specific elements, such as definitions that may be considered too sensitive to list on the open web. As shown in the figure, the linked data database is envisioned to include authoritative sets of competencies and credentials as well as data schemas. Those schemas define the structure of data in different systems (e.g., the "header" rows and data structures in various course catalog database), which supports automated detection and alignment of data sources with TLA data standards and controlled vocabularies. This is an important concept when connecting legacy training and education systems into the DoD's future learning ecosystem, and it allows for the incremental adoption of, and migration to, TLA data standards.

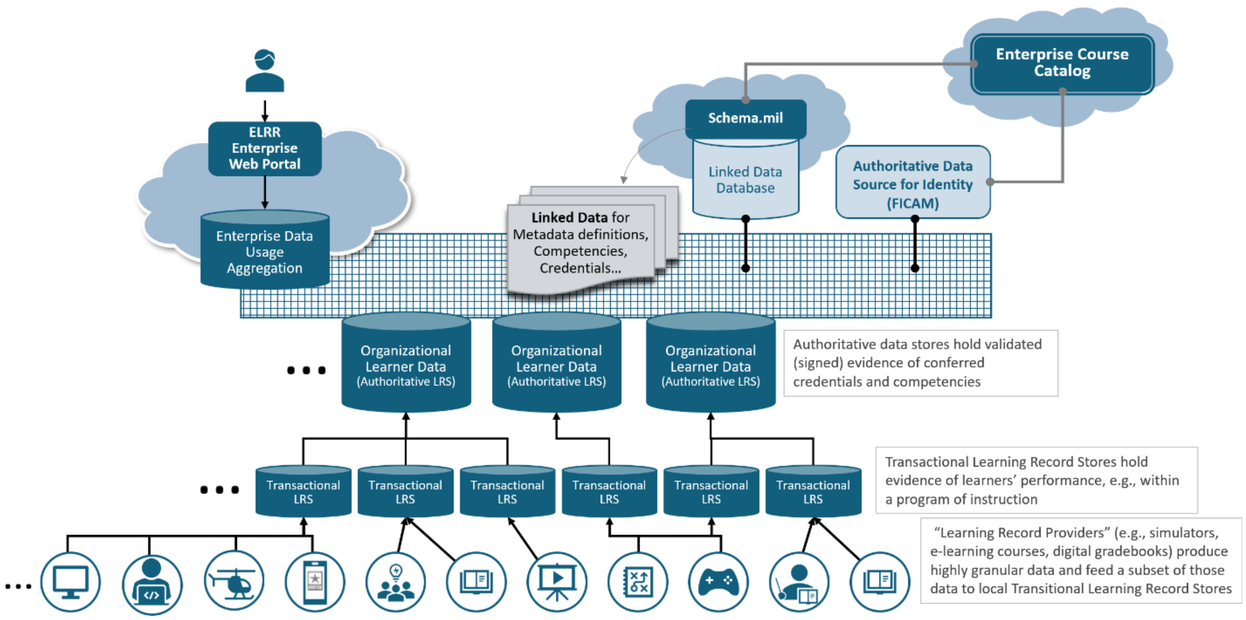

Figure 5. Enterprise Digital Learning Data Fabric. *The TLA's federated data structures maintain data ownership with local DoD components while ensuring all data is accessible by other connected systems.*

As part of the this linked approach, TLA profiles and controlled vocabularies are also required. A profile is a structured template of information that describes a data container; for example, profiles define the alignment between TLA data and the different types of software systems used for business functions, such as human capital management or talent development. A key strategic concept for a profile is that it contains minimal information requirements to assure a container is sufficiently described for self-identification to support any enterprise query or data sharing need. This information is provided to enable valid enterprise consumption of the associated metadata. One example of how this is implemented in the TLA's involves the Learning Experience Metadata standard, which creates a core standard and profiles for the different types of learning experiences (e.g., course, simulation, conference). This approach makes the governance of the TLA data strategy more flexible by enabling the training and education community to create profiles that best suit their needs without having to modify the baseline standard.

Controlled vocabularies are used to populate each profile's data elements, and these also inform the architectural design patterns applied to develop different types of systems that consume TLA data. In software engineering, a design pattern is a reusable solution to a common task within a given context. These work in concert with the TLA microservices to allow different systems to publish and subscribe to different types of TLA data.

Implement Secure Authentication, Access Management, Encryption, Monitoring, and Protection

Implementing DoD cybersecurity is a perennial challenge, and the accelerating pace and scope of digital transformation will only exacerbate this issue. A modern, data-centric DoD requires a cloud-enabled IT infrastructure that scales to meet the breadth of DoD's data needs and speed of changing requirements—albeit without sacrificing information security. The TLA framework is incorporating several mechanisms to address these issues.

Infrastructure as Code

The Defense Information Systems Agency's (DISA) Cloud Computing Program Office is working to expedite the time-consuming processes required to design, provision, configure, assess, and authorize cloud-hosted services through the DoD Cloud Infrastructure as Code (IaC) initiative. The existing DoD Cloud IaC baselines provide scripted processes to generate preconfigured, preauthorized, Platform as a Service (PaaS) focused environments. These baselines exist in the form of templated instructions built into the deployment scripts that automate many of the processes typically performed by humans. This allows different DoD organizations to perform the complex series of tasks required to activate and deactivate cloud resources. The environments can be immediately consumed for development and test workloads, with concurrence from a local Authorization Official (AO). This also provides the basis for accelerated Authority to Operate (ATO) approval for production workloads, by allowing individual AOs to accept IaC assessment and authorization (A&A) from other AOs, enabling mission owners to achieve an Authority to Operate (ATO) using an inheritance scenario. The ADL Initiative is incorporating these DoD cloud IaC tools into the TLA component development process (e.g., for Experience Indices, a course catalog portal, and metadata aggregation services). This type of automated infrastructure will be a key enabler for the next generation of training and education activities.

Policy-Based Access Control

TLA-enabled systems will also need to meet detailed access control requirements that establish who can access certain data elements, when data can be accessed, and how to maintain records for the accessed data. There are different approaches for handling such access control.

Identity-Based Access Control (IBAC) provides each actor a unique identifier, which is used to non-reputably assert the actor is who they say they are. Role-Based Access Control (RBAC) focuses on job function. It assigns a role for every organizational position or system desiring to access the data, and it manages each role's access to certain records. This approach is inflexible and doesn't scale effectively. Attributed-Based Access Control (ABAC) uses different characteristics for each actor (person or system) accessing the data and, based on the attributes assigned, determines if data access is granted, and the types of operations permitted. Finally, Policy-Based Access Control (PBAC) provides context-driven access control. It leverages IBAC, RBAC, and ABAC to automatically manage the set of policy rules for enabling access. This is the approach being integrated into the TLA framework.

PBAC is essential to developing viable, dynamic approaches to protect the privacy and security of learner data across interconnected TLA systems. The ADL Initiative is working to integrate a PBAC permission engine into the TLA Reference Implementation by late 2021. This capability ensures that access control permissions can restrict data to authorized users, no matter where data are used or which connected systems have access to it.

CONCLUSION AND NEXT STEPS

Data underpins digital modernization. The *DoD Data Strategy* describes an ambitious approach for transforming the Department into a data-driven organization, and TLA aligns DoD's training and education community with this broader strategy.

This paper highlighted the TLA's four key data standards, and while these standards will continue to evolve (as part of a good governance process), DoD education and training communities are encouraged to adopt them now. TLA data standards define entity names, data element names, descriptions, definitions, and formatting rules. TLA standards are created through an international voluntary consensus standards process, via working groups comprised of industry, academia, and government. And while standardized data has potential to become inflexible and overly constrained in time, the TLA approach is carefully designed to enable flexible, emergent patterns, allowing a TLA-enabled system to mature in a managed way without restricting the data within it. In other words, TLA standards are not overly prescriptive in how the data are defined, thereby enabling interoperability, facilitating the exchange of information across systems, and reducing the time spent cleaning and translating data—without imposing unnecessary rigidity.

More work on the TLA remains. For example, a comprehensive data governance strategy, acquisition guidance, and additional research are needed to realize its full benefits. Additionally, profiles and controlled vocabularies need to be defined, and the central software services that support these components must be fully tested and deployed.

While the TLA approach is complex, its promised rewards encourage this investment. Through implementation of the TLA, adaptive systems will be able to optimize individual career progression by identifying opportunities to improve technical skills, accelerate professional development, and acquire credentials for career advancement. These data will also support other DoD programs (e.g., Joint Services Transcript, Credentialing Opportunities On-Line) to help optimize talent development or ease the transition from one job to another (e.g., for transitioning veterans). TLA-enabled data can also better empower senior leader, helping them glean more insights into how to best prepare our military personnel, upskill the DoD workforce, or support DoD personnel as they adapt to an unexpected future.

ACKNOWLEDGEMENTS

We'd like to acknowledge the technical insights provided by staff from the Office of the DoD's Chief Data Officer. Special thanks to Jordan Gottlieb and Dr. Nathaniel Fuller for their technical expertise in reviewing and analyzing the TLA and its alignment to the DoD Data Strategy. This paper is also informed by the work performed by the ADL Systems Engineering and Technical Assistance (SETA) contractors, our vendors, the DoD stakeholders whom we work with, and DoD's training and education community.

REFERENCES

113th Congress (2013-2014). H.R.3979 - Carl Levin and Howard P. "Buck" McKeon National Defense Authorization Act for Fiscal Year 2015. https://www.congress.gov/bill/113th-congress/house-bill/3979/text

American Legion, Credential Engine, & SOLID LLC (2021, March 15). *The Military Credentialing Advancement Initiative Report.* https://credentialengine.org/2021/03/15/new-report-explores-the-importance-of-understanding-credential-pathways-for-veterans-and-servicemembers-and-notes-the-ctdl/

Barr, A. B., Fletcher, J. D., & Morrison, J. E. (2020). *The Advanced Distributed Learning's Total Learning Architecture: IDA Report 2020.* Institute for Defense Analyses Alexandria United States.

DoD (2014, March 14). Cybersecurity (DoD Directive 8500.01). https://www.esd.whs.mil/portals/54/Documents/DD/issuances/dodi/850001_2014.pdf

DoD (2015, August 3). Implementing the Sharing of Data, Information, and Information Technology (IT) Services in the Department of Defense (DoD Instruction 8320.07). https://fas.org/irp/doddir/dod/i8320_07.pdf

DoD (2019, July 12). *Digital Modernization Strategy: DoD Information Resource Management Strategic Plan FY19-23.* https://media.defense.gov/2019/Jul/12/2002156622/-1/-1/1/DOD-DIGITAL-MODERNIZATION-STRATEGY-2019.PDF

DoD (2019, September 19). *DoD Enterprise DevSecOps Reference Design – version 1.0.* DoD Enterprise DevSecOps Reference Design v1.0_Public Release.pdf (defense.gov)

DoD (2020, November 18). Engineering of Defense Systems (DoD Instruction 5000.88). https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500088p.PDF

DoD (2020, October). *Preserving our Competitive Advantage: Personnel and Readiness Strategy for 2030.* https://prhome.defense.gov/Portals/52/Documents/Strategy/PR_Strategy_FINAL_.pdf

DoD (2020, September 30). *DoD Data Strategy.* https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DOD-DATA-STRATEGY.PDF

DoD (2021, May 5). *Creating Data Advantage*. Deputy Secretary of Defense Memorandum. https://media.defense.gov/2021/May/10/2002638551/-1/-1/0/DEPUTY-SECRETARY-OF-DEFENSE-MEMORANDUM.PDF

DoD and the Office of Personnel Management (2019). Memorandum of Agreement between the Department of Defense and the Office of Personnel Management. https://federalnewsnetwork.com/wp-content/uploads/2019/05/OPM-DOD-USALearning-MOA.pdf

Gallagher, P. S., Folsom-Kovarik, J. T., Schatz, S., Barr, A., & Turkaly, S. (2017). Total Learning Architecture development: A design-based research approach. *In Proceedings of the I/ITSEC.*

GAO (2013). *DoD Business Systems Modernization: Further Actions Needed to Address Challenges and Improve Accountability* (GAO-13-557). Washington, DC: GAO.

Gordon, J., Hayden, T., Johnson, A., Smith, B. (2020). 2019 Total Learning Architecture Report. https://adlnet.gov/assets/uploads/2019%20Total%20Learning%20Architecture%20Report.pdf

Hicks, Kathleen. (2021, May 5). *Memorandum for Senior Pentagon Leadership.* https://media.defense.gov/2021/May/10/2002638551/-1/-1/0/DEPUTY-SECRETARY-OF-DEFENSE-MEMORANDUM.PDF

Ireland, Jim. (2021, March 23). *Profiles for Standards Assessments.* HR Open Standards Spring Virtual Meeting. https://www.hropenstandards.org/news/profiles-for-standards-assessments--spring-virtual-meeting

Joint Task Force Transformation Initiative (2013). Security and Privacy Controls for Federal Information Systems and Organizations. *NIST Special Publication, 800*(53). https://doi.org/10.6028/NIST.SP.800-53r4

Kitchens J., Sutton S., Barker P., (2021, May 27). *Credential Engine Schemas Handbook.* Credential Engine. https://credreg.net/ctdl/handbook

Open Skills Network (2021, June). *Draft Rich Skill Descriptor Standards Implementation Recommendations.* https://rsd.osmt.dev/

Reardon, A. & Gordon, J. (2020) From Silos to Manifolds: Strategies for Improved Learner Record Administration. In Volume 2020 (Eds.), *Proceedings of the 2020 IITSEC*. Arlington, VA: NTSA.

Reardon, A., Smith, B., Hayden, T., & Liu, Y. (2020). Enterprise Course Catalog Pre Alpha Prototype. Orlando, FL: Advanced Distributed Learning Initiative.

Roberson, D. L., & Stafford, M. C. (2017). *The Redesigned Air Force Continuum of Learning: Rethinking Force Development for the Future.* Air University: LeMay Center for Doctrine Development and Education. https://www.airuniversity.af.edu/Portals/10/AUPress/Papers/LP_0001_ROBERSON_STAFFORD_REDESIGNED_AIR_FORCE.PDF

Senate Committee on Armed Services (2012, Apr. 18). Hearing 112–658, Section "Importance of Business Enterprise Architecture and Reengineering Business Processes," pp. 34.

Sims, K., Schatz, S., Brewer, V., Rogers, A., McMahon, S. (2020). Enterprise Digital Learning Modernization: What, Why, and Who Says So? In *Proceedings of I/ITSEC*.

Theophile G., Koedinger K., Schneider J., & Mitchell T. (2020, October). When is Deep Learning the Best Approach to Knowledge Tracing, *Journal of Enterprise Data Mining, 12*, pp. 31-54. https://jedm.educationaldatamining.org/index.php/JEDM/article/view/451

Third, A., & Domingue, J. (2017, April). Linked data indexing of distributed ledgers. In *Proceedings of the 26th International Conference on World Wide Web Companion* (pp. 1431-1436).

Walcutt, J.J., & Malone, N. (2019). Lifelong Learning. In Walcutt, J.J., & Schatz, S. (Eds.), *Modernizing Learning: Building the Future Learning Ecosystem* (pp. 61-81). Washington, DC: Government Publishing Office.

White, R.W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*(5), 297–333.