

T3 INNOVATION NETWORK

Work Group 2 Report

Exploring Sustainable Data
Standards Convergence

JULY 2018



U.S. CHAMBER OF COMMERCE FOUNDATION
Center for Education and Workforce

INTRODUCTION

The T3 Innovation Network organized Work Group 2 to explore sustainable data standards convergence which included the membership of the Credential Data Ecosystem Mapping Team and other interested stakeholders from the T3 Innovation Network.

The Mapping Team is a group of volunteers representing several North America-based data standards organizations that convene to map competency and credential data standards. The overall goal of the Mapping Team is to promote collaboration, interoperability, and harmonization among standardization initiatives related to the credential ecosystem by cataloging, sharing, and mapping data models, standards, and schemas. The credential ecosystem includes credentials, credentialing organizations, and individual learner achievement records that contain information on credentials and/or competencies (e.g., transcripts, resumes, portfolios, comprehensive student records).

This credentialing ecosystem mapping project includes—A4L, CASS, CEDS, Credential Engine, DCMI/LRMI, Ed-Fi, HR Open Standards Consortium, IMS Global, IEEE - LTSC, MedBiquitous, PESCS, Schema.org and is open to other standardization organizations.

The work group was asked to:

- Review the stakeholder use cases identified by Work Group 1 and find additional use cases that require data standardization and/or improve data standards harmonization.
- Develop plans and identify pilot projects for improving data standardization.

Work Group 2 held two webinars. The first webinar focused on reviewing the relevant stakeholder use cases from Work Group 1 along with the objectives of Work Groups 2, 3, and 4. Members of Work Group 2 were also asked to identify additional use cases, challenges, and potential solutions. The second webinar focused on potential solutions, pilot project ideas, and implications for a public-private data infrastructure. This information was collected using a Google Sheet shared across Work Groups 2, 3, and 4.

This work group report first summarizes the most important stakeholder use cases related to data standardization. Next, the report summarizes input from the work group on major issues and technological challenges and enumerates potential solutions that have emerged through the work group's discussion.

STAKEHOLDER USE CASES

Each work group was asked to contribute use cases that require advanced technologies that converge Semantic Web, Artificial Intelligence, machine actionable competency frameworks, and Distributed Ledger Technologies (.e.g. blockchain) across talent supply and demand domains. This is based on the premise that emerging technologies and competency data can provide significant impacts to the ever-changing talent marketplace.

Work Group 2 reviewed the use cases provided by Work Groups 1, 3, and 4 and considered additional use cases. Work Group 2 expanded the previously provided use cases to include technical challenges and solutions. This use case led to a primary metadata specification technical challenge in which metadata for describing entities in the talent pipeline ecosystem



comes from diverse providers using different languages of description (schemas) that do not interoperate or harmonize.

TECHNICAL CHALLENGES & POTENTIAL SOLUTIONS

Work Group 2 identified technical challenges and potential solutions. The second webinar for Work Group 2 was scheduled to be the last webinar of all of the work groups so members could review the input provided by Work Groups 1, 3, and 4. As a result, the list below summarizes input across all four work groups.

Technical Challenges

- Exchange of data across the talent pipeline ecosystem to be reused, compared, and analyzed is challenging because the data is frequently expressed using different languages of description for the same ecosystem entities.
- Semantic mapping, comparing, and translating data from providers throughout the talent pipeline ecosystem is challenging because data providers:
 - rely on different information models to convey content and meaning, including tabular data (e.g. spreadsheets, CSV); relational data (e.g. database tables); hierarchical; tree-like structures (e.g. XML, JSON, markup languages); and graph structured data (e.g. RDF, JSON-LD); and
 - rely on different syntactic bindings including XML, RDF/XML, Turtle, JSON-LD.
- Exchange and compare metadata from different providers is challenging because the data being compared has been expressed in languages of description designed with different focal points and granularity (e.g. one provider's competency framework data may be expressed using a general purpose language such as CASE or ASN and the other provider is using a domain focused language such as MedBiquitous).
- Comparisons of data generated using multiple languages of description is challenging because different underlying information models have been used making it necessary to aggregate and disaggregate semantic units before meaningful comparisons can occur.

Potential Solutions

- Supporting Standards and Vendor Communities: extend the community of standards mapping into the future with finding libraries, programming standards, and best practices for translating between standards.
- Metadata interoperability: the ability of two or more systems—or components—to exchange descriptive data and to interpret the descriptive data that has been exchanged in a way that is consistent with the interpretation of the creator of the data.
- Metadata harmonization: the ability of two or more systems—or components—to exchange combined metadata conforming to two or more metadata specifications and to interpret the metadata that has been exchanged in a way that is consistent with the intentions of the creators of the metadata.



POTENTIAL NEXT STEPS

The following potential next steps were derived by the project team from the input provided by all work group members. These steps represent an expansion of the original charter of the Credential Data Ecosystem Mapping Team:

- **Enable Ongoing Support and Expansion**
Provide support for staffing, hosting online mapping to harmonize tools, and convening.

- **Schema Crosswalks**
Advance the current schema cross-walking of the *Ecosystem Mapping Team* beyond their cross-walks of credential and competency framework description languages to include other languages necessary to a T3 Innovation Network-enabled talent ecosystem including:
 - transcripts
 - resumes
 - profiles
 - portfolios
 - comprehensive student records
 - job descriptions and job postings
 - product operation and maintenance specifications from where task and competency frameworks can be derived to improve training and credentialing

- **Schema Development**
Identify languages of description (schemas) that are necessary to the talent ecosystem that are underdeveloped or missing and promote their development. This work should begin with the Mapping Team on topics that may include:
 - pathways
 - employment and earnings records
 - assessment rubrics
 - evidence of achievement

- **Pathway to Harmonization**
Develop a pathway for applications across the talent ecosystem to exchange data with full data harmonization. When metadata shared between two or more applications is done in the same or compatible ways, the applications' metadata are interoperable. When metadata shared between two or more applications is done differently, but mapped to a common semantic model, the applications' metadata are harmonized.

