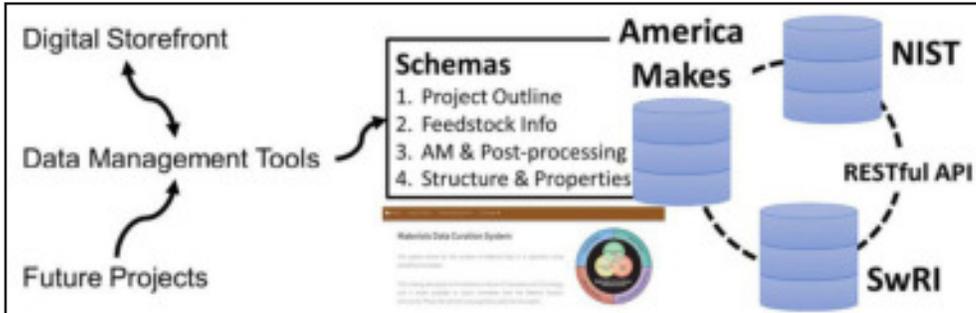


**SUCCESS STORY**

*Python-based data management tool parses deliverables, prepares XML files, and saves information to CDCS instance.*

## Schemas/data management tools were tested using deliverables from projects which were then entered to CDCS.



*Ideal data infrastructure adopting the Materials Data Curation System (MDCS) and Representational State Transfer Application Programming Interface (RESTful API) to archive project deliverables and connect to the other databases.*

**PROBLEM**

America Makes project deliverables contain high volumes of complicated data making data management tools more critical. To this end, an interoperable information system is crucial. A well-designed database and visualization software are the key tools of an information system to understand the present knowledge and point out the gaps. An ideal database contains schemas to label data with detailed descriptions (metadata) which specify the differences among the datasets for subsequent visualization and analyses. An extensible database is preferred for maturing the technology. The needs of schemas and a database for additive manufacturing (AM) development correspond to one of the strategic goals of the Material Genome Initiative (MGI).

**OBJECTIVE**

The objective of this project was to develop a human-readable, searchable, and extendable XML-based database to systematically archive various types of data from America Makes projects. The project enabled the America Makes membership and government stakeholders to cross search the project deliverables for evaluation of project performance and identification of technology gaps.

**TECHNICAL APPROACH**

The project approach leveraged the data management tools developed by Southwest Research Institute (SwRI) and the open source data infrastructure from the National Institute of Standards and Technology (NIST) to improve the searchability and interoperability of the America Makes Digital Storefront. A preliminary data management tool was developed to automatically parse different types of raw data to XML format. A pedigree chart of XML schemas was created to cure the project data and organize the deliverables into a hierarchical structure of materials, processing strategies, structures, and property measurements. The schemas were developed based on the success of the NIST AM-Bench program. (AM-Bench provides a continuing series of controlled benchmark measurements, in conjunction with a conference series, enabling modelers to test their simulations against rigorous, highly controlled AM benchmark test data.) During the project period, the sample XML files were archived in an MDCS instance inside the SwRI firewall for additional cyber protection. Through the built-in functions and the RESTful API, MDCS enabled the interconnection of instances for data federation among the allied institutes. This framework provides a consistent mechanism to cure and share data.


**AMERICA MAKES  
TECHNOLOGY  
DEVELOPMENT  
ROADMAP**

This project aligns to:


**ASTM PROCESS CATEGORY:**

- Material Jetting
- Binder Jetting
- Material Extrusion
- Powder Bed Fusion
- Directed Energy Deposition

**MATERIAL:**  
Multiple Materials

## ACCOMPLISHMENTS

This project adopted and refined the public XML schemas which were developed for AM benchmark and AM material database projects at NIST. These schemas accommodated the data including the project background, feedstock information, AM and post-building schedules, microstructure characterization, property measurements, and simulation codes. To achieve the project objective, a comprehensive data management workflow was developed to parse information, enter data, and validate the schemas. To identify the project features and enable the searching function, a Python tool was implemented to parse the project deliverables into reusable formats. Since these deliverables contained files in different formats such as pdf, docs, and pptx, the Python tool employed several libraries including PyMuPDF, tabula, docx2python, and pptx.

This project developed an XML-based database, archived the selected project deliverables, and demonstrated Configurable Data Curation System (CDCS), which is an open-sourced software providing the data federation function. The CDCS instance for this project was then connected to other databases to demonstrate the mechanism of data federation. The data federation function allows the America Makes members to cross search project deliverables using scientific terminologies for future applications.

## PROJECT END DATE

July 31, 2021

## DELIVERABLES

- XML schemas
- Automated data parsing and entry software
- Visualization software
- Final project report and suggestions for future developments of data infrastructure

*All downloadable deliverables are available to members of America Makes via the Digital Storefront.*

## FUNDING

**\$50,000 total project budget**  
(\$50,000 public funding)

## PROJECT PARTICIPANTS

### Project Principal:

Southwest Research Institute

### Public Participants:

U.S. Department of Defense

5001.001.002.003 Development of a Pedigree Informatic System for America Makes Materials

#### NCDMM Headquarters

486 Cornell Road  
Blairsville, PA 15717  
Phone: (724) 539-8811

[ncdmm.org](http://ncdmm.org)

#### Letterkenny Offices

4755 Innovation Way  
Chambersburg, PA 17201  
Phone: (717) 553-0068

#### America Makes Offices

236 West Boardman Street  
Youngstown, OH 44503  
Phone: (330) 622-4299

[AmericaMakes.us](http://AmericaMakes.us)