**The Materials Data Curation System**

Mary Brady\(^1\), Carelyn Campbell\(^1\), Alden Dima\(^1\), Philippe Dessauw\(^1\), Marcus Newrock\(^1\), Adele Peskin\(^2\), Pierre-Francois Rigodiat\(^1\), Guillaume Sousa Amaral\(^1\), Zachary Trautt\(^1\), Sharief Youssef\(^1\)

\(^1\)National Institute of Standards and Technology
Gaithersburg, Maryland 20899

\(^2\)National Institute of Standards and Technology
Boulder, Colorado 80305

**Introduction**

The NIST Materials Data Curation System (MDCS) provides a means for capturing, sharing, and transforming materials data into a structured format based on the Extensible Markup Language (XML) that is amenable to transformation to other formats such as those used by existing computational tools. The data are organized using user-selected community-developed templates encoded in XML Schema used to create data documents that are saved in a non-relational (NoSQL) document database. Each project, group, or organization can run as many MDCS instances as needed. Individual MDCS repositories can be interconnected for federated searches and data sharing.

---

**How the MDCS fits into NIST’s overall MGI strategy**

**MDCS REST API**

The MDCS provides a Representational State Transfer (REST) API that allows other software to directly interact with it over a network. MDCS functions are available via the API, allowing for full automation.

**Enabling Modular Data Models**

Complex data models can be composed from reusable templates allowing for groups of users and communities to develop data formats tailored to their needs. This example shows a new schema created using three schemas as types.

**Capturing Metadata**

Metadata can be captured and packaged with the data allowing for the data to be fully described to enable its reuse.

**Community Developed Templates**

Communities can standardize templates for common use cases, leading to reusable data that is consistently formatted and described.

**Reusable Data Types**

The MDCS provides for modularity down to the lowest levels where reusable data types offer the promise of standard representations to simplify data access.

**How are these Templates Created?**

- Researchers can start with an existing template and modify it or use existing collection of component templates to create new documents
- Future development will leverage public material registries to share templates and types across materials communities

A Template Composer hides the complexity of the underlying XML Schemas allowing for users to quickly and intuitively design new templates.

**Data Exploration**

1. Select Template
   - Global Templates
   - User Defined Templates
2. Select Fields
   - Specific fields to search against
3. Perform Search
   - Query by Example
   - SPARQL Queries

The MDCS uses its templates to guide the search data. More complex searches can be accomplished using a SPARQL endpoint.

**MDCS on GitHub**

The MDCS is available from the NIST GitHub repository.

https://github.com/usnistgov/MDCS

---

Certain commercial equipment or software are identified in this poster to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.