MATDB ONLINE—A STANDARDS-BASED SYSTEM FOR PRESERVING, MANAGING, AND EXCHANGING ENGINEERING MATERIALS TEST DATA

TSP Austin* and HH Over

European Commission Joint Research Centre Institute for Energy and Transport, Postbus 2, 1755 ZG, Petten, The Netherlands
**Email:* timothy.austin@ec.europa.eu

ABSTRACT

With ICT Standards playing a key role in support of research and development in many disciplines, the European Commission Institute for Energy and Transport is keen to promote the development and adoption of ICT Standards for engineering data. In this respect, its MatDB Online facility is a Standards-based system for preserving, managing, and exchanging engineering materials test data. While MatDB Online has evolved over more than 30 years to incorporate the latest innovations in data preservation and exchange, such as XML-based data transfer and data citation using digital object identifiers, it continues to rely on a robust data model developed more than 30 years ago through the joint efforts of the National Research Institute for Metals (the predecessor to NIMS, the National Institute for Materials Science), the European Commission Joint Research Centre, and the National Institute of Standards and Technology. While this data model has endured over many years, there is no corresponding Standard. Similarly, related efforts by the engineering materials community to deliver a Standard representation for engineering materials, such as MatML, have failed to be ratified. In consequence of the continued absence of a Standard representation for engineering materials data, there is no common mechanism for preserving and exchanging materials data and no formal means of maintaining a data model to support advances in materials technology, such as the emergence of nanomaterials. It is for these reasons that the European Commission Institute for Energy and Transport is supporting SERES, a CEN Workshop on Standards for Electronic Reporting in the Engineering Sector. As one of more than thirty organisations supporting the SERES Workshop, the Institute for Energy and Transport will make the MatDB XML schema available as one of several resources that will be taken into consideration when the prenormative Standard for representing engineering materials data is formulated. With the participation of the Institute for Energy and Transport in the SERES Workshop taking place in parallel with a related project with Oak Ridge National Laboratory, there is good reason to expect that a Standard representation for engineering materials, which has so far eluded the materials community, will be realised. This paper describes MatDB support for engineering materials Standards and related innovative features.

Keywords: Matdb, Schema, Test data, Testing standards, Datacite, Data transfer

1 INTRODUCTION

Scientific and technical communities recognize the importance of access to data, and yet a recent study into access versus importance of information indicates that of the various resources generated, from patent information to research articles, data sets remain the least accessible irrespective of region, organisation type, or discipline (Publishing Research Consortium, 2010).

Critical review indicates that researchers have yet to be convinced of the merits of embedding data preservation and exchange into mainstream research. One problem identified by a recent review of information use and exchange in the life sciences is that the technologies and tools developed by data management practitioners often fail to meet the requirements of the broader research community whom they are intended to serve (Research Information Network & the British Library, 2009). Coupled with continued skepticism of the merits of conserving and sharing data (Wiley, 2009), embedding effective data management practices into mainstream research is a problem that remains unresolved. The delivery of effective data management solutions is thus as much about engaging with the end-user and providing services that add value to the research process as it is about technology.

In the engineering disciplines, there are additional obstacles to promoting the effective capture, preservation, and exchange of data, not least of which is that any suggestion of sharing data may compromise long-established but fragile links with the industrial sector (Ball & Neilson, 2010; Howard, Darlington, Ball, Culley, & McMahon, 2010). However, while there are certainly challenges associated with preserving engineering data, they are unlikely to prove insurmountable. Looking to the life sciences for example, it can be argued that patient confidentiality and competition in the pharmaceutical sector pose greater barriers to developing practices for preserving and sharing data, and yet the life sciences are at the vanguard of the emerging disciplines of web and data science. Further, while ease of use and security are high on the requirements list of researchers, there is certainly an interest to share and reuse data. Given this interest, and considering the very significant resources invested in generating engineering materials test data, it is reasonable to assume that the availability of well-designed data management systems has an important role to play in the realization of effective data management practices in the engineering sector.

2 MatDB ONLINE

At https://odin.jrc.ec.europa.eu the European Commission Joint Research Centre Institute for Energy and Transport (EC-JRC-IET) hosts an Online Data and Information Network (ODIN) in support of energy and transport research. The facility consists of a collection of online databases organized into four main categories: documents, engineering, nuclear, and product information. In the engineering category, MatDB Online is a sophisticated database for engineering materials test data that has benefited from many hundreds of person years invested in the development of a robust data model, comprehensive test support, and an intuitive user interface. As shown in Table I, the database covers mechanical properties, thermo-physical properties, and corrosion data of engineering alloys generated in accordance with international material testing Standards and recommendations.

Table 1. MatDB Online Test Type Support.

MECHANICAL PROPERTIES	IRRADIATION
CRACK GROWTH & FRACTURE	Irradiation creep
Creep crack growth	Swelling
Cyclic creep crack growth	In-pile relaxation
Fatigue crack growth	TENSILE
Fracture toughness	Compression
Impact	Multiaxial tensile
CREEP	Uniaxial tensile
Cyclic creep	Small punch tensile
Multiaxial creep	THERMO-PHYSICAL PROPERTIES
Torsional creep	Density
Uniaxial creep	Electrical resistivity
Small punch creep	Emissivity
RELAXATION	Linear thermal expansion
Multiaxial relaxation	Poisson's ratio
Uniaxial relaxation	Specific heat
FATIGUE	Shear modulus
High cycle fatigue	Thermal conductivity
Low cycle fatigue (load control)	Thermal diffusivity
Low cycle fatigue (strain control)	Young's modulus
Thermal fatigue	CORROSION
Thermo-mechanical fatigue	High temperature corrosion
Creep-fatigue interaction	Complex test

At the present time, MatDB Online contains over 40.000 data sets coming mainly from European R&D projects and provides a web-interface for data content, data entry, data retrieval, and analysis routines. As well as serving the data management needs of the research community, MatDB Online is available commercially through a technology transfer agreement with a partner company.

3 ADDED-VALUE SERVICES FOR MANAGING ENGINEERING MATERIALS DATA

Although the delivery of well constructed and robust repositories for data is a necessary component of a sustainable data management solution, it is simply the foundation, and added-value services are required to ensure that researchers benefit in a tangible way from conserving and sharing the fruits of their labour. In recognition of this issue, MatDB Online has been enabled to deliver innovative services on top of its basic data management features, which together with community building efforts are designed to promote the adoption of effective data management practices in the engineering materials research sector. These services include support for ICT Standards for materials test data delivered by CEN/WS ELSSI-EMD (Economics and Logistics of Standards-compliant Schemas and ontologies for Engineering Materials Data) (CWA 16200, 2010), RESTful Web Services in support of systems integration and data pipelining, and DOIs (digital object identifiers) for data (Brase, 2010; Starr & Gastl, 2011).

4 COMPLIANCE WITH TESTING STANDARDS

The underlying MatDB database schema defines five entities (source, materials, specimen, condition, and test) and for each test type the database structure reflects international testing Standards (Nagy, Over, & Wolfart, 2005; Ojala, & Over, 2007). Beyond its compliance with such Standards, MatDB supports the ICT Standards for materials test data delivered by CEN/WS ELSSI-EMD, a recently completed CEN Workshop that has demonstrated the feasibility of treating documentary Standards for mechanical testing as software specifications from which data formats (schemas and ontologies) can be derived. In this respect, CEN/WS ELSSI-EMD delivered data formats compliant with ISO 6892-1, the documentary Standard for ambient temperature tensile testing (ISO 6892-1, 2009(E)). It is contended that such data formats enable systems interoperability.

In support of CEN/WS ELSSI-EMD and to help verify the claim that Standards-compliant data formats enable interoperability, MatDB Online has been extended to demonstrate the feasibility of transferring data directly from the test facility (MatDB interoperability Console, 2010). This is significant insofar as delays storing test data result in their value diminishing as the sample pedigree, test conditions, and results become disassociated. The prototype demonstrates how the use of Web Services in combination with the ELSSI-EMD data formats has enabled MatDB Online to integrate with testing facilities without the need to modify OEM software.

5 STANDARDS FOR ELECTRONIC REPORTING IN THE ENGINEERING SECTOR

In consequence of the success of CEN/WS ELSSI-EMD, EC-JRC-IET is supporting CEN/WS SERES (Standards for Electronic Reporting in the Engineering Sector), a new CEN Workshop that aims to deliver prenormative ICT Standards for representing and reporting engineering materials test data. CEN/WS SERES acts on the key recommendations of the ELSSI-EMD CWA to develop a unified, Standards-based architecture for representing engineering materials data and to undertake an eReporting proof-of-concept (PoC) to demonstrate the viability of a transition to electronic reporting (CWA 16200, 2010).

CEN/WS SERES aims to develop prenormative ICT Standards for engineering materials test data, leveraging existing resources in the engineering materials sector. Resources relevant to ICT Standards for representing engineering materials data include MatML (Material Markup Language, 2006), JRC MatDB (MatDB XML Schema), NMC-MatDB (Ueno, Sakai, Sakaida, Isonishi, Shuto, & Oyatsu, 2011), ISO 10303-235 (2009), ISO 15926-2 (2003), Material Ontology V1.1 (Ashino & Fujita, 2006), MatOWL (Zhang, Hu, & Li, 2009), and the ISO terminology database at https://cdb.iso.org. Resources relevant to ICT Standards for reporting engineering materials data include the ECISS/TC 100 documentary Standards for reporting engineering materials data, such as EN 10204 (2004) and EN 10168 (2004). As depicted in Fig. 1, the PoC is intended to demonstrate the viability of a transition to eReporting using the prenormative ICT Standards for representing and reporting in combination with eBusiness Standards and best practices.

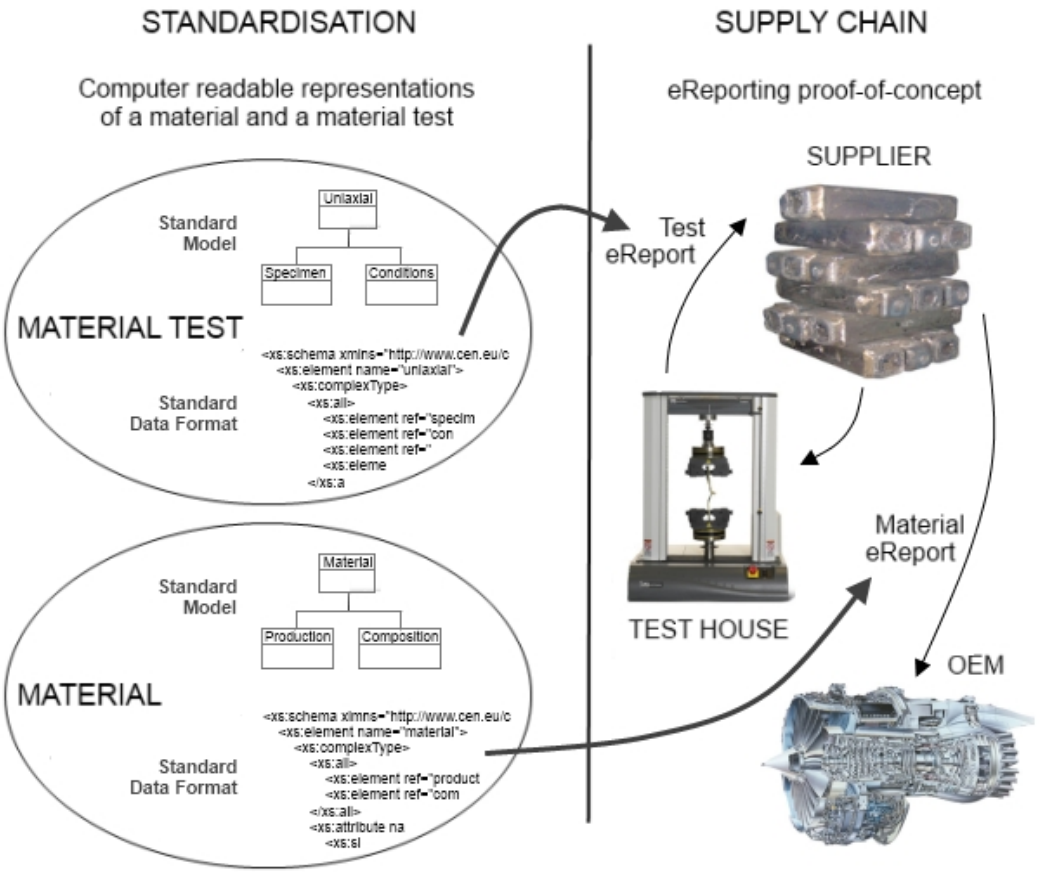

Figure. 1 Schematic overview of the CEN SERES Workshop.

The Project Team appointed to undertake the work of CEN/WS SERES consists of four information engineers to develop the data formats, an expert with a specialist knowledge of standardisation practices, a business analyst and eInvoicing specialist, and an industry manager to oversee the introduction of the PoC in the workplace. To ensure that CEN/WS SERES deliverables are aligned to the requirements of the engineering materials community, a broad group of registered participants reviews and approves the work of the Project Team. For the duration of CEN/WS SERES, stakeholders can enroll at any time as registered participants at http://www.cen.eu/cen/Sectors/Sectors/ISSS/Activity/Pages/SERES.aspx.

6 DATA CITATION

While adherence to testing Standards is of particular interest to industry and large research organizations, researchers working in the academic community need additional incentives to preserve and share their data. Motivations for researchers to share their data include establishing a means to acknowledge their contribution to the research process in a tangible way. Making data citeable in much the same way as publications offers just such an opportunity, and the DataCite initiative makes this possible. Consequently, MatDB Online supports assigning DOIs to individual data sets, thereby allowing researchers in the engineering materials sector to cite their data in much the same way as a publication. The following example shows a reference where the owner, title, host, and the DOI corresponding to a data set are referenced:

1. H.H. Over, Data for a Strain-Controlled LCF test at 650°C and 0.053 Hz on Udimet, EC-JRC Institute for Energy and Transport, http://dx.doi.org/10.5290/1000000010021

The DOI in the example is resolved (by the DOI lookup service) to a page hosted by MatDB Online that displays a summary of the data set and links to the full data set. In this way, researchers can make their data available to be

used for derivative work by other groups while at the same time benefiting from a direct citation or an acknowledgement of their contribution.

7 SUMMARY

MatDB Online is an internationally recognized facility for managing engineering materials test data. It incorporates intuitive data entry and retrieval features, a robust data model, and extensive test type support. An instance of MatDB Online hosted at the European Commission Institute for Energy and Transport serves the interests of the European research community while a technology transfer agreement with a partner company makes the application commercially available. To promote the preservation of consistent and reliable data, its database structure reflects international mechanical testing Standards. Beyond this compliance with existing Standards, MatDB Online supports the Standards-compliant data formats delivered by ongoing CEN Workshops, both by making its underlying database structure available and by demonstrating the potential of the data formats. Its added-value services include a RESTful test facility to database data pipelining service, standards-based systems interoperability, and citeable data (using DOIs).

8 REFERENCES

Ashino, T. & Fujita, M. (2006) Definition of a Web Ontology for Design-Oriented Material Selection. *Data Science Journal* Vol.5.

Ball, A. & Neilson, C. (2010) Curation of research data in the disciplines of Engineering, *SCARP Case Study 7*, Digital Curation Centre. Retrieved from the WWW, October 17, 2012: http://www.dcc.ac.uk/sites/default/files/documents/publications/case-studies/SCARP_B4812_EngCase_v1_2.pdf

CWA 16200:2010 (2010) *A Guide to the Development and Use of Standards-compliant Data Formats for Engineering Materials Test Data*. Retrieved from the WWW, February 27, 2012: ftp://ftp.cen.eu/CEN/Sectors/List/ICT/CWAs/CWA16200_2010_ELSSI.pdf

Brase, J. (2010) DataCite - A global registration agency for research data, *RatSWD Working Paper no. 149*, German Data Forum. Retrieved from the WWW, October 17, 2012: http://www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_149.pdf

ECISS, European Committee for Iron and Steel Standardization. Retrieved from the WWW, February 27, 2012: http://www.cen.eu/CEN/sectors/sectors/materials/Pages/eciss.aspx

EN 10168:2004(E) (2004) Steel products — Inspection documents — List of information and description.

EN 10204:2004(E) (2004) Metallic products — Types of inspection documents.

Howard, T., Darlington, M., Ball, A., Culley, S., & McMahon, C. (2010) Understanding and Characterizing Engineering Research Data for its Better Management, *ERIM Project Document erim2rep100420mjd10*, University of Bath, Bath, UK.

ISO 6892-1:2009(E) (2009) Metallic materials — Tensile testing — Part 1: Method of test at ambient temperature, International Organization for Standardization.

ISO 10303-235:2009(E) (2009) Industrial automation systems and integration — Product data representation and exchange — Part 235: Application protocol: Engineering properties for product design and verification, International Organization for Standardization.

ISO 15926-2:2003 (2003) Industrial automation systems and integration — Integration of life-cycle data for process plants including oil and gas production facilities — Part 2: Data model, International Organization for Standardization.

MatDB Interoperability Console (2010) Retrieved from the WWW February 27, 2012: https://odin.jrc.ec.europa.eu/dataentry/XMLConsole.html

MatDB XMLSchema. Retrieved from the WWW, February 27, 2012: https://odin.jrc.ec.europa.eu/alcor/XMLSchema.html

Material Markup Language (2006) Public Review Draft 01, 06 June 2006. Retrieved from the WWW, February 27, 2012: http://docs.oasis-open.org/materials/materials-matml-spec-pr-01.htm

Nagy, M., Over, H.H., & Wolfart, E. (2005) XML related data exchange from the test machine to a web-enabled MAT-DB. *Data Science Journal* Vol.4.

Ojala, T. & Over, H.H. (2007) Approaches in using MatML as a common language for materials data exchange. *Data Science Journal* Vol.7

Publishing Research Consortium (2010) Access vs. Importance, A global study assessing the importance of and ease of access to professional and academic information, Phase I Results. Retrieved from the WWW, February 27, 2012: http://www.publishingresearch.net/documents/PRCAccessvsImportanceGlobalNov2010_000.pdf

Research Information Network and the British Library (2009) Patterns of information use and exchange: case studies of researchers in the life sciences. Retrieved from the WWW, February 27, 2012: http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/disciplinary-case-studies-life-sciences

Starr, J. & Gastl, A. (2011) isCitedBy: A Metadata Scheme for DataCite. *D-Lib Magazine* Vol.17. Retrieved from the WWW, February 27, 2012: http://www.dlib.org/dlib/january11/starr/01starr.html

Ueno, A., Sakai, T., Sakaida, A., Isonishi, K., Shuto, T., & Oyatsu, Y. (2011) Current JSMS Factual Fatigue Database and Database Systems Including VHCF Data. *Fifth International Conference on Very High Cycle Fatigue.* Berlin.

Wiley, S. (2009) Why Don't We Share Data? *The Scientist* Vol.23, p 33-36.

Zhang, X., Hu, C., & Li, H. (2009) Semantic Query on Materials Data Based on Mapping MatML to an OWL Ontology. *Data Science Journal* Vol.8.

(Article history: Received 31 August 2012, Accepted 10 October 2012, Available online 2 November 2012).