Proc. of the 10th Model-Based Enterprise Summit (MBE 2019), Gaithersburg, Maryland, USA, April 2-4, 2019

Securing, Authenticating, and Visualizing Data-Links for Manufacturing Enterprises

William Z. Bernstein
Systems Integration Division, NIST
Gaithersburg, MD, USA

Sylvere Krima
Engisis LLC
Bethesda, MD, USA

Laetitia Monnier
Universite de Bourgogne Franche-Comte
Dijon, France

Mehdi Shahid
TELECOM Nancy
Nancy, France

This publication is available free of charge from: https://doi.org/10.6028/NIST.AMS.100-24

ABSTRACT

Managing digital resources generated from product design, manufacturing, and sustainment activities has become a significant burden for enterprises. In response, we introduce a prototype implementation of the Securing and Authenticating Data-Links (SADL) Interface, which interacts with a manufacturing handle registry to facilitate traceability of digital resources for engineering projects. This paper outlines the intended use of SADL and the handle registry by laying out hypothetical questions from potential users. Additionally, we map the core concepts of key standard data representations in manufacturing to a popular data type taxonomy. Future work will include the design and testing of data visualizations based on our mapping protocols.

INTRODUCTION

Manufacturing operations produce an immense amount of data, estimated at two exabytes of data annually in 2013 [1]. Considering the emergence of more complex electro-mechanical devices to the market, e.g., fully electric automobiles, the amount of code and manufacturing data is expected to grow significantly [2]. As a result, managing the associated digital resources has become, and will continue to become, a burden. An efficient and robust approach for labeling, categorizing, and curating diverse data is an essential first step for visualizing trends and deriving actions. In this paper, we present a prototype implementation of the manufacturing handle system, aimed at recording appropriate meta-data ("data about data") for digital resources related to the product lifecycle. Additionally, we present initial guidelines for developing data visualizations for the handle system to facilitate data exploration.

Each phase of the product's lifecycle incorporates its own data representations, organizational functions, and business processes. Each of these functions and processes uses tools and methods (e.g., computer-aided design (CAD) applications and requirement formalization). Though there have been efforts in improving information exchange across the various lifecycle phases (e.g., design and manufacturing) [3], there has not yet been a conclusive and robust demonstration at scale to achieve the so-called "digital thread" [4]. The metaphor "digital thread" conveys the seamless exchange and flow of data between engineering, manufacturing, business process and across supply chains [5].

However, in practice, the dearth of interoperability has led to gaps in information flow across manufacturing enterprises contributing to a number of challenges, including communicating across multi-tiered supply chains, reacting to engineering changes, and responding to customer requirements. For small-to-medium sized

DOI:https://doi.org/10.6028/NIST.AMS.100-24#page=59

This publication is available free of charge from: https://doi.org/10.6028/NIST.AMS.100-24

enterprises, these challenges are even more difficult to overcome, since most solutions targeted at the Digital Thread are expensive, expert-driven one-off prototypes. In response to these challenges, we present the following research contributions: (1) a prototype interface, coined the Securing and Authenticating of Data-Links (SADL) Interface that allows users to register digital resources, add meta-data, and query the registry, (2) a classification scheme of lifecycle data from manufacturing phases, e.g., as-designed, as-planned, as-executed, and as-inspected, based on data nature and type [6], and (3) requirements on the further improvement of the SADL interface. It is our hope that this work facilitates a better digital resource certification management in a product's lifecycle for end-users, including plant managers, design teams, and supply chain managers.

Our efforts aim to facilitate the realization of a Model-Based Enterprise (MBE) that can quickly respond to product lifecycle disruptions, e.g., engineering change requests, weather events affecting suppliers, and machine degradation. From this perspective, we focus on the design, manufacturing, and inspection phases that incorporate a standard data representations, which have been primary focuses of the MBE journey. Managing these representations as digital resources, and changes to them, poses a significant challenge.

THE SADL INTERFACE AND THE MANUFACTURING HANDLE REGISTRY

SADL (Securing and Authenticating of Data-Links) is an application serving as a middleware between digital objects hosted on a Handle.Net registry and its end-users. Its goal is to offer a customizable overview of a product's lifecycle digital objects, the product data, by providing additional meta-data (e.g. lifecycle phase or product category) from which users can query, rank, order, classify, and construct links between objects.

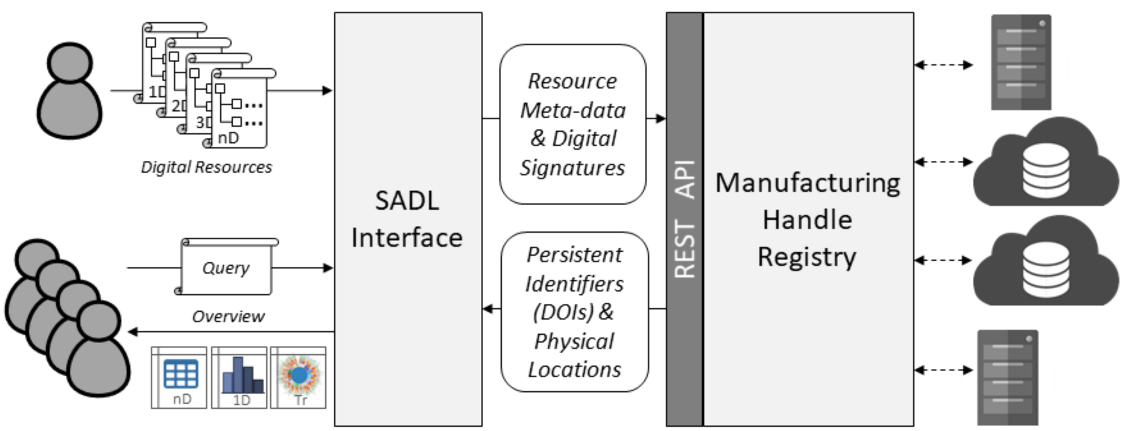

Figure 1: Vision of the SADL Interface and its interactions with users and the Handle Registry.

Figure 1 introduces the vision of the work presented in this paper. Initially, users label meta-data for digital objects by interacting with the SADL Interface. The SADL interface then leverages the Representational State Transfer (REST) Application Program Interface (API) of the Handle.Net registry to assign digital signatures and meta-data to the objects. The pipeline leverages existing technology that creates persistent identifiers (Digital Object Identifiers or DOIs [7]) to manage the physical locations of the resources. Other users can then submit queries through the SADL Interface to access report summaries of a collection of digital resources relevant to a manufacturing-oriented use case.

The pipeline presented above relies on the concept of the Handle System. Released by the Corporation for National Research Initiatives (CNRI) in 1994, the Handle System offers a means to locate, track, and manage data even in the face of constant modification [8]. In particular, manufacturing represents a domain that faces constant data modification and improvement. For instance, given engineering needs, it is very common that a

This publication is available free of charge from: https://doi.org/10.6028/NIST.AMS.100-24

product's design goes through several iterations before being approved for manufacturing, which leads to the creation of multiple design files. It is critical to construct a digital footprint outlining the complete history of each digital resource for various scenarios, e.g., liability investigations and engineering change requests. By providing a unique identifier for each digital object, the Handle System Registry acts as a DOI Repository providing a unique access point to all the digital resources inside an enterprise. Each "handle" (i.e., the name given to each of the DOIs) contains a set of meta-data about a specific digital object and allows for modifications to incur on the source file without compromising the validity and integrity of the handle.

By using the digital repository provided by the Handle System as a gateway to a broader view of all the different data objects inside the product's lifecycle, users can visualize the digital objects and link them together. This broader view describes the concept of a digital map encompassing critical digital resources given a particular use case. The Handle System provides digital record-keeping with a way to also offer efficient feedback to end-users and to facilitate better understanding of complex interactions with a product's lifecycle. Additionally, it is important that the current view in the registry reflects the current status in the real world. A proof-of-concept for a "System Lifecycle Handler" [9] confirmed the utility of these idea. In our pipeline, we include a means to store digital signatures to certify the validity of the stored data.

Throughout the rest of this paper, we explore potential means for visualizing a large collection of digital resources through the SADL Interface. We examine the expected range of data nature and type assuming that the registry is well-seeded with manufacturing-oriented data. This includes mapping expected business functions, from well accepted data representations to possible visualization schemes. We then conclude by addressing hypothetical, explorative questions from envisioned users to demonstrate the impact of the SADL Interface.

TOWARDS VISUALIZING A LARGE COLLECTION OF DIGITAL RESOURCES

In practice, the collection of digital resources for a given handle system would be quite large and a challenge to navigate. In response, we perform a data type mapping between the business processes and information embedded in standard data representations, namely STEP AP242 [10], STEP AP238 [11], MTConnect V1.4 [12], and QIF 3 [13]. STEP AP 242 provides an exchange format for design data including fully characterized Product and Manufacturing Information (PMI). STEP AP238 is a descriptive data representation for machine instructions, providing an additional layer of semantic descriptions compared to traditional G-code. MTConnect is a read-only communication protocol for capturing execution data from machine tool controllers. QIF is a semantically rich data format for representing, exchanging, and storing inspection plans, rules, and results. We applied the seven data type taxonomy used in Shneiderman's task-by-type taxonomy [6] keeping in mind that future work will involve designing and implementing data visualizations to respond to user queries to the SADL Interface. Nomenclature and a brief description of all seven data types are below.

***1-dimensional* (1D):** linear data types that are organized in a sequential manner, e.g., textual documents that only contain alphabetically ordered strings. From the perspective of engineering design, 1D data types can relate to the rules and requirements needed to apply STEP AP242. Considering the rules and requirements elements, a Mural visualization in the background of the scrollbar [14] might be appropriate. This visualization would highlight different parts of the text that are related to a particular requirement.

***2-dimensional* (2D)**: planar or map data including maps, floorplans, or newspaper layouts. Manufacturing-oriented examples could include entire factory layouts or plans for individual cell layouts. In a robotic factory, a production cell layout could highlight the distance between the robotic arm and other assets. A higher-level

This publication is available free of charge from: https://doi.org/10.6028/NIST.AMS.100-24

instance of 2D data could represent the entire plant factory decomposed into multiple cellular maps each relating to one another.

***3-dimensional* (3D)**: real-world objects or representations that represent volumetric data, such as solid models or computer-aided design (CAD) files. From a visualization perspective, the challenge is to find a balance between the view of the real-life object and the information within it. The aspect of the object must correlated with the data in a way that the user understands its spatial positioning in a larger context, e.g., a complex assembly.

***Multi-dimensional* (nD):** relational and statistical databases manipulated as multidimensional data in which items with *n* attributes become points in a *n*-dimensional space. QIF inspection results, including probe information are an example. One visualization technique to represent a database table is a cube visualization [15]. Each face of the cube is composed of one attribute of the table and layers of each face correspond to a possible value of the face attribute. The intersection of two adjacent faces represents a data point.

Temporal* (Ts)*: time series data, such as historical presentations or future projections. The difference between 1D and temporal data is rather nuanced. Both can be simple text documents, but the main difference is that Ts data is anchored through a timeline. Process plans in STEP AP238 and sample type data in MTConnect streams are examples of Ts data types.

Tree* (Tr)*: hierarchies with each item having a link to one parent item except the root. For example, an MTConnect Device model is organized based on the design of devices. Using a tree visualization, like a Tree-map [16], could provide a snapshot of capabilities and characteristics of available devices.

Network* (Nk)*: items linked to an arbitrary number of other items not following rules of trees. For example, the STEP AP242 Assembly structure contains parts and subassemblies as items, and multiple type of links can exist between these items. Some links might have numerical values or represent specific actions or modifications. A matrix diagram [17] is a way to expose different items and their associated links. The matrix can be color coded so that the user has a better understanding of the differences in type and meaning of the links.

			Data Types						
Representation	**Business Function**	**Concept Description**	**1D**	**2D**	**3D**	**nD**	**Ts**	**Tr**	**Nk**
STEP 242 (as-designed)	Specification, Breakdown & Configuration	Assembly Structure						●	●
		Transformations, Geometry, & Coordinate System		●	●				
	... ...								
STEP 238 (as-planned)	Model-Based Manufacturing Process	Generic Toolpaths		●	●				
		Parameters (feeds, speeds, etc.)	●						
	... ...								
MTConnect (as-executed)	Historical Machine Operations	Samples					●		
		Conditions	●						
	... ...								
QIF (as-inspected)	Model-Based Definition	Computer-aided design (CAD) data			●			●	
		Product manufacturing information (PMI) data				●			
	... ...								

Figure 2: Classification of data type per business function and concept description within each studied data representation. The complete table[1] can be accessed here: https://goo.gl/Zbkqmb.

[1] An initial draft of the mapping. We expect it to evolve as we dive deeper into each data representation.

This publication is available free of charge from: https://doi.org/10.6028/NIST.AMS.100-24

Figure 2 illustrates the structure of the data type classification completed for each business function of studied standard representations. We conducted this classification to identify design opportunities for custom visualizations for the SADL Interface. Besides individual visualizations per business function, we also envision potential for effective overview visualizations. In other words, given that an organization registered a large amount of digital resources within the handle system, we can present, for example, a hierarchical representation based on a prominent concept description, e.g., the assembly structure of a product. In the future, we plan to implement and test the effectiveness of sunburst plots, cartesian node-link diagrams, and matrix views [18] using accepted information visualization principles [19].

FROM DATA SETS TO VISUAL INDICATORS

In the previous section, we introduced a mechanism to uniquely identify, locate, authenticate, and navigate through a product's lifecycle digital objects using the Handle.Net and our SADL interface. We also described different data visualization types. In this section, we will discuss some steps towards generating and visualizing insight and performance metrics from trusted product data. While our area of focus is limited to as-designed, as-planned, as-executed, and as-inspected lifecycle data, the ideas and methods introduced here can be applied to other data.

The first step consists of identifying and categorizing the different product data generated and available to the organization. During this step, one must ensure that (1) the data is complete and fits under one of the four lifecycle stages previously mentioned, (2) the data is available in an open-standard format to reduce the cost of processing and enable interoperability, and (3) the data concepts are properly identified (as in the third column in Fig. 2).

The second step consists of identifying metrics or indicators based on key organizational characteristics and derived from the data concepts previously identified. These metrics and indicators should provide answers to questions related to organizational resources, activities, and performance. In this project, we have focused our questions on basic organizational components, namely processes, products and people (or the 3Ps). The following is a set of hypothetical questions illustrating some basic metrics and indicators that can easily be computed and/or inferred from the data sources/standards we use:

1. **Process**:
    - 1.1. How long did it take to execute process X during the past 10 days?
    - 1.2. How many parts a day are handled during process X?
    - 1.3. Was there a quality improvement between V2 and V1 of process X?
2. **Product**:
    - 2.1. What was the assembly structure of Product Y?
    - 2.2. How many parts were affected after changing feature X on product Y?
    - 2.3. Was the new design of Product X actually ready to move to production on November 2, 2018?
3. **People**:
    - 3.1. Who inspected the version of part Z that was built on November 2, 2018?
    - 3.2. What was the chain-of-command for Product X through its lifecycle?

The third and last step maps the different questions from the second step to the right source of data, data concepts, and visual data types to address each question. The output should be similar to Table 1 and be used as a guideline for the solution implementer(s). Our recommended output is a Table with the following columns:

This publication is available free of charge from: https://doi.org/10.6028/NIST.AMS.100-24

1. **Question**: the question whose answer is a metric or indicator regarding a key organizational component
2. **Representation** or data source: the format of the data that will be used to compute the metric or indicator in response to the question
3. **Key concepts**: a list of the data elements that need to be extracted from the data source to compute the metric or indicator
4. **Shneiderman's Data Type**: the data type used to visualize the metric or indicator based on the list of types identified in the previous section

In this section, we presented a three-step process to guide a user from identifying the type of data sources available, derive performance metrics and indicators, and map them to a visual data type using the Shneiderman's classification. This process would facilitate the design and development of a visualization dashboard providing insight to engineering teams through open data representations.

Table 1: Recommended output to describe requirements for presenting relevant data through the SADL Interface based on an enterprise-driven question.

Question	Representation(s)	Key concept(s)	Shneiderman's Data Type
1.1	MTConnect	Events	Ts
1.2	MTConnect	Part Count, Samples	1D
1.3	MTConnect, QIF	Events, Measurements data	1D
2.1	AP242	Assembly structure	Tr
2.2	AP242	Assembly structure	Tree/1D
2.3	AP242	Meta-data entered at SADL*	Tr
3.1	AP242, QIF	General Mgmt. Information	1D
3.2	AP242	Meta-data entered at SADL*	Ts/Tr

*Not part of the standard representation itself. The digital signatures would be appended to the digital resource once the user enters the information in the SADL Interface.

CONCLUSIONS

We presented progress towards the SADL, an interface designed to secure and authenticate data-links within a standardized handle system. In doing so, we stressed the importance of implementing effective and interactive visualizations that aid users in querying a large collection of digital resources. We demonstrate this process across four leading standards for the model-based enterprise: STEP AP242, STEP AP238, MTConnect, and QIF. We expect that others can follow the same process for other standards as well. Relating the underlying domain-specific data to a taxonomy of domain-agnostic data types eases the integration of state-of-the-art visualizations. Such visualizations are expected to be integrated within the SADL interface. Future work will consider the feasibility of generalized visualizations so that a variety of digital resources can be represented to enhance organizational decision-making.

This publication is available free of charge from: https://doi.org/10.6028/NIST.AMS.100-24

DISCLAIMER

No endorsement of any commercial product by NIST is intended. Commercial materials are identified in this report to facilitate better understanding. Such identification does not imply endorsement by NIST nor does it imply the materials identified are necessarily the best available for the purpose.

ACKNOWLEDGEMENTS

We thank Moneer Helu, Robert Lipman, and Tom Kramer for their valuable feedback that improved the paper.

REFERENCES

[1] Auschitzky, E., Hammer, M., & Rajagopaul, A. (2014). How big data can improve manufacturing. *McKinsey & Company*, *822*.

[2] Yin, S., & Kaynak, O. (2015). Big data for modern industry: challenges and trends [point of view]. *Proceedings of the IEEE*, *103*(2), 143-146.

[3] Panetto, H., & Molina, A. (2008). Enterprise integration and interoperability in manufacturing systems: Trends and issues. *Computers in industry*, *59*(7), 641-646.

[4] West, T. D., & Blackburn, M. (2017). Is Digital Thread/Digital Twin Affordable? A Systemic Assessment of the Cost of DoD's Latest Manhattan Project. *Procedia Computer Science*, *114*, 47-56.

[5] Hedberg, T., Lubell, J., Fischer, L., Maggiano, L., & Feeney, A. B. (2016). Testing the digital thread in support of model-based manufacturing and inspection. *Journal of computing and information science in engineering*, *16*(2), 021001.

[6] Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Visual Languages, 1996. Proceedings., IEEE Symposium on*. IEEE, 1996.

[7] Paskin, N. (2010). Digital object identifier (DOI®) system. *Encyclopedia of library and information sciences*, *3*, 1586-1592.

[8] Kahn, Robert, & Wilensky, Robert, "A framework for distributed digital object services," May 13, 1995. URL: http://www.cnri.reston.va.us/home/cstr/arch/k-w.html. Accessed December 1, 2018.

[9] Bajaj, Manas, and Thomas Hedberg Jr. "System Lifecycle Handler—Spinning a Digital Thread for Manufacturing." *INCOSE International Symposium*. Vol. 28. No. 1. 2018.

[10] ISO (2014). 10303-242: 2014, Industrial automation systems and integration—product data representation and exchange—Part 242: Application protocol: Managed model based 3d engineering. *Geneva (Switzerland): International Organization for Standardization (ISO)*.

[11] ISO (2007). 10303-238: 2007, Industrial automation systems and integration—product data representation and exchange—Part 238: Application protocol: Application interpreted model for computerized numerical controllers. *Geneva: International Organization for Standardization (ISO)*.

[12] Sobel, W. (2015). MTConnect standard. Part 1—overview and protocol. *Standard—MTConnect.* URL: http://www. mtconnect. org/standard. Accessed December 1, 2018. .

[13] Dimensional Metrology Standards Consortium. (2018). Part 1: Overview and Fundamental Principles in Quality Information Framework (QIF)—An Integrated Model for Manufacturing Quality Information. *Dimensional Metrology Standards Consortium*. URL: http://qifstandards.org/. Accessed Dec 1, 2018.

[14] Jerding, Dean F., and John T. Stasko. "The information mural: A technique for displaying and navigating large information spaces." *IEEE Transactions on Visualization and Computer Graphics* 4, no. 3 (1998): 257-271.

[15] Stolte, Chris, Diane Tang, and Pat Hanrahan. "Multiscale visualization using data cubes." *IEEE Transactions on Visualization and Computer Graphics* 9.2 (2003): 176-187.

[16] Shneiderman, Ben. "Tree visualization with tree-maps: 2-d space-filling approach." *ACM Transactions on graphics (TOG)*11.1 (1992): 92-99.

[17] Van Ham, Frank. "Using multilevel call matrices in large software projects." *Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on*. IEEE, 2003.

[18] Heer, J., Bostock, M., & Ogievetsky, V. (2010). A tour through the visualization zoo. *Commun. Acm*, *53*(6), 59-67.

[19] Carpendale, M. S. T. (2003). Considering visual variables as a basis for information visualisation.