This is an Accepted Manuscript of an article published by ASME in Journal of Computing and Information Science in Engineering on 14 JUL 2016, available online: http://doi.org/10.1115/1.4034132.

Towards a Lifecycle Information Framework and Technology in Manufacturing

Thomas Hedberg, Jr.*, **Allison Barnard Feeney, Moneer Helu**
Systems Integration Divison, Engineering Laboratory
National Institute of Standards and Technology, Gaithersburg, Maryland 20899

Jaime A. Camelio
Grado Department of Industrial and Systems Engineering
Virginia Tech
Blacksburg, Virginia 24061

ABSTRACT

Industry has been chasing the dream of integrating and linking data across the product lifecycle and enterprises for decades. However, industry has been challenged by the fact that the context in which data is used varies based on the function / role in the product lifecycle that is interacting with the data. Holistically, the data across the product lifecycle must be considered an unstructured data-set because multiple data repositories and domain-specific schema exist in each phase of the lifecycle. This paper explores a concept called the Lifecycle Information Framework and Technology (LIFT). LIFT is a conceptual framework for lifecycle information management and the integration of emerging and existing technologies, which together form the basis of a research agenda for dynamic information modeling in support of digital-data curation and reuse in manufacturing. This paper provides a discussion of the existing technologies and activities that the LIFT concept leverages. Also, the paper describes the motivation for applying such work to the domain of manufacturing. Then, the LIFT concept is discussed in detail, while underlying technologies are further examined and a use case is detailed. Lastly, potential impacts are explored.

1 Introduction

Contextualizing data from the product lifecycle to make design decisions is very difficult. Different data in the product lifecycle is stored in different locations with different people using the data in different ways and in different contexts. The significant difference in data across the lifecycle is the reason why industry anecdotally says, “the lifecycle is drowning in data, but starving for information.” A solution is needed to link all the disparate systems of the lifecycle and cultivate information for decision support. We propose a Lifecycle Information Framework and Technology (LIFT) concept to develop and integrate technology and standards to enable a novel and straightforward product lifecycle management (PLM) implementation that is intelligent, self-learning, and self-aware. The LIFT concept would stretch and/or replace current PLM paradigms with innovative processes and technologies to remove the “silo effect” between organizations. The intent is to create the “Google” for engineering and manufacturing data that supports data curation and information cultivation in an efficient and effective manner. The LIFT concept supports a “data observatory” wherein a user of the PLM system-of-systems would

*Corresponding Author, tdh1@nist.gov

be able to search, discover, and retrieve information from throughout the enterprise when the information is needed. Our viewpoint of a data observatory, synonymous to an astronomical observatory, is a technology that supports the study of engineering and manufacturing phenomena and events through the use of data from the product lifecycle. The LIFT concept is a framework for lifecycle information management and the integration of emerging and existing technologies, which together form the basis of a research agenda for a common model to support digital-data curation in manufacturing industries.

This paper first provides a discussion of the existing technologies and activities that the LIFT concept leverages and integrates in novel ways. Then, the paper describes the motivation for applying such work to the domain of manufacturing. Then, the proposed LIFT concept is described. Underlying technologies are further examined. A use case is detailed. Lastly, potential impacts are explored.

2 Background

The concept of linking cross-domain data is not new. The contribution of the LIFT concept is the implementation and integration of multiple existing technologies, paradigms, and concepts in a novel way to form a framework that would provide significant benefit to manufacturing industry groups. In this section, we provide background information about concepts and technologies that have informed our development of the LIFT concept. In each sub-section, we review similar efforts being deployed and/or tested in scientific domains and discuss how the similar concepts informed our work.

2.1 Open Data

An open-data culture is growing globally. Technology is enabling the open-data revolution – governments, academic institutions, and industries are using data to create knowledge about the world and make decisions [1]. Open data is, "data and content that can be freely used, modified, and shared by anyone for any purpose" [1]. The Open Knowledge Foundation launched CKAN [2], which is an open-source data portal platform to streamline publishing, sharing, finding, and using data [3]. CKAN provides the infrastructure sufficient for sharing data across the globe, but it lacks the ability on its own to provide trust and traceability in the data. The Open Data Institute provides a solution to fill the trust and traceability gap.

In addition, the Open Data Institute (ODI) [4] strives to create economic, environmental, and social value through open data. The ODI is developing the "open data certificate" to help data publishers explain what the open data is about and deliver open data that people can trust [5]. The ODI certificate is similar to the X.509-based [6] digital certificates used widely across the Internet today. The goal of the ODI certificate is to provide sufficient meta-data about the open data so users can vet the fidelity of the data [5]. The ODI certificates may act as a "seal of approval" to build trust in open data.

A framework that includes the CKAN and ODI certificate technologies would deliver significant impact to the public. The U.S. General Services Administration [7] claims that "open government data is important because the more accessible, discoverable, and usable data is the more impact it can have. These impacts include, but are not limited to: cost savings, efficiency, fuel for business, improved civic services, informed policy, performance planning, research and scientific discoveries, transparency and accountability, and increased public participation in the democratic dialogue."

2.2 Open Data Initiatives in the Sciences

The expansion of the open-data culture is enabling scientific communities to develop a deep understanding needed to reform data collection and sharing in the sciences. Three successful applications of open-data concepts are the Physical Science Informatics (PSI) system [8], the Ocean Observatories Initiative (OOI) [9], and EarthCube [10]. These examples show how open data can accelerate scientific discovery through a "force multiplier" effect gained through the ability to reuse or expand upon shared datasets. Further, the PSI, OOI, and EarthCube examples may provide the foundation for the eventual development of a data observatory.

Scientists have been conducting micro-gravity experiments on the International Space Station (ISS) since 2001. The National Aeronautics and Space Administration (NASA) developed a repository to store all of the raw or minimally processed data from scientific experiments conducted on the ISS. The NASA PSI system [8] makes the stored data available to the public. The NASA PSI system provides an informatics repository that may extend data from a single research project to multiple research opportunities.

OOI determined studying the complex interaction of biological, chemical, physical, and geological processes in the ocean is limited severely by available technical infrastructures [9]. In response, the OOI developed an interactive cyber-infrastructure with common design elements to enable integrating multiple scales of ocean observations [9]. The OOI cyber-infrastructure brings researchers together for data curation, discovery, and sharing of experimental data captured throughout the OOI system. The OOI infrastructure provides persistence, adaptability, interoperability, and community related to the data in the system.

EarthCube [10] uses a concept similar to the OOI for monitoring geo-sciences data. EarthCube discovered common themes that called for the integration of data across different domains. EarthCube recognizes that the complexity of different domain models presents a key challenge to the geo-sciences community. But EarthCube believes, "to cultivate future generation of researchers and also for EarthCube to be of use to policy and decision makers, geo-scientists must be able to retain access to and communicate the results and uncertainties of information and research that advances knowledge to society" [10].

2.3 Manufacturing Data in the U.S. Department of Defense

In 1988, Defense Logistics Agency (DLA) implemented the Military Engineering Data Asset Locator System (MEDALS). MEDALS is a globally accessible networked system where engineering drawings and technical documents reside [11]. DLA [11] describes the system as a research tool to act as a "first-discovery mechanism to assist a user with finding engineering documents when the user does not know where the document might reside. MEDALS also houses the specific repository in which the documents reside. As of 2014, MEDALS tracked the location of 44 million engineering data assets across 45 different data repositories [11]. However, use of MEDALS has decreased recently because a significant amount of human input is required to keep the system up to date. DLA recognized MEDALS needed a web service interface to enable automating data-asset management. DLA subsequently added Extensible Markup Language (XML) ingestion capabilities and enhanced batch querying capabilities to support a MEDALS web service interface [11].

The Department of Defense (DOD) also supports collaborative design research and development. The Defense Advanced Research Projects Agency (DARPA) Manufacturing Automation and Design Engineering (MADE) program [12] was one of the first collaborative design research projects to use a digital infrastructure. The goal of MADE was to develop Internet-based tools, services, protocols, and design methodologies to support the design activities of geographically distributed expert teams [12]. The project tested the ability of a diverse and geographically dispersed team to design a product collaboratively. The MADE program recognized the important need to curate the design process as much as the actual design. The project team developed the "Madefast Web," which was a set of web pages, shared via the World Wide Web (WWW), that acted as a repository for the computer-aided design (CAD) models, notes, test results, calculations, and other information relating to the design [12]. Because the MADE tools are Internet-based, information and applications were simply accessed through the point and click of a mouse.

The MADE program provides an interesting example of a knowledge base. Cutkosky et al. [12] said, "Madefast uses the WWW as a corporate memory, sharing design information across the design team, and preserving it for downstream tasks such as maintenance and redesign. Such information is, of course, useless if it cannot be found." In some aspects, the MADE project was ahead of its time. The benefits of the Madefast concept became evident when the project team decided to make a second version of the design product. The design information was readily available to accelerate decisions for the new design version. The MADE program also identified a few key issues that still exist today. Data organization within the Madefast Web was difficult. Although three major overhauls were made in an effort to keep data organized, newcomers remained unable to find information quickly [12]. Human interaction and different workplace cultures across the team also remained a challenge. However, the most significant issue in the MADE project related to systems and services integration.

2.4 Semantic Web

The Oxford Dictionary [13] defines *semantic* as, "relating to meaning in language or logic." In the early 2000's, several papers and magazines [14, 15, 16] wrote about adding semantic definition to the WWW and developing the concept of the *Semantic Web*. The World Wide Web Consortium (W3C) declares the Semantic Web to be the integration of linked data, vocabularies, query, inference, and vertical applications [17]. The WWW started out as a cluster of documents published via the Internet and linked to other documents using hypermedia methods [15]. Markup languages (e.g., HTML) are used to deliver content to people in human-readable form. The traditional document-centric WWW provides little focus on the representation of the data contained in the documents [15]. The purpose of the Semantic Web is to provide machine-readable data that represents the semantics of the concepts presented in traditional WWW documents.

The Semantic Web functions through machine-interpretable access to structured collections of information [14]. The first step in building the Semantic Web was to start building higher-order relationships between data. The concept of *Linked Data* provides a foundation for defining high-order relationships between sets of data. Berners-Lee [16] proposed four rules for Linked Data, which are: (1) use uniform resource identifiers (URIs) [17] as names for things, (2) use Hypertext Transfer Protocol (HTTP) URIs so names are discoverable, (3) provide useful information using standards (e.g., Resource Description Framework (RDF) [18], SPARQL [19]), and (4) include links to other URIs so n-order links are discoverable.

However, linking data is simply not enough to enable the Semantic Web. It is important that the data is organized and sufficient structure and context exist to support the evolution from data to information to knowledge. Vocabularies, or *ontologies*, support the query and inference components of the Semantic Web. Ontologies improve the accuracy of WWW searches at the basic level [14]. In more complex systems (e.g., the Semantic Web), ontologies relate the information of one system to the associated knowledge structures and inference rules used to apply semantic representation using data [14].

In 2009, Khilwani et al. [20] presented a survey of 14 different ontologies relevant to the manufacturing domain. Since then [21, 22, 23] have discussed additional ontology studies. While having only one ontology to encapsulate the entire product lifecycle would be ideal, achieving consensus on the definition of that ontology is unlikely because of the many different viewpoints that exist across the lifecycle. Therefore, what matters more is the ability to link data through the product lifecycle and apply the necessary context to the data to ensure the right information and knowledge is available to the roles and functions when they need it. Enabling linked data by normalizing the process for linking different ontologies in the Semantic Web enables a powerful data curation and discovery mechanism. Thus, integrating the Semantic Web with the product lifecycle provides users (e.g., humans, machines) the ability to build queries for discovering complex information, which is paramount to discovering and extracting data for the user.

Inference are also an important component of the Semantic Web. Inference enables automatic analysis of data. The automatic analysis, built with reasoning and other artificial intelligence algorithms, supports managing knowledge in the Semantic Web. Further, inference provides automatic procedures to generate new relationships based on the data and on additional information from ontologies [24].

Previous work in Semantic Research [22, 25, 26] provide some examples for the manufacturing domain. However, much of the Semantic Web research in manufacturing is limited to a single component of the Semantic Web such as inference systems or ontology development. There has been little investigation into fully integrating all of the Semantic Web components (i.e., linked data, vocabularies, query, inference, and vertical applications) into manufacturing and across the product lifecycle. Further, the ontologies discussed in [20, 21, 22, 23] do not address integrating cross-domain (e.g., design, manufacturing, quality) ontologies to generate an ontology for the entire product lifecycle. Design information may not be linked directly to quality information (e.g., [21]) and the manufacturing information may only be linked to a small portion of design information that is directly relevant to the manufacturing viewpoint (e.g., [22]). In other words, the Semantic Web in manufacturing remains as silos of domain-specific data linked only within the phase of the product lifecycle that generated the data. Therefore, the manufacturing domain remains an untapped opportunity for implementing the Semantic Web to provide significant impact with a full-integration of the product lifecycle.

2.5 Digital Thread for Manufacturing Domains

Information technology advances (e.g., data analytics, service-oriented architectures, and networking) have triggered a digital revolution [27] that when coupled with operational technology (e.g., hardware and software for sensing, monitoring, and control of product and processes) holds promise for reducing costs, improving productivity, and increasing output quality. Modern manufacturing enterprises are both more globally distributed and digital, resulting in increasingly complex manufacturing system networks [26, 28]. Manufacturers are under mounting pressure to perform digital manufacturing more efficiently and effectively within these distributed manufacturing systems. To do so, industry is changing how product definitions are communicated – from paper to models. This transition is being called model-based enterprise (MBE).

MBE has introduced new requirements on data usage in manufacturing systems. MBE calls for each phase and function of the product lifecycle to adopt model-based data standards to effectively integrate data for efficient reuse and exchange between product lifecycle phases. The need for automated methods to collect, transmit, analyze, and act on the most appropriate data is gaining attention [29, 30, 31, 32]. However, the MBE strategy must also ensure model-based-data interoperability between design activities (e.g., product and assembly design) and manufacturing activities (e.g., fabrication, assembly, and quality assurance). ISO 10303-242:2014 [33], MTConnect [34], and Quality Information Framework (QIF) [35] are three emerging standards that show promise for enabling linked data throughout the product lifecycle.

2.5.1 STEP AP242

ISO 10303-242:2014 [33] titled "Managed Model Based 3D Engineering," commonly known as Standard for the Exchange of Product Model Data Application Protocol 242 (STEP AP242), is an international standard that supports a manufacturing enterprise with a range of standardized information models that flow through a long and wide "digital thread that makes the manufacturing systems in the enterprise smart [36]. Digital data plays a central role in achieving the goal of STEP AP242. STEP AP242 contains extensions and significant updates to other Standard for the Exchange of Product Model Data (STEP) Application Protocols (APs) for product and manufacturing information (PMI), kinematics, and tessellation [37]. PMI is the presentation and representation of geometric dimensions and tolerances (GD&T), material specifications, component lists, process specifications, and inspection requirements within a three-dimensional (3D) product definition [38].

2.5.2 MTConnect

MTConnect is an open-source, read-only data-exchange standard for manufacturing equipment and applications developed by the Association for Manufacturing Technology (AMT) [34]. It is based on XML and HTTP and provides information models and communications protocols to enhance the data acquisition capabilities of manufacturing equipment and applications and to enable a plug-and-play environment. While other communication protocols may exist for data transfer, the information models defined in MTConnect are the only common vocabulary and structure created for manufacturing equipment data. Perhaps the most important type of data addressed by the standard is real and near-realtime data from the equipment (e.g., current speed or position, program blocks). This ability is critical in enabling the standard to support the digital thread by providing data and information on the as-built condition of a part.

2.5.3 Quality Information Framework

The QIF [35] is an American National Standards Institute (ANSI) standard sponsored by the Dimensional Metrology Standards Consortium (DMSC). QIF defines an integrated set of XML information models that enable the effective exchange of metrology data throughout the entire metrology process. QIF handles dimensional metrology use cases, including feature-based dimensional metrology, quality measurement planning, first article inspection, multi-part statistics and discrete quality measurement. QIF supports defining or importing the product definition and reusing data for inspection planning, execution, analysis, and reporting. A full range of quality measurement terminology and semantics is represented in the QIF standard. The QIF information models are normalized in XML Schema Definitions (XSD). The QIF XSDs are organized into six application areas for metrology: model-based definition (MBD), Rules, Resources, Plans, Results, and Statistics.

3 Motivation for Manufacturing

Manufacturing organizations are increasingly using digital engineering artifacts in the product lifecycle. Industry is calling for a digital thread to stitch the phases of the product lifecycle together [38]. In practice, PLM describes several phases within the product lifecycle. For the purposes of this paper, we define the product lifecycle to be the design, analysis, manufacturing, quality assurance, and customer and product support phases. A digital engineering artifact is an object created and managed using primarily software tools (e.g., CAD, computer-aided manufacturing (CAM)) [39]. MBD is a type of digital engineering artifact. To manage the data within the software systems, the tools implement proprietary data formats for storing the data. Due to various data format changes aligned historically with new product introductions from the software vendors, industry consortia developed various standard open-data formats (e.g., Juniper Tesselation (JT) [40], Portable Document Format (PDF) [41] / Product Representation Compact (PRC) [42], STEP [33]). In addition, industry has adopted domain-specific ad-hoc format specifications (e.g., ACIS [43], Stereolithography (STL) [44]) that are published openly by software vendors.

A major challenge for the product lifecycle is the existence of various data format standards, and the existence of few practice standards and no lifecycle information standards. While there are many documented mapping efforts to create interoperability between domain-specific data standards, the majority of the mappings have no way of determining what data and context are required for each phase of the product lifecycle. This leads to information being lost with every data translation – starting with the first translation out of the software tool (e.g., CAD) where the data originated. Moreover, data coming from the authoring CAD system are typically only shape representations [39]. To make sense of the data, the data users in the product lifecycle may also require information about the provenance [45] of the data, feature semantics [46] of the product, and/or activities within an organizational workflow [39].

In addition, the Bill of Materials (BOM) differs between the various phases of the product lifecycle because downstream functions (e.g., manufacturing, quality, product support) require additional information that engineering does not include in the original BOM. The downstream functions take the BOM coming from engineering and modify it or generate new BOMs to meet functional needs. This has led to creating multiple BOMs with different lifecycle viewpoints. The most commonly found BOMs are the engineering/design BOM (eBOM), the manufacturing BOM (mBOM), and, most recently, the maintenance and support BOM (sBOM). This highlights an interoperability problem similar to the interoperability issues in CAD systems. While there are common elements in each of the BOMs, keeping all the disconnected data flows synchronized is difficult.

The lack of a common-elements model for the product lifecycle is the key driver for the LIFT concept. The QIF

This is an Accepted Manuscript of an article published by ASME in Journal of Computing and Information Science in Engineering on 14 JUL 2016, available online: http://doi.org/10.1115/1.4034132.

standard [35] started to tip the scales toward common elements. However, QIF is not broad enough for all of the information in the entire product lifecycle because the standard was developed to address only the metrology domain-specific issues. Historically, industry using the available commercial PLM solutions [47, 48, 49] suggested managing all of the lifecycle data in a singular (homogeneous) system as a workaround to the lack of a common-elements model.

Estimates [50] point out that data interoperability costs the automotive supply chain over $1 billion[1], but the interoperability cost of knowledge transfer and PLM is immeasurable currently. In practice, PLM requires the consideration of many product-data forms beyond the simple inclusion of CAD and BOM information. However, solution providers and industry often conflate singular large-enterprise product-data management (PDM) tools and PLM.

The industry's and the commercial PLM vendors' suggestion for a homogeneous system across the enterprise is unrealistic because it is cost prohibitive for small-to-medium enterprises (SMEs). Furthermore, PDM, enterprise-resource planning (ERP), manufacturing execution (MES), and quaility management (QMS) systems were built to solve different problems in the lifecycle. Trying to integrate all of those requirements into a homogenous system would result in a system that is a, "Jack of all trades, master of none," [51]. The motivation of this paper is to introduce a concept for a lifecycle information framework that supports a common-elements model of the product lifecycle, which would enable the development and implementation of technology built around the information needs of the product lifecycle while utilizing a heterogeneous system-of-systems.

More recently, the PLM domain has shifted towards "platformization" [52]. The platform concept calls for a foundational infrastructure that represents a PLM operating system. This system would provide a baseline set of capabilities to the end-user. Then, applications, or "Apps," could be plugged into the system to extend the system's capabilities. The platform concept has been successful in the smart-phone domains – Apple iOS and Google Android™ are two examples. Currently, each of the major commercial PLM vendors are developing platform solutions. While, platforms work well in the smart-phone domain, platforms may not be the answer for manufacturing. Much like the single-vendor homogeneous PLM solutions, what platform should industry select? Industry needs a universal plug-and-play solution that would enable native integration of the many systems across the product lifecycle. In Section 4 and Section 5, we present our novel approach for integrating systems across the product lifecycle regardless of the type (e.g., homogeneous, platform, pure heterogeneous) of system.

4 The Framework

Figure 1 presents the structure of the LIFT concept integrated with the product lifecycle. The framework consists of three layers: (1) product-lifecycle data, (2) data certification and traceability, and (3) data-driven applications. This section describes each layer in detail.

Product lifecycle data makes up the first layer of the framework. Recall, the design, analysis, manufacturing, quality assurance, and customer and product support phases define the product lifecycle for the purpose of this paper. The design

[1] Though over 15 years old, the estimate is still relevant. In fact, the authors propose the costs are much more significant than the original estimate due to the increase in the complexity of new products and systems used to design those products.

This is an Accepted Manuscript of an article published by ASME in Journal of Computing and Information Science in Engineering on 14 JUL 2016, available online: http://doi.org/10.1115/1.4034132.

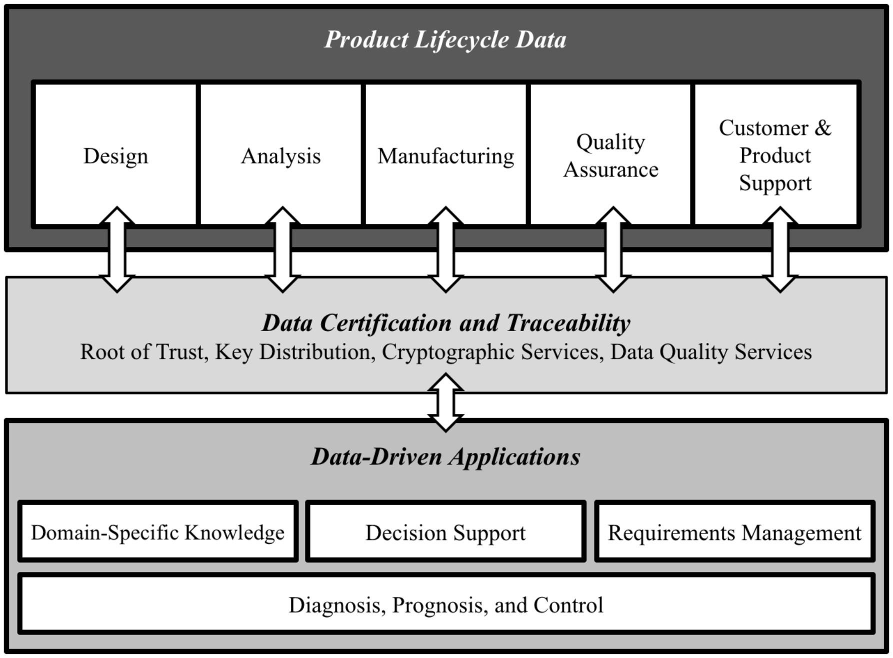

Fig. 1. Essential supporting structure of the LIFT concept and the relationship of the three layers that make up the framework.

phase encompasses design processes and functions, which require data be approved by a person with the appropriate authority, certified by providing confirmation the data is what it is declared to be, and then authorization provided for how the data may be used. The analysis phase of the lifecycle analyzes a product using computer-aided engineering (CAE) tools (e.g., simulation-based finite-element analysis (FEA) and computational fluid dynamics (CFD)). The manufacturing phase is where a product is fabricated. The quality assurance phase deals with the inspection and measurement of a product. Lastly, the customer and product support phase manages the end-user support through maintenance services, technical publications, and other support activities. Each of these phases have its own set of data that it keeps for recording-keeping. This data would benefit other phases of the lifecycle too, but linking the data today is difficult and requires significant human capital.

The design development activity is often supported by a design knowledge base. The knowledge base contains metadata, rules, standards, correlations, and design dictionaries. The knowledge base consists of different types of information transformed into knowledge to support decision making. Those decisions may be manual or automated. Regardless of how the decisions are made, typically the design process must follow some type of knowledge base to be able to design a product.

The product can be a new product introduction, for research and development, or a revision of an existing product. The knowledge base should support the decision-making for all product types based on the knowledge collected over time. However, current industrial knowledge-based methods require manual input and manual interpretation of the knowledge – usually through documents in large policy books or on-line repositories of text-based documents. If a knowledge base has automation, the automation is typically rules-based. In the majority of cases, a human must be able to read documents to extract information from the knowledge base. To make matters worse, there are often multiple domain-specific knowledge bases (e.g., design, manufacturing, quality) throughout the lifecycle and those knowledge bases are not always in agreement with all the lifecycle requirements because they all have different viewpoints.

As evident in the DOD MEDALS [11], the amount of human capital required to keep the data repositories up-to-date is significant. The lifecycle cannot afford to deploy the needed amount of human capital to maintain near-dynamic knowledge bases and still deliver products to market. Industry needs a way to discover data relationships and link the data across the lifecycle. This would enable near-real-time dynamic updating of domain-specific knowledge bases in the lifecycle using machine learning and artificial intelligence methods. In the LIFT concept, the only human capital required to enable to automated updating is a data administrator needs to register a domain-specific knowledge base with a “registry” to ensure data is discoverable. Once that registration is complete, the knowledge base would receive near-real-time dynamic updates as users complete their day-to-day activities.

In the middle of the framework is the *Data Certification and Traceability* layer. The data certification and traceability layer supports building trust throughout the product lifecycle. Throughout the whole lifecycle there are different requirements that come in and out of the lifecycle. A lot of those requirements from different phases of the lifecycle often contradict or compete with each other. This raises challenges for the lifecycle to be able to manage all of those requirements, to understand the requirements, and to use the requirements effectively. Misunderstanding and/or not complying with all the requirements leads to distrust of the data.

To enable and ensure trust, the framework needs cryptographic and product-date quality (PDQ) services available through the data certification and traceability layer. The PDQ services would ensure the product data is verified and validated against the multiple sets of requirements that are constraints on the product in the lifecycle.

The PDQ services could interface with a *Requirements-Management* application in the *Data-Driven Applications* layer of the framework. The data-driven applications layer would support integrating applications through using plug-and-play methods. Initially, our framework could include applications to support domain-specific knowledge management, decision support, requirements management, diagnosis, prognosis, and control.

A requirements-management application could work closely with a knowledge-management application to ensure all of the lifecycle requirements are captured, understood, and available for reuse. Once the PDQ services complete the verification and validation activities, the cryptographic services embed digital certificates in the product data to create *digital fingerprints* that enable authentication, authorization, and traceability through the lifecycle. The certificates assure the product data is what it says it is (i.e., authentication) and the product data can be used how it is intended to be used (i.e., authorization). Moreover, the certificates support traceability by capturing who did what to whom and when it was done. Overall, the certificates bring seamless authentication, authorization, and traceability to the product data.

Having always up-to-date knowledge bases would support a *Decision Support* application. Working in concert with a knowledge-base application and a requirements-management application, the decision-support application could provide near-real-time feedback to a user as decisions are made. For example, the design knowledge base and requirements manager could build design for manufacturing (DFM) rules based on diagnostic and prognostic data feedback from manufacturing and quality assurance. Those DFM rules would notify the design engineer of tool reach issues based on the data from manufacturing, which would ensure that the design engineer develops a product that manufacturing can produce effectively and efficiently.

Furthermore, with the creation of the QIF standard [35], engineers now have the ability to conduct quality analytics. Quality analytics would allow engineers to look at all the different results, resources, rules, and statistics coming out of a quality organization. The amount of data that can be readily available from the quality-assurance phase of the lifecycle supports the ability to run automated analytics. The analytics can mine data that engineers can turn into design information by applying engineering context to the quality data. Quality analytics is key in generating correlations between the virtual and physical worlds because quality is often the first point in the lifecycle where data is generated from both worlds and compared to each other.

For example, a CAD model could be used during the inspection process to verify the conformance of the physical products. Inspection reports and digital-data sets would be generated that hold valuable data representing the physical product in the cyber-space. This would enable a cyber-physical systems view of the product lifecycle, where the quality data could be used for the purposes of quality assurance, product acceptance, and analyzing what happens throughout the lifecycle. The data-driven application layer of framework would enable the ability to leverage statistical process control, prognosis, and health monitoring methods in novel ways – such as, feed-back and feed-forward between design and supply chain to control the entire product lifecycle.

Customer and product support historically have large amounts of performance data, maintenance records, and customer feedback data that are stored in some location. But a lot of times the data is represented within paper-based records or some disparate database system. This makes getting feedback to the design knowledge base or design engineering role very difficult, if feedback happens at all. The manufacturing functions provide some feedback to engineering, but a lot of the feedback is through ad-hoc discussions. The discussions are started typically because a part cannot be built or manufacturing is having difficulty with a requirement from the design. When formal communication is used between manufacturing and engineering, it is through the use of problem and corrective action reports. But in reality, an e-mail or phone call between the manufacturer and designer are the primary communication methods.

The LIFT concept would enable the input of all the disparate pieces of data from the lifecycle-phase silos into a self-learning and self-aware system supported by the data-driven application layer of the framework. Eventually, we would like the system to utilize self-learning algorithms for looking at semantic, syntactic, linguistic, and statistic data that comes from all the different data and information repositories of analysis, manufacturing, quality, and customer and product support. The stream of data and information into these self-learning algorithms would enable the system to learn dynamically from streams of data based on product experience. The learning supports near-real-time dynamic updating of design knowledge bases in an effective manner. The effective and traceable information flow, self-learning methods, dynamic updating of the design knowledge base, and real-time decision support form the framework of the LIFT concept.

5 The Technology

The backend of the LIFT concept is a derivative of the Handle System – a digital object architecture – developed by the Corporation for National Research Initiatives (CNRI) [53]. The Handle System defines three components: an identifier system, meta-data registries, and digital object repositories [54]. The Handle System architecture enables interoperability

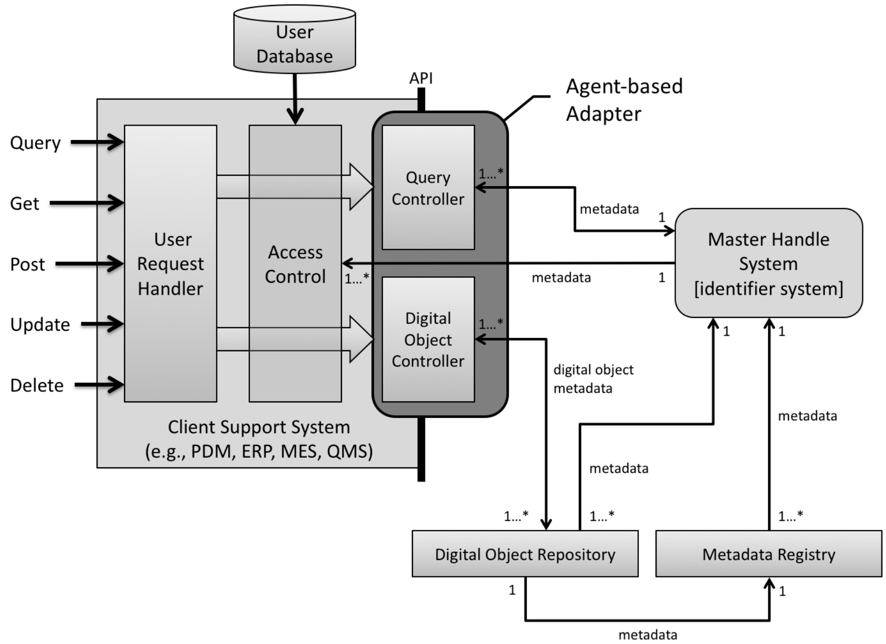

Fig. 2. Schematic overview of the LIFT concept

for resolution purposes among a varying set of implementations. The Handle System consists of global root servers, local handle servers, clients, and proxy servers – making the system scalable, reliable, and secure for identifier resolution [54]. The purpose of the Handle System is to manage digital objects. A digital object is an abstract entity that has a persistent and unique identifier [54]. The digital object may point a user to a physical, virtual, and/or abstract thing. The most popular implementation of the Handle System is the Digital Object Identifier (DOI®) system [55] supported by the International DOI Foundation [56] and defined in ISO 26324 [57].

The LIFT concept provides a master handle system to act as a "traffic cop" of data. The LIFT concept builds upon the concept of ISO 26324 [57] to develop a manufacturing-centric product-lifecycle extension to CNRI's Handle System [53]. A potential outcome of this work is the development of a standard for a manufacturing master handling system. The master handle system is the identifier system and resides between all the different databases and/or repositories of data. Figure 2 shows the technology architecture of the LIFT concept. The master handling system is an "index-of-indexes that would understand and be able to inform a user where to go to find particular pieces of information and assist the user with retrieving the appropriate information. The goal of the master handle system is to replace the extensive burden for maintaining links between data that existed in the DOD MEDALS [11] system.

Each piece of information stored in the various repositories across the lifecycle is considered a digital object. Each of the existing databases and/or repositories is considered a digital-object repository. The master handle system does not store or manage the digital objects. The master handle system controls an index of different digital-object repositories and what types of data those digital-object repositories contain. The meta-data registries assist the master handling system with determining the types of data in the digital-objects repositories. In addition, the meta-data registries support access control – an often overlooked kind of meta-data [54] – by controlling who or what can use the digital object and/or how the user can use the digital object. The access control, supported by the data certification and traceability layer of the LIFT framework, enables authentication and authorization capabilities.

Typically, an enterprise utilizes more than one type of data management system that has its own repository and acts as a client support system. These client support systems are the solutions mentioned earlier at the end of Section 3. Engineering and design organizations use the PDM system, manufacturing and supply chain organizations use the MES and ERP systems, and the quality assurance organizations use the QMS systems. Each of those systems already have query, get, post, update, and delete capabilities integrated. The goal of the LIFT concept is to leverage each of the existing client support systems and develop a common element model and integration technologies to support the flow of information between the systems with little or no customization required in the client support systems.

We propose an agent-based adapter method for integrating systems in the product lifecycle. Bond and Gasser [58] suggest agents should be used when all of the following are true:

- Data, control, expertise, or resources are inherently distributed.
- The system is naturally regarded as a society of autonomous cooperating components.
- The system contains legacy and new components, which must be made to interact with each other.

The product lifecycle meets all three elements of Bond's and Gasser's agent-based rationale. Our agent-based adapter method would wrap services to support integrations across the lifecycle. The initial services to target are query control and the digital-object control. Each control service would be a "micro-service" wrapped by the agent-based adapter, which would act as an application programming interfaces (APIs) gateway between a client-support system and other product lifecycle systems. Wrapping the control services in an agent-based adapter is an example of a micro-services architecture.

Micro-services architecture requires splitting applications into a set of smaller interconnected services instead of building singular monolithic applications. A service would implement a set of distinct functionality in our case, query or digital-object control. Each functional area of our application could be implemented by its own micro-service. The result is a set of simpler applications making it easier to deploy distinct experiences for specific users, devices, or specialized use cases [59].

Using micro-services architecture supports our key goals for the agent-based adapter and integrating various systems across the product lifecycle. Micro-services support scalability and state-less services. Fast plug-and-play would be achievable, which supports low-friction deployment of solutions. The architecture also supports the management of common concerns with infrastructure and framework standards. Deployments only need to worry about business logic. We could enable automated service discovery and routing through decoupled services. Flexible integration and registration of systems with the rest of the product lifecycle supports SMEs with minimal IT resources. Lastly, micro-services would enable a universal Lifecycle Object Identifier (LOI) schema for the handle system by supporting the mapping from vendor-specific and/or proprietary systems to the rest of the product lifecycle.

Most, if not all, enterprise systems have accessible APIs that allow external systems to interact with the enterprise system using a set of routines, protocols, and/or tools defined by the enterprise systems. Having no standard APIs for integrating enterprise systems is a major challenge today. The solution providers each develop their own APIs to be as open or closed as their business models allow. Using the LIFT concept for integrating the enterprise systems, the query controller and digital object controller concept enables a pseudo-universal plug and play method for integrating all of the systems in the enterprise.

Additionally, we would be able to leverage existing service-oriented architecture (SOA) solutions to enable PLM functionality instead of having to develop new SOA capabilities. The existing SOA could be wrapped in the agent-based adapter. The same can be done for existing solution stacks or emerging PLM platforms. We could integrate those stacks and platforms using an integrated agent-based adapter. The agent-based adapter method supports integrating homogeneous systems, platforms, and heterogeneous systems alike – all built with plug-and-play functionality.

The master handling system, through the agent-base adapter, would work in concert with the query controller to enable the building of queries to search for information from the different digital object repositories. The query controller is the link between the user and the master handling system. The query controller would pose a query to the master handling system based on the user's input from the client-support system. The master handling system forms the proper query that is interoperable with all the repositories in the lifecycle. The link between the query controller and master handling system ensures a query across an enterprise is transparent to the user without requiring multiple query inputs from the user. Overall, the queries allow the indexing of all the data that already exists in databases. Then, the indexing supports communicating the data digitally through the lifecycle to the roles and functions that have a need to know based on queries that can be posed by the master handling system.

Once the data or information is discovered in the enterprise, the digital-object controller (Figure 2) takes over. The digital-object controller is essentially a data-flow manager. Another goal of the LIFT concept is to not duplicate the data throughout the enterprise. The LIFT concept would work to enable and support data and/or information discovery. The digital-object controller moves data between the source and requesting digital-object repositories as the data is needed. The digital-object controller lets the system clean the data, determine the data quality, and apply the correct context to the data to transform the data into the needed information. This would ensure the data is not just duplicated across the enterprise, but is put to effective and efficient use.

The integration of technology in the LIFT concept forms a semantic web [17] for manufacturing and the entire product lifecycle. The LIFT concept leverages existing ontology research for manufacturing by reusing, expanding, or modifying the ontologies based on the needs for the full product lifecycle. In addition, the extensive research on linked data, described in Section 2.4, is reused throughout the master handling system. The query and inference portions of the semantic-web architecture defined by the World Wide Web Consortium (W3C) [17] is a research output of the query controller, digital-object controller, and machine-learning algorithms in the data-driven applications layer that are described in this paper. Lastly, a vertical application of the LIFT concept using semantic web methods is described as a use case example in the next section.

6 Use Cases Descriptions

To validate the LIFT concept, we apply the concept to the use cases of the engineering change request (ECR) and dynamic scheduling processes. We believe the ECR and scheduling processes are high-quality use cases because all portions

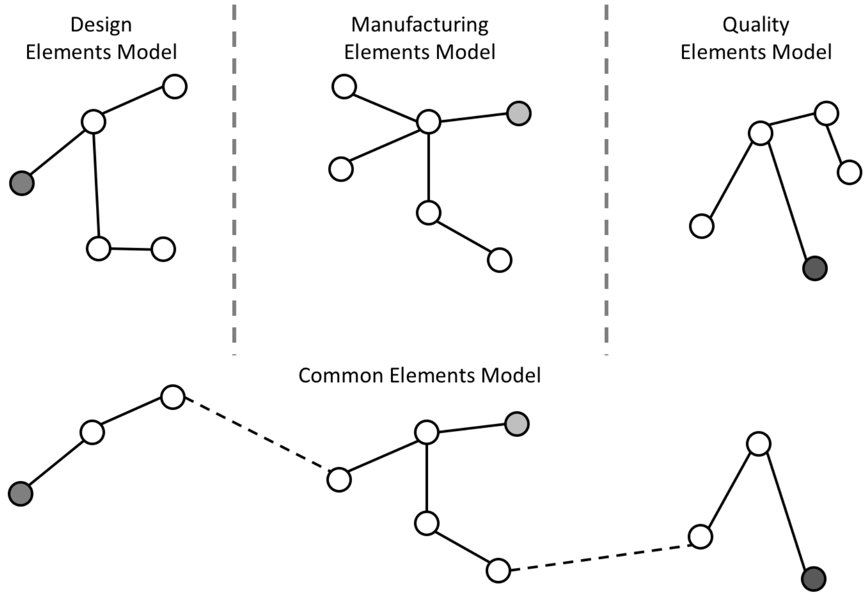

Fig. 3. Example Common Elements Model formed by linking domain-specific information models

of the product lifecycle have the ability to influence a product change and process improvements. A challenge for industry is to determine when a change is needed in a product or process. Industry struggles with determining when a product or process change is needed because discovering enough information from the product lifecycle is a costly activity.

The LIFT concept supports the information discovery activity with the implementation of the common elements model. Figure 3 shows an example of how the common elements model is formed by linking domain-specific element models. The goal of the common element model is to ensure linkage between the minimum amount of information the product lifecycle needs in order to be successful. The LIFT concept does not produce new domain-specific models, but leverages previous research and development to extend existing domain-specific models by linking them together to encapsulate the entire product lifecycle information needs. For example, in Figure 3's common elements model, information may flow between the shaded node in the design elements model to the shaded nodes in the manufacturing and quality elements models. The links between the domain-specific models enables data with context to flow between product lifecycle phases and roles by allowing access to various cross-domain nodes required by each phase and role.

Figure 4 shows a hierarchical model of the data flow in ECR and dynamic scheduling processes. At the top of the hierarchical model are the lifecycle processes (e.g., design, manufacturing, quality assurance). As you move down the model, there is an abstraction of data between the process layer and the product data, common elements model layer, and the decision layer. The design, manufacturing, and quality assurance activities are part of a process layer. In the product data layer exists the virtual and physical data related to the product. The product definitions (e.g., 3D models), process monitoring data (e.g., MTConnect [34] output), and quality measurement data (e.g., QIF results files) are examples of the type of data in the product data layer. The next layer is the common elements model layer, which is where data combined with context to generate actionable information. That last layer takes into account knowledge built upon the linked data and information

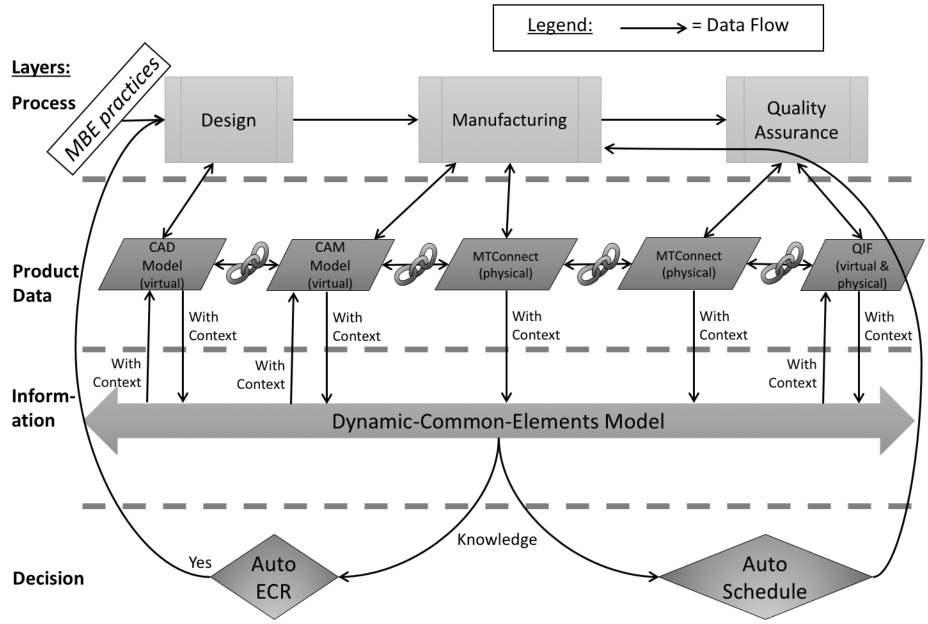

Fig. 4. Example data flow for engineering change requests to design and automated/dynamic scheduling to manufacturing. Each dotted line, represents a different abstraction layer. Here, we illustrate how data exchange manifests itself. First, as data is collected out of the process layer, context is added to the data layer to generate the information layer. Then, knowledge is built upon the information layer to support decision making. In this case, deciding when an engineering-change request is required or to determine dynamic scheduling within manufacturing.

flows through the product lifecycle and supports the actual decisions for ECR and dynamic scheduling processes.

Thus, the data flows out of the process layer into the common elements model by applying the appropriate domain-specific context to the data from each activity. This is the transformation of data into information. That information can be analyzed upon which to build knowledge. The gained knowledge would support decisions that need to be made in the product lifecycle.

The LIFT concept is implemented in all layers of the ECR and scheduling processes shown in Figure 4. The common elements model would act as a universal information communicator in the framework portion of the LIFT concept. In the process level, MBE recommended practices are implemented to ensure an effective level of product-data quality through the activities. However, the majority of the LIFT concept is implemented in product data and decision layers of the use cases. The data-driven application layer of the LIFT framework (Figure 1) is the link between all the process activities and their associated product data. The applications would leverage various product definitions (e.g., CAD models), MTConnect [34] output, and QIF [35] to implement linked data methods. The domain-specific knowledge base update engines, decision support engines, and requirements management engines from the LIFT framework are enabled in the decision layer of the ECR use case.

Figure 5 describes an example process for automating the ECR process. The data-driven application layer of the LIFT framework interfaces with the knowledge-base engines, decision-support engines, and the requirements-management engines to generate and/or discover information that supports automating the ECR process. The first step in the automated ECR

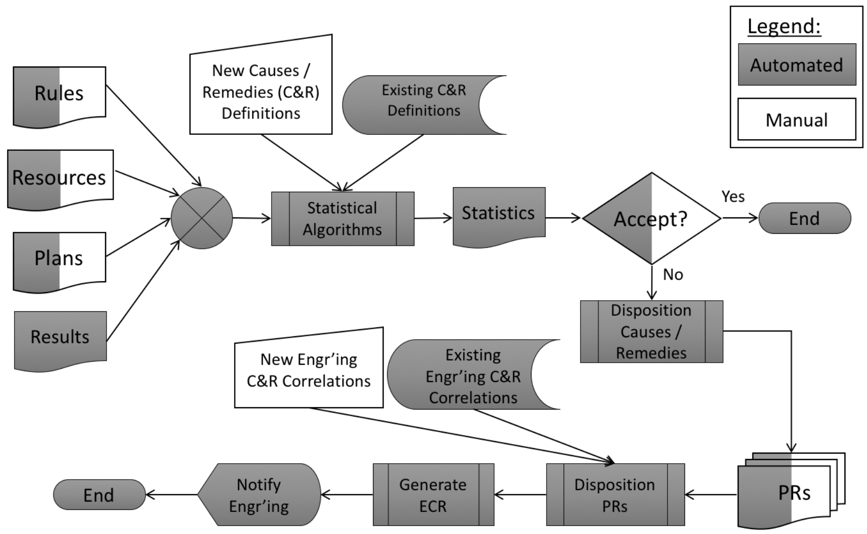

Fig. 5. Example process for Automated Engineering Change Requests to show how data is first aggregated together, statistical anomalies detected, engineering-design problems discovered, and engineering changes requests generated.

process is to aggregate all the virtual-world and physical-world data. The lifecycle (e.g., design, manufacturing, and quality) rules, resources, plans, and results are aggregated and links are established. In this use case, only statistical algorithms are considered to simplify the research task, but semantic, syntactic, linguistic, as well as statistics methods must be utilized to automate the ECR decision-making process fully. Aggregating and linking all the lifecycle data supports the identification of product issues. The cause of the issues must be understood and a remedy identified.

Part 8 (Statistics) of QIF [60], describes a schema for defining cause and remedy (C&R) databases. Using the schema could enable supervised machine-learning to automatically identify causes of product issues and apply a recommended remedy. When a product issue is identified and the system cannot apply a cause and remedy definition, a human would be notified to teach the machine-learning system the cause and recommend remedy for the identified product issue. The automated C&R process could alleviate the human-resources burden on the supply chain and product cycle-time by only bringing the human in the loop when the machine-learning system needs supervision. Once a C&R is assigned to a product issue, the decision-support engines could make a determination on the acceptance of the product.

The requirements-management engines would support decision-support engines in making the acceptance decision. If a product is not accepted, the reason for non-acceptance needs to be understood. Therefore, the automated ECR system would need to analyze the C&R assignment and other product lifecycle information to develop a potential problem report.

If enough problem reports are generated for a particular problem, the problem needs to be dispositioned and correlated to a phase of the product lifecycle (e.g., design, manufacturing, quality assurance). In the use case described in this paper, our work is demonstrating statistical process control on the entire product lifecycle by correlating the problem back to the appropriate lifecycle phase. The information used for the development of the problem reports is captured and fed back to the appropriate knowledge bases to ensure that real-time data is accurate and available to decision makers. Just like the C&R definitions databases, C&R correlation databases are created to assist in the supervised machine-learning. For our study, we are testing the use of design engineering C&R correlations. When a problem is correlated back to a design engineering issue, the problem would be dispositioned and an ECR generated. The final step in the process would be to notify engineering of the ECR.

The National Institute of Standards and Technology (NIST) is developing the Smart Manufacturing Systems Test Bed (SMS Test Bed) [31] to support validating use cases and concepts. The SMS Test Bed is comprised of two components: (1) a computer-aided technologies (CAx) laboratory and (2) a manufacturing laboratory. Both components focus primarily on machining processes, but the goal of the test bed is to generate a systems-level perspective applicable beyond specific manufacturing process domains. The CAx laboratory utilizes various design, manufacturing, inspection, data management, and verification and validation tools. The manufacturing laboratory contains various milling and turning machine tools, inspection equipment, production management systems, and data acquisition tools. The SMS Test Bed collects manufacturing-related data and publishes it to the public via a web service[2]

We validated each layer of the LIFT concept. We plan to continue testing the LIFT concept and use cases to further validate the full integration of the LIFT concept. The next step is we are planning to run single-blind experiments. The idea is to develop a product definition with a known design defect that causes manufacturing and quality deficiencies. The system designers, and therefore the system algorithms, would not have knowledge of the design defect. The product definition would be generated in CAD systems and translated to STEP AP 242 [33] for exchange with the manufacturing and quality functions. We would collect manufacturing and quality data using the MTConnect [34] and QIF [35] standards while producing the product with the known design defect in the manufacturing laboratory of the SMS Test Bed. We would then aggregate the data using a process like the one outlined in Figure 5. If our system is able to automatically detect the defect and deficiencies and correlate them back to engineering, then we have shown it is feasible to automate engineering change request generation and supported our use case.

7 Potential Impacts

The Office of Science and Technology Policy (OSTP) directed Federal agencies with research and development expenditures to develop a plan to support public access to data and results of research funded by the Federal Government [61]. The purpose of the directive was to help Federal agencies comply with the America COMPETES Reauthorization Act of 2010 [62]. In addition, various science and technology communities [8, 9, 10, 63] provide guidance to managing and publishing domain-specific data.

However, the engineering and manufacturing domain may not share the same view of data sharing as the physical sciences domain. Due to intellectual property (IP) and data-security concerns, most manufacturing data remains behind lock and key – much less discoverable. The creation and rollout of the National Network of Manufacturing Innovation (NNMI) would help change the current thinking for the management of manufacturing data. The NNMI aims to support the

[2]The SMS Test Bed web service is accessible at `http://smstestbed.nist.gov`.

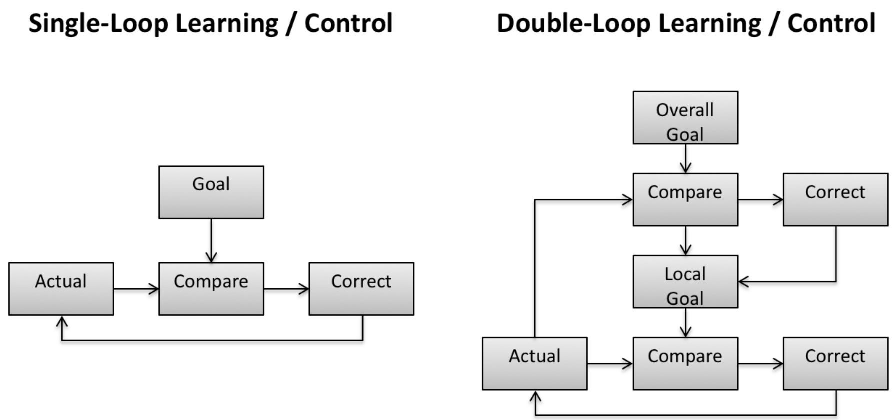

Fig. 6. Single-loop learning versus double-loop learning – derived from [68]

development of a United States infrastructure for manufacturing research [64] and may be subject to the OSTP mandate.

For the NNMI to be successful, a data management plan acceptable to all parties and consistent across the network must be available. The LIFT concept provides a common data management plan that supports data discovery activities and manages the authentication and authorization control around the data. The reference infrastructure from the LIFT concept and demonstrated in the NIST SMS Test Bed [31] provides deployable data management guidance to the various institutes that make up the NNMI. By deploying a common infrastructure across the NNMI, the manufacturing data generated in the network may be discoverable across the network. Making the data in the NNMI discoverable and/or available via a common infrastructure benefits the public, industry, and academia by enabling researchers to understand and exploit existing manufacturing data for innovation and discovery – in much the same way witnessed in the forecasting industry with open weather data or the biotechnology sector with public access to genome sequences [61]. The exploitation of manufacturing data would help accelerate industry into the the third industrial revolution [65] where data becomes the commodity through the digitization of manufacturing.

The LIFT concept also provides a reference architecture for a manufacturing application of semantic web technologies – further driving the commoditization of data. This enables industry to build a grid of product data across the product lifecycle to support real-time decision support and knowledge building. From organizational learning theory, people have mental maps with regard to how to act in situations [66]. The mental maps people form are driven by people's experiences in previous situations and involves the way people plan, implement, and review their actions [66]. The maps guide people's actions rather than the theories the people espouse explicitly causing a split between theory and actions [67]. In control theory, the mental map is a single-loop control system.

Figure 6 shows a single-loop control system and a double-loop control system. Single-loop control is the repeated attempt at same problem with no variation of method and without ever questioning the goal [68]. Whereas, double-loop control is after attempting to achieve a goal on different occasions, the goal is modified in light of experience or – in the worst case – the goal is rejected all together [68]. Argyris [67] argues double-loop control can be applied to organizations to break the pattern of, colloquially, "this is the way we've always done it." Argyris [67] calls that "double-loop learning."

The LIFT concept has the potential to provide the information needed to reduce the double-loop learning theory to practice and would provide effectively a complete perspective of the product-lifecycle impact of a decision. The LIFT concept would put information into the hands of the roles and functions that have a need to know for decision support and knowledge-management purposes. The LIFT concept would support the cost and quality influencers in the product lifecycle by educating them through double-loop learning on how to make designs better and more producible, which may lead to removing the disconnect between theory and action.

A more far-reaching impact of the LIFT concept is the ability to deploy prognosis, health monitoring, and control methods to all aspects of the product lifecycle (e.g., the product, operations, activities) and not just to the product as prognosis and health monitoring (PHM) is applied historically by industry. The use case described in Section 6 only scratches the surface of capabilities enabled by the LIFT concept. The Smart Manufacturing Operations Planning and Control (SMOPAC) program [69] at the NIST is investigating how the LIFT concept applies to all aspects of the product lifecycle from the perspective of the manufacturing domain. The SMOPAC program is investigating how feed-back (i.e., diagnostic) and feed-forward (i.e., prognostic) information can be applied to the design operations, machine tool, manufacturing operations, quality-inspection tool, quality-assurance operations, and product-support operations activities in the product lifecycle by ingesting manufacturing data and information using the LIFT concept. The goal is to accelerate the benefits and impacts for the public, industry, and academia by maturing the LIFT concept through reference infrastructure to demonstrate linking of all the data in the product lifecycle similar to the method describing the links between the process and product data layers in Figure 4 from Section 6.

8 Conclusion

The framework and technology together make up the LIFT concept. This paper provided an overview of the emerging information framework and needed technology infrastructure to support semantic product lifecycle management, showed how the framework and technology relates to a real engineering problem and process, and suggested additional research topics to mature the concept. This paper discussed previous and related research to describe how the LIFT concept is aiming not to reinvent the wheel, but leverage and expand the work from other domains to springboard success within the manufacturing domain. This paper used the engineering change request process to illustrate the potential applications of the dynamic knowledge base, decision-support engine, and requirements-management engine components of the framework. The discussion throughout the paper highlighted areas of research that need expansion and posed questions that need to be answered before the concept can realize its full impact.

A few research questions remain. These questions pertain to calculating uncertainty and variation in the product lifecycle, semantics and domain-specific knowledge, and collaboration and interoperability. Uncertainty quantification is understood at the domain-specific levels, but how do those uncertainties aggregate up into a total uncertainty understanding for the entire product lifecycle? Also, what are the modes of variation in the product lifecycle and where are those models most likely to occur? To support applying control theory to the complete product lifecycle, we must be able to understand how uncertainty and variation across the lifecycle relates. We must know how to predict the uncertainties of lifecycle phases, how those aggregate together, and how to identify where variation could be introduced. From there we must determine how we minimize the uncertainties and variation through decision-support systems and selection mechanisms.

In addition, the LIFT concept can be successful only if semantic links are generated between data, which would support computer-processable knowledge bases. The problem with defining semantics is the definer often requires domain-specific knowledge of the things that the semantics are defining. Therefore, generating links and relationships between domain-specific data and information may be domain-specific itself. The question remains, how would semantic links between data be generated using automated methods?

Lastly, intra-domain (e.g., design to design) and inter-domain (e.g., design to manufacturing) interoperability must be increased. Studies [37, 38, 70] show communicating product data using model-based methods is feasible, but barriers in standards and interoperability remain. For MBE and our concept to be successful, closing the interoperability gaps is paramount.

In closing, the research currently suggests the LIFT concept could manage all the information within the lifecycle by supporting linkages between currently disconnected information. The LIFT concept could also support organizational learning and support the removal of inter-organizational socio-technical barriers. While additional research is required, the outcome of the LIFT concept is a novel solution to the problem that the lifecycle is drowning in data, but starving for information.

Disclaimers

Acknowledgement

We want to acknowledge Mr. Paul Huang (U.S. Office of Naval Research) and Dr. William Regli (Defense Advanced Research Projects Agency) for the helpful insight and feedback they provided. Lastly, we wish to thank Mark Carlisle (NIST), Vijay Srinivasan (NIST), and the journal peer-reviewers for their review and comments.

References

[1] Open Knowledge Foundation. About, 2015. URL `https://okfn.org/about/`.

[2] Open Knowledge Foundation. CKAN, 2015. URL `https://okfn.org/about/our-impact/ckan/`.

[3] CKAN Association. About, 2015. URL `http://ckan.org/about/`.

[4] Open Data Institute. About the ODI, 2015. URL `http://theodi.org/about-us`.

[5] Open Data Institute. Open data certificate, 2015. URL `https://certificates.theodi.org/`.

[6] Telecommunication Standardization Sector of ITU. Information technology – open systems interconnection – the directory – part 8: Public-key and attribute certificate frameworks, 2014. URL http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=64854.

[7] U.S. General Services Administration. Impact, 2015. URL http://www.data.gov/impact/.

[8] National Aeronautics and Space Administration. Physical science information: History and purpose, 2014. URL http://psi.nasa.gov/History.aspx.

[9] Consortium for Ocean Leadership. Final network design. Report, Ocean Observatories Initiative, 2010. URL oceanleadership.org/wp-acontent/uploads/2009/04/1101-00000_FND_OOI_2010-04-22_ver_2-06_public1.pdf.

[10] EarthCube. Earthcube strategic science plan. Report, National Science Foundation, 2015. URL http://earthcube.org/document/earthcube-strategic-science-plan.

[11] Logistics Information Service. Military engineering data asset locator system (MEDALS). Report, Defense Logistics Agency, 2014. URL https://www.dlis.dla.mil/medals/pdfs/MEDALSHomepageOverview.pdf.

[12] Mark R. Cutkosky, Jay M. Tenenbaum, and Jay Glicksman. Madefast: collaborative engineering over the internet. Communications of the ACM, 39(9):78–87, 1996. ISSN 0001-0782. doi: 10.1145/234215.234474. URL http://dl.acm.org/citation.cfm?id=234474.

[13] Angus Stevenson and Maurice Waite. Concise Oxford English dictionary. Oxford University Press, Oxford ; New York, 12th edition, 2011. ISBN 9780199601080.

[14] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. Scientific American, May, 2001. URL http://www.scientificamerican.com/article/the-semantic-web/.

[15] Jacco van Ossenbruggen, Lynda Hardman, and Lloyd Rutledge. Hypermedia and the semantic web: A research agenda. Journal of Digital Information, 3(1), 2002. ISSN 1368-7506. URL https://journals.tdl.org/jodi/index.php/jodi/article/view/78/77.

[16] Tim Berners-Lee. Linked data, 2009 2006. URL http://www.w3.org/DesignIssues/LinkedData.html.

[17] World Wide Web Consortium. Semantic web, 2006. URL https://www.w3.org/standards/semanticweb/.

[18] World Wide Web Consortium. Resource description framework (RDF), 2014. URL http://www.w3.org/RDF/.

[19] World Wide Web Consortium. SPARQL 1.1 overview, 2013. URL http://www.w3.org/TR/sparql11-overview/.

[20] Nitesh Khilwani, Jennifer A. Harding, and Alok K. Choudhary. Semantic web in manufacturing. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 223(7):905–925, 2009. doi: 10.1243/09544054JEM1399. URL https://dspace.lboro.ac.uk/2134/9519.

[21] Rui P. Fernandes, Ian R. Grosse, Sundar Krishnamurty, Paul Witherell, and Jack C. Wileden. Semantic methods supporting engineering design innovation. Advanced Engineering Informatics, 25(2):185–192, 2011. ISSN 14740346. doi: 10.1016/j.aei.2010.08.001. URL http://dx.doi.org/10.1016/j.aei.2010.08.001.

[22] Farhad Ameri and Lalit Patil. Digital manufacturing market: A semantic web-based framework for agile supply chain deployment. Journal of Intelligent Manufacturing, 23(5):1817–1832, 2012. ISSN 09565515. doi: 10.1007/s10845-010-0495-z. URL `http://dx.doi.org/10.1007/s10845-010-0495-z`.

[23] Farhad Ameri and Christian McArthur. Semantic rule modelling for intelligent supplier discovery. International Journal of Computer Integrated Manufacturing, 27(6):570–590, 2014. ISSN 0951192X. doi: 10.1080/0951192X.2013.834467. URL `http://dx.doi.org/10.1080/0951192X.2013.834467`.

[24] World Wide Web Consortium. Inference, 2008. URL `http://www.w3.org/standards/semanticweb/inference`.

[25] Paul Witherell, Sundar Krishnamurty, Ian R. Grosse, and Jack C. Wileden. Improved knowledge management through first-order logic in engineering design ontologies. Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM, 24(2):245–257, 2010. ISSN 08900604. doi: 10.1017/S0890060409990096. URL `http://dx.doi.org/10.1017/S0890060409990096`.

[26] Dazhong Wu, J. Lane Thames, David W. Rosen, and Dirk Schaefer. Enhancing the product realization process with cloud-based design and manufacturing systems. Journal of Computing and Information Science in Engineering, 13(4), 2013. ISSN 15309827. doi: 10.1115/1.4025257. URL `http://dx.doi.org/10.1115/1.4025257`.

[27] Dazhong Wu, David W. Rosen, Lihui Wang, and Dirk Schaefer. Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation. Computer-Aided Design, 59(0):1–14, 2015. ISSN 0010-4485. doi: 10.1016/j.cad.2014.07.006. URL `http://www.sciencedirect.com/science/article/pii/S0010448514001560`.

[28] Xun Xu. From cloud computing to cloud manufacturing. Robotics and Computer-Integrated Manufacturing, 28(1): 75–86, 2012. ISSN 0736-5845. doi: 10.1016/j.rcim.2011.07.002. URL `http://www.sciencedirect.com/science/article/pii/S0736584511000949`.

[29] Energetics Inc. Measurement science roadmap for prognostics and health management for smart manufacturing system. Report, National Institute of Standards and Technology, 2015.

[30] R. Gao, L. Wang, R. Teti, D. Dornfeld, S. Kumara, M. Mori, and M. Helu. Cloud-enabled prognosis for manufacturing. CIRP Annals - Manufacturing Technology, 64(2):749–772, 2015. ISSN 0007-8506. doi: 10.1016/j.cirp.2015.05.011. URL `http://www.sciencedirect.com/science/article/pii/S000785061500150X`.

[31] Moneer Helu and Thomas Hedberg Jr. Enabling smart manufacturing research and development using a product lifecycle test bed. Procedia Manufacturing, 1:86–97, 2015. ISSN 2351-9789. doi: http://dx.doi.org/10.1016/j.promfg.2015.09.066.

[32] Min Li, Shuming Gao, and Charlie C. Wang. Real-time collaborative design with heterogeneous CAD systems based on neutral modeling commands. Journal of Computing and Information Science in Engineering, 7(2):113–125, 2006. ISSN 1530-9827. doi: 10.1115/1.2720880. URL `http://dx.doi.org/10.1115/1.2720880`.

[33] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 242: Application protocol: Managed model-based 3D engineering, 2014.

[34] MTConnect Institute. Mtconnect standard, 2014. URL http://www.mtconnect.org/media/39542/mtc_part_1_overview_v1.3.pdf.

[35] Dimensional Metrology Standards Consortium. Part 1: Overview and fundamental principles in quality information framework (QIF) – an integrated model for manufacturing quality information, 2014. URL http://www.qifstandards.org.

[36] Allison Barnard Feeney, Simon P. Frechette, and Vijay Srinivasan. A portrait of an ISO STEP tolerancing standard as an enabler of smart manufacturing systems. Journal of Computing and Information Science in Engineering, 15(2):021001–021001, 2015. ISSN 1530-9827. doi: 10.1115/1.4029050. URL http://dx.doi.org/10.1115/1.4029050.

[37] Asa Trainer, Thomas Hedberg Jr, Allison Barnard Feeney, Kevin Fischer, and Phil Rosche. Gaps analysis of integrating product design, manufacturing, and quality data in the supply chain using model-based definition, 2016. 2016 ASME Manufacturing Science and Engineering Conference.

[38] Thomas D. Hedberg Jr, Joshua Lubell, Lyle Fischer, Larry Maggiano, and Allison Barnard Feeney. Testing the digital thread in support of model-based manufacturing and inspection. Journal of Computing and Information Science in Engineering, 2016. ISSN 1530-9827. doi: 10.1115/1.4032697. URL http://dx.doi.org/10.1115/1.4032697.

[39] William C. Regli, Joseph B. Kopena, and Michael Grauer. On the long-term retention of geometry-centric digital engineering artifacts. Computer-Aided Design, 43(7):820–837, 2011. ISSN 0010-4485. doi: http://dx.doi.org/10.1016/j.cad.2010.11.012. URL http://www.sciencedirect.com/science/article/pii/S0010448510002228.

[40] International Standards Organization. Industrial automation systems and integration – JT file format specification for 3D visualization, 2012.

[41] International Standards Organization. Document management – portable document format – part 1: PDF 1.7, 2008.

[42] International Standards Organization. Document management – 3D use of product representation compact (PRC) format – part 1: PRC 10001, 2014.

[43] Spatial Corporation. ACIS R25, July 2014. Proprietary Specification.

[44] Marshall Burns. The StL Format, pages xxv, 369 p. PTR Prentice Hall, Englewood Cliffs, N.J., 1993. ISBN 0131194623.

[45] Sudha Ram and Jun Liu. Understanding the Semantics of Data Provenance to Support Active Conceptual Modeling, volume 4512 of Lecture Notes in Computer Science, book section 3, pages 17–29. Springer Berlin Heidelberg, 2007. ISBN 978-3-540-77502-7. doi: 10.1007/978-3-540-77503-4_3. URL http://dx.doi.org/10.1007/978-3-540-77503-4_3.

[46] R. Bidarra and W. F. Bronsvoort. Semantic feature modeling. Computer-Aided Design, 32(3):201–225, 2000. ISSN 0010-4485. doi: http://dx.doi.org/10.1016/S0010-4485(99)00090-1. URL http://www.sciencedirect.com/science/article/pii/S0010448599000901.

[47] Dassault Systemes. Achieving better, faster, smarter innovation. Waltham MA, 2012. URL

http://www.3ds.com/fileadmin/PRODUCTS-SERVICES/ENOVIA/Resources-center/Achieving-Better-Smarter-Innovation-WP.pdf.

[48] PTC Inc. Windchill (R): Managing the complete product lifecycle – from concept to service. Needham MA, 2014. URL http://www.ptc.com/product/windchill/new.

[49] Siemens PLM. Teamcenter: Simplifying plm. Plano TX, 2014. URL http://www.plm.automation.siemens.com/en_us/products/teamcenter/#lightview%26uri=tcm:1023-79817%26title=Teamcenter-Brochure-4680%26docType=.pdf.

[50] Gregory Tassey, Smita B Brunnermeier, and Sheila A Martin. Interoperability cost analysis of the us automotive supply chain. Report 7007-03, Research Triangle Institute, 1999.

[51] Gregory Y. Titelman. Random House Dictionary of Popular Proverbs and Sayings. Random House, New York, 1996.

[52] Peter A. Bilello. The next step in PLM's evolution: Platformization. Product Design and Develoment, 2014-12-16 2014. URL http://www.pddnet.com/article/2014/12/next-step-plms-evolution-platformization.

[53] Robert Kahn and Robert Wilensky. A framework for distributed digital object services. International Journal on Digital Libraries, 6(2):115–123, 2006. ISSN 1432-5012. doi: 10.1007/s00799-005-0128-x.

[54] Sean Reilly and Robert Tupelo-Schneck. Digital object repository server: A component of the digital object architecture. D-Lib Magazine, 16(1/2), January / February 2010. doi: doi:10.1045/january2010-reilly.

[55] International DOI Foundation. DOI Handbook. Report, International DOI Foundation, 2012. URL http://www.doi.org/hb.html.

[56] International DOI Foundation. International DOI Foundation Members, June 2014. URL http://www.doi.org/idf-member-list.html.

[57] International Standards Organization. Information and documentation – digital object identifier system, 2012.

[58] Alan H. Bond and Leslie George Gasser. Readings in distributed artificial intelligence. M. Kaufmann, San Mateo, Calif., 1988. ISBN 093461363X.

[59] Chris Richardson. Introduction to microservices, 2015-05-19 2015. URL https://www.nginx.com/blog/introduction-to-microservices/.

[60] Dimensional Metrology Standards Consortium. Part 8: QIF statistics information model and XML schema files in quality information framework (QIF) – an integrated model for manufacturing quality information, 2014.

[61] John P. Holdren. MEMO for increasing access to the results of federally funded scientific research. Report, Executive Office of the President, Office of Science and Technology Policy, 2013. URL http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

[62] US Law. America COMPETES Reauthorization Act of 2010, 2011. URL http://www.gpo.gov/fdsys/pkg/PLAW-111publ358/pdf/PLAW-111publ358.pdf.

[63] Open Data Institute. Implementation plan 2102 and beyond. Report, Open Data Institute, 2012. URL http://theodi.org/about-us.

[64] U.S. Department of Commerce. Three takeaways from national network for manufacturing innovation (NNMI) day, 2014. URL `http://www.commerce.gov/blog/2014/09/22/three-takeaways-national-network-manufacturing-innovation-nnmi-day`.

[65] Economist. Manufacturing: The third industrial revolution. The Economist, page 14, 2012. URL `http://www.economist.com/node/21553017`.

[66] Chris Argyris and Donald A. Schön. Organizational learning. Addison-Wesley OD series. Addison-Wesley Pub. Co., Reading, Mass., 1978.

[67] Chris Argyris. Teaching smart people how to learn. Harvard business review classics series. Harvard Business Press, Boston, Mass., 2008. ISBN 9781422126004.

[68] Donald G. Reinertsen. Managing the design factory : a product developer's toolkit. Free Press, New York, 1997. ISBN 0684839911.

[69] Allison Barnard Feeney and Brian Weiss. Smart manufacturing operations planning and control program, 2014. URL `http://www.nist.gov/el/msid/syseng/smopc.cfm`.

[70] Kevin Ficher, Phil Rosche, and Asa Trainer. Investigating the impact of standards-based interoperability for design to manufacturing and quality in the supply chain. Report, National Institute of Standard and Technology, 2015. URL `http://www.nist.gov/manuscript-publication-search.cfm?pub_id=920033`. NIST GCR 15-1009.

List of Figures