STANDARDIZATION: JOURNAL OF RESEARCH AND INNOVATION

JSRI 2022;1(5):1–10

WORLD STANDARDS DAY WINNING PAPER

The Next Generation of Industry Standards: A Proposed Solution for Digital Transformation

By Andrew Bank; Rupert Hopkins; Bob Solomon

(NOTE: This is the winning paper in the 2022 World Standards Day Paper Competition.)

Abstract

Organizations of all kinds, in nearly every industry, are making huge investments in digital transformation. It would be difficult to pick one industry that is not affected. But there is one area across all these industries that has lagged behind.

In our experience, engineering documents—industry standards included—have been neglected in digital transformation. Most of these documents and the data contained within are still "locked up" in dead-text PDF or Word format, making them difficult to use in engineering workflow, the very process that companies are working to modernize with digital transformation. Most product design data containing geometric dimensions and tolerances (GD&T) were transformed into "digital model" artifacts (like CAD files) starting in the 1990s. Around the same time, engineering documents went from paper to PDF and have remained there ever since. PDF was never meant to be an engineering information tool and it produces more problems than it solves.

Manufacturing companies are hit the hardest by this challenge. For American manufacturers to continue thriving and to remain competitive at home and abroad, the American public-private partnership of standards developers must address these challenges with a new solution. Taking swift and thoughtful action to develop a "SpecX" system – the next generation of industry standards – will help organizations gain the full benefit from their digital transformation investments, contribute to the continued leadership of the American standards system, and ensure interoperability with commercial end-users and international standards developers.

THE DIGITAL TRANSFORMATION BANDWAGON IS MISSING A PASSENGER

Digital transformation (DX) is underway in a big way. "The process by which companies embed new technologies across their businesses to drive fundamental change"[1] has permeated nearly every industry, and all signs point to acceleration. Organizations of all sizes are making long-term DX investments to create—or modify existing—business processes, culture, and customer experiences to adapt to evolving needs of customers and markets. With the advent of industry 4.0, the manufacturing industry has witnessed a spike in demand for automation technologies including machinery,

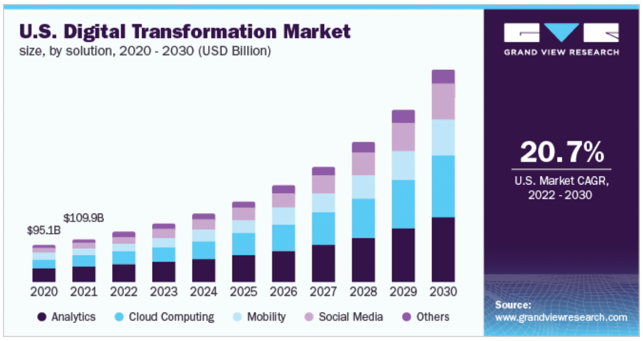

Figure 1. *The US digital transformation market is expected to grow at a compound annual growth rate (CAGR) of more than 20% from 2022 to 2030.*

1

robots, software, controls, and IoT. Global investment in DX is forecast to increase from $600 billion in 2022 to nearly 1.6 trillion in 2027 (Figure 1).[2]

The U.S. public-private sector partnerships that develop standards have played an important role in this transformation effort by identifying areas of innovation and providing consensus-based market-driven standards. ASTM developed important standards for additive manufacturing.[3] ICC now provides an open API into its building code content.[4] IEEE is developing standards on ethical design of AI.[5] ASME recently published a report on the potential benefits of big data in oil and gas workflows.[6] These standards and related products will undoubtedly help U.S. companies keep pace with technology and maintain a competitive edge.

However, the method by which standards are written, published, and delivered to end-users—plain-text documents delivered as PDF files—has not kept pace with technology and has already begun to cause more problems than it intended to solve. One of the biggest obstacles to efficient and productive DX is having to deal with analog artifacts in digital engineering workflow. This challenge represents a giant industry opportunity hiding in plain sight: developing a solution can be a powerful force multiplier for the very standards and drivers of DX itself.

For companies to fully leverage the benefits of DX, they must be able to consume and integrate the standards and specifications into their modern engineering workflow. This does not mean developing standards for integration or engineering; it means writing and delivering standards information in such a way that the individual data elements—text, tables, equations, graphs, images, and more—are described and tagged with their meaning, their place in a standardized ontology, and their contextual relationship to other pieces of data. It means that the information contained in standards must be interoperable with common enterprise applications (e.g., PLM, MES, ERP), and maintained in a master repository that permits access to the data with proper licensing and authentication. Furthermore, DX depends to a large extent on automating as much as possible; therefore, standards of the future must be written for humans but also for machines.

This sounds complex and perhaps a bit like fairy dust, but there are powerful reasons behind these requirements and powerful forces driving them. This paper will recall the events that led us to this moment, describe some of the difficulties in using static PDF files in digital engineering workflow, and then propose a set of principles to govern the future development, publishing, and delivery of "SpecX standards," the standards of the future.

IT'S ALL ABOUT CHANGE—AND WHAT HAS NOT CHANGED

The *MIT Sloan Management Review* says, "Digital transformation is better thought of as continual adaptation to a constantly changing environment."[7] Today's process of DX takes transformation to a whole new level primarily because it's all about change, and change is happening more rapidly now than ever before.

Applying new technologies to business with the goal of improving products, culture, and customer experience is not new. In the 1990s, there was a paradigm shift in which engineering and industrial companies migrated their paper documents, drawings,

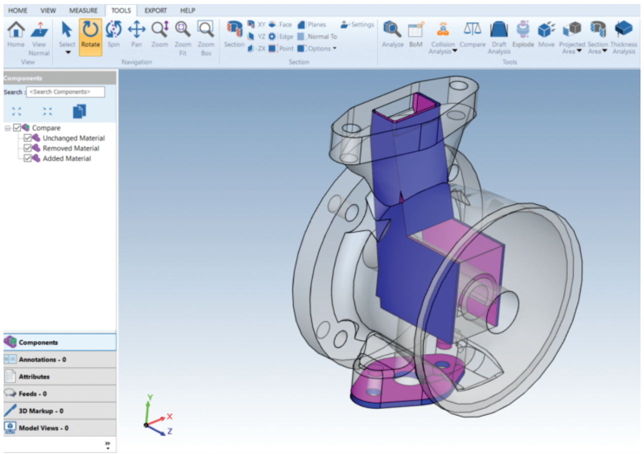

Figure 2. *In the 1990s, physical parts moved from paper to virtual 3D CAD models. Engineering documents never*
2 *made the same transition.*

and blueprints to electronic format. With the growing adoption of model-based engineering (MBE), product designs—those with geometric dimensions and tolerances, or GD&T data—were "freed" from 2D versions and placed directly into 3D models, ostensibly making the 3D CAD (computer aided design) model a single source of design information (See Figure 2).

At the same time, the underlying engineering documentation—non-GD&T information such as industry specs and standards, as well as internal corporate standards—moved from paper to PDF. At the time, it was a revolution that increased productivity and speed, enabled simple sharing, and reduced human error.

Amazingly though, not much has changed with engineering documents or industry standards formats since the mid-1990s. All the text, graphs, tables, equations, images, and references that comprise manufacturing notes, materials and finishing requirements, test methods, work instructions, supplier instructions, customer requirements, and more are often trapped in legacy paper, static PDF, PowerPoint, or Word files. Manufacturing engineers, procurement professionals, estimators, factory-floor workers, testing labs, and others struggle to find the precise information they need, extract the relevant requirements for the task at hand, and feel confident they are using the right versions and references in a sea of disconnected, static documents.

The CHALLENGES OF USING STATIC DATA

These are some of the concrete problems and inefficiencies that technical professionals encounter when using industry standards in static PDF format.

Little or No Interoperability

Most engineering documents contain dozens of references and explicit and implicit connections to other documents, many of them belonging to third-party owners (e.g., an IEEE standard linking to an ASTM standard or a Boeing internal spec linking to an SAE standard), but there is no way for users to easily click between them.

It's also very difficult to move between documents and systems. For example, product life cycle management (PLM) systems (a common desktop application used by manufacturers) often contain hundreds or thousands of references to internal and external standards. But rather than clicking directly from PLM into the relevant document or piece of data, users must go silo-hopping, leaving one system and hunting down the document in another. Not only do they have to find the document itself, but they also must find the exact piece of data inside that document (which could be 5 or 50 or 500 pages long), then somehow export that data into the desired application.

Copy/Paste and Manual Rekeying

Companies use bits and pieces of engineering standards throughout their operation. Some requirements here, an equation there. One very common task is to copy/paste or even manually rekey data from these documents into other documents or applications (e.g., to create work instructions or test plans or supplier specs). Copying/pasting and manually rekeying data is tedious manual labor and heightens the risk of potential human errors that occur when manually rekeying equations or recreating tables from PDF files.

Impact Assessment is Tedious and Risky

Nearly every company that builds a product must deal with changing standards and regulations. As mentioned, parts and pieces of standards are embedded in their operation and indeed in the DNA of their product design and manufacturing. When changes are made to the source material (for example, a new or revised standard), companies have a difficult time monitoring and assessing the impact of that change on their operation and communicating the impacts to stakeholders up and downstream. For companies that build products with very long lives—50 years or more for aircraft, refineries, and other

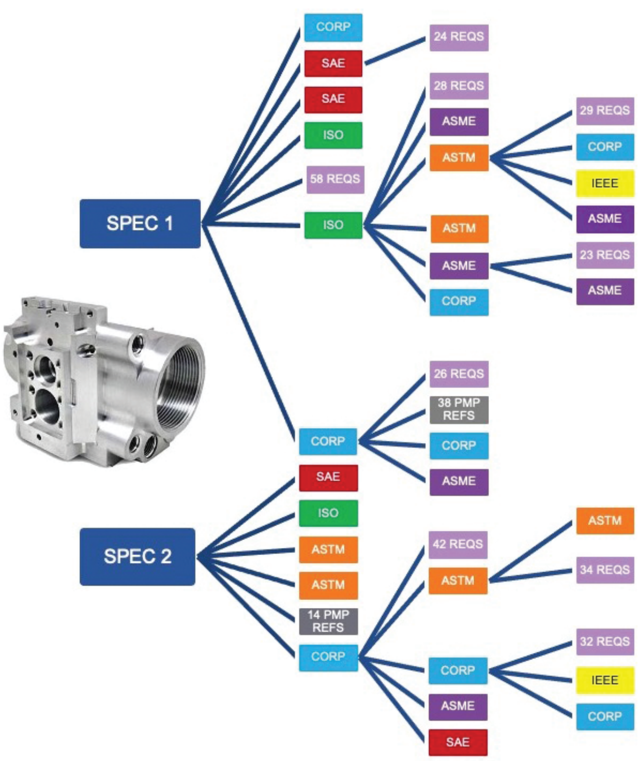

Figure 3. *One small project involving just 2-3 primary standards can quickly cascade into dozens of references with requirements distributed across many different documents.*

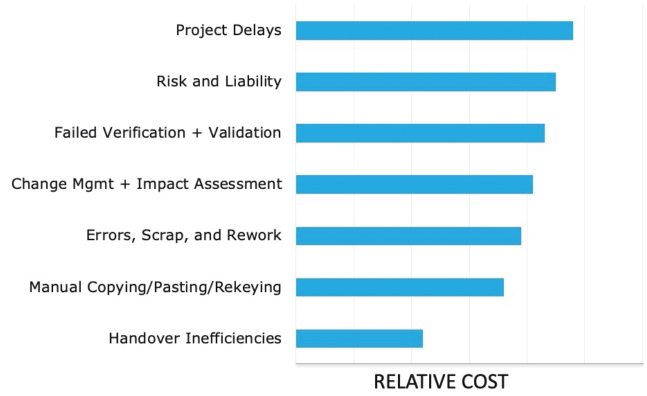

Figure 4. *The relative costs and inefficiencies of using static documents in modern engineering workflow.*

infrastructure—the 2–3-year life of standards poses an even bigger challenge than for companies with very short product shelf lives (like mobile phones or medical devices).

Finding Requirements

Engineers often spend hours, days, or weeks analyzing specs, standards, customer orders and other documents to identify the specific requirements that they need to do the job at hand. Extracting those requirements and classifying them according to the part or material or process is painstaking manual labor. One small project involving just 2-3 primary standards can quickly cascade into dozens of references, with requirements distributed across many different documents (See Figure 3).

THE MODEL-BASED REVOLUTION, AKA "SPECX"

We previously mentioned the 1990s conversion of 2D drawings and GD&T design data into CAD files. This type of digital artifact is called a digital model. A digital model is a data-centric model of a physical object that describes the form, character, and context of that object. A CAD model provides a three-dimensional visual representation of a physical object along with geometric dimensions, textures, materials, tolerances, and more. Likewise, a digital model, or "digital twin," of a building or a part or a system of parts provides all the critical data associated with the object or system.

A PDF file was never meant to be an engineering information tool, let alone a digital model. It is merely a static and non-editable and non-descript copy of the source document. Sometimes it is nothing more than an image of the physical document. It does not describe the form, character, and context of the information contained therein. It does not know when changes at the source occur and how they might affect the content in the PDF file, the associated product(s), or the broader operation of the business. And any data taken from the document and used elsewhere remains as unintelligent, unaware, and disconnected as the source document. ***A PDF is a poor container of data, and a PDF document can only be understood and acted upon by having a human read it*** (See Figure 4).

If the design and manufacturing information chain is defined by its weakest link, engineering documents and industry standards are now often the weakest links. The public-private industry standards partnerships (i.e., SDOs) have done a good job embracing the DX of its members' products and services. We believe it is time now for the industry to embrace the process by which members consume standards information and help them leverage the full benefits of their DX investment. It's time for engineering documents to get the same royal transformation treatment that 2D drawings got in the 1990s and have benefited from ever since.

Static industry standards must now become "SpecX" digital models. A SpecX model is more than just a document; it's a collection of actionable, intelligent, and dynamic data elements that are organized in a database and connected to each other in relevant ways. A SpecX model enables a whole new world of capabilities and value-added benefits that are impossible to achieve with PDF files. We believe this information needs to be freed (though not free of charge) from PDF documents and moved to interoperable, change-aware, machine-readable digital models.

THE EIGHT CORE PRINCIPLES OF SPECX DIGITAL MODELS

To convey more clearly what SpecX is, consider these core principles that we believe should be at the heart of any digital transformation strategy involving engineering documents.

1. Engineering data should be available from a single, authoritative source of truth, and changes should be communicated to all stakeholders. The primary means of communication for this information must move from stand-alone static documents to interoperable and change-aware SpecX models. This information must be held in a common set of digital models managed in a query-capable Semantic Web database or knowledge graph in which each IP owner manages and controls their respective data. As a result, authorized human and machine stakeholders will have the current, complete, authoritative, and consistent information for use over the product life cycle. When changes are made to the source material by IP owners, the changes can be automatically communicated up- and downstream to all relevant stakeholders.

2. Advanced AI technology should be used to enable scalable transformation to SpecX. The transformation of engineering artifacts to digital models cannot scale with manual modeling of the semantic information described by the engineering documents. Expecting different people in different organizations to tag documents in a consistent way is a fool's errand; even expecting organizations to have the resources to tag is a tall order. An ensemble of AI tools—from ontology-based to modern language models based on neural networks—can now be deployed to model syntactic and semantic information from various engineering sources inside and outside an organization.

These tools can act as a human expert to characterize and contextualize the information at scale. For example, a document containing dozens of finishing requirements under various conditions can be summarized and classified according to those conditions, the types of finishes, or other selectable criteria. A user seeking finishing requirements for a specific material and use case can simply click and extract the requirements needed rather than wading through the entire document.

3. Data must be interoperable across IP owners, repositories, and applications. Information users spend too much time "silo-hopping" from disconnected repositories and applications to find the information they need. Nearly every engineering document references many others, often from third parties, but users are left to their own devices to obtain those references and verify their status and currency. This is time-consuming and repetitive manual labor that could easily be eliminated using the linked-data structure of digital models and the automation capabilities of AI.

Much like the Crossref system enables reference linking between journal articles, AI and machine learning can automatically identify and activate reference links from document to document (with proper licensing from the IP owners, of course), and each data element can be stored in a unique location in the semantic web or knowledge graph. Combined with an open API (application programming interface), the result is a consistent method for users to find, cite, link to, assess, and reuse data elements contained within SpecX models. Most importantly, users can easily discover requirements distributed across documents, their dependencies, and the impact of change up and down the requirements' value stream (See Figure 5).

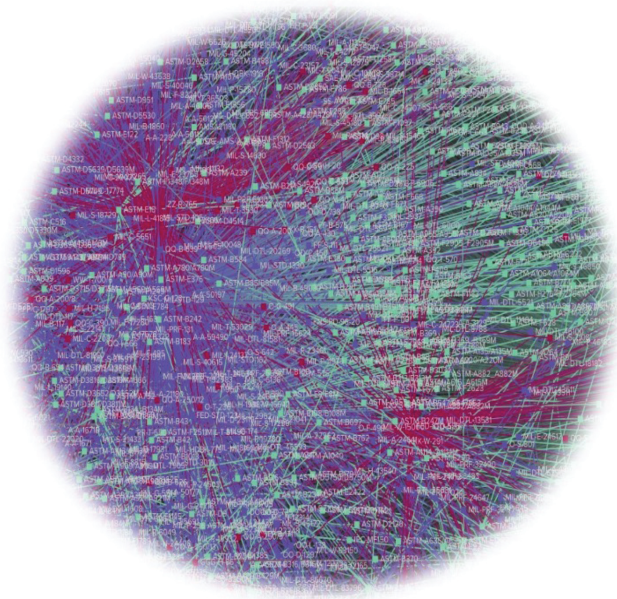

Figure 5. *Data stored in a knowledge graph can activate contextual connections across content sets.*

Note that IP owners of the native engineering documents (whether OEMs, SDOs, or government) will continue to own their data and documents. SpecX models and a knowledge graph simply make their data more usable and more readily accessible to more users from more applications—arguably making their information more valuable. With proper licensing and access and e-commerce tools layered upon the knowledge graph, IP owners can benefit from wider usage and any number of new business models and new net revenue.

4. Engineering information should be delivered to the point of use in existing applications (e.g., PLM and authoring tools) using an open API. Engineers don't need more workbenches or applications. The information engineers need from ASTM, DoD ASSIST, or their own corporate standards can be brought via API into MS Word, PowerPoint, PLM systems, manufacturing execution systems (MES), authoring tools, and more. Developers using a robust, open API with a variety of microservices can quickly develop apps such as shop floor wizards, executable calculations, quality checklists, table look-ups, and PLM plugins.

Delivering content in this way also aligns with the principles of the Digital Thread currently being adopted by the engineering and manufacturing communities. Data used in an engineering workflow must be available at the point of need, but it should always be connected to its authoritative source. Changes made at the source should be communicated up- and downstream throughout the Digital Thread so that the single source of truth is always known to all.

5. Legacy documents cannot be left behind; the creation of SpecX models must accommodate new and old content. Microservices using AI/ML, machine vision and other technologies can be dedicated to analyzing and transforming legacy content. That includes the transformation of MS Word-based specifications or PDF files into digital models, the extraction of key semantic concepts from 2D PDF drawings (including legacy raster drawings), and the transformation of ISO-STS/NISO-STS XML standards, and other XML formats.

6. While internal corporate content must be interoperable with external industry standards, security and IP protection cannot be sacrificed. Cyber threats evolve rapidly. Cyber security standards try to keep pace with the threat, so they evolve continuously. There are no global security standards for model-based standards; each enterprise has its own security standards, and the overlap between enterprises is not complete. NIST 800-171 (rev2), in compliance with DoD requirements, provides a baseline, and the security solutions will evolve toward the government cloud. Corporations must have confidence that storing their proprietary information alongside external industry information does not mean that unauthorized users will have access to their data.

Similarly, specs and standards have been targets of piracy for decades. The paradigm shift from PDF to SpecX must also bring about a paradigm shift in IP protection. Heightened security around copy/paste and document/data sharing should be combined with new licensing models that enable users to quickly and easily obtain licensed and protected content that is aligned with the needs of their enterprise.

7. Engineering content must evolve from documents to models to digital assistants. AI is evolving rapidly. Neural networks and other machine learning tools are improving by orders of magnitude. There is little reason to doubt that engineering content will be more interactive, in the sense that an engineering document (now model) should eventually be the users' digital assistant. A user should be able to ask the document/model a question and get a contextual reply that may be the result of the system crawling through a graph of requirements threaded across multiple authoritative sources. This 'Alexa-like' functionality will enhance the productivity and accuracy of humans and could also enhance the power of VR/AR tools throughout the product life cycle.

8. Engineering information must be consumable by machines. Many years ago, a major aerospace company fed the torque specifications for various fasteners into a software application used by a fleet of bolt-tightening robots. It was a complex set of specs, organized by the type of bolt (material, thread type, thread class, etc.), the substrate, and many other criteria. Because all the source data was stored in a PDF file, it was a tedious job that required hard-coding data, and any updates also had to be updated manually. The result was a semi-autonomous machine that applied the proper torque based on the situation. This was a monumental achievement at the time, but it was only the seed of what was to come.

A growing number of industrial machines rely on data from specifications and standards to run their processes. Today, humans are the bridge—and the bottleneck—between the data and the machines. For example, a machine operator might enter torque specifications into the machine interface, copying the numbers from a MIL Spec or National Aerospace Standard. This manual process consumes too much time, is prone to human error, and is certainly not scalable. For manufacturers to leverage the full benefits of the digital thread and autonomous manufacturing, the information in engineering documents must be readable and interpretable by machines without human intervention.

Documents, from their general structure down to the most granular data elements (equations, tables, etc.), must be quantified in standardized formats that machines can understand and act upon. Further still, the content must be written and tagged in such a way that machines can analyze the full scope and context of information, combine it with data from other sources, and automatically make decisions and take actions based on the machine's interpretation. In the not too distant future, machines, not humans, will be the primary consumers of engineering data.

YOUNG PROFESSIONALS HATE TEDIOUS MANUAL LABOR

A young aerospace engineer told us of their recent experience fresh from their master's degree and right into the QA department of a major aircraft manufacturer–they were baffled and demoralized by the volume of tedious manual labor. At university, they learned about using AI and analytics tools to process information that would aid and accelerate decision making. In their job, they spend time copying/pasting from document to document. In the interviews leading up to the job, they were told about the modern engineering workflow and the company's passion for digital transformation. Although the company was investing heavily in DX, they were still forced to deal with analog wrenches in the digital thread.

If technology companies and the standards industry in general hope to attract young, skilled professionals to replace the exodus of retiring older professionals, we must walk the talk of DX. Initiatives like SpecX will satisfy a new generation of standards users and standardization professionals. Aside from eliminating non-value-added manual labor, SpecX offers dozens of new ways to take advantage of intelligent standards data to improve productivity, increase speed, and reduce errors.

IF WE DON'T, THEY WILL

The move from flat-text engineering documents to machine-readable, data-centric artifacts is already happening elsewhere. CEN/CENELEC, ISO, and IEC are developing their own independent system, called SMART, that will elevate EU and international standards above U.S. and regional standards. In their own words, "SMART will ensure that IEC and ISO products and services remain the most attractive and relevant to markets and societies, today and tomorrow."[8]

SpecX can certainly be compatible with whatever the EU and international community develop. In fact, core principle #3 dictates such interoperability. Creating a strong independent US-based SpecX—just as we created a strong independent SDO ecosystem—will help American businesses accelerate DX, gain the full benefits of standards adoption, and maintain a competitive edge.

Although there is no dominant or fully formed system anywhere in the world yet, there are seeds of similar solutions in the U.S.

NISO has adopted the ISO STS-XML schema, which dictates a standardized set of syntactic elements of documents such as section headers, tables, lists, etc. But more work needs to be done to characterize and tag the semantic elements such as parts, materials, processes, requirements, test methods, and more.

The Defense Standardization Printing Office (DSPO), an agency of the Department of Defense, has started a project called "Documents as Digital Data,"[9] but it is in very early phases.

SWISS (Semantic Web for Interoperable Specs and Standards)[10] is a neutral, AI-enhanced semantic platform that provides interoperable, reusable, and machine-readable digital-model data that they call "digital twins of engineering documents." SWISS pioneered the concept of SMART standards and SpecX digital models several years ago and has an operational commercial platform available today. The platform facilitates the transformation of static documents to digital models and then provides ongoing AI enhancement, data management, and an open API to provide enterprise access to companies and users. SWISS is the world's largest engineering knowledge graph, has partnered with the DoD/DLA, ASTM, ASME, and others, and counts three of the five largest aerospace and defense companies as their customers.

ASME formed a Model-Based Engineering (MBE) Standards Committee[11] that oversees the development of rules, guidance, and examples for the creation, use and reuse of model-based datasets, data models, and related topics within a Model-Based Enterprise (MBE).

CONCLUSION: LET'S BEGIN

For over 100 years, industry standards have been the backbone of commerce and the basis for quality and interoperability,[12] driving borderless trade, technical innovation, and digital transformation. Now, standards need to contribute to the success of its industries, not only by the content inside the documents, but by the way the content is consumed outside the document.

"America's leadership in standards has harnessed private sector speed and flexibility to help companies keep pace with technology, foster economic growth, and save government partners and taxpayers billions."[13] We could not have said it better ourselves!

Continuing American leadership for the next 100 years will require the same speed, flexibility, and innovation that we used to get here. But using the same tactics that got us here will not get us where we need to go. Engineering documents have not changed since the 1990s, and they are long overdue for their own DX.

Gerald Kane, professor of information systems at Boston College, said, "The more I study digital transformation, the more I realize that it's not mostly about either 'digital' or 'transformation.'

"Instead, digital transformation is about how technology changes the conditions under which business is done, in ways that change the expectations of customers, partners, and employees."[14]

Technology—indeed, DX itself—has changed the conditions by which companies operate, and in return, companies now have changed expectations of our industry. We must respond thoughtfully but swiftly to meet their expectations.

NOTES

1. Accenture. 2022. "What is Digital Transformation?" https://www.accenture.com/us-en/insights/digital-transformation-index.
2. Grand View Research. 2021. "Digital Transformation Market Size, Share & Trends Analysis Report By Solution (Analytics, Cloud Computing, Social Media, Mobility), By Service, By Deployment, By Enterprise, By End Use, By Region, And Segment Forecasts, 2022 –2030." Dublin, Ireland.
3. Maxwell, Jack. 2020. "The Next Industrial Revolution." *Standardization News*, May/June.
4. International Code Council. "ICC Code Connect API." https://solutions.iccsafe.org/codeconnect.
5. https://standards.ieee.org/initiatives/artificial-intelligence-systems/standards/
6. ASME. (2020). "STB-1-2020: Guideline on Big Data/Digital Transformation Workflows and Applications for the Oil and Gas Industry." 2020. ASME International.
7. Kane, Gerald C. 2017. "'Digital Transformation' is a Misnomer." *MIT Sloan Management Review*, August, 7.
8. https://www.iso.org/smart
9. Chambers, Jesse. 2022. "Documents as Digital Data." Washington, D.C.: Defense Standardization Program Office, August 3.
10. www.xsb.com/swiss
11. American Society of Mechanical Engineers. "*Model-Based Enterprise*." https://www.asme.org/codes-standards/about-standards/the-future-of-asme-standards/model-based-enterprise-(mbe).
12. 2022 World Standards Day Paper Competition Rules, https://www.ses-standards.org/page/WSD2022
13. 2022 World Standards Day Paper Competition Rules, https://www.ses-standards.org/page/WSD2022
14. Kane, Gerald C. *2017*. "'Digital Transformation' is a Misnomer." *MIT Sloan Management Review*, August, 7.

Andrew Bank leads strategy and business development for the digital twin documents platform, SWISS, created by XSB. He is the co-founder of Techstreet, Larky, and i2B Strategies.

Rupert Hopkins is president and founder of XSB, Inc. He guides the company's vision and business strategy for automated tools that acquire, organize and enrich data using data mining and artificial intelligence. Rupert has served as an adviser to the Defense Office of Economic Adjustment, the Defense Standardization Program Office, the Council on Foreign Relations, the Soros Foundation and others.

Bob Solomon is a board member of XSB and the principal of Software Platform Consulting, Inc. (SPCI). Bob helps software companies and their investors develop multi-sided platforms and networks.

Authors Queries

Journal: **JORI**
Paper: **JORI-01-05-05_1XO**
Title: **The Next Generation of Industry Standards: A Proposed Solution for Digital Transformation**

Dear Author
During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof. Many thanks for your assistance

Query Reference	Query
1	AU: Please provide text callout of all figures.
2	AU: Please provide caption of Figure 2.