

Drowning in data, but starving for insights

Starting the digital supply network journey
with legacy systems

Drowning in data, but starving for insights

Deloitte Consulting LLP's Supply Chain and Manufacturing Operations practice helps companies understand and address opportunities to apply Industry 4.0 technologies in pursuit of their business objectives. Our insights into additive manufacturing, the Internet of Things, and analytics enable us to help organizations reassess their people, processes, and technologies in light of advanced manufacturing practices that are evolving every day.

CONTENTS

Not a leap into the deep end | 2

Start with what you already have | 4

Leveraging existing assets and systems | 6

Step 1: Locate the data among your assets and systems, and organize and prepare it

Step 2: Organize and validate the data for analysis

Step 3: Turn insights into action

Start your journey | 11

Endnotes | 11

Not a leap into the deep end

The digital supply network (DSN) can be a powerful tool for companies, allowing them to harness data and information to make more effective decisions in the physical world via their assets, machines, and people. The benefits can be myriad: deeper visibility into the supply network; greater connectivity with suppliers, partners, and customers; smarter factories; and the ability to act, respond, and adapt intelligently to shifts in the ecosystem. Whatever the result, data—and the ability to analyze and derive insights from it—lies at the heart of the DSN.

COMPANIES may think they need to make significant infrastructure investments to realize these benefits, given the typical cost, complexity, and time to “rip and replace” existing applications. However, the road to a fully realized DSN does not necessarily equate to a wholesale replacement of IT assets. In fact, legacy systems are often able to support more DSN capabilities than previously imagined—and the data they already generate often contains significant potential. If the inefficiencies of the traditional supply chain came from lack of information and poor data, the transformational outcomes of DSN can spring from timely access to and analysis of the *right* data.

While data exists in every organization, it is often not fully organized or understood. Rather, it may be housed in disparate sources, data marts and warehouses, trapped in systems, and in various formats with limited context. The challenge, then, for some companies is they may not know what data they already have, where it lives, what may be useful, or how to turn it into meaningful insights that they can act upon.

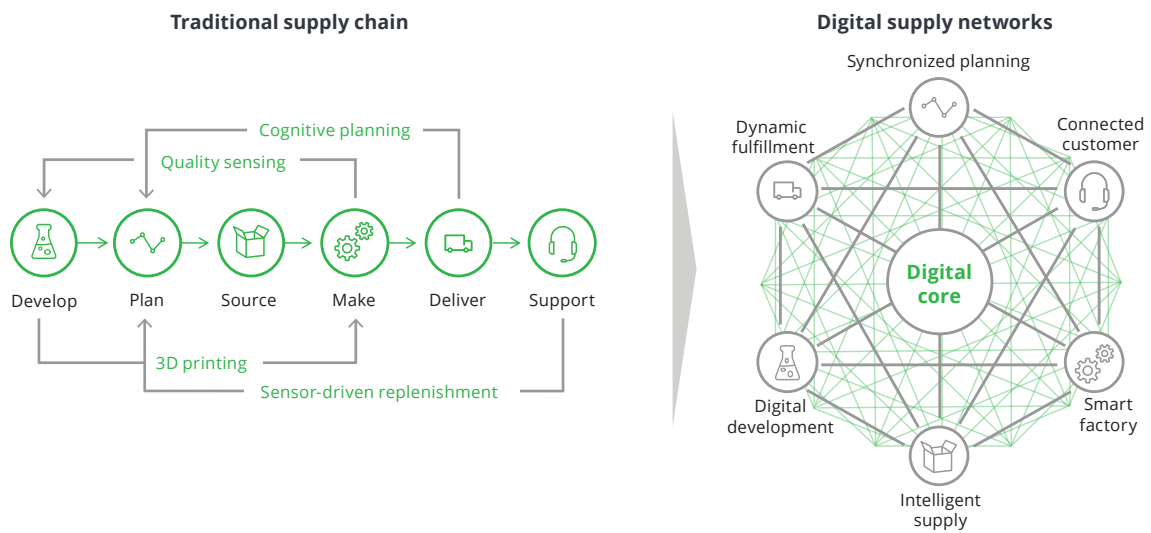
To get started on building a DSN road map with the assets already in place, some companies may need to change their mind-set: Rather than taking a system-centric approach to investing in new physical assets, a data- and insight-centric approach can leverage the underlying legacy technologies—but also allow for application modernization as the company progresses along its DSN path. In this way, organizations can gain new insights today and add intelligence to their operations and processes over time, with less up-front financial investment.

This paper explores how businesses starting their journey to DSNs can leverage assets and systems they already have, by locating, exposing, and analyzing the data those systems are already generating. Managing data correctly generally enables businesses to explore its untapped potential and significantly improve efficiency, generate insights, and even monetize the data—with minimal investment in new technologies and capabilities.

A BRIEF LOOK AT THE DIGITAL SUPPLY NETWORK

In Deloitte’s first publication of this series, *The rise of the digital supply network*, we examined how supply chains traditionally are linear in nature, with a discrete progression of design, plan, source, make, and deliver.¹ Today, however, many supply chains are transforming from a static sequence to a dynamic, interconnected system—a *digital supply network*—that can more readily incorporate ecosystem partners and evolve to a more optimal state over time. DSNs can integrate information from many different sources and locations to drive the physical act of production and distribution.² In figure 1, the interconnected lattice of the new digital supply network model is visible, with digital at the core. There is potential for interactions between each node and every other point of the network, allowing for greater connectivity among areas that previously did not exist. In this model, communications are multidirectional, creating connectivity among traditionally unconnected links in the supply chain.

Figure 1. Shift from traditional supply chain to digital supply network



Source: Deloitte analysis.

Deloitte Insights | deloitte.com/insights

For more information, see [The rise of the digital supply network](#) on Deloitte Insights.

Start with what you already have

PERHAPS by definition, data is the lifeblood of the DSN and the technologies that make it possible. Fortunately, most companies are already awash in the data they need to create a DSN, whether through networked systems on the factory floor or back-office databases. The key is to tap into these data sources to optimize the DSN. This data can take many forms:

- **Master data:** Business-critical data that is consumed by applications to enable business processes. It largely relates to material master, supplier master, customer identities, product material specifications, etc.
- **Transactional data:** Post-business-process information such as purchasing inventory records or sales volumes by region.
- **Sensor data:** Unstructured data that characterizes the conditions of the enterprise's physical assets, from voltage to vibration.
- **Other unstructured data:** Data existing within the organization such as spreadsheets, emails, engineering schematics, drawings, and beyond.

What may be a challenge, however, is that while companies may have all the data they need to become a truly digital enterprise, some may not be able to access such data particularly well—much less understand how to turn the data into actionable insights. Much of the data could be unstructured and stored in formats not easily amenable to automation and insightful analysis. This can be further complicated by organizations continuously

adding new unstructured data sources—and new data points, especially with respect to sensors that exist throughout the DSN.³

As manufacturing processes become ever more complex and data rich, the data sources may become more dispersed, duplicative, and siloed—and ultimately without a coherent strategic purpose.⁴ This can lead to uncertainty about what information is stored and where. The obvious solution would be to pull all this disconnected data together so that it can be seen and analyzed in one place. However, these “data lakes” can be difficult to locate, let alone pull together;⁵ moreover, simply pooling all the data together is not always the best solution.

Simply pooling all the data together is not always the best solution.

An illustration of these issues may be found in the case of predictive maintenance, where data gathered from connected, smart equipment can predict when and where failures could occur. Manufacturing at capacity and managing downtime tend to be critical metrics for

plants, mills, and factories; enabling insight into possible downtime translates to millions of dollars saved or new revenues generated. In a digital factory, machine sensors that measure a variety of upper and lower control limits for voltage, vibration, humidity, and other conditions are a new source of data; they can be linked with traditional master data sources to deliver insights that can improve plant productivity and profitability. The predictive maintenance process ends with the preventive repair of the troubled asset. This means integrating events and data for a failure event and coordinating with plant maintenance systems, including scheduling,

planning, resourcing, and maintenance, repair, and operations (MRO) inventory, coupled with maintenance bills of material, to not only identify the issue in advance but repair it before unplanned downtime can occur.⁶

While most organizations have recognized the need to adopt key data management capabilities—including data governance, master data management, and data quality, among others—the complexity of managing and integrating data across the enterprise may also require new technologies, de-

scribed in table 1. These advanced tools can help to uncover data, make it useful to businesses, and help ensure that it is available in the right place at the right time to support decision-making in the DSN.

But despite the help that these tools can offer, the question remains: How does an organization that is drowning in data but starving for insights make sense of it all and embark on a journey to build a DSN? That organization should start by clearly understanding the data it already has. We explore how in the next section.

Table 1. Cognitive technology categories and select application examples

Categories	Descriptions	Application examples
Robotic process automation (RPA)	“A combination of artificial intelligence and automation” that’s able to “sense and synthesize vast amounts of information and can automate entire processes or workflows, learning and adapting as it goes”	<ul style="list-style-type: none"> • Process automation and configuration • Graphical user interface (GUI) automation • Advanced decision systems
Cognitive—language technologies	A set of statistical techniques that enable the analysis, understanding, and generation of human languages to facilitate interfacing with machines in written and spoken contexts, that is, to convert human (natural) languages into machine languages and vice versa	<ul style="list-style-type: none"> • Natural language processing and generation • Semantic computing • Speech recognition • Speech synthesis • Sentiment and text analytics
Cognitive—machine learning (ML)	A set of statistical techniques that automate analytical model-building using algorithms that iteratively learn from data without the need for explicit programming	<ul style="list-style-type: none"> • Supervised learning • Unsupervised learning • Deep learning
Cognitive—computer vision	Automatic extraction, analysis, and understanding of useful information from a single image or a sequence of images, thereby modeling, replicating, and, more importantly, exceeding human vision using computer software and hardware	<ul style="list-style-type: none"> • Image recognition • Video analysis • Handwriting recognition • Voice recognition • Optical character recognition

Source: David Schatsky, Craig Muraskin, and Ragu Gurusurthy, *Demystifying artificial intelligence*, Deloitte University Press, November 4, 2014; Tiffany Dovey Fishman, William D. Eggers, and Pankaj Kishnani, *AI-augmented human services*, Deloitte University Press, October 18, 2017; and Deloitte analysis.

Leveraging existing assets and systems

A DSN can be enabled by leveraging existing data sources, aligning them with desired business outcomes, and applying the insights in a scalable way. This journey has three primary steps:

1. Locate the data among your assets and systems, and organize and prepare it
2. Organize and validate the data for analysis
3. Turn insights into action

It is important to note, however, that the journey through these steps is typically not a one-time process. Rather, it can be an ongoing feedback loop, in which organizations continually feed learnings back into the process to get smarter with each cycle, inform how data is collected and analyzed, and gain new insights.

Step 1: Locate the data among your assets and systems, and organize and prepare it

Before the organization begins its journey on the road to digital transformation, it should take stock of the data it already has. In so doing, it should organize the data such that it becomes actionable.

START WITH YOUR DATA SOURCES

Taking inventory of the data existing throughout the organization can uncover previously unknown sources of information, potentially allowing untapped data to be explored, categorized, and leveraged. But not all data is created equal; it is important to focus on the data that is relevant to solving your business challenge. (See the sidebar “Understanding the goal and asking the right questions” to learn more.)

INGEST THE DATA

Just because a data lake can hold immense volumes of data, that doesn’t mean it should. In fact, storing too much data (especially in isolated data lakes) can obscure important signals in the data, much like trying to find a needle within a haystack. But, within the haystack, how can leaders determine what data is relevant? Interactive data preparation (IDP) is an approach that can help to truly understand where to start and what value is hidden in the corralled data lakes.⁷ Using a set of interactive tools, IDP can empower business SMEs to explore, evaluate, and transform the gathered data for downstream use. IDP can also minimize reliance on traditional data processing tools and help organize existing data sets.

It’s important to note that in addition to the data that exists in your current systems, other new data sources, both internal and external, could be explored and leveraged for analysis.

Storing too much data (especially in isolated data lakes) can obscure important signals in the data, much like trying to find a needle within a haystack.

UNDERSTANDING THE GOAL AND ASKING THE RIGHT QUESTIONS

While getting a handle on the data that already exists within an organization is a key prerequisite for starting on the path to a DSN, that step alone is not enough to begin the journey. First, it is important that the organization identifies the reason it is pursuing the DSN journey—what it wants the data to reveal. Data can support any number of different objectives. For example, organizations can leverage sensor data for factory asset intelligence to drive predictive maintenance. Yet that same data, combined with other information—such as voltage, vibration, humidity, and other conditions—can be used to ensure a plant is manufacturing at capacity or to gain a better line of sight into managing downtime, overall equipment effectiveness (OEE), MRO inventory, and productivity, among other insights. Further, organizations can leverage operations excellence skill sets that likely already exist within the supply chain/manufacturing organization, and deploy them against this new volume of data. This “digital lean” approach can put companies on a new productivity curve.

A clear understanding of the desired outcomes can help companies remain focused on the data that they need to make key decisions and avoid becoming bogged down in a morass of irrelevant details. From there, organizations can expand to additional objectives, eventually scaling up to a full DSN.

READY THE DATA FOR USE

Once all the relevant data hiding within an organization is found and brought together, it should be cleaned, formatted, and processed so that it can be put to work. In a world of artificial intelligence and machine learning, very often this process of “data wrangling” is about preparing sample data sets to “train” algorithms to provide the business value that a company seeks.

By cleansing the data, data wrangling can save valuable time of data scientists, increasing their output significantly.⁸ According to cross-industry surveys, many data scientists spend the majority of their time removing outliers, matching and merging, and cleaning up data values.⁹ In Deloitte’s experience, 80 percent of a business analyst’s time is spent looking for data—and only 20 percent of the time in trying to gain insights.¹⁰ Inventorying and wrangling data up front can help to invert this pyramid and allow analysts to spend less time with spreadsheets and more time finding insights.

Step 2: Organize and validate the data for analysis

Data in isolation is essentially useless. If companies are to use their assets in the DSN, their data should be organized, structured, and then analyzed—all while keeping the identified business goals firmly in mind.

ORGANIZE AND ADD STRUCTURE TO THE DATA

For optimal analysis, the data should be classified and categorized along with its metadata. With so many sources and types, context is necessary for data to be properly understood, searched, and analyzed. Modern companies are awash in data coming from Internet of Things sensors. In our predictive maintenance scenario, for example, data is captured from supervisory control and data acquisition, the programmable logic controller, historical plant maintenance, and other sources. These new sources

One key is to add structure to prevent data lakes from becoming data swamps.

often promise a wealth of insight, but absent the ability to make sense of the data to support specific business decisions, they will likely not prove helpful. These data sources should be gathered, organized, and prioritized. Bringing all available data together is a key first step, but it is not generally sufficient by itself. By design, big data—and data lakes—are inherently unstructured. One key is to add structure to prevent data lakes from becoming data *swamps*.

One approach to providing structure is to ensure that the data is classified and organized by the business use case the organization is looking to optimize. The needs of the business use case can provide effective guidance on how data should be organized as well as the urgency for analysis. Use cases that

require immediate action force data to be processed and organized in real time, while less time-sensitive use cases can store and analyze data at rest. For example, a safety-sensitive application may require real-time data processing, while a reporting or invoicing use case can store and pull data from static databases. After all, if goals such as “always on” or “instant access” do not support a specific business decision, they may just mean “bad decisions—faster.” Further, it is important to consider that in many cases, near-real-time analysis may be required but can translate to higher costs and greater complexities.

VALIDATE DATA INTEGRITY

Unqualified data yields limited insights, so getting reliable data is also important, as well as being able to establish data integrity checks for the data’s validity. Early data integrity checks are one way to enable trusted outcomes and thus quicker return on investment.¹¹ During data acquisition and preparation, a set of operations on the “training” data set should ensure the data is “clean” enough to test algorithms efficiently. One way to test and validate data integrity is to use similar technologies that leverage machine learning to better understand the usage and context behind the data sets profiled.¹²

IN REAL LIFE: COMBINING EXISTING AND NEW DATA SOURCES

Access to data does not need to stop at the front door of the organization. Today, numerous free data sets are available on a wide range of topics. In fact, there can be so much data available that focusing on a core business problem can be even more important—and challenging—than with internal data.

For example, if a consumer products company wanted to gain insight into how its product was performing, its options were mostly restricted to customer complaint lines. However, these would only capture a small portion of product users—and only those who felt strongly enough to call in—potentially giving a misleading impression of the product. Using application programming interfaces or sensing services, data can be gathered from social media, giving access to a broader range of customers than just those angry enough to pick up the phone. Pairing this data with natural language processing allows the company to potentially draw nuanced conclusions about how customers use the product, satisfaction levels, and product performance. This data could not only help improve design on future product iterations, but could also help marketing understand which customer segments are buying the product and where. Furthermore, existing data sets from production and quality could also be correlated to understand the root cause of negative customer sentiment.

Another way is to establish rules, standards, and policies for master data across suppliers, customers, materials, MROs, and others, and profile the master data to help ensure it is complete, consistent, and clean.

PRIORITIZE THE DATA

The needs of the business use case can help to organize the collection, processing, and analysis requirements for different types of data. These requirements, in turn, can determine the specific technology needed and how they all could come together in a data management plan. Another possible benefit of such use-case-focused planning is that it can highlight where holes exist in the data portfolio. In addition to the data that exists in current systems, new data sources, both internal and external, can be explored and leveraged for analysis. These new data sources also do not need to be expensive solutions; they can be publicly available data sets such as those on traffic or weather, or they can be from inexpensive “lick and stick” sensors applied to legacy, disconnected assets.¹³ Further, understanding where holes exist in the data—and thus in the ability to meet business goals—can help organizations understand where investments in new assets may,

in fact, be necessary. In this way, the company can make these investments more strategically.

Step 3: Turn insights into action

Once the data has been prioritized and validated, the arguably most valuable part can begin: drawing insights and taking action. But linear spreadsheets and endless columns don’t always translate to a clear path forward. The human workers who must take action on findings do not easily consume rows upon rows of figures.¹⁴ By making data visualizations and insights compelling, teams can deliver data in useful, actionable ways.

In our predictive maintenance scenario, for example, when existing and new data sources are analyzed, organizations can effectively predict plant performance and potential failures. Simply having a flashing red light alerting a prediction of failure isn’t the real value, however. Rather, managers can dispatch teams to be available, on site, with the right spares and inventory, and the right skill sets to fix the equipment in advance of failure. Further, by tying analysis with action, rather than



IN REAL LIFE: WHEN IT ALL COMES TOGETHER

A consumer products company was curious about the potential for DSNs, but hesitant to make a large investment in new technology without understanding the resulting value. To test out some capabilities, the company decided to run a small proof-of-value project.

First, it identified the business goal for the project: determining a link between product quality and operating costs. To achieve this goal, the company located existing but unused machine data on its production assets within the plant, and combined that data with information from other sources, including downtime, quality, and production, along with new data from new sensors—and then bringing all the data together in a central repository.

By combining different types of data, the company was able to expose previously hidden information about the production process. Specifically, it pinpointed key material handling and operator engagement points that impacted the performance of the production line and product quality. Not only could the company remedy these areas immediately to increase performance, but it could also use the data to create predictive models to proactively alert teams when machine or process failures were likely to occur in the future.



completing only reactive work orders, organizations can address operational requirements such as inventory optimization and safety stocks, provide for intercompany trading of spare parts, improve procurement insights, and focus on greater throughput.

IDENTIFY ANY GAPS THAT MAY
NECESSITATE FURTHER INVESTMENT

Creating the connections between data is the beginning of the DSN journey. Creating the same connections in operational technologies would be

the end of that journey—the arrival at new value. At the same time, however, as organizations get a handle on the data they already have and use it to best inform their business goals, they can begin to understand where gaps exist in their knowledge—or where new capabilities may be necessary. In this way, organizations can be smarter about the investments they make in new assets or systems, and make more targeted, strategic acquisitions as they advance further toward DSN transformation.

Start your journey

DSNs can drive more strategic decisions, enable organizations to make fewer trade-offs in their processes, and take a more flexible, adaptive approach to shifts in the marketplace or among stakeholders. As organizations seek to build a functional DSN, the good news is that they do not have to start from scratch with massive investments in wholly new assets or systems before they can begin to see real value. Rather, companies can leverage the data their legacy assets and systems are already producing to gain greater insight, inform

their business goals, and begin their journey toward a connected, integrated network.

At the end of the day, it is not just about the assets, but the insights that can be gained from the data you already have. Indeed, while the journey to the DSN may feel intimidating, the reality is that it may be attainable by everyone. After all, most companies often have more data on hand than they even realize. By tapping into that data—organizing, processing, and making use of it—organizations can get smarter, more strategic, and agile on their road to a fully realized DSN.

ENDNOTES

1. Adam Mussomeli, Doug Gish, and Stephen Laaper, *The rise of the digital supply network*, Deloitte University Press, December 1, 2016.
2. Rick Burke et al., *The smart factory*, Deloitte University Press, August 31, 2017.
3. Mussomeli, Gish, and Laaper, *The rise of the digital supply network*.
4. Edd Wilder-James, "Breaking down data siloes," *Harvard Business Review*, December 5, 2016.
5. Anja Lambrecht and Catherine Tucker, "The 4 mistakes most managers make with analytics," *Harvard Business Review*, July 12, 2016.
6. For further information about predictive maintenance, see Chris Coleman et al., *Making maintenance smarter*, Deloitte University Press, May 9, 2017.
7. Yan Zhao, "On interactive data mining," *Encyclopedia of Data Warehousing and Mining*, 2nd ed. (Hershey, PA: IGI Publishing, 2009).
8. Erhard Rahm and Hong Hai Do, "Data cleaning: Problems and current approaches," *Data Engineering* 23, no. 4 (2000).

9. Sean Kandel et al., "Enterprise data analysis and visualization: An interview study," *IEEE Transactions on Visualization and Computer Graphics* 18, no. 12 (2012).
10. Steve Lohr, "For big-data scientists, 'janitor work' is key hurdle to insights," *New York Times*, August 17, 2014.
11. Nate Hindman, "How Singular delivers blazingly fast app ROI analytics," Singular Resource Center, September 18, 2017.
12. Tobias Cagala, "Improving data quality and closing data gaps with machine learning," Bank for International Settlements, May 2017.
13. Robert Schmid, "Turnkey IoT: Balancing customizability and ease of implementation," *IoT Agenda*, January 9, 2018.
14. Alan R. Tate et al., "The effect of format on parents' understanding of the risks and benefits of clinical research: A comparison between text, tables, and graphics," *Journal of Health Communication* 15, no. 5 (2010).

ABOUT THE AUTHORS

JOHN FERRAIOLI

John Ferraioli is a Consulting managing director in Deloitte's Analytics and Information Management service line, and the Data leader of Deloitte's Digital Supply Network practice. Ferraioli brings over 25 years of business process and enterprise data management experience to enable supply networks, and procurement and inventory optimization; streamline new product introduction; and B2B/C collaboration across multiple industries.

RICK BURKE

Rick Burke is a specialist leader in Deloitte Consulting's Digital Supply Networks practice, helping to guide clients through their supply chain digitization journeys. He has more than 20 years of experience in supply chain management, primarily at the intersection of business, technology, and people.

ACKNOWLEDGEMENTS

The authors would like to thank **Pawan Verma** of Deloitte Consulting LLP, and **Brenna Sniderman, Joe Mariani**, and **Jonathan Holdowsky** of Deloitte Services LP.

CONTACTS

Adam Mussomeli

Digital Supply Networks leader
Principal
Deloitte Consulting LLP
+1 203 905 2646
amussomeli@deloitte.com

Stephen Laaper

Digital Supply Networks leader
Principal
Deloitte Consulting LLP
+1 312 513 7900
slaaper@deloitte.com

John Ferraioli

Managing director
Deloitte Consulting LLP
+1 631 902 4572
jferraioli@deloitte.com


Rick Burke

Specialist leader
Deloitte Consulting LLP
+1 713 494 9559
rburke@deloitte.com

Deloitte.

Insights

Sign up for Deloitte Insights updates at www.deloitte.com/insights.

 Follow @DeloitteInsight

Deloitte Insights contributors

Editorial: Aditi Rao, Nikita Garia, Preetha Devan

Creative: Anoop K R, Sonya Vasilieff

Promotion: Shraddha Sachdev

Artwork: J. F. Podevin

About Deloitte Insights

Deloitte Insights publishes original articles, reports and periodicals that provide insights for businesses, the public sector and NGOs. Our goal is to draw upon research and experience from throughout our professional services organization, and that of coauthors in academia and business, to advance the conversation on a broad spectrum of topics of interest to executives and government leaders.

Deloitte Insights is an imprint of Deloitte Development LLC.

About this publication

This publication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms, or its and their affiliates are, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your finances or your business. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

None of Deloitte Touche Tohmatsu Limited, its member firms, or its and their respective affiliates shall be responsible for any loss whatsoever sustained by any person who relies on this publication.

About Deloitte

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. In the United States, Deloitte refers to one or more of the US member firms of DTTL, their related entities that operate using the "Deloitte" name in the United States and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting. Please see www.deloitte.com/about to learn more about our global network of member firms.

Copyright © 2018 Deloitte Development LLC. All rights reserved.
Member of Deloitte Touche Tohmatsu Limited