

White Paper

The EMEA Datasphere: Rapid Growth and Migration to the Edge

Sponsored by: Seagate

David Reinsel
January 2019

Archana Venkatraman

John F. Gantz

John Rydning

EXECUTIVE SUMMARY

The Global Datasphere, a measure of how much new data is created and replicated each year, will grow by more than five times over the next seven years. The total amount of new data created in 2025 is forecast to increase to 175ZB from 33ZB in 2018.

The major drivers of this growth are largely consistent across the world's various regions but occur at different rates. Entertainment data and video surveillance footage have long been (and continue to be) significant drivers of the Global Datasphere. However, signals from the Internet of Things (IoT) devices, metadata (vital for analytics, contextualization, and artificial intelligence [AI]), and productivity data are showing even faster growth in today's increasingly digitized world.

Nevertheless, amid the similarities across various regions, there are subtle differences. These differences are based on technology adoption and digital transformation across a region's population of consumers and enterprises.

The Europe, the Middle East, and Africa (EMEA) Datasphere is growing slightly slower than the overall Global Datasphere (a 2018-2025 CAGR of 26.1% versus 27.2%, respectively). The EMEA Datasphere will increase from 9.5ZB in 2018 to 48.3ZB in 2025, or from 28.8% to 27.6% of the Global Datasphere, respectively. Nearly a third of the Global Datasphere will be driven by growth of video surveillance, signals from IoT devices, metadata, and entertainment. For example, user-created and user-consumed online video like YouTube is one of the top 5 fastest-growing segments of data creation.

Also contributing to growth of the EMEA Datasphere is the sheer number of users getting online for the first time to begin their own digital journeys that result in the consumption, creation, and sharing of data. In MEA countries alone, only 31% of the population in 2018 is using the internet compared with 86% in Western Europe countries and 53% globally¹. This dynamic places pressure on enterprises and governments to upgrade infrastructures to accommodate the growing base of users.

In the more technologically advanced EMEA countries, the edge is an important intermediary between the core and the endpoints to help facilitate the creation and consumption of online video, as well as real-time, on-the-go decisions. Hence the percentage of data in the EMEA Datasphere emanating from or replicated in the edge will nearly double – from 11% to 21% of the region's total Datasphere – as IoT devices increasingly drive processing and analytics closer to the point of origin of the data itself.

¹ IDC's 4Q18 New Media Market Model

Data is at the heart of this digital world, and we are increasingly becoming an information economy. The value is moving to data so that we can create a new world of smarter products, better customer experiences, and self-learning and always improving digital services. In fact, 43% of EMEA organizations executing on digital transformation initiatives have put data capitalization as the top priority to progress. Data is also the heartbeat of modern user experiences and services built using next-generation technologies such as cognitive, IoT, AI, and machine learning (ML).

This unprecedented data growth combined with the pressures of deriving value from data for digital transformation will create imperatives for IT and business organizations across all regions over the next decade to develop a fitting data storage, management, and capitalization strategy and drive a new level of engagement with consumers using data-informed services and products. Whether it's surveillance growth in the United Kingdom and France, manufacturing in Germany, or growth in Russia's mining industry, data is playing a much more critical role.

METHODOLOGY

IDC has been studying the size and nature of the Global Datasphere – all the data created and replicated in one year – for more than a decade. The data creation numbers are driven by IDC forecasts of installed devices and their data creation or capture capacities across more than 70 categories. The analysis takes into account duty cycles and compression techniques. IDC also sizes the amount of data that is stored. It is driven by IDC's ongoing market analysis of the storage market in more than 80 countries.

IN THIS WHITE PAPER

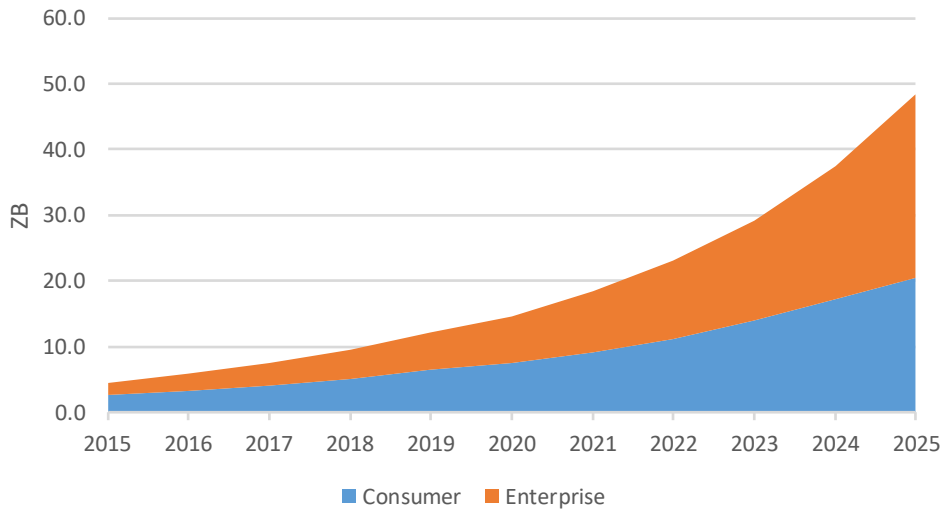
This White Paper is a regional companion document to *The Digitization of the World – from Edge to Core* (IDC #US44413318, October 2018). It summarizes the trends and dynamics of the Datasphere and of data storage as they relate to the Europe, the Middle East, and African countries.

CREATING DATA IN THE DATASPHERE

A decade of growth of the EMEA Datasphere is shown in Figure 1, which also shows the changing share of the Datasphere generated by consumers and enterprises. The enterprise Datasphere is poised to grow from 42% of the EMEA Datasphere in 2015 to 58% in 2025, accounting for more than half of the total EMEA Datasphere and growing at an accelerated rate from 2022.

FIGURE 1

EMEA Datasphere: Originator Segmentation, 2015-2025



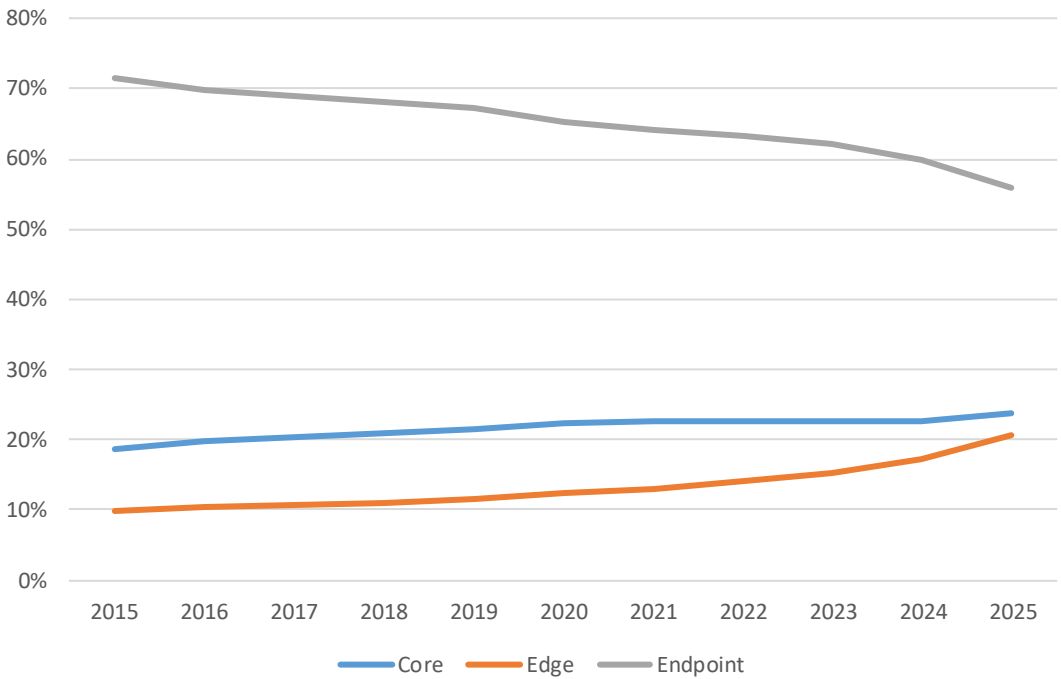
Source: Data Age 2025, sponsored by Seagate, November 2018

Growth of the enterprise share of the Datasphere is driven by several dynamics, including the completion of the migrations of analog to digital TV and movies; the growth of big data and analytics; the migration of consumer data storage from local devices to the cloud; the proliferation of applications, edge devices, and IoT sensors that act as systems of engagements with customers and collect data; and data retention for test and dev, analytics, or compliance purposes.

Figure 2 shows the EMEA Datasphere based on where data originates or is replicated. Basically, while most data will originate at endpoints, more and more will be replicated, transferred, or backed up at the edge for local analytics and retention purposes. Some of that data replicated at the edge will, in turn, be replicated at the core, so enterprises can execute wider enterprise-scale analytics at the core datacenter level.

FIGURE 2

EMEA Datasphere: Location, 2015-2025

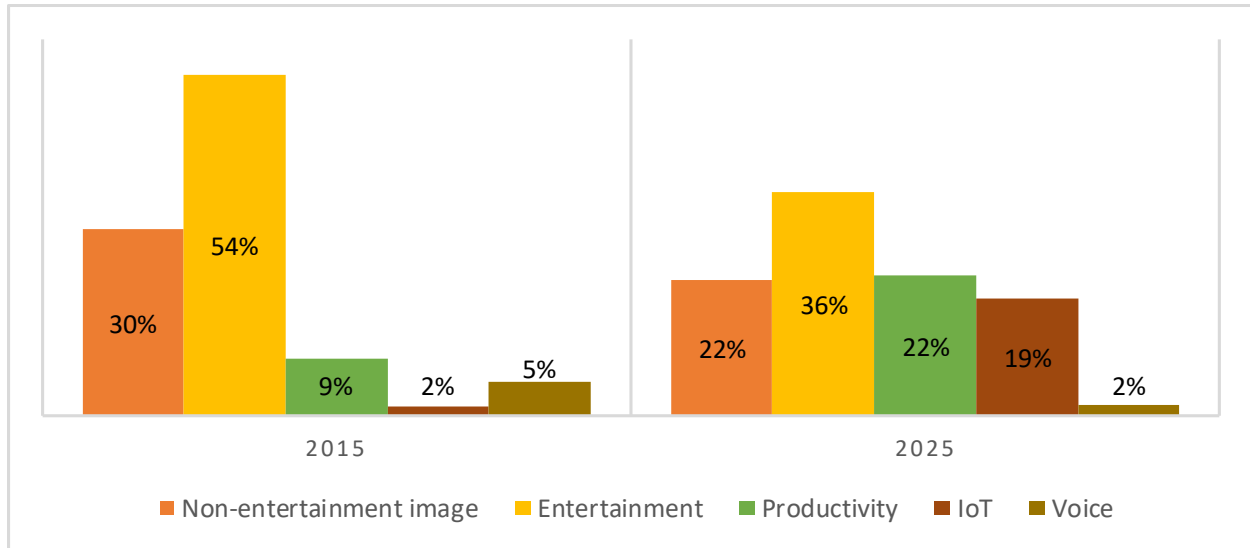


Source: Data Age 2025, sponsored by Seagate, November 2018

The *type* of data in the EMEA Datasphere will experience tangible shifts, as shown in Figure 3. While entertainment data will grow by a factor of 7.1 from 2015 to 2025, it will be outpaced by growth in productivity (big data and metadata) and IoT such that its total share in the EMEA Datasphere will shrink from 54% in 2015 to 36% in 2025. Although entertainment will still be the largest data type, IoT and productivity data types will see exponential growth through 2025 and beyond. Growth of the image sector, which includes surveillance, medical imaging, and entertainment-related content, slows as the conversion from analog to digital in the media and entertainment industry winds down.

FIGURE 3

EMEA Datasphere: Data Type Share, 2015 and 2025



Source: Data Age 2025, sponsored by Seagate, November 2018

There are other types of data that present challenges (and opportunities). For example, mobile data continues to grow at the same pace as the Global Datasphere today (28%); hypercritical data that can affect human lives or property – like telemetry of self-driving cars, real-time medical imaging, and antiterrorism facial recognition – is growing 32% annually; and data touched by artificial intelligence is increasing at a rapid CAGR of 68% for 2015-2025. Use of data in life-critical areas such as automated cars, defense, and healthcare will become viable only if the systems are in place to make sure the data is secure, available, accurate, and contextualized.

One example of growth in productivity data is how Audi UK leverages IoT, APIs, and analytics to deliver dynamic, contextual marketing campaigns. Its system analyzes transport APIs and combines with weather data in real time to deliver contextual advertising to drivers on the road. For instance, when the weather conditions are adverse (rain, sleet, hail, or snow), the digital ads highlight the four-wheel drive in Audi cars; when the traffic on the road is heavy, the ad displays the predictive safety feature in Audi cars. This first data-driven campaign at Audi started in 2017 and the brand continues to combine data and 3rd Platform technologies to create "idea-led communication," which is yielding results.

Another example is how Dutch bank ING has put data management as the cornerstone initiative for its global transformation initiative to become the bank of the future. It is emphasizing on the quality of data and focusing on the accuracy, context, and granularity of data as well as security to become a "data-minded company that enables its customers to manage all their finances anywhere, anyway, and at any time via a single personalized visual dashboard."

Siemens too has developed a robust two-step process: collecting data and making use of the data through meaningful analytics. Its CEO Joe Kaeser believes that customers would "pay for information

that makes life easier, better, less costly, or more valuable." Kaeser also believes that data "changes the relationship massively because data analytics gives a company a lot of information it can use to optimize and shorten the value chain. You can make products faster, more cost efficiently, and more flexibly. You can produce in lot sizes of one."

Storing Data from the Datasphere

Much of the data in the Datasphere will not be stored in any permanent way. In fact, the EMEA installed storage capacity compared with the EMEA Datasphere in 2018 was less than 15%. By 2025, it will be a little more than 8%. This is because most of the data in the Datasphere will evaporate after use – such as digital TV signals not stored on DVRs, multiplayer gaming uploads and downloads, IoT sensor signals that don't trigger alarms, and surveillance images that get recorded over.

The value of data is increasing, and it is imperative for companies to understand the value of the data that they store. Data is the lifeblood of modern life for businesses and individuals – it can support customer-facing activities, operations, research and development, retention of intellectual property, and financial and employee records. Increasingly, data is leveraged in automation, AI, and IoT. Data can also be sold, creating new data-as-a-service opportunities and revenue streams. Hence companies must manage their data properly.

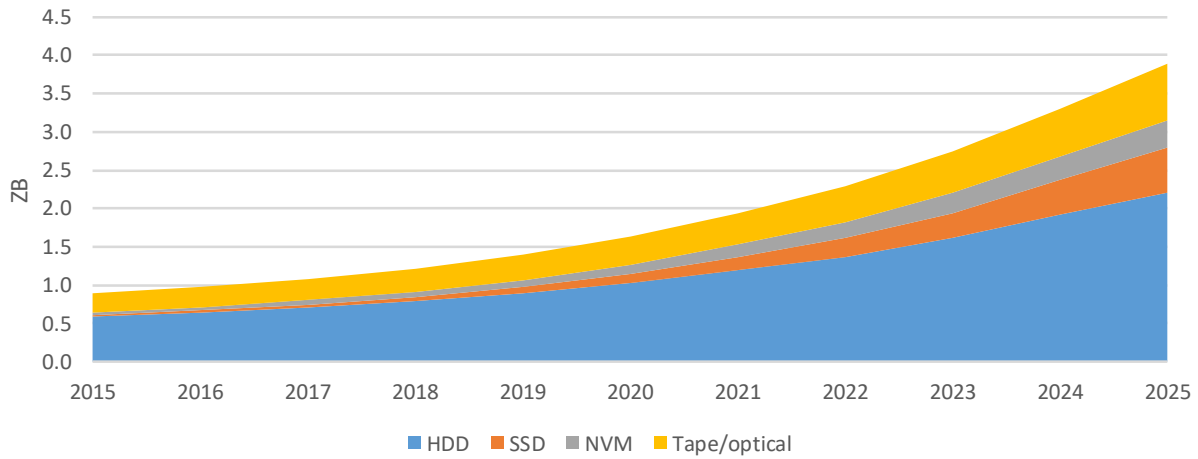
With the value of data in mind, companies, as well as governments and countries, should take note of how much data and storage capacity they have. For example, in 2018, EMEA countries had 24.4% of the world's 5.0ZB installed base of storage capacity². This share of storage capacity decreases to 23.5% in 2025, when the world is expected to have an installed storage capacity of roughly 16.5ZB.

Figure 4 shows growth of installed storage capacity that will be available in the EMEA countries from 2015 to 2025 by storage type. Perhaps surprisingly, tape and optical will continue to have an important role, even as movie viewing and music listening migrate to streaming from DVDs and CDs. Archival storage and backup on tape and optical systems will still be needed years from now. Some organizations may view tape as a higher-maintenance yet safer data storage option to protect from ransomware.

² Installed base of storage capacity is the sum of used and unused bytes across hard disk drives, flash memory, and tape and optical media.

FIGURE 4

EMEA Datasphere: Installed Storage by Media Type, 2015-2025

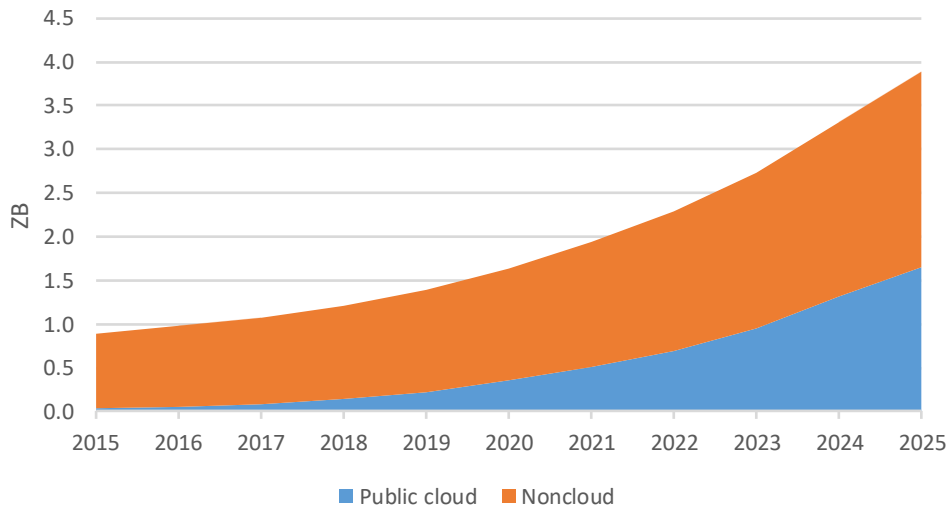


Source: Data Age 2025, sponsored by Seagate, November 2018

There *will* however be a major change in the storage environment as storage migrates from the enterprise and end-user devices to the cloud. Public cloud installed bytes will grow from 4% in 2015 to 42% in 2025 (see Figure 5). In fact, IDC believes that cloud adoption is entering the next phase and enterprises are moving away from an "all in" or public cloud-only adoption model to a much more nuanced and best-fit approach – making multicloud and hybrid cloud infrastructures a reality. In fact, 90% of European enterprises admit to having multicloud strategies by 2019 but only a handful are confident about their multicloud environments being well designed or orchestrated, meaning that they have complete visibility of data to ensure effective analytics or information governance.

FIGURE 5

EMEA Datasphere: Installed Public Cloud Storage, 2015-2025



Note: Noncloud equals all HDD, flash, tape, and optical bytes not residing in a public cloud infrastructure.

Source: Data Age 2025, sponsored by Seagate, November 2018

IDC GUIDANCE

The 11 times growth of the EMEA Datasphere from 2015 to 2025 by itself will create challenges for enterprises – in management, security, storage, and utilization. Businesses across EMEA are in the throes of digital transformation – using new technologies, applications, and innovation accelerators such as IoT, robotics, AI/ML, blockchain, and 3D printing to transform their operations. These innovation accelerators leverage data to deliver valuable insights to improve products and services or to improve processes and operations for efficiency.

The developed economies within the EMEA region must deal with rules of compliance; legacy information technology, processes, and organizations; and global competition. They also need to be cognizant of the data privacy and sovereignty sentiment in the region and work toward building a digital trust and brand reputation while leveraging advancements in analytics to deepen customer engagements. The emerging countries have an opportunity to integrate privacy and compliance technology from the start.

But the challenges (and opportunities) won't come from data growth alone. Consider just three aspects common across all regions:

- **Security:** On a global basis, IDC estimates that in 2018, 56% of the data in the Datasphere required security protection, from simple protection of account information to full "lockdown" protection of bank deposits, critical infrastructure, and user identities. By 2025, that 56% will increase to 66%. And yet the percentage of the data needing protection that *is* protected will barely move from 45% to 50%. At this rate, unprotected data in EMEA that *needs* protection

will grow faster than its own Datasphere between now and 2025 and account for a full one-third of the Datasphere, or 16.1ZB, by 2025.

- **Real-time data:** Again on a global basis, the real-time percentage of the Datasphere will grow from 12% in 2015 to 29% in 2025, driven by growth of IoT. Apply these percentages to the EMEA region and you get more than 25 times growth in real-time data by 2025, or 28% of the world's real-time data. This will not only drive automation to edge computers that aren't already in place but also introduce more interrupt-driven traffic into IT organizations as they begin to inherit responsibility for computing – once the province of operational organizations.
- **Data fragmentation:** As multicloud infrastructure proliferates and as organizations develop a continuum of IT services and applications from edge to core to cloud, the data is highly fragmented across multiple platforms and applications, making it harder to identify, classify, manage, secure, and utilize this data. Organizations need to evaluate the complete data pipeline and develop a data-first strategy to mitigate risks. Issues and regulations around GDPR, data privacy, and the right to forget complicate data management. Protecting privacy is often achieved by secure data silos, which often stifle data management innovation, monetization, and advancements in analytics.

In fact, IDC's 2018 European multicloud and storage research on enterprises revealed that the top 4 priorities for organizations are managing data growth, creating information governance strategies, focusing on analytics, and leveraging cloud for data storage, backup, and DR.

And it won't just be enterprises facing challenges of the Datasphere. By 2025, the percentage of the global population interacting with data will approach 75% – and will surely be higher in EMEA. Gigabytes per day per capita will grow at 21% a year between now and 2025. For example, in 2025, the average connected consumer in the world can expect to have one digital data engagement every 18 seconds, which translates into almost 5,000 interactions per day.

CONCLUSION

Data growth and the growing value attached to data are changing the landscape for EMEA consumers and businesses, and data is shaping how consumers, governments, emergency service providers, and businesses work.

In the public sector, the Dubai government, in its bid to create a Smart City, is launching an open data initiative to share and collaborate around data to create a new data economy for the city.

Businesses are using data to reach new markets, better serve and retain existing customers, innovate, improve efficiency, streamline operations, and even create new revenue streams selling it. Data may not be on a balance sheet, but data is a company's most valuable intangible asset, which can lend a competitive edge in digital transformation. One should just remember Caesars Entertainment's (one of the world's largest casino entertainment companies) bankruptcy: the data it had collected and analyzed became its most valuable asset (worth \$1 billion) for creditors.

Web inventor Sir Tim Berners-Lee says that data is a precious thing and will last longer than the systems themselves.

Companies looking to be relevant between now and 2025 will need to understand the role data plays in their organization and how the Datasphere will evolve during that period. They will need to embrace their role as data guardians, leverage the cloud, and take a global approach to their data. But

organizations need to make data management and innovation priorities to remain competitive in the digital era as digital disruptors across all verticals are making data-driven innovation their key priority.

Take the United Kingdom-based online grocer Ocado that is investing in robotics, IoT, AI, and data science to create automated warehouses. Ocado is now investing in data professionals (data scientists, AI experts, data engineers, etc.) to develop and implement a "vision for how Ocado makes effective use of vast amounts of data produced every minute by our platform: web and mobile analytics, demand trends, optimization algorithms, warehouse and vehicle telemetry, etc." IDC believes that leading-edge organizations that have put data at the heart of their transformation will define the future market dynamics of their sectors as well as the adjacent sectors.

Consumers are building more and deeper connections to data and accessing products and services more easily – and at the time and place of their choosing. They are also benefitting from advances in medical technology, enjoying new forms of entertainment, and living in smarter homes and cities. They are also beginning to expect products and services that are smarter and learn over time and provide accurate, personalized experiences without breaching data privacy sentiments.

The Datasphere is large, dynamic, and complex and increasingly intersecting and driving the physical world – a far cry from the past century where data was something kept in records and files, analyzed over time (if at all), and assisting with running – but not running – factories, automobiles, home appliances, or toll systems. In the digital era, model airplanes have morphed into self-navigating, video-capturing rescue drones; 100,000s of books can be downloaded onto a single digital device smaller than a paperback book; and vacuum cleaners can vacuum a home, navigate around obstacles, and return to their own charging stations all on their own.

Ultimately, in this new age of data, businesses are not just technology or software businesses first, but they are also information businesses first and will need to become data savvy to optimize the opportunities and mitigate risks using technologies and become data thrivers and gain a competitive advantage in the data-driven economy.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com

Copyright Notice

External Publication of IDC Information and Data – Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2019 IDC. Reproduction without written permission is completely forbidden.

