IDC
A N A L Y Z E   T H E   F U T U R E

Becoming Data Driven to Boost Innovation in the Digital Age With Cloudera on Microsoft Azure

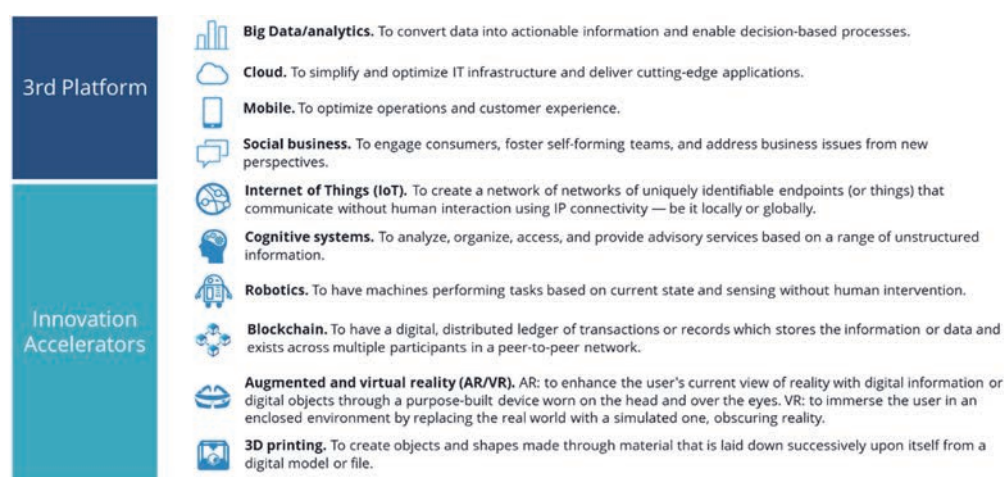## Digital Transformation and Building Blocks of Business Innovation

Digital transformation (DX) has become inevitable for success. IDC estimates that by 2022, over 60% of global GDP will be digitized, with growth in every industry driven by digitally enhanced offerings, and almost $7 trillion in IT-related spending up to that time.

DX is essentially a technology and data-led business strategy. Businesses want to improve customer experiences and operational efficiencies, identify new revenue streams, and build a digitally trusted brand, making data the lifeblood of the business. As a result, data transformation is essential for digital transformation.

The foundational technologies of digital transformation, which IDC calls 3rd Platform Pillars, are cloud, Big Data, mobility, and social business (collaboration). These pillars have helped organizations transform their infrastructures and applications in the past few years.

But now we are entering the next phase of DX where organizations want to use "Innovation Accelerators" such as IoT, cognitive systems (AI/ML/DL), robotics, blockchain, AR&VR, 3D printing, and next-gen security on top of the 3rd Platform to innovate at scale.

Figure 1
Game-Changing Technologies

Big Data/analytics. To convert data into actionable information and enable decision-based processes.

Cloud. To simplify and optimize IT infrastructure and deliver cutting-edge applications.

Mobile. To optimize operations and customer experience.

Social business. To engage consumers, foster self-forming teams, and address business issues from new perspectives.

Internet of Things (IoT). To create a network of networks of uniquely identifiable endpoints (or things) that communicate without human interaction using IP connectivity — be it locally or globally.

Cognitive systems. To analyze, organize, access, and provide advisory services based on a range of unstructured information.

Robotics. To have machines performing tasks based on current state and sensing without human intervention.

Blockchain. To have a digital, distributed ledger of transactions or records which stores the information or data and exists across multiple participants in a peer-to-peer network.

Augmented and virtual reality (AR/VR). AR: to enhance the user's current view of reality with digital information or digital objects through a purpose-built device worn on the head and over the eyes. VR: to immerse the user in an enclosed environment by replacing the real world with a simulated one, obscuring reality.

3D printing. To create objects and shapes made through material that is laid down successively upon itself from a digital model or file.
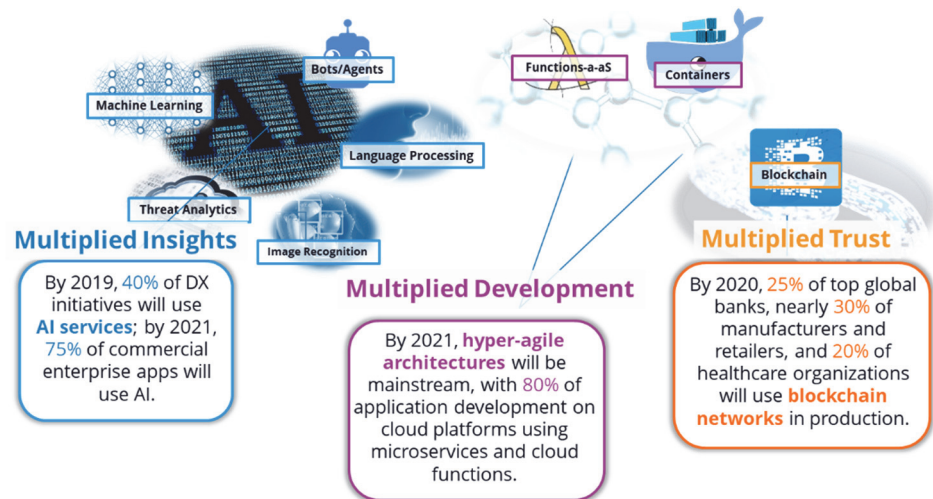
Source: IDC, 2019

A data analytics platform that works across multiple cloud platforms, has the core capabilities for end-to-end data management, security, and analytics, and implements modern data warehousing features can help build the right foundation for a company's data strategy and help it progress in its future IoT, AI/ML, and data science initiatives in a sustainable way.

This second phase of DX is an era of "multiplied innovation" in which hyperscale infrastructure, AI, IoT, containers, functions as a service, and blockchain will be the cornerstone technologies for multiplying innovation and development. This means it is critical to reinvent IT and data strategy around a distributed cloud infrastructure, public cloud software stacks, agile and cloud-native app development, and data sciences.

Figure 2
"Multiplying Innovation" Key Technologies

Source: IDC, 2019

Imperatives to Become Data Driven

Data is evolving from a mere business background to becoming life critical. By 2025, nearly 20% of the data in the global datasphere will be critical to our daily lives and nearly 10% of that data will be hypercritical. In the digital era, it is said that businesses are software and technology businesses first. It is also fair to say that companies are information businesses first too.

IDC has identified some key trends that are intensifying the role of data:

- **IoT and embedded systems.** Connected digital devices will generate vast amounts of data that help refine and improve systems and processes in previously unimagined ways. New data-specific roles will emerge. Big Data and metadata (data about data) will eventually become very important.

- **Hyper-availability of data and mobile and real-time data.** Increasingly, data will need to be instantly available whenever and wherever anyone needs it. Industries around the world are undergoing digital transformation (DX) motivated by these requirements. By 2025, more than a quarter of data created in the global datasphere will be real time in nature, and real-time IoT data will make up more than 95% of this.

- **Artificial intelligence (AI) systems that change the landscape.** The flood of data enables a new set of technologies such as ML, natural language processing, and deep learning — collectively known as AI systems — to turn data analysis into a proactive driver of strategic decision and action.

IDC estimates that the amount of the global datasphere subject to data analysis will grow by a factor of 50 to 5.2ZB in 2025; the amount of analyzed data that is "touched" by AI systems will grow by a factor of 100 to 1.4ZB in 2025.

- **Security as a critical foundation.** Data from new sources opens new vulnerabilities to private and sensitive information. There is a significant gap between the amount of data being produced today that requires security and the amount of data that is actually secured, and this gap will widen — a reality of our data-driven world.

In Western Europe alone, IDC forecasts that the Big Data and analytics (BDA) software market will grow at 9.2% CAGR by 2022. Within this, the business intelligence (BI) and analytic tools segment, as well as the analytic data management and integration platform segment, are expected to grow at 10% CAGR.

The BDA market is seeing robust growth because of data-driven initiatives and the three significant shifts happening in the market:

- The accelerated growth of data management software based on nonrelational databases as well as the trend toward IoT embedded with analytics and ML capabilities to facilitate the processing and action on IoT data for impactful decision making. EMEA's IoT Big Data and analytics software tools market is set to grow by 30.7% CAGR to 2021 as it moves to mainstream.

- The migration to public cloud. Moving data closer to where it is born to reduce overheads is forcing organizations to explore cloud deployments as modern applications are in the cloud.

- The growing demand for automation reflected in the strong appetite for cognitive/AI software. Automation, analytics, and machine learning are promoted as "must haves" by companies because the ability to automate, standardize, and build more intelligence into repetitive and routine tasks can help enhance productivity and free up resources for strategic and high-value tasks.

Monetizing data is becoming critical to create a competitive advantage. Real-time analytics is one of the fastest-growing workloads, according to IDC's European end-user multicloud research in 2018. But for European organizations, this needs to happen in the context of GDPR and data privacy obligations. This is not easy.

IDC's 2017 data value survey on organizations' ability to harness data reveals that a minority of organizations were "data thrivers" (11%), while nearly three-quarters (74%) were "data resisters," "data survivors," or "data responders" — leaving a wide gap between organizations that use data for competitive advantage and those that don't.

Customer Challenges

IDC estimates that the global datasphere will reach 163ZB by 2025. The large volume and variety of data sources, tougher regulatory requirements, stagnant budgets, and data fragmentation across multicloud are proving challenging to turn data into insights. In conversations with IDC, large enterprises admit that vertical scalability is proving unsuccessful, so they are grouping machines into clusters and handling data with distributed processing systems.

There is a need to have a state-of-the-art data repository and streaming platform and break away from multiple siloed traditional data warehouses, so that queries are instantly, more frequently answered, and management is simplified. Beyond this, the role of ML in providing better insight when it comes to operational excellence, business processes, marketing, and security applications is becoming a key next step.

Beyond technology, the skills gap and the limited number of data scientists are cited as another major challenge, as are the limitations in assigning IT budgets for innovation and change.

The early adopters of Big Data analytics are also facing the challenges of traditional analytics strategies from the pre-multicloud era:

1. **Data silos:** data stored across multicloud infrastructures

2. **Data quality:** missing data and unverified data

3. **Shadow IT:** data scientists and Big Data professionals selecting the tools of their choice, leading to new vulnerabilities

4. **Heterogeneity of analytics tools:** multiple single-purpose tools relevant for single infrastructure or tasks requiring integration efforts

5. **IT transformation:** the changing expectations of business users from IT in the era of cloud, and making that cloud transition quickly and with minimum risk

It is in this context that IDC views the partnership between Microsoft Azure and Cloudera Hortonworks as significant in terms of bringing a combination of enabling technologies and business models to help European organizations transition to the public cloud and improve their data competitiveness. The alliance has the potential to help companies leverage the analytics platforms they have already invested in and take them to a public cloud platform and use it as an opportunity to modernize their analytics strategies to transform themselves into data stewards in the digital era.

How to Succeed in Data Capitalization Strategies

Cloud is the launchpad for innovation and accelerated speed to market — a key objective for businesses. With the evolution in cloud and growing value of data, Big Data itself has evolved, from data being a bastion of the IT decision maker who also decided the technology, infrastructure, and systems for the business. But with cloud

> *"Cloudera on Azure is our solution of choice, even for our emerging initiatives such as IoT," says Sven Loeffler, Deutsche Telekom's data-driven initiatives executive*

commoditizing and simplifying infrastructure access, application owners, business users, data scientists, and developers have become key influencers of IT decisions.

First and foremost is to develop a businesswide culture to treat data as a valuable asset and invest in skills and data sciences to leverage the right data with security, privacy, and data protection as core hallmarks.

One key consideration in succeeding in data capitalization and analytics is to empower the data scientists and give them the choice and freedom of tools. These new personas who create insights from business data influence what tools and services a business should invest in. In conversations with IDC, data scientists prefer tools that empower them, give them the flexibility of deployment, and deliver broad capabilities from the edge to the cloud, adding AI capabilities to become future-ready.

Alongside this is creating the technical foundation with an agile, scalable, secure, open, and API-driven cloud infrastructure that integrates with the analytics platforms that customers choose.

Digital Platform With an Intelligent Core

To succeed at innovating at speed and scale, enterprises need to think differently about their architectures. IDC predicts that by the end of next year, 60% of all enterprises will be implementing a new IT foundation.

Figure 3 is a conceptual representation of what and how a digital platform is organized. This platform is software defined, API driven, automated, hyperscale, and data driven. It facilitates the use of external data feeds and combines that with internal data and insights, using technologies such as AI and ML to digitally respond to those feeds. Security, automation, integration, and orchestration wrap this core.

Figure 3
Digital Platform with an Intelligent Core
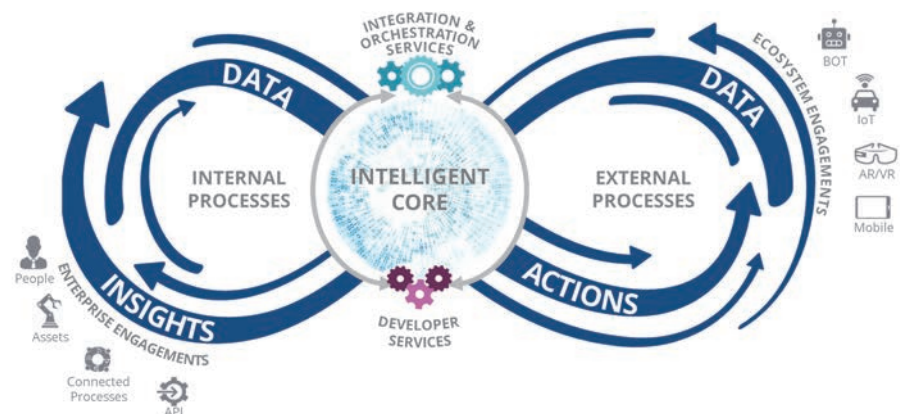
Source: IDC, 2019

At the heart of the DX platform is the "intelligent core," where data is managed and utilized and where the algorithms, code, and models live. The intelligent core houses data that promises value to the enterprise beyond single systems and

processes. Because innovation happens more at the edge, it is critical to make data seamlessly accessible from everywhere in the ecosystem.

The flood of data enables a new set of technologies such as ML, natural language processing, and deep learning — collectively known as AI systems — to turn data analysis into a proactive driver of strategic decision and action.

Another key step is to evolve data warehousing and Big Data strategies to overcome the challenges of data fragmentation, lack of holistic view, and rising analytics costs. Companies need to evaluate modern data warehousing tools that help analyze data from new sources and unstructured data to provide a complete view of the relevant data to turn it into insights.

Organizations also need to develop long-term strategies to become data driven and assess the advanced analytics features, real-time analytics, and the full spectrum from IoT to AI, and how data science will evolve within their organization.

The foundation businesses choose today for their data transformation will determine their success tomorrow. IDC believes that agility, platform independence, openness, rich ecosystem, deep integrations, technology innovation, vendor credibility, and even customer stories are important considerations.

As data explodes in volume and variety, there is a growing realization that all data created is not the same and those that can effectively identify, analyze, and take advantage of the critical subset of data will be the ones that improve user experience, solve complex problems, and create new economies of scale.

To capitalize on data, CIOs and IT leaders need to:

- Develop and implement an industrial-class technical foundation for data capitalization

- Clarify the role and expectations of IT organizations in relation to data management and data monetization

- Involve key stakeholders such as data scientists, engineers, BDAs, security, and the developer community to architect the data strategy

Making Sense of the Fast-Evolving Big Data and Analytics World

There are many different terms thrown about in the analytics space that are quite similar. Clarifying these terms can provide a useful foundation and context. IDC believes that, ultimately, enterprises need to understand the problems they are trying to solve or identify the use cases and then select the data capitalization initiative to achieve the objective, rather than simply jumping onto the Big Data bandwagon.

1. **Analytics.** The broad practice of deriving insight from information using tools, technologies, and methods, which can be descriptive, predictive, or prescriptive analytics.

2. **Data science.** Similar to analytics, data science lies at the intersection of statistics and computer science. It refers to tools, techniques, and methods

of analytics used by data scientists — professionals with a deep understanding of the algorithms and technologies used in analytics.

3. **Machine learning.** This is a set of tools and techniques used to train computer systems to generate predictions without being explicitly provided with expert instructions. ML can be applied to a variety of techniques, ranging from regression analysis and clustering to neural networks.

4. **Data mining.** This is the extraction of patterns, relationships, and knowledge from large amounts of data. It resides at the intersection of statistics, machine learning, and database systems.

5. **Artificial intelligence.** This is a term that was once used to describe the use of neural networks but is now used as a more generic term to refer to everything involved in making machines intelligent and automated.

Let's assess how the offering from Microsoft and Cloudera is opening up new opportunities for customers to be data driven.

Value Proposition of Microsoft Azure and Cloudera

All eyes were on Microsoft when five years ago its CEO, Satya Nadella, pledged a cloud-first, mobile-first, and an open source friendly strategy. Since then, it has come a long way. On the open source front, Microsoft opened up its patent portfolio by joining the Open Invention Network (OIN), it has forged deeper alliances with open source technology distributors, and has engineered deeper integrations between Azure and open source OSs, databases, analytics platforms, and workloads.

In tune with the rapid innovation in the public cloud market, Microsoft is racing ahead on multiple fronts, including expanding products and features covering security, developer tools, DevOps capabilities, AI and IoT capabilities, and a wide variety of hardware and service improvements in the Microsoft Azure Cloud. It is also continuing to integrate its broad portfolio of software products in conjunction with its cloud services to provide hybrid cloud solutions for customers. At its recent Ignite conference, it tried to allay public cloud concerns with improvements and expansions to the Microsoft toolset, making it practical to adopt a hybrid strategy and grow confidence in public cloud.

Microsoft has also invested in deep integrations with best-of-breed complementary platforms such as Cloudera. This alliance gives customers the flexibility to leverage familiar data analytics tools and progress in their AI, ML, and data-driven initiatives and extend it to the public cloud. The combination of Azure features such as Azure Data Factory, Azure data services, security features, data management tools, and technical support, along with Cloudera's modern data warehousing capabilities, AI-driven analytics, and end-to-end data management, can help enterprises make progress in their data objectives.

One of the hallmarks of the partnership is the simplicity where customers have a single point of contact for their Cloudera Data Platform on Azure, helping them accelerate their speed of business. The simplicity is also extended to bringing

together widely fragmented data sets into a single analytics platform for impactful analyses.

From a business perspective, the scale and economics of cloud infrastructure enable data scientists to experiment faster, with data scientists creating models with small subsets of data on the Cloudera platform and then scaling the model using Azure's scale, data services, and security features.

Cloudera Data Platform (Enterprise Data Hub)

The solution brings together multiple analytics disciplines — data warehouse, data science, data engineering, search, and streaming and real-time analytics — into a single platform. This improves shared data experience and simplifies policy changes, security, information governance, and data life-cycle management.

What differentiates the Cloudera platform is the multifunction capabilities where multiple functions can be performed on the same data repository, essentially moving away from single vendor use case and giving customers the ability to reuse data across platforms. What IDC is also seeing is how multiple factors such as compliance, governance, cost, security, and proximity to workloads are determining where data resides (called location optimization), forcing companies to evaluate technologies that can work across these heterogeneous infrastructures of multicloud and hybrid cloud environments.

Cloudera has transitioned beyond Hadoop into a data platform software vendor, placing data intelligence at the core, competing to lead in AI, data science, and machine learning for the enterprise. It is also disrupting the data warehouse for analytical workloads with modern data warehousing capabilities that include structured, semistructured, and unstructured data analytics, and driving platform innovation to simplify complex data environments.

In IDC's opinion, a data analytics platform that works across multiple cloud platforms as well as on-premise infrastructure can lend a unified management and data visibility experience to customers. Cloudera's core capabilities for end-to-end data management on-prem and in public cloud, along with security and analytics, and in modern data warehousing features, can help build the foundation for a company's data strategy and help it progress in its future IoT, AI/ML, and data science initiatives. This unified experience is key as it gives enterprises the flexibility to move workloads on-prem or to the cloud as their needs change without compromising or affecting user experience.

Cloudera, with its investment, engineering efforts, mergers, and partnerships, is moving in this direction to become the data platform of choice for customers. For one, its merger with Hortonworks brings complementary technologies that give it capabilities from IoT to analytics to AI, simplifying the data transformation journey for its customers. The merger also increases Cloudera's value proposition to customers from edge to AI — a unified data management and analytics platform. From an innovation standpoint, it has added features of modern data warehousing and data management capabilities.

Its partnership with Azure demonstrates its commitment to integrating its technologies with fast-growing hyperscale platforms, showing how it is capitalizing on cloud migration trends.

Long-Term Partnership

Microsoft Azure and Cloudera began their partnership in 2014 and have increased the integration in the past four or five years. There are dedicated partner management teams and both companies consider each other as a tier 0 partner indicating deeper engagement and commitment to make technologies work seamlessly for improved customer experience.

One example of this is the availability of Cloudera on the Azure marketplace with a one-click enablement of Cloudera on Azure.

The broad pillars where the partnership shines are:

- Azure IaaS. Joint customers can leverage basic public cloud advantages such as vast compute capabilities, Azure storage, and its Express Route network to have closed access to the hyperscale environment and to instantiate the Cloudera platform for analytics.

- Delivering data science experience across hybrid cloud.

- Integrating innovation from each vendor faster. New offerings such as Cloudera ML are on the road map to be available on Azure, as is integration with the Azure Kubernetes Service container app platform. IDC sees an inherent synergy in Azure's containerization strategy and Cloudera's cloud-native application landscape vision.

- Complementary vision. Both vendors have similar open source strategies, with Microsoft demonstrating its commitment to open source firmly in the past two years. For its part, Cloudera runs on enterprise-grade Linux environments and the merger with Hortonworks further bolsters its open source strategy. The core objective of the Microsoft-Cloudera alliance is to make the customer journey seamless in this process, and the vendors jointly fund the PoCs and migration enablement. The open source technologies such as Hadoop, Linux systems, and Cloudera are deeply integrated in Azure. Microsoft is making a strong impact within the open source community and bringing more Azure features to open source platforms on it.

Cloudera was named the open source data and AI partner of the year for Microsoft in 2018. IDC believes there is a natural convergence in customer bases between the two companies because of the large enterprise focus both vendors have.

Case Study

Here is how one European customer used Cloudera and subsequently the platform on Microsoft Azure to leverage analytics-driven insights for internal benefits, deliver services to its customers, and increase revenues.

Deutsche Telekom

The German telco uses Big Data and analytics for two key reasons — for internal processes and systems efficiency to cut costs, and to facilitate new capabilities for employees and customers, and improve use cases such as fraud detection.

Network fraud is a key challenge for telecommunications players, and to prevent it huge volumes and a range of network data need to be collected and analyzed to identify suspicious events.

The fragmentation of network data was becoming a key hurdle for holistic analytics, which limited visibility and made machine learning at scale impossible. It selected Cloudera to modernize its data platform because of the platform's ability to analyze vast datasets and the opportunity to apply ML and real-time analytics and deliver a shared data experience.

It used Cloudera Altus Data Warehouse, Cloudera Data Science Workbench, and SDX (Shared Data Experience) to establish its Telekom Data Intelligence Hub: a trusted, fully governed platform and ecosystem where its users are empowered to exchange and analyze data and develop multifunction, data-driven applications securely.

The modern data intelligence platform is yielding results, and Deutsche Telekom has been able to overcome its key challenge of network fraud. The large scale, fast data processing of complete datasets, and the interactive querying tool in Cloudera have helped improve network quality and detect fraudulent activities in real time, saving revenue losses by 10%–20%.

The telco also created an enterprise view of data, ranging from network data to CRM data, and the analytics instantly helped it better serve customers and improve service quality. The data-driven initiative also brought operational efficiencies of 50%.

It all started in 2013–2014 when DT selected Cloudera for its European presence, support, and advanced data warehousing features in the platform. It started with some Hadoop distribution on a private dedicated bare metal infrastructure but when public cloud started gaining strong momentum, DT assessed the hyperscale market and found a natural partner in Microsoft Azure in 2016. Two things that made the alliance a natural fit were Deutsche Telekom's T-Systems data trustee agreement with Microsoft Azure for the European market, which met data sovereignty and privacy regulations, and secondly, the integration of its existing data platform, Cloudera on Azure, and the combined benefits of cloud economics and modern analytics and data warehousing.

"Cloudera on Azure is our solution of choice, even for our emerging initiatives such as IoT, now that Cloudera on Azure is even stronger with the Hortonworks merger," said Sven Loeffler, DT's data-driven initiatives executive. Deutsche Telekom is now

looking to grow its Azure environment with a dedicated cluster for its analytics, rather than just persistent clusters. Based on its expertise and experience, Deutsche Telekom uses Cloudera on Azure to provide managed data analytics services to its customers and help them in their data transformation strategies. "Data transformation is becoming a key priority for even SMBs in the German and European markets," said Loeffler. IDC agrees that there is a growing appetite for becoming data driven but smaller and medium-sized enterprises lack the in-house skills to adopt modern data warehousing solutions and leverage AI and ML capabilities effectively, opening up a big market for managed data analytics services from vendors such as Deutsche Telekom. This in turn increases the total market opportunities for Cloudera and Azure. In conversations with IDC, Cloudera and Azure emphasized the importance of the channel and ecosystem in delivering managed data analytics services.

"The three-way partnership between DT, Microsoft Azure, and Cloudera is very strong and beta participation is ongoing for all new services," said Loeffler.

The main advantage of using Cloudera on Azure is the remarkable improvement in speed to market for Deutsche Telekom.

> "*Speed to market is hugely important to us. It used to take between one and three months to provision Cloudera on-premise on our bare metal infrastructure. But now with Cloudera on Azure, it is available on-demand.*"
>
> Sven Loeffler, Deutsche Telekom

The on-demand version in the cloud has served a specific use case well at the telco. It had ramped up one of its workloads to exchange data across the enterprise and had to make data persistent to use it to build models. But setting up persistent data each time was proving costly because of the unpredictability of the workload. With the on-demand Cloudera solution on Azure, the company can build models as per the flexibility of the data scientists and developers. "We are seeing positive cost implications with this," said Loeffler.

DT's managed data analytics services are popular among automotive companies as well as SMBs.

The Growing Use Case and Business Outcome Scenarios for Unified Data Analytics and Data Science

More and more workloads lend themselves to AI-driven analytics in cloud and data science. Core prediction applications such as logistics, anomaly detection, and pattern detection are growing, as are healthcare and precision medicine, predictive maintenance, and fraud detection in the automotive and manufacturing sectors, especially the manufacturing aftermarket segment. Financial services organizations trying to optimize investments and market movement are also heavily leveraging AI and ML to understand and manage multiple scenarios. Beyond this, IoT-heavy industries such as manufacturing, retail, utility, and energy sector players are modernizing and transforming their workloads for cost and resource optimization and for applying analytics on sensor data to drive ML models.

IDC also sees GDPR as a trigger that has forced organizations to assess their data landscape and increase investment in managing data. Regulation has made data privacy and management a boardroom priority and a key investment area for CIOs. IDC believes that organizations that use GDPR as an opportunity to secure investment to develop data hygiene and build their analytics strategy — such as Barclays Bank — are the ones that have a competitive edge. A security-defined and modern data platform hosted on regional hyperscale environments offers enterprises the ability to meet the regulations while ensuring they leverage public cloud capabilities. IDC's GDPR-focused survey revealed that enterprises' appetite for cloud is not diminishing but it is becoming strategic and well calculated to overcome GDPR compliance risks. In our opinion, data analytics platforms across multicloud can act as a foundation for modern IoT, ML, and AI initiatives and new workloads.

Cloudera's Shared Data Experience feature helps companies unify, secure, and govern data across multiple silos, simplifying compliance and privacy management. What is significant is the encryption and metadata management in the Cloudera platform along with Azure's security features, which help meet specific clauses in GDPR, giving European enterprises the confidence to execute on their cloud- and data-driven strategies without compromising on data privacy.

IDC observes how many automotive players in Europe are looking at data exchange use cases for predictive maintenance to optimize their businesses cases. These car makers are quickly ramping up data analytics use from TB to petabyte scale.

Other use cases such as asset optimization, capex allocation, and network optimization are all emerging fast and strong with advanced, AI-driven analytics.

A combination of internal and external network optimization, as with Deutsche Telekom, is also proving popular. IDC also sees more need for analytics on combined internal and external data sets (processing data from across the ecosystem) in retail and manufacturing to improve supply chain management, make the estimated time of arrival for logistics efficient, and predict the ETA.

IDC believes that unified data analytics platforms such as Cloudera on Azure simplify analytics and broaden the use cases. Many organizations bring together several data warehouse appliances to support complex analytic requirements such as market surveillance and compliance analyses, but the results are patchy because of the complex integration. The varying scalability of the warehouse solutions also forced enterprises to silo the data by market.

Consolidating all datasets into one analytics platform, leveraging the scale and on-demand capabilities of the cloud, the security and data management features, and modern AI can bring huge value to analytics strategies.

Eliminating silos and running deep queries at scale were key considerations for European enterprises when investing in data transformation strategies.

Future Outlook

IDC believes that the ongoing work on the alliance from both vendors is making the offering compelling. For example, in the road map, there is a plan to explore containers, edge to AI on cloud, and more managed data analytics services for a new group of customers such as SMBs.

Although the vendors offer support and technical advice and joint PoCs to help customers adopt and use these technologies effectively, Microsoft and Cloudera need to assess how they can further help overcome the skills gap either through partner training, best practices, and sharing PoV (proof of value) to illustrate what the technology brings to the customer in terms of business outcomes.

The vendors also need to be mindful of rising competition from other hyperscalers and innovate faster as this is a critical turning point when many large customers are moving from the testing phase to production environments.

Becoming Data Thrivers and Successfully Competing in the Digital Era

We are entering an era of "sense, actuate, compute" — to turn data into value. It is fair to say that enterprises are becoming information businesses first. Data and intelligence represent a unique opportunity to create unimaginable value. Enterprises seeking to differentiate themselves in the digital economy need to build their intelligent core at the heart of a digital platform. IDC predicts that by 2020, 80% of enterprises will create data management and monetization capabilities, thus enhancing enterprise functions, strengthening competitiveness, and creating new sources of revenue. Moreover, by 2023, 95% of entities will have incorporated new digital KPI sets, focusing on product/service innovation rates, data capitalization, and employee experience.

End users need to be aware that technology alone won't do it. They need to identify the key problems they intend to solve, develop a data-driven culture, and be well aware of the characteristics of workloads so that they can match the right workload to the right analytics platform and use the features that matter the most.

With tremendous innovation happening in the data warehousing space, hyperscale, and data management in the cloud, IDC believes that now is the time to set up the technical foundation to become a data thriver in the digital age.

As companies do that, they should evaluate technology options such as the Cloudera Data Platform on Microsoft Azure.

IDC UK

5th Floor, Ealing Cross,
85 Uxbridge Road
London
W5 5TH, United Kingdom
44.208.987.7100
Twitter: @IDC
idc-community.com
www.idc.com

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.