

Sign up for our newsletter and get the latest big data news and analysis.

[Home](#) » [Big Data](#) » Big Data Industry Predictions for 2022

## Big Data Industry Predictions for 2022

December 15, 2021 by [Daniel Gutierrez](#) [Leave a Comment](#)

[tweet](#) [share](#) [share](#) [email](#)

Welcome to insideBIGDATA's annual technology predictions round-up! The big data industry has significant inertia moving into 2022. In order to give our valued readers a pulse on important new trends leading into next year, we here at insideBIGDATA heard from all our friends across the vendor ecosystem to get their insights, reflections and predictions for what may be coming. We were very encouraged to hear such exciting perspectives. Even if only half actually come true, Big Data in the next year is destined to be quite an exciting ride. Enjoy!



Daniel D. Gutierrez – Editor-in-Chief & Resident Data Scientist

### Analytics

Predictive analytics will drive new, emerging use cases around the next generation of digital applications. The technology will become more immersive and embedded, where predictive analytics capabilities will be blended seamlessly into the systems and applications with which we interact. Predictive analytics will drive use cases in next-gen apps like metaverse applications (convergence of digital and physical worlds, powered by technologies such as IoT, digital twins,

AI/ML, and XR) and the next generation of composable applications. – Nelson Petracek, CTO of [TIBCO](#)

Self-Service Analytics – Every company wants to become a data company or at least a data-driven company. This naturally will lead to the rise of self-service analytics. Currently, business leaders and their teams are heavily dependent on dedicated analytics teams within the organization who have a large backlog of analysis and dashboards they need to build and work on. With more businesses that want to put their customer at the center, it is imperative that data & insights are available across the company and access is democratized. – [NWO.ai](#)

Most people think about analytics as being about periodic queries such as querying my data warehouse to get a snapshot of my business. Most people don't think about data analytics as code—as the core of an always-on application that is interactive; incorporates fresh, real-time data; and supports thousands of internal and external end-users at the same time. Most of all, when people think about analytics, they think of data analysts, IT analysts or security analysts using a business intelligence or operational intelligence tool; they don't think about software developers and the power of the applications they build. They don't ask what would be possible if analytics were part of the core toolkit of a software developer. In 2022, a new mindset around analytics will emerge—one that is less about periodic questions

### INDUSTRY PERSPECTIVES

#### From Mystery to Mastery: How to Get the Most from Artificial Intelligence

In this special guest feature, Lloyd Humphreys, Principal Product Manager for Data & Analytics, Tradeshift, suggests that you change your mindset, forget about relinquishing complete control, and you'll soon discover that AI can give you a competitive advantage in areas of your business you barely gave a moment's thought to before, like the sometimes unglamorous world of invoice processing.



[\[READ MORE...\]](#)

### WHITE PAPERS



#### Garbage in, Garbage Out – How We Got Here and Why We Must Get Out Now

This whitepaper from our friends over at Profisee, reflects on why the state of data in most organizations is as dismal as it is, and why there is such a challenge involved in demonstrating the value of trusted data available across mission-critical operations and analytics in an enterprise.

[Download](#)

[See More White Papers »](#)

### FIND US ON:



and more about always-on, interactive applications. We will see these applications being built by every digital and SaaS company as they seek to gain insights about their business and operations. Equally important, we will see these applications emerge as digital and SaaS companies seek to share insights with their customers and partners—a trend that in 2022 will be the new norm and the new imperative for every digital and SaaS company. – Fangin Yang, founder and CEO, [Imply](#)

The CAO will eclipse the CDO. While many companies today have a chief data officer, in 2022 we will see more enterprises establish “chief analytics officer” or “chief data and analytics officer” roles. Elevating analytics reflects an evolving understanding of data science and machine learning as the ultimate functions that turn data into business value, and increasingly core to company-wide strategy. – [Domino Data Lab](#) CEO Nick Elprin

In 2022, with powerful technologies available to them, organizations will invest more in unstructured analytics. To date, most business intelligence has been conducted using structured data; however, there are countless problems that cannot be answered by these clean-cut numbers. Burgeoning people analytics teams are offered a new means of assessing uniquely human situations—talent acquisition, workforce sentiment, productivity, etc.—by analyzing the textual, conversational, and communicative data created by the workforce each day. These emails, files, and collaboration data speak to the human side of the enterprise that has long remained out of reach. – Ryan Splain, [ZL Technologies](#)

The Rise of the “Just in Time” Data Analytics Stack – There’s a small, but fast growing, segment of the data analytics space that is focused on new approaches to the enterprise stack, including continuing to move all the things to the cloud. However, the hybrid multicloud imposes requirements of its own most notably the ability to manage and analyze data no matter where it lives in the hybrid multicloud environment. Startups like Starburst, Materialize.io, Rockset, and my own company Stardog develop platforms that are designed to query, search, connect, analyze, and integrate data where it lays without moving or copying it, in a just-in-time fashion. In a world where the number of places that data may be residing in storage is increasing, rather than decreasing, expect to see enterprises reach for data analytics solutions that are not coupled to where data lives. This trend will accelerate in 2022 as data movement between storage systems will continue to be removed from the stack in order to accelerate time to insight. – Kendall Clark, Founder and CEO at [Stardog](#)

Small and wide data analytics begin to catch on – AI/ML is transforming the way organizations operate, but to be successful, it is also dependent on historical data analytics, aka big data analytics. While big data analytics is here to stay, in many cases this old historical data continues to lose its value. In 2022, organizations will leverage small data analytics to create hyper personalized experiences for their individual customers to understand customer sentiment around a specific product or service within a short time window. While wide data analytics is comparatively a new concept and yet to find widespread adoption – given the pace at which organizations are making use of unstructured and structured data together – expect to see small and wide data analytics to gain better traction across organizations as we enter 2022. – Ravi Shankar, SVP and CMO at [Denodo](#)

The Most Transformational Analytics Use Cases will Come From “Citizen Analysts” – Due to their domain expertise, proximity to the business, and availability of new (tools|technologies), citizen data analysts will become the most important and influential individuals who work with data. This will lead to an explosion of new ideas and practical applications for data, marking the next big turning point for the industry. – Matthew Halliday, Co-founder and Executive Vice President of Product at [Incorta](#)

Organizations will redefine what it means to build a “culture of analytics.” For too long, business leaders have assumed that up-skilling their workforce with data classes/certifications and investing in self-service tools would lead to a data-driven organization. They are finally ready to admit that it’s not working. Self-service BI does not “close the skills gap.” Not everyone has time or interest in becoming a data analyst or data literate, especially now in today’s post-COVID landscape where teams are understaffed and people are valuing their time differently in and outside of work. In 2022, organizations will redefine what it means to build a “culture of analytics” and change the paradigm by bringing insights to workers in a more digestible way – turning to methods and solutions like embedded analytics that won’t require them to learn new skills or invest additional time. – Ashley Kramer – Chief Product and Marketing Officer, [Sisense](#)

The onus on data analysts and engineers to treat data like a product, in other words, embed dashboards, data platforms, and self-service data workflows with the same diligence as we treat SaaS products. This boils down to ensuring that that data and associated data products

are securely administered, accessible to the right individuals, trustworthy, and scalable across different domains. Data leaders who figure out how to scale this mindset while keeping data debt at bay will be the real winners. – Barr Moses, CEO & Co-founder of [Monte Carlo](#)

The analytics engineer displaces the data scientist as the world's sexiest job – For years, data science has been the craze for companies looking to tap into the value of digital transformation. However, the role of the data scientist has since lost its luster in recent memory. Companies have failed to operationalize models, universities and certificate programs have churned out coders who cannot apply their skills in a business context, and data scientists spend countless hours on the drudgery of dealing with messy, disparate data – all of which has tarnished data science's sheen. For that reason, I expect 2022 to see the rise of a new role in the industry that replaces data scientists: the analytics engineer. Paired with the ability for transformations to be done within cloud platforms on all data, analytic engineers will be essential to controlling transformation logic and leverage the full capabilities of the modern data stack. – Cindi Howson, Chief Data Strategy Officer at ThoughtSpot

In 2021, analytics deployments grew at a crazy pace as businesses mined for gold nuggets within their data. However resources usually followed two different trajectories: either for web-scale analytics or for core business analytics. Web-scale utilized the power of the cloud, and business analytics remained in the data center. The reliability and performance of the cloud data infrastructure was key to driving the wedge between these two. In 2022, we will see the convergence of analytics environments as new performance infrastructure in the cloud for compute, networking and storage is built out. As a result, many companies will migrate their core business applications and database environments to the cloud, uniting their data in a central resource. From BI, database analytics and into the AI/ML environments, it's now entirely possible for near-real time analysis of data to be done in the cloud, using cloud engines together with the web-scale data platforms. – Jeff Whitaker, Vice President of Products, [Excelero](#)

### **Artificial Intelligence**

More women will move to AI – there has been a global push to involve more women in science and technology careers and AI is one of the fields in which women can experience tremendous success. In order for organizations to achieve the highest AI maturity level by ensuring data is unbiased and represents the entire population, women will become part of all enterprise endeavors in artificial intelligence, from research to product launch. – Leah Forkosh Kolben, Co-founder & CTO at [cnvrg.io](#)

With the proliferation of unstructured text, knowledge workers are struggling to gain insights from the volume of information they must comb through. In 2022, organizations will look for AI technologies that remove the barriers of traditional supervised learning models so that they can more easily and quickly turn these troves of data into usable information. AI vendors will flip the script and deliver solutions that do not require the time, resources, and expense required for supervised learning models. They will deliver solutions that provide highly relevant and context-driven information with unprecedented speed and precision so that humans are empowered to do their most meaningful work. Rather than replacing human intervention, these modern — and evolving — AI technologies will allow people to analyze and use unstructured as well as structured data in a smarter, faster, and more natural way. Today, approximately 80% of the information data in both public and private sector organizations is unstructured text. To put this in perspective, most enterprises have over 1 petabyte of data, and, according to McKinsey, a petabyte is the equivalent of about 20 million four-drawer filing cabinets full of text. As a result, companies cannot gain actionable insights from, and find the hidden intelligence in, text needed to ensure effective decision making. Worse, most struggle to analyze and leverage the growing volumes of data they've already collected. In 2022 and beyond, businesses will move on from the supervised learning approach to a natural language enabled model that enable humans to identify opportunities and threats and to take more immediate action. – Ryan Welsh, Founder and CEO of [Kyndi](#)

Enterprises Will Discover the Big AI Lie – 92% of companies are invested more in AI in 2021, yet just 12% are deploying it at scale, down from last year. What's going on? How can companies be spending MORE on AI but getting LESS from it? There are many non-obvious factors at play: culture, tools, bias concerns, fear, and automation grace the top of the list. In 2022, firms must meet these challenges head-on with a cultural approach to model operationalization to better manage, track and optimize algorithms. Only then will data science move from the playground to the battleground. – Mark Palmer, SVP of data, analytics and data science products, [TIBCO](#)

Over the years, AI has gone from buzzword to game-changer technology, and it is revolutionizing how developers work. From productivity to quality and speed increase, the benefits are unmeasurable; however, the developer community continues to face a challenge: the implementation of AI. With the AI market expected to blow past \$500 billion by 2024, next year is bound to be the steppingstone toward an AI-centric software market. For one thing, AI will alter how code is written, updated, and released – DevOps will become increasingly automated and responsive. Software developers will need to learn how AI will fit within their own tasks – with AI-empowered to make changes to itself, the focus for developers will shift to a more creative, strategic level. For example, developers will need to learn how to “talk AI” to provide insights and drive core business operations; integrate different APIs using AI to build a better product and provide faster go-to-market time frames. Lastly, they will have to focus on the aspects of the software that are not so easily automated, such as finding ways that multiple software systems could work together. Developers will likely shift away from the practice and process of development and into building highly customized solutions for a wide range of challenges. – Jonathan Grandperrin, CEO of [Mindee](#)

AI may be one of the most hyped technologies in recent years, but from where I sit, it is one of the most effective technologies to determine future behavior. According to [Forrester](#), brands will flock to AI-powered audience solutions, fueling 20% of media and advertising category growth in 2022. With AI and machine learning, marketers can gain insights in real time and at scale, providing them with the ability to better understand their audience, what they need and where they're looking for it. This empowers them to create better online experiences, improve business performance and build brand trust through true relevance. If a marketer or advertiser is not using AI-driven solutions to enhance their campaigns, they are missing out on insights, new audiences, and productivity gains. Those that embrace AI and machine learning now will gain a long-term competitive advantage. – Konrad Feldman, Co-Founder and CEO of [Quantcast](#)

Early adopters of rudimentary enterprise AI embedded in ERP / CRM platforms are starting to feel trapped. In 2022, we'll see organizations take steps to avoid AI lock-in. And for good reason. AI is extraordinarily complex. When embedded in, say, an ERP system, control, transparency and innovation is handed over to the vendor not the enterprise. AI shouldn't be treated as a product or feature: it's a set of capabilities. AI is also evolving rapidly, with new AI capabilities and continuously improved methods of training algorithms. To get the most powerful results from AI, more enterprises will move toward a model of combining different AI capabilities to solve unique problems or achieve an outcome. That means they'll be looking to spin up more advanced and customizable options and either deprioritizing AI features in their enterprise platforms or winding down those expensive but basic AI features altogether. – Doug Gilbert, CIO and Chief Digital Officer at [Sutherland](#)

More Open Source behind Analytics & AI. As the momentum behind the Open Data Lake Analytics stack to power Analytics & AI applications grew over the past year, we'll see a bigger focus on leveraging Open Source to address the limitations around flexibility and cost when it comes to traditional enterprise data warehouses. Open source cloud-native technologies like Presto, Apache Spark, Superset, and Hudi will power AI platforms at a larger scale, opening up new use cases and workloads that aren't possible on the data warehouse. – Dipti Borkar, Co-founder and Chief Product Officer (CPO), [Ahana](#)

AI fatigue will reach a breaking point. AI has long been positioned as the solution to all of our problems, especially for customer experience. 2022 will be the year that the technology will lose some of its shine. Some organizations have already realized that AI solutions, like chatbots, do not deliver on CX the way they were sold, often frustrating customers more than they help. More organizations will become tired of how AI is positioned to them in the year ahead. To combat this, AI companies will shift how they sell. Instead of positioning AI as a silver bullet, it will be portrayed for what it truly is — a supporting tool to help humans, like CX agents, do their jobs more effectively and help organizations uncover valuable customer insights. If handled properly, these insights have the potential to move past commodification to improve overall business outcomes. The more AI companies sell solutions as being able to generate data-driven insights, as well as embedding these findings and closing the feedback loop, the more they'll win over buyers. – Jeff Gallino, CTO of [CallMiner](#)

In 2022, artificial intelligence (AI) will increasingly move from software simulation closer to the real world. This technology, coupled with machine learning (ML) will only continue to advance as 5G advances to 6G and beyond. With mobile networks becoming more complex, AI/ML design will be increasingly utilized to optimize communication systems and networks. Faced with the challenging of squeezing every bit of bandwidth from available spectrum,

improving latency, and creating energy efficiency – the design of advanced AI/ML systems requires new data sets and new training techniques. We know these technologies will not arrive to their optimal places this year but do expect to see big strides. – David Hall, Global Go-To-Market Director, Semiconductor & Electronics at [NI](#)

The future of AI technology and its credibility will rely on organizations mitigating AI bias through technology diversity initiatives. As AI is adopted for an increasing number of business functions and data analysis, AI bias has become increasingly concerning for experts. Bias can impact AI algorithms in numerous ways to skew results and provide information that's not fair or objective by proxy. This is damaging to the credibility of AI technology and has the potential to stifle its growth and the consumer trust needed to advance it forward. The future of AI technology and its credibility will rely on organizations mitigating AI bias through technology diversity initiatives. For example, a dataset that used to be considered the benchmark for testing facial recognition software had data that was 70% male and 80% white – not representative of the holistic population. Even if sensitive variables such as gender, ethnicity and sexual identity are excluded, AI systems learn to make decisions based on training data, which may contain skewed human decisions or represent historical or social inequities. While diversity and inclusion are discussed from a hiring and corporate perspective, it must also be a critical component of product development. To get ahead of this issue, CTOs of organizations using facial recognition technology should be asking their technology providers how their algorithms are trained. This will put pressure on identity vendors to ensure their solutions' AI algorithms are built to represent the broader population. – Labhesh Patel, CTO of [Jumio](#)

The need for localized AI/ML models will significantly increase – AI and ML models are only as “intelligent” as the data they are fed. When you rely on these models to grow your business, they need to be malleable to the myriad of external factors that will affect your desired outcome. That's why experimenting with localized AI/ML models is becoming more necessary for businesses to have a clear understanding of their demographics. When you're implementing AI/ML in your business, typically what happens is that with the first few versions of the models, you can see a lot of change. You're able to quickly move from zero to 60 percent of the way in your AI journey, with just a few tweaks to the algorithm. Going from 60 to 90 percent gets much harder; when you're trying to expand, you must also start thinking more about the differences among your various use cases. Capitalizing on localized models can provide a wider optic and vital insights for businesses to meet their goals and stay at the forefront of competition. – Harish Doddi, CEO, Datatron

AI investments shift from generic models to more precise industrial AI – 2022 will see AI's maturation into industrial AI reach full bloom, graduating to real-world product deployments with concrete time-to-value. To achieve this, we'll see more industrial organizations make a conscious shift from investments in generic AI models to more [fit-for-purpose, precise industrial AI applications](#) that help them achieve their profitability and sustainability goals. This means moving away from AI models that are trained on large volumes of plant data that can't cover the full range of potential operations, to more specific industrial AI models that leverage domain expertise for interpreting and predicting with deep analytics and machine learning. Industrial data will be transformed into real business outcomes across the full asset lifecycle. This shift will have the dual benefit of also facilitating new best-of-breed alliances built around industrial AI. Previously, partnerships were very tech-centric, driven by services or one large vendor. The more specialized focus of industrial AI will require a larger set of solutions providers, pooling together their independent and customized expertise. Not only does this help evolve partnerships away from more generic AI projects, it will also place a greater focus on time-to-value partnerships as opposed to do-it-yourself approaches, helping to lower the barrier to AI adoption more than ever. – Bill Scudder, SVP and AIoT General Manager at [AspenTech](#)

AI in marketing is here to stay – “We're seeing AI technologies playing a bigger and bigger role across the entire customer journey, from digital self-service, where AI-powered chatbots are able to offload repetitive cases for agents, to Voice of the Customer technology, where AI is used to constantly listening to survey responses and identify important trends and themes, enabling brands to proactively address customer satisfaction issues. The role of AI in marketing and CX overall will only continue to be more pervasive as brands work to provide better customer experiences at a bigger scale across multiple channels. – Yuval Ben-Itzhak, President and Chief of Strategy, [Emplifi](#)

Responsible AI shifts from an aspiration to a foundational requirement for most AI projects. In 2021, responsible AI was one of the hottest topics in the AI industry, but adoption remained relatively low. According to the Appen 2021 State of AI report, concern around AI

ethics remained at just 41% among technologists and 33% among business leaders. In 2022, however, the stakes become much higher, as businesses recognize that responsible AI leads to better business outcomes. The principles of responsible AI are now well-established: unbiased data, fair treatment on the data collection and labeling side of the industry, and a recognition that AI projects should promote the social good (or at least avoid the potential for social harm). Implementing these principles ensures that AI projects work as expected and protects the brand. In addition governments are beginning to recognize the potential harm that can come from the irresponsible use of AI. So in the same way that data privacy has moved from concern to regulation, responsible AI will begin the same journey. Gartner expects that by 2023, all personnel hired for AI development and training work will have to demonstrate expertise in responsible AI. – Mark Brayan, Appen

AI/ML drive the citizen experience: Smart Government applications will look more like consumer apps and less like corporate intranets. The smartest cities will have integrated ML and AI in recommendation engines, support natural language interactions, deliver everything digitally and consider citizen experience the top requirement. – Brian Gilmore, [InfluxData](#), PM IoT

More responsible AI will bridge the gap from design to innovation. While companies are starting to think about and discuss AI ethics, their actions are nascent, but within the next year we will see an event that will force companies to be more serious about AI ethics. An increasing number of companies will get more serious about AI ethics with transparent explainability, governance and trustworthiness at the center. – David Sweenor, Senior Director of Product Management at Alteryx

Synthetic 3D Data for the Next Era of AI: The rate of innovation in AI has been accelerating for the better part of decade, but AI cannot advance without large amounts of high quality and diverse data. Today, data captured from the real world and labeled by humans is insufficient both in terms of quality and diversity to jump to the next level of artificial intelligence. In 2022, we will see an explosion in synthetic data generated from virtual worlds by physically accurate world simulators to train advanced neural networks. – REV LEBAREDIAN, Vice President of Simulation Technology, Omniverse Engineering, [NVIDIA](#)

Synthetic Data Will Be a Requirement to Build the Metaverse. The metaverse cannot be built without the use of synthetic data. To recreate reality as a digital twin, it's necessary to deeply understand humans, objects, 3D environments, and their interactions with one another. Creating these AI capabilities requires tremendous amounts of high-quality labeled 3D data--data that is *impossible* for humans to label. We are incapable of labeling distance in 3D space, inferring material properties or labeling light sources needed to recreate spaces in high-fidelity. Synthetic data built using a combination of generative AI models and visual effects (VFX) technologies will be a key enabler of the AI models required to power new metaverse applications. – Yashar Behzadi, CEO and Founder of [Synthesis AI](#)

Companies will lean more on human-powered AI to avoid "Garbage In, Garbage Out" algorithms. As AI continues to evolve at a breakneck pace, companies often overlook the importance of keeping humans actively involved in the AI implementation process, creating a scenario where tech's obsession with the newest, biggest thing neglects basics that make AI actually useful: plugging in useful data and teaching it how to deal with outliers. For AI to truly be useful and effective, a human has to be present to help push the work to the finish line. Without guidance, AI can't be expected to succeed and achieve optimal productivity. This is a trend that will only continue to increase. Ultimately, people will have machines report to them. In this world, humans will be the managers of staff (both other humans and AIs) that will need to be taught and trained to be able to do the tasks they're needed to do. Just like people, AI needs to constantly be learning to improve performance. A common misconception is that AI can be deployed and left unsupervised to do its work, without considering the reality that our environments are always shifting and evolving. Would a manager do this with a human worker? The answer is no. – Varun Ganapathi, Ph.D., co-founder and CTO at [AKASA](#)

AI introduces software development teams to the age of augmented analytics. AI's next shining moment will be empowering humans with data-driven recommendations for business decisions, across industries, in the form of augmented analytics. With an increased emphasis on governance and risk, we are going to see AI predict risk around software release schedules and tell companies *why* that release is at risk, providing deeper insights and allowing companies to avoid detrimental errors like the ones Facebook and Twitch could not. – Florian Schouten, VP of Product Management at [Digital.ai](#)

More and more enterprises are going to stop forecasting and doing Quarterly Business Reviews (QBRs) and will instead, rely on real-time data and AI to provide accurate intelligence on what is working, what isn't working and what will be working. – Art Harding, Chief Operating Officer at [People.ai](#)

AI will become a key part of remote collaboration: Recently, Zoom added a new feature fueled by AI which allows organizations to directly upload their meeting recordings into a lockbox powered by the company. In the new year, we can expect to see even more growth when it comes to collaborative tools as we continue to work from home, however, AI will be utilized much more to ensure more efficiency and collaboration in the workplace. – Doug Wilson, CPO at [OnBoard](#)

If enterprises want their investments in AI to pay off – and according to [PwC](#), 86 percent of 1,032 business and technology executives now consider AI a “mainstream technology” – they need to embrace a new standard that ensures AI is used in a way that is explainable, ethical, and most importantly, responsible. Fortunately, such a standard now exists. [IEEE 7000](#), released on September 15, 2021, provides businesses with a systematic, transparent, and traceable framework to developing AI platforms, ensuring they address ethical and regulatory obligations every step of the way. I've long believed that transparency and *ethics by design* is the only way for businesses to responsibly optimize their investments in AI. As we ring in 2022, IEEE 7000 is a big step in the right direction. – [FICO](#) Chief Analytics Officer Scott Zoldi.

AI regulation will start to look like data privacy regulation – AI is predicted to change just about everything. However, there is a lot of debate about whether those changes will be for the good. Data bias in machine learning models is one of the hottest topics in the AI industry for good reason; an AI model that rejects loan applications or increases insurance premiums for the wrong reasons will have a very deleterious effect. And there are other concerns. Many companies, especially in social media, are essentially in the business of collecting personal information. What can they do with that information? What are they allowed to learn about people and what are they allowed to do with that knowledge? The [EU already has a draft AI regulation in place](#), and in 2022 we can expect to see many other countries move in that direction. Once again, compliance will demand an ability to know what data you have, where it is, and who has access to it. – Nick Halsey, CEO, [Okera](#)

Businesses will finally derive value from AI through contextual experiences – Though the benefits of AI have been lauded for years, it's challenging to find use cases of AI providing true organizational value. Non-contextual AI/ML can only go so far. For content creators, contextualized intelligence will be a game-changer – especially when it comes to managing and searching for images. The CMS will eventually learn business lexicon to add context, provide warnings about sensitive content and provide guidance in terms of any content that needs to be trimmed down. – Nishant Patel, CTO, [Contentstack](#)

The interesting thing about ethics is that while they've never been more present than they are now, there doesn't seem to be much progress. Ethical challenges around AI are visible but workstreams aren't being substantially changed. Some early-stage companies are building things to help detect model drift and people in the industry are being asked to police themselves. However, for the most part, we don't. It's a difficult problem and I foresee more regulation around AI and the use of personal data. To progress and create more ethical AI, there needs to be governance within the industry that doesn't rely on self-monitoring. – Kevin Goldsmith, CTO of [Anaconda](#)

In contact centers, consumers and agents will no longer fear AI, they will become reliant on it – AI has quickly gained widespread acceptance in the business world and has proven to be an important element in business processes. As the labor shortage continues to persist, businesses can't risk burning out their agents so they will look to AI technology to help offload mundane tasks agents dislike while augmenting their capabilities to solve customer issues. Without a doubt, in 2022, consumers will also embrace artificial intelligence to help make their lives easier while preserving their ability to speak to humans. – Patrick Ehlen, VP of AI at Uniphore

New emerging AI models will deliver more individualized shopping experiences in 2022. AI and machine learning are invaluable to creating better shopping journeys which are individualized for each customer and vital to determining customer intent and influence behavior in real time. We'll see AI-powered personalization evolve to deliver more highly customized experiences in 2022. This entails machine learning models that consider all customer behaviors as well as various data sources that can be fed, ingested, and leveraged to better understand consumers at scale. We've already moved past the one-size-fits-all AI

model, to algorithms that meet the demands of individual customers without having to “test and learn” each time. For example, new AI models can be used to power product carousels that are based on your location, or behavior, or even the weather. AI will continue to play an important role in advancing personalization and optimizing the product discovery journey in 2022. – Tracey Ryan O’Connor, Group Vice President at Qubit, recently acquired by market-leading AI-powered relevance platform [Coveo](#)

While datasets are only getting bigger, many companies are now starting to understand and learn from the data they’ve acquired, with AI playing a critical role in surfacing key insights. However, in order to continue capitalizing on the incredible value of these vast datasets, AI practitioners will need the right tools and compute power to maximize productivity and deliver even faster time to insight. In 2022, we expect organizations to explore hybrid remote computing models, as well as end-to-end solutions in order to drive efficiency and turbo charge productivity in the workflows that challenge AI practitioners the most. – Mike Leach, Sr. Manager, Worldwide Solutions Lead for the Lenovo Workstation & Client AI Business

Data over algorithms – Expert opinion is coalescing around the idea — championed by AI pioneer Andrew Ng — that the best way to improve AI performance is with better data, not better algorithms. That’s not to say algorithms aren’t important, but we’ve reached a point of diminishing returns. Research suggests organizations can improve AI performance much more, and much faster, by training existing algorithms on wider data that’s carefully curated. In 2022, we’ll see access to external data emerge as a strong competitive advantage. Where before businesses might have raced to be first with AI, now they’ll aim to outperform competitors by training their AI on the most up-to-date, relevant data. – Omer Har, Co-Founder and CTO of [Explorium](#)

In 2022, AI will continue to grow as a valuable and critical workload for enterprise organizations across industries. We will see a larger number of teams investing in world-class AI computing to accelerate their research and business. With this, the need for faster, more power efficient, and purpose-built AI compute will continue to grow rapidly along with applications, models, and datasets. Companies leveraging AI as a key strategy for their business growth will need faster time-to-solution from their AI computing infrastructure, more scalability, and broader accessibility through diverse consumption models. In terms of AI models and use cases, we anticipate a continued expansion and use of large language models for text and other sequence data modeling problems, with increased attention being paid to more parameter- and data-efficient models and methods. In computer vision, we will see increased use of high-resolution 2D and 3D image datasets and video, which will lead to greater demand for purpose-built AI compute platforms with greater performance and efficiency at scale. We also expect to see continued development and greater adoption of graph neural networks for industry applications ranging from drug discovery to finance to social network analysis. – Andy Hock, Head of Product at [Cerebras Systems](#)

Startups focusing on AI-driven software development will continue to see increased investments and generally ML and AI will start playing a bigger role in all aspects of the software delivery supply chain. While ML and AI today is largely siloed inside various pieces of the supply chain, more connected analysis across the toolchain will see an increase. Perhaps the biggest driver for this will be VSM, and its goal to collect and correlate data and metadata from across the supply chain. – Shawn Ahmed, CMO, [CloudBees](#)

Full-Stack, Problem-Specific AI Thrives as Generic AI Fades – Pre-pandemic, AI was a nice-to-have for many industrial companies but over the last two years they were forced to rely on AI and other digital technologies to solve urgent, real-world problems in supply chains and production. As a result, investment focused on full-stack AI solutions (which includes the hardware required to gather data as well as the Machine Learning models using the data) that can solve specific problems fast, rather than more generic AI tools that have to be trained and customized by customers before they show value. – Artem Kroupenev, VP of Strategy at [Augury](#).

In 2022, AI will no longer be one big, complicated tech and instead a network of hundreds. In recent years, artificial intelligence (AI) has become a technological behemoth. With so much speculation surrounding the technology, its implementation and business use cases, many organizations have yet to scratch the surface of its far-reaching capabilities. Expect this to change in 2022 as business leaders increasingly realize that the path forward for successful AI is with multiple, narrow use cases of human-led technology that is designed and deployed to accomplish specific tasks. This emerging approach to and application of AI will spark the start of projects designed to have the different AIs communicate and coordinate with each other, rather than relying on one large, monolithic initiative. At the end of the day, I predict



this will generate more seamless and integrated experiences across the entire landscape. – Joshua Feast, Co-Founder, CEO at [Cogito Corp.](#)

In 2022, more brands will use conversational AI as their first point of contact with customers. The explosive growth of conversational commerce will require even more automation so brands can keep up, handling everything from product recommendations and purchases to customer service complaints and returns. – [LivePerson](#) CEO Rob LoCascio

AI-driven assistants will largely take over the troubleshooting process in networks. They say video killed the radio star, and now artificial intelligence (AI), natural language processing (NLP) and natural language understanding (NLU) are going to kill the “dashboard star.” The days of hunting and pecking or looking at charts will go to the wayside when you can literally just type in a question and get an answer, or have issues flagged for you and in some cases actually fixed on their own – known as self-driving. You’re going to see a trend around AI-driven assistance replacing dashboards and changing the way we troubleshoot, essentially eliminating the “swivel chair” interface. – Jeff Aaron, VP of Enterprise Marketing for Juniper Networks

A small data approach to AI will gain even more momentum in 2022. People are finally asking the right questions on the data used to power AI. When things such as the Metaverse arrive, this is going to be even more important. Just think of all the data that will be used to build that environment and how it will influence everything that happens there. Based on what we’ve seen so far do we think this will be done right? Given the track record of the social media giants over the last decade I’m not confident. It is more important than ever for us to move beyond the problematic big data approach where there is no control or accountability in what is being fed to AI models. – Dr. Lewis Z. Liu, [Eigen Tech](#)

Within the next year, AI companies will continue to improve data collection methods and develop processes that avoid bias in algorithm training and, in turn, performance in the intended population. Specifically, improved clinical study design will foster more heterogeneous and representative patient populations, resulting in algorithms that reduce bias. On the technical side, methods will develop to provide greater insight into the “black box” of AI algorithm decisions, which will guide understanding into whether these decisions represent bias based on factors including race, gender and age. – Mark Day, EVP of research and development at [iRhythm](#)

Data Consumers Will be Augmented by AI – While some are skeptical about the notion of a citizen data scientist, it is easier to predict the rise of a citizen data analyst: someone who consumes and leverages data, metrics, and insights as a natural extension of their job description. Today, nearly everyone – knowingly or implicitly – leverages data. Better APIs and low-code platforms are a great step in the right direction, but a breakthrough of scale in the use of data requires that every human – of all data skill levels – should expect some help from the systems they use. In the past, it was help functions and training. Increasingly it must be AI, and that AI must seem natural or be invisible. AI will increasingly empower those who lack the technical skills of a “power user”, by creating a bionic analyst. To achieve this, software vendors will need to improve their understanding of how data consumers use and gain advantage from data. – Kylligence CEO and co-founder Luke Han

## **Big Data**

The unrelenting pace of innovation will continue in 2022, and the gap between the ‘haves’ and ‘have nots’ will likely increase. The ‘haves’ understand that everything is being powered by software and they’re mastering the software development process with quality, speed and high levels of collaboration. Companies like Facebook, Apple, Amazon, Netflix, and Google (FAANG) have been teaching us this for years. In August of this year, the FAANG had a combined market cap of \$7.1T and made up approximately 19% of the S&P 500. These companies understand that great achievements come from the continuous release of software improvements – not ‘digital transformation projects’. To ramp up release momentum, mainstream companies will be forced to build stronger continuous software development muscles with companies like GitLab and incorporate a DevSecOps/GitOps approach to designing, building, testing, deploying and managing their applications at scale. These innovations will need to be released onto an autonomous and composable infrastructures like Upbound, a Telstra Ventures investee and open source Crossplane, which leverage the power of Kubernetes to deliver high levels of flexibility, automation, resilience and speed. – Steve Schmidt, General Partner, Telstra Ventures

Collaboration across the supply chain will become the norm. Collaborative systems that bring in data from all points of the supply chain will tell us where to apply our effort to make

changes and improvements. Decentralized technologies will allow us to scale and enable us to pinpoint pain points and surgically solve problems. – [Higg](#) CTO, John Armstrong

Data will become even more dynamic. As Greek philosopher Heraclitus once said, “There is nothing permanent except change.” A big change in 2022 will be—change. Data will change faster and more frequently than ever before. It will no longer be acceptable to analyze massive amounts of static data once per month, once per week, or even once per day. Organizations will need to glean insights from streaming data in real time to find new patterns and discover and act on them. Navigating data is like running whitewater, where you need to adapt instantly to a changing environment. Those that learn to run the rapids will succeed. – Aerospike’s Chief Strategy Officer Lenley Hensarling

The Rise of Data Fabrics – From theoretical architecture to actual implementation. Data and Analytics leaders ahead of the curve will begin to evaluate and build data fabric architectures – an approach to data integration that focuses on data agility across a complex and distributed environment. Data fabrics make data discoverable and accessible in real-time to data consumers, regardless of where that data physically sits – with centralized security and governance policies built right in. This set of functionality will allow organizations to adapt to consumer needs in real-time, build a more cohesive analytics experience, and power operational AI applications. However, today, the term “data fabric” often refers to a hypothetical wishlist for an enterprise data strategy – one that combines best practices from data governance, operations, security, analytics, and orchestration – but has yet to physically manifest. But with data management technologies maturing, data compliance policies forming, and data agility becoming a recognized competitive edge for enterprises, 2022 will see the first generation of implemented data fabric architectures. – Brian Platz, co-CEO and co-founder of [Fluree](#)

The role of the chief data officer needs to keep up with the dynamic evolution of traditional data functions, such as storage, architecture, modeling, forecasting, business intelligence, and analytics, especially with the current digital-first landscape. As companies continue to implement this critical role in their operations, it will increasingly need to be involved in the additional areas of strategy, product, ethics, and legal. – Dave Costenaro, Chief Data Officer, [Capacity](#)

Business leaders around the world will see the subjective qualities of data and realize the need to treat it as an asset to successfully enable businesses. The pandemic accelerated the need for businesses to digitally transform and rely on data to increase operational efficiency and remain competitive in the market. Yet according to a recent survey, 78% of executives have challenges making data-driven decisions and 60% don’t always trust the data they use. Treating data as an asset that can be measured, trusted, and acted on will provide healthy data for businesses to make critical decisions that drive business outcomes. – Christal Bemont, CEO, Talend

One of the biggest challenges DevOps engineers will continue to face is data gravity – the pull of the accumulating data black-hole and resulting lack of data mobility. This slow delivery of data is expected to double annually from now until 2024. Data gravity thwarts data movement, and our ability to keep up with dynamic customer demands, deploy applications seamlessly and produce efficient CI/CD pipelines. Moving data is costly and wastes valuable time. Unlike transporting apps, in cloud environments transporting data takes hours or days, and can create massive egress charges. Data gravity threatens the entire value proposition of elasticity. It’s harder to move the data required by applications than it is to move the actual applications. In 2022, IT professionals will need to implement innovative data services solutions to combat data gravity and the disruption of DevOps pipelines. This is the final piece needed to achieve freedom from data gravity and see gains in reduced complexity, cost and management. I predict we’ll see a collective move to advanced container native storage that can eliminate data gravity by enabling instant movement of data to and from any cluster anywhere and providing instant access to any point in time. These solutions can offer data the freedom to move as fast and easily as applications. – Kirby Wadsworth, CMO of [ionir](#)

The Cloud’s Growing Data Gravity Attracts Data Protection Solutions:

While organizations have been using data protection solutions to back up their on-premises data for years, many have been slow to use these solutions to protect their SaaS application data and other types of data they have stored in the cloud. However, as more and more organizations move both a higher percentage and more important data to the cloud, the cloud’s data gravity – its power to attract solutions, services and other data – has increased exponentially. The growing data gravity of the cloud is now attracting data protection solutions to it, as organizations seek to be able to backup and rapidly restore cloud-based data

after a cyberattack, misconfiguration, or other disaster. Further fueling data protection solutions growing attraction to the cloud is the fact that IT professionals are waking up to the fact that, under the SaaS and cloud service providers' shared responsibility model, they are responsible for all the data they store in their SaaS applications and elsewhere on the cloud. As [recent successful cyberattacks](#) on cloud-based data demonstrate, when organizations do not protect this data by creating a pristine, verified backup copy of it that they can restore after a successful ransomware or other cyberattack, the results can be devastating. Insurance companies are also causing the cloud to attract more data protection solutions. Faced with high ransomware payouts, insurance companies are now requiring their customers to put in place data protection strategies for their SaaS and other cloud-based data before they write or renew cyberattack insurance policies. On top of all of this, cybercriminals are launching more and more sophisticated ransomware attacks, and the damage caused by successful attacks is becoming more visible. All these reasons are why we can expect that in 2022 the cloud's growing data gravity will attract more data protection solutions, with practically every organization with SaaS or other cloud-based data having implemented a strategy to backup and rapidly restore this data before the end of the year. – Manoj Nair, General Manager at [Metallic](#), a [Commvault](#) Venture

Data and Analytics leaders ahead of the curve will begin to evaluate and build data fabric architectures – an approach to data integration that focuses on data agility across a complex and distributed environment. Data fabrics make data discoverable and accessible in real-time to data consumers, regardless of where that data physically sits – with centralized security and governance policies built right in. This set of functionality will allow organizations to adapt to consumer needs in real-time, build a more cohesive analytics experience, and power operational AI applications. However, today, the term “data fabric” often refers to a hypothetical wishlist for an enterprise data strategy – one that combines best practices from data governance, operations, security, analytics, and orchestration – but has yet to physically manifest. But with data management technologies maturing, data compliance policies forming, and data agility becoming a recognized competitive edge for enterprises, 2022 will see the first generation of implemented data fabric architectures. – Brian Platz, co-CEO and co-founder of [Fluree](#)

Data is not the new oil; data is a renewable energy source. Data, after it is transformed into useful insight through analytics, continues to increase in value and that value will exist in perpetuity – the more that is extracted from it, as opposed to oil, which is burned and then gone. – David Sweenor, Senior Director of Product Management at Alteryx

A day of reckoning will come for organizations using data centralization. The concept of data centralization for threat detection and response had a chance of working when data volumes were small, housed on-premises, and protected by a security perimeter – but, even then, it was a lofty goal. In today's world, it's impossible. There are new technologies producing different data types, formats, and sources; data lives in disparate silos across many different environments, including on-premises, on the cloud, and within SaaS apps; and data volumes have skyrocketed – all of which have eradicated the reality of universal data centralization and a single pane of glass. Today, organizations must modernize their security operations to deal with decentralized, distributed data from a variety of tools and platforms, and this means thinking outside the box. – Andrew Maloney, co-founder of and chief operating officer at [Query.AI](#)

Continuous Intelligence for More Agile Business Decision-Making: Businesses have more data and more data sources to handle than ever before. As manufacturers and other businesses are pushed to deliver new product ideas with greater efficiency, new data analytics models such as augmented analytics and continuous intelligence (CI) will be essential to ideation and critical thinking for advancement. For instance, with CI, real-time analytics are integrated into business operations, enabling users to get the most out of their data. Since CI exists in a “frictionless state,” businesses can leverage these continuous, AI-driven insights based on automated calculations and specific recommendations to make actionable, forward-thinking decisions, right as data events unfold. This more accurate information model benefits those business areas that need timely response, including supply chain, fraud detection, customer experience, and IoT-enabled manufacturing. – Sam Mahalingam, CTO of [Altair](#)

Unstructured Data Will Continue to Shape Data Management in 2022: Unstructured data continues to remake the data management landscape at a time when there not only is an unprecedented amount of data being generated, but it's also being collected, stored, processed and analyzed in multiple places (on premises, in the cloud and at the edge) and moved between those environments. Enterprises are using videos, images, IoT sensor data,

social media and similar information as foundations for much of the analytics, machine learning and business intelligence tasks they perform. It won't be a surprise to see unstructured data continue to be a focus of enterprises' data management efforts as we roll into 2022. – Krishna Subramanian, President, COO and Co-founder of [Komprise](#)

The democratization of real-time data follows upon a more general democratization of data that has been happening for a while. Companies have been bringing data-driven decision making out of the hands of a select few and enabling more employees to access and analyze data for themselves. As access to data becomes commodified, data itself becomes differentiated. The fresher the data, the more valuable it is. Data-driven companies such as DoorDash and Uber proved this by building industry-disrupting businesses on the backs of real-time analytics. Every other business is now feeling the pressure to take advantage of real-time data to provide instant, personalized customer service, automate operational decision making, or feed ML models with the freshest data. Businesses that provide their developers unfettered access to real-time data in 2022, without requiring them to be data engineering heroes, will leap ahead of laggards and reap the benefits. – Dhruva Borthakur, Co-Founder and CTO of [Rockset](#)

Increasingly data is becoming the currency of competitive advantage. The size of data packets, speed and frequency of data transmission and update, and the “intelligence” of data handling, are critical factors for successfully generating revenue expansion opportunities. In 2022, the amount of data to be harnessed and managed will grow exponentially. Intelligent data platforms will be a requisite to facilitate innovative architectures that can handle the escalating streaming data volume. – Sean Bowen, CEO of [Push Technology](#)

Data Intensity will be new KPI – The concepts of data intensity and complexity will be widely adopted in the coming years to measure digital dexterity, as organizations need to drive data intensity without adding complexity. Data intensity increases naturally as more constraints are connected to the data: variety, volume or velocity, geographic distribution, diverse types and structure, diverse use cases, automation privacy, security, number of producers and consumers. Data intensity is positive, but if not properly managed will lead to complexity that adds cost and friction. While data intensity today is mostly an attribute of applications, I predict that by 2024 the majority of organizations will have objectives, key results and KPIs tied to data intensity to capture their digital maturity. – Oliver Schabenberger, Chief Innovation Officer, [SingleStore](#)

We live in a world where more digital businesses recognize that leveraging automation and analytics to support human-centric engagement will improve the quality of customer relationships and drive empathetic loyalty. In 2022 this trend will accelerate. Companies will prioritize the digitalization of big data, and in the process transform customer support from being a cost center to a growth driver. – Somya Kapoor, CEO of [Theloops](#)

Digital twins grow up – 2022 is going to be the year for companies to stop talking about the definition of a “digital twin” and start deploying twins. We're already seeing a move from experimentation to production at leading operators, and digital twins have evolved from historical data based static models to dynamic representations of real-time operations. Dynamic digital twins move beyond historical analysis to provide better predictions or even support simulation-based learning. Looking forward, we'll see more physics-based models deployed, and more use of AI for adaptive control with advanced digital twins. – Andy Bane, CEO of Element Analytics

We're reaching a tipping point with data centralization and automation. The early days of “big data” are over, and companies that have built systems to use data well are outcompeting those who can't. The winners are shifting gears to operationalize data and the corresponding insights to deliver business value: making better decisions and creating more personalized experiences. AI and other forms of automation will accelerate this trend as they're able to deliver step function increases in value, and lean, upstart teams are able to build world-class customer experiences. This will ultimately lead to wins for consumers and brands as they build long-term relationships. – Kevin Wang, SVP of Product, [Braze](#)

Today, there is more data, from more sources, spread across more clouds than ever before – nearly 80 percent of organizations store more than half of their data in hybrid and multicloud infrastructures. On top of that, data is fragmented and siloed, making it more difficult for leaders to discover, manage, and control their data – 79% of organizations are using more than 100 data sources, with 30% using more than 1000 sources. In 2022, we anticipate that data fragmentation will be the biggest pain point for CDOs and CIOs, and that companies with end-to-end solutions that can manage all types of data and make it

interoperable across siloed environments, will emerge as winners in the data landscape. – Jitesh Ghai, Chief Product Officer, [Informatica](#)

Automated Context-Rich Data Classification Goes Mainstream – Every piece of data within an organization represents a unique combination of business value and level of risk. As privacy concerns, cybersecurity threats, and compliance mandates gain intensity, the need for effective data classification is more urgent than ever. Classification systems help organizations set boundaries around data access, use, and modification, acting as a natural next step to protect data once discovery efforts are complete. But many organizations find the process challenging because the system is too cumbersome to gain widespread adoption. The sheer volume of data makes the concept of manual classification untenable, and just getting started seems daunting. In the new year, organizations can start simply by focusing on automation to better understand the value of their organization's data. – Kevin Coppins, CEO of Spirion

Life after Hadoop – In 2022, we can expect the continued decline of the Hadoop platform, even though like some tough weeds in your garden the roots and trailers of Hadoop will be hard to completely eradicate. Expect CIOs and data teams to continue to de-emphasize Hadoop and to continue the process of removing it from their production data stack. Also look for IT departments to continue to make their on-premise implementations look and function like the public cloud. In the near term, organizations may continue to use the Hadoop File System (HDFS) as a storage platform until a better private cloud storage solution can be devised. In reality, to protect existing investments, and to comply with local government regulation, organizations can't simply move all existing workloads and applications built on top of on-premise Hadoop to the public cloud. The on-premise data stack will continue to exist. A hybrid solution across the public cloud and private cloud will be a more practical approach. – Kyligence CEO and co-founder Luke Han

### **Business Intelligence**

Collaborative Mining – Collaboration and BI have been inseparable since the start of the pandemic. As employees started working remotely, there was an urgent need to quickly embed BI within workstreams and productivity apps like Teams, Slack, and Zoom. This, in turn, expanded opportunities for more collaboration with outside stakeholders, further breaking down the barriers of data silos while revealing the need to collaborate sooner. In striving to improve the way we come together around data, networks, and processes, we'll see the advent of "collaboration mining," enabling decisions to be tracked. This provides crucial auditability while simultaneously boosting trust with multiple stakeholders. – Dan Sommer, Senior Director, Global Market Intelligence Lead at Qlik (and former Gartner analyst)

In order to be a successful business analyst in 2022, you'll have to operate like a data scientist:

The recent uncertainty we've faced has been a catalyst for many organizations to adopt cloud solutions out of necessity. Departments like finance and legal that were once lagging in cloud adoption have been increasingly accelerating their use of automation and collaboration technology to keep up with the faster and ever-increasing pace of business. As a result, business leaders are not only collecting more data and at higher levels of granularity, but also finding streamlined ways to glean better insights and thus make more frequent strategic decisions. It used to be said that those who went to business school would become business analysts. However, the advancements in Artificial Intelligence and Machine Learning (AI/ML) have changed that, as we've learned just how critical the contextual peculiarities of the underlying data can be. Cloud technology providers are now enabling AI/ML to work out of the box, making it more accessible and intuitive than ever before. Moving forward, we can expect to see data science continue to extend through every corner of the business and increasingly leveraged by non-data scientist employees. – Sanjay Vyas, Chief Technology Officer at [Planful](#)

Metrics Stores Eclipse Business Intelligence – Although business intelligence will continue to grow at a healthy clip in 2022, it is not for everyone. Its use is skewed toward decision makers, dashboards, and reporting. But data is for everyone, and everyone is tracking key metrics (KPIs) one way or another. While tracking KPIs can also be a BI use case, the ubiquity of key metrics and their relevance to virtually every facet of the business and every user begs the creation of a metrics store. If you consider that BI is bounded by the notion of creating and depending on a single source of truth to analyze business operations, sales, and marketing, the broader more universal use of metrics can be thought of as a single source of reality. This reality applies to every user, not just executives and leaders. Therefore, it will be

metrics stores that will drive digital transformation, not BI. – Kyligence CEO and co-founder Luke Han

## Chatbots

As consumers are more willing to engage with smarter chatbots who can solve their issues faster in many cases, rather than waiting for a customer service agent, we will see a dramatic rise in use cases for conversational Artificial Intelligence chatbots, or “next-gen” chatbots in 2022. As it stands, [new research shows that intelligent chatbots are already table stakes in the realm of customer support](#). Not only are intelligent chatbots reshaping consumer expectations, they're also reshaping the future of customer support – and companies who aren't adopting the latest technologies are being left behind. Secondly, the research shows that a new segment of high-value chatbot users has emerged. “Power Users” seek out chatbots for advanced and highly personalized issues, such as managing a subscription, looking up an account balance, or initiating a return. These users will continue to grow in 2022, as next-gen chatbots become more and more mainstream and technology-savvy consumers turn to them for assistance. Finally, with the pandemic, live agent teams have had to become increasingly remote, which in turn has accelerated the pressure on organizations to recruit and retain talent. In 2022, chatbots will increasingly be used to relieve this pressure by improving agent efficiency while also boosting agent job satisfaction by allowing deeper focus on more complex and engaging issues. – Mahesh Ram, Founding CEO of [Solvvy](#).

Today's chatbots have proven beneficial but have very limited capabilities. Natural language processing will start to be overtaken by neural voice software that provides near real time natural language understanding (NLU). With the ability to achieve comprehensive understanding of more complex sentence structures, even emotional states, break down conversations into meaningful content, quickly perform keyword detection and named entity recognition, NLU will dramatically improve the accuracy and the experience of conversational AI. This will have two results: (i) Increase the amount of automation and human augmentation . It will be capable of real-time human assistance, such as supporting an employee through language translation or recommending responses based on behavior or based on skill level; (ii) Change how, for example, a customer or client perceives how they are being treated, with NLU delivering a more natural and positive experience. – Doug Gilbert, CIO and Chief Digital Officer at [Sutherland](#)

The implementation of AI has transformed service desk agents from password reset experts into automation engineers. As a result of the extra capacity, AI has provided to service desk agents, service desk agents are now reaching into tier two (support involves technical knowledge and is staffed by technicians who have troubleshooting capabilities beyond the tier one) and tier three (requires a person who has specialized skills over and above the work the techs do in tier two) to identify additional automation opportunities. This is resulting in resources across all of IT being freed up, while up leveling the skills of the service desk agents. In 2022, more companies will adopt an open platform where service desk agents can create their own no-code automations versus building their own chatbot or solutions that rely on vendors. This will make it very easy for any agent to create content and maximize service desk capacity. – Pat Calhoun, CEO and founder of [Espressive](#)

As voice technology grows, business leaders are deploying a string of new applications, which is leading to an increasing number of consumers turning to voice engagement. Voice-enabled chatbots provide several benefits like faster responses and zero time wait, better two-way interactions, enhanced customer experiences, and fraud detection, to name a few. – Joe Hagan, chief product officer, at [LumenVox](#)

Conversational AI: Last year, I predicted conversational AI will be used to make video games more immersive by allowing real-time interaction to flesh out character-driven approaches. This year, conversational AI is all work and no play. Companies will race to deploy new conversational AI tools that allow us to work more efficiently and effectively using natural language processing. Speech synthesis is poised to become just as emotive and persuasive as the human voice in 2022, which will help industries like retail, banking and healthcare better understand and better serve their customers. – BRYAN CATANZARO, Vice President of Applied Deep Learning Research, [NVIDIA](#)

Empathetic customer service will be the standard – Above all else, consumers crave a more personalized experience – particularly given the feeling of isolation and separation from the past year and a half. Whether it's chatbots powered by AI, self-service or traditional customer service agents, for companies to achieve brand affinity and customer loyalty, they need to find not only the right mix of customer communications – enough to help but not so much that it wastes time – but also identify the right mix of technology and human

communications to express the empathy necessary to address the complex and unique emotions of each customer. – Chris Bauserman, Vice President of Marketing, NICE CXone

## Cloud

We predict that 2022 will see the first public cloud vendor make their services available on another public cloud. This will trigger an arms race to disaggregate the most valuable capabilities from the overall service – from analytics to databases and AI/ML frameworks like NLP. We also predict that it won't be AWS to break the seal – they have no incentive here. The net result will be good for customers as it will have the effect of accelerating the trend of commoditization of cloud infrastructure and will pressure economics across the board. – MinIO co-founder and CEO Anand Babu Periasamy

Hybrid cloud adoption at enterprises will accelerate as companies embrace support for core systems. To satisfy the need for faster digital transformations enterprises need to embrace their core (legacy) systems in new ways. Hybrid cloud infrastructure that incorporates both cloud aspects and legacy systems working together as a unified whole without users needing to care about where one finishes and the next begins. This unified solution only works when the legacy logic and data is easily accessible through cloud-native services in an automated way. Many enterprises now recognize the need to take this hybrid approach and simplify how they work with legacy systems. – Zeev Avidan, Chief Product Officer at [OpenLegacy](#).

Revenge of the Rushed Migration—The pressure of the business imperative to adopt cloud at rapid speed during the pandemic will begin to unravel as it becomes apparent security slipped through the cracks in rushed migration. As a result, we will witness the rise of huge breaches due to simple cloud security misconfigurations and permissions errors. This will fuel the mushrooming of startups based on automation of cloud configuration, permission analysis and remediation platforms. – Archie Agarwal, Founder and CEO of [ThreatModeler](#)

The hybrid cloud conversation is now driven by public cloud vendors rather than infrastructure/on-premises vendors. For the last few years, hybrid cloud was championed by technology vendors who sold on-premises technologies, but now public cloud vendors are offering cloud-like experiences on premises. This is not a good or bad thing, but as companies decide how they will approach their hybrid cloud strategy they need to consider how much control they want to maintain. By handing their private cloud to a public cloud vendor, companies may lose some control and ability to customize, but they will gain a unified, consistent private cloud experience. Companies need to decide what will be best for their business, but overall the conversation has shifted with public cloud vendors taking the wheel. – Jesse Stockall, Chief Architect, Cloud Management at [Snow Software](#).

## Database/Data Warehouse/Data Lake

Data warehousing and its hipper offspring analytics and data science have once again put database technology in the limelight. What does this mean for enterprise customers? Going forward, we can expect accelerated commoditization driven by managed cloud offerings. Ever broader audiences will be able to analyze and query data without explicitly maintaining database servers and without any database administration skills. 2022 will accelerate this trend and see even risk-averse organizations moving to the cloud. Commoditization can also be expected to drive consolidation. The database market is woefully crowded with over 100 commercially successful database products on offer. Instead of maintaining specialized silos, more and more users will want to use just a few central services for all their data needs. Being able to move existing applications to a consolidated environment with fewer database offerings will translate into a better and simpler way of using data. IT leaders will appreciate the positive impact of consolidation on their bottom line. Top leaders will be able to take advantage of this market trend, leverage the agility it affords them, and turn it into an immediate competitive advantage. – Mike Waas, CEO of Datometry

The days of relegating graphs to specialty analytics projects and continuing to use relational databases for transactional systems will cease to be reality. Graph technology has gained the performance necessary to execute real-time transactions at scale, enabling graphs to replace relational databases as the central System of Record (SOR) for enterprises. By 2030 we will see leading enterprises creating a single data fabric consisting of multiple interwoven graphs, document and time-series databases that are used for real-time transactions as well as predictive, machine learning analytics as well as real-time transactions and the system-of-record (SRO). – Dr. Jans Aasman, CEO of [Franz Inc](#)

The first generation of databases were the Oracles and Informix and DB2. The second was this database sprawl where you saw the influx of DB2, Couchbase went public, and the other 300. The next generation of databases is the consolidation of these data platforms and types

into a database that can handle modern data, and do it in a hybrid, multi-cloud manner with extremely low latency. – Raj Verma, CEO of [SingleStore](#)

New stack both in the storage and the compute layer keeps innovating. Data Lakes are rising to prominence and structured data is transitioning to new formats. In 2022, open-source projects like Apache Iceberg or Apache Hudi will replace more traditional Hive warehouses in cloud-native environments, enabling Presto and Spark workloads running more efficiently on a large scale. – Haoyuan Li, Founder and CEO, [Alluxio](#)

OpenFlake – the Open Data Lake for warehouse workloads. Data warehouses like Snowflake are the new Teradata – they're locking people into proprietary formats. As users start feeling the burden of higher costs as the size of their cloud data warehouse grows, they'll start looking for cheaper AND open options that don't lock them into a proprietary format or technology. In 2022 it'll be all about the Open Data Lake Analytics stack, the stack that allows for open formats, open source, open cloud – and absolutely no lock-in. – Dipti Borkar, Co-founder and Chief Product Officer (CPO), [Ahana](#)

The rise of cloud-native databases: As the pandemic drove increasing use of online services, traditional database systems struggled to keep up with all the requests and new data that flooded in. In 2022, more organizations will remedy this by transitioning to cloud-native databases. Cloud-native databases provide improved agility, scalability, reliability and availability compared to traditional databases. Adoption of cloud-native databases will pick up particularly among enterprises in the e-commerce and finance sectors, which must support a massive number of customer transactions and rapidly expanding data volumes while having to create new apps in order to deliver new services. – Shen Li, Head of Global Business, [PingCAP](#)

Unleashed data lakes for business users: business users have long been able to visualize data in relational databases and cloud data warehouses, but data lakes have been restricted to advanced data analysts and data scientists for machine learning. The power of data lakes will become accessible to business users as analytics applications tap into them and simplify advanced analysis to become available to non-data experts. – Tellius – Ajay Khanna, CEO and founder

Accelerated Data Science Platforms Thaw Enterprise Data Lakes: Much has been written about data lakes forming the foundation for enterprise big data strategies. Enterprise data lakes are effective for large scale data processing, but their broader usefulness has been largely frozen for the past few years, as they are isolated and decoupled from machine learning training and inference pipelines. In 2022, data lakes will finally modernize through end-to-end data pipelines because of three inflection points: centralized infrastructure, the agility of Kubernetes-based applications, and best-in-class, fit-to-task storage. – SCOTT MCCLELLAN, Senior Director of the Data Science Product Group, NVIDIA

Databases 3.0: The Great Database Consolidation – The first generation of databases were the Oracles and Informix and DB2. The second was this database sprawl where you saw the influx of DB2, Couchbase went public, and the other 300. The next generation of databases is the consolidation of these data platforms and types into a database that can handle modern data, and do it in a hybrid, multi-cloud manner with extremely low latency. – Raj Verma, CEO, [SingleStore](#)

Data warehouses are dead! Hello open data architectures – We hear it again and again: data warehouses are expensive, and costs are out of control. Newer technologies like data lakehouses will gain even more traction in 2022 because they have more to offer the enterprise than older data warehouse models that lock them in and drive up costs. Companies are more budget conscious than ever and will be reevaluating their data management systems. With a lakehouse architecture, there's no need to ETL data from the lake into the warehouse. In its Dec 2020 report "Market Guide for Analytics Query Accelerators", Gartner noted that analytics query accelerators are working to "shrink the performance impact of the zone of confusion" and enable the data lake to provide sufficient optimization on the data, making it suitable for an increasing percentage of workloads. – Tomer Shiran, Founder and CPO of Dremio

Organizations will need a new purpose, vision, and mission for their data warehouse – Data warehouse users have traditionally been data engineers, data scientists, and analysts who are interested in complex analytics. These users typically represent a relatively small percentage of an organization's employees. The power and accessibility of a data platform capable of running not just in the data center but also in the cloud or at the edge will invariably bring in a broader base of business users who will use the platform to run simpler



queries and analytics to make operational decisions. Accompanying these users will be new sets of business and operational requirements. To satisfy this ever-expanding user base and their different requirements means a new purpose for the data warehouse (why it exists), a new vision (what it hopes to deliver), and a new mission (how will it achieve the vision). – Teresa Wingfield, Director of Product Marketing at Actian

### **Data Center**

The pandemic put an emphasis on digital transformation and the importance of cloud-based services. As we look to the year ahead, massive intra-data center traffic is multiplying the need for additional bandwidth and faster networking interconnection speeds. Current data consumption trends suggest an increasing demand for data and compute, and we are seeing a convergence of infrastructure for data centers and wireless as data centers move toward edge compute models that are tied directly into 5G networks. Meeting those demands requires advanced, reliable technologies that provide scalable, high-performance interconnectivity. Optical interconnect technology will be key in supporting the shift to next-generation data centers by enabling higher speeds with low latency and lower cost per bit. – Dr. Timothy Vang, Vice President of Marketing and Applications for Semtech's Signal Integrity Products Group.

Data Center Is the New Unit of Computing: Applications that previously ran on a single computer don't fit into a single box any more. The new world of computing increasingly will be software defined and hardware accelerated. As applications become disaggregated and leverage massive data sets, the network will be seen as the fast lane between many servers acting together as a computer. Software-defined data processing units will serve as distributed switches, load balancers, firewalls, and virtualized storage devices that stitches this data center scale computer together. – KEVIN DEIERLING, Senior Vice President of Networking, [NVIDIA](#)

“Capacity at cost” will become a key factor when determining any CIO's success. In years past, overhead has not been considered the primary business driver. The math used to be Platform Capacity/units-of-work; now it's (Platform+Operations Overhead)/units-of-work. This significantly raises the cost per unit of work analysis, and the pressure is on CIOs to drive that cost down over time. In the past, it was quite common to treat operational costs as a fixed burden, and as capacity grew so did the overhead at the same rate. Today, the overhead portion must decrease as economies of scale are expected. The same reliability and performance of the infrastructure is expected, but the operational plans must become smarter, to do more with less. – Song Pang, SVP customer engineering, NetBrain.

### **Data Engineering**

By 2024, data technology will have evolved to allow organizations to optimize for frictionless digital transformation rather than optimize for read/write of transactions or efficient scans of large datasets. Databases will be an active participant and orchestrator of decision support. Analytic assets such as model pipelines, networks, business rules will be a common form of metadata just as structural or descriptive metadata is today. Over the next two years we will also see more innovations that bring the data science and data engineering communities closer together. To reduce data movement and duplication, more data science workloads will execute in and near the database. The analytic database instance will support highly performant offline model training, and the operational database will support real-time model inference for monitored and continued online training. – Oliver Schabenberger, Chief Innovation Officer at [SingleStore](#)

In 2022, to enable more resilience, more data engineers and citizen data scientists will be needed. Streaming data technologies and real-time analysis of data streams enable automated sensing and decision-making to respond to, and even predict, needed adjustments to elements of a given supply chain. But getting value from real-time data in an automated way will require more building of models, tuning of models, and more data governance. Data scientists, data engineers, and citizen data scientists will continue to be in demand for those organizations seeking resilience. – Lori Witzel, Director of Product Marketing at [TIBCO](#)

ModelOps is hot. Working from home in the pandemic has accelerated collisions and collaborations between teams of data scientists, devops and model ops developers – to get data science apps into production. Emerging from this is a focus on converting ad-hoc processes into a controlled environment – for managing low code and code first components, processes for data flows and model connections, along with rules, actions and decisions. Continuous analysis of models actually in operations is also in focus – to assess

ROI of the data science app, model drift and model rebasing. ML Engineers are now in the middle of this – configuring deployment scenarios in hybrid cloud environments, working with data scientists, data engineers, business users and devops teams; and with app dev and design teams. – Michael O’Connell, Chief Analytics Officer at [TIBCO](#)

The Synthetic Data Revolution Will Create a New ‘Synthetic Data Engineer’ Vocation to Become of the Most In-Demand Jobs – In 2022, a new position will surface — the ‘synthetic data engineer’ — data scientists who handle the creation, processing, and analysis of large synthetic datasets in an effort to support the automation of prescriptive decision-making through visuals. This new vocation, a natural evolution of the computer vision engineer, is already emerging in larger companies, where synthetic data teams have sprouted. The synthetic data engineer will become one of the most sought-after professionals in the AI market as more enterprises and startups alike will need the skills to support their simulated data initiatives. Expect to see such job postings soar and more training courses to become available, to fill the 22% rise in computer and information research scientist jobs over the next 10 years ([US Bureau of Labor statistics](#)), of which CV (and synthetic data) engineers are a subset. In addition, we will see other data-related professionals reposition themselves as synthetic data engineers to take advantage of expanding opportunities. – [Datagen’s](#) executive team

AI Technologies Associated with Data Science Will be Used Increasingly by Data Engineers – Data engineers will increasingly use AI-based tools in their day-to-day work. To support this, more analytics vendors will incorporate AI programmatic capabilities in their platforms, opening up new opportunities for data engineers. This will also blur the line between data engineering and data science, providing new opportunities for innovation. – Matthew Halliday, Co-founder and Executive Vice President of Product at [Incorta](#)

### **Data Governance**

For decades, data governance has resided in a wonted state of managing official corporate records while relegating the bulk of enterprise data to be kept in the dark. However, modern analytics, legal, and privacy projects require access to all enterprise data. To meet these needs, organizations will need to expand the scope of their data management practices and technologies to include in-place governance. – Matt Adams, [ZL Technologies](#)

Data Governance Will Rely on MLOps – The best ML technologies have well-defined training sets and MLOps techniques to identify data at the right time, from the development process through training and testing. This MLOps transition parallels what we see in DataOps and what we saw with DevOps: you need to have good metadata to accomplish those processes. In the coming year, we will begin to see more crossover between data governance and MLOps because you need not just high-quality source data but also metadata to describe the data to feed into the MLOps process for development, training, and testing of those algorithms. – Matthew Monahan, Director of Product Management at [Zaloni](#)

Preparing for Data Governance: The Rise of Industry Clouds - A recent IBM [study](#) found 64% of C-Suite respondents agree industry-related regulatory compliance is a significant obstacle to cloud adoption. As organizations grapple with security and compliance – especially highly regulated industries such as the financial services sector and government agencies for example – cloud adoption is evolving towards specialized clouds. As these industries strive to meet the demands of today’s digital-first customers and constituents, industry-specific platforms will be key to helping them balance innovation and functionality with stringent compliance protocols. By choosing the right platform - one with built-in controls - they will be able to innovate at the pace of change, ensuring they don’t get left behind while their industry puts new regulations into place or modifies existing ones. – Hillery Hunter, GM, Cloud Industry Platforms & Solutions; CTO IBM Cloud

As we’re seeing more and more large organizations fully embrace the modern data stack, many are now grappling with how to govern what’s there. In 2022, there will be a huge amount of work within the industry to help organizations solve data governance. Expect to see advances in interoperability between tools and APIs that expose metadata. There will be an early push towards standardizing our understanding of what data is and what compliance policies apply to it. Software projects and vendors that don’t collaborate on governance are going to be left behind. Doing your own isolated thing won’t solve problems at scale for Enterprises. – Fraser Harris, VP of Product, [Fivetran](#)

A data governance ecosystem aligns to drive usage and adoption – Data governance has historically been looked at as a necessary burden, something imposed on an enterprise that limited agility and slowed innovation. This is no longer the case. Enterprises are now waking

up to the reality that data governance is a key building block of agility and innovation. As a result, in 2022, data governance will no longer be a mere checkbox in vendor solutions. Instead, an ecosystem, including data governance platform providers, compute vendors, and platform vendors, will align around delivering data governance capabilities as a way to drive usage and adoption. We are already beginning to see this emerge, and it will accelerate in the coming year. – CTO, Nong Li, [Okera](#)

The data governance/access control market will accelerate – In 2022, the demand will skyrocket for scalable, automated ways to author and evolve complex data access control policies, the need to simplify data policy management, and the desire to efficiently scale cloud data and analytics initiatives to an ever-growing number of internal and external data consumers. As data volumes grow and usage expands, it has become impossible to control who has access to what data, ensure proper compliance, and enable safe data sharing. Similar to how the separation of compute and storage was the foundation for data processing in modern data stacks, it's becoming necessary to separate the policy layer to scale data access. With the decoupling of the policy layer with automated data access control, there are many factors to consider in how this critical security and governance solution gets deployed. Without automatic data access control, organizations have no way of monitoring who is accessing what data, when, and for what purpose, jeopardizing the data's security and privacy. Data teams need and want the ability to deploy row access and column masking policies, as well as leverage object tagging while benefiting from universal cloud policy authoring and highly scalable and evolvable attribute-based access controls, and 2022 is the year to do it. – Matt Carroll, CEO, [Immuta](#)

An emphasis on operational data governance: too many organizations suffer from a lack of data lineage (in particular column-level lineage) and operational governance associated with proper data modeling. Both items have become a result of too much manual coding and poorly architected tools. – Armon Petrossian, CEO and co-founder of [Coalesce](#)

#### **Data Science**

MLOps will relieve data scientists from tedious tasks – Many labor-intensive tasks, such as preparing data, feature engineering, and training models, that involve repetitive, tedious, and extremely time-consuming functions, will begin to be automated in 2022. – Leah Forkosh Kolben, Co-founder & CTO at [cnvrg.io](#)

Continued focus on data equity: Societal biases and inequities can be present whenever data is used. I expect individuals and organizations to continue discovering errors, omissions, and blunders in their data where biases in collection and storage led to incorrect, misleading, and harmful outcomes. Continued focus on identifying and resolving these issues is important for both accuracies of conclusions and equity in data use. – Andrew Kasarskis, Chief Data Officer at [Sema4](#)

The standard in which organizations prepare data for analytics is set to be streamlined in 2022 as companies embrace analytics at the source. The traditional ETL (extract, transfer, load) process for preparing data relies on creating data copies, scrubbing them, and exporting them to an external platform. However, modern analytics technologies are enabling organizations to conduct analytics on original data stored in their native environments. By cutting out the analytics middleman, organizations can dramatically reduce data preparation times while also increasing the control and governance they have over data throughout the analytics process. – Ryan Splain, [ZL Technologies](#)

In 2022, companies will need to embrace the role of the “citizen data scientist,” which are employees who work with predictive/prescriptive analytics models but whose primary job function lies outside the field of data science and analytics. The data science field is one of the fastest growing, and with the workforce currently experiencing “The Great Resignation,” companies will need to make data science more accessible in order to help fill gaps on their teams. – Alicia Frame, Director of Product Management for Data Science at Neo4j

Previously, Python was perceived to be a hobbyist language used by R&D, data science, and machine learning groups within an enterprise. However, as data processing and machine learning have become core to the business, so too Python becomes core. As data sets increase in size, so too must Python be able to increase in handling datasets of size. There is now tremendous demand for Python at scale. Fortunately, there are several good answers to this question in development, coming both from traditional technologies like Spark and more Python native technologies like Dask. There are half a dozen different efforts today to solve this problem. – Matthew Rocklin, CEO of [Coiled](#)

Industrial data scientists emerge to facilitate industrial AI strategy – The generational churn occurring in the industrial workforce will inspire another trend: the widespread emergence of industrial data scientists as central figures in adopting and managing new technologies, like industrial AI – and just as importantly, the strategies for deploying and maximizing these technologies to their full potential. [New research](#) revealed that while 84% of key industrial decisionmakers accepted the need for an industrial AI strategy to drive a competitive advantage – and 98% acknowledged how failing to have one could present challenges to their business – only 35% had actually deployed such a strategy so far. With one foot in traditional data science and the other in unique domain expertise, industrial data scientists will serve a critical role in being the ones to drive the creation and deployment of an industrial AI strategy. – Bill Scudder, SVP and AIoT General Manager at [AspenTech](#)

Enter the Age of the Engineer/Data Scientist: As more manufacturing companies invest in AI and machine learning to tap the benefits of their application in simulation design, we're seeing a growing need for an unfamiliar skillset in the engineer's wheelhouse: data science. While engineers have the design knowledge, they are typically not equipped to handle data at scale or properly interpret its meaning. However, engineers have the analytical mindset to embrace data science. As hiring trends reveal, the time has come for the engineer/data scientist to emerge, where having expertise in data analysis along with engineering is essential to tapping the competitive advantages of AI. By leveraging insights from historical as well as real-time data, engineers can make quicker and smarter design decisions earlier in the process, resulting in shorter time to market and more innovative and efficient products. – Brett Chouinard, CTO at Altair

With the continued democratization of analytics, data scientists need to evolve from 'problem solvers' to 'teachers.' Organizations are now looking to fill these roles with someone who can articulate and explain – not just code to encourage people to be creative and think critically. However, there is an existing skills gap between data scientists as practitioners and those as teacher. – Alan Jacobson, Chief Data & Analytics Officer at Alteryx

Data science expands beyond the elite organizations. As the shortage of qualified data science graduates continues, expect to see the rise of the "citizen data scientist." Companies will scramble to find existing internal resources to champion the next-stage intelligent systems. The lack of qualified candidates also reinforces the need for AI and machine learning driven solutions driven by savvy business users for all but the largest organizations. [Data Science as a service](#) is expected to flourish in 2022 as more companies recognize the value of data to identify actionable insights but also as a key component of digital transformation initiatives. Big companies will invest in developing their own teams to support these initiatives, but smaller organizations will not be standing idle. These services will provide a diverse offering that companies will be able to work with depending on their level of maturity with analytics and data science. For larger companies, the more in-house data scientists will be involved in pricing segmentation optimization, thus increasing the need for "bring your own science" models. – Gabriel Smith, Pricing Expert and Chief Evangelist for [Pricefx](#)

### **Deep Learning**

As the toolset for AI applications continues to evolve, machine learning and deep learning platforms have entered the mainstream and will attain the same level of maturity as specialized data analytics. Just like we currently see a plethora of fully integrated managed services based Apache Spark and Presto, in 2022, we will see vertical integrations emerging based on the likes of PyTorch and Tensorflow. MLOps for pipeline automation and management will become essential, further lowering the barriers and accelerating the adoption of AI and ML. – Haoyuan Li, Founder and CEO, [Alluxio](#)

Concerted Efforts by Vendors to Reduce Bias in Speech Tech – Voice is the most natural form of communication. However, machines have historically been locked out of listening and analyzing conversations. In 2022, machines will be able to do more than just describe which words were said, but how they were said. This will enable users to truly understand what their customers want and empathize with their needs. Reducing bias in speech infrastructure will also be a top priority for vendors so that their customers can more accurately understand the voices of various backgrounds, genders, and languages of their users. – Scott Stephenson, CEO and co-founder of [Deepgram](#)

### **Graph**

In the past few years, organizations have experienced the advantages of combining Graphs with Artificial Intelligence. In 2022 and beyond, leading companies will apply Machine

Learning's advanced pattern matching to Graph Neural Networks (GNNs), which are complex high-dimensional, non-Euclidian datasets. By fusing GNN reasoning capabilities with classic Semantic inferencing available in AI Knowledge Graphs, organizations will get two forms of reasoning in one framework. Automatically mixing and matching these two types of reasoning is the next level of AI and produces the best prescriptive outcomes. This 'Total AI' is swiftly becoming necessary to tackle enterprise scale applications of mission-critical processes like predicting equipment failure, optimizing healthcare treatment, and maximizing customer relationships. – Dr. Jans Aasman, CEO of [Franz Inc](#)

Fervor for knowledge graphs/graph databases soared in 2021, especially at the business level. Adoption has increased across the board, from small businesses to large enterprises due to the ease of implementation. In 2022, this trend will not only continue but sprout new use cases in fields such as digital twin technology, patient journey analytics, biomarker detection and root cause analysis. – Maya Natarajan, Senior Director Product Marketing at Neo4j

Gartner indicates that data fabric is the foundation of the modern data management platform with capabilities for data governance, storage, analytics, and more. Relying on traditional integration paradigms involving moving data and manually writing code is no longer acceptable as data scientists and data engineers spend almost 80 percent of their time wrangling data before any analytics are performed. Shrewd organizations looking to adopt this approach are realizing that the centerpiece of a properly implemented data fabric is an enterprise knowledge graph, which compounds data fabric's value for better, faster, lower cost analytics while hurdling the data engineering challenges obstructing them. 2022 will be the year organizations adopt enterprise knowledge graph platforms to support their data fabrics that use a combination of graph data models, data virtualization, and query federation—along with intelligent inferencing and AI—to eliminate this friction by simplifying data integration, reducing data preparation costs, and improving the cross-domain insights generated from downstream analytics. – Kendall Clark, Founder and CEO at [Stardog](#)

#### **Hardware**

The chip shortage will impede overall IoT market growth by 10-15%. The chip shortage will not be resolved until mid-2023 and IoT devices will be hit worse. With many IoT-based "smart" products like appliances, automobiles, and consumer electronics being unavailable or overpriced, this will increase demand for "less smart" equivalents. – Forrester

#### **IoT and Edge Computing**

There will be a demand for highly scalable distributed asset management. The number of connected devices is [set to triple](#) by 2026 and no one is paying any attention. Enterprises are about to face one of their greatest IT challenges and are woefully unprepared. Large percentages of employees working from home on their personal computers will begin requesting that their employers provide them the hardware they need to continue working. Distributed asset management will become a major pain-point for large enterprises as they tackle the requirements of anywhere operations and the proliferation of connected devices. – Andrew Sweeney, co-founder and co-CEO of [ReadyWorks](#)

IoT, Edge Computing Solutions for Collaborative Remote Work Are Poised for a Breakthrough – IoT and Edge computing solutions that bridge the physical and digital worlds and make it easier for workers to collaborate remotely. Examples include tools such as digital whiteboards and the use of AR/VR to better replicate the experience of face-to-face interaction and remote monitoring and management of industrial infrastructure to minimize the need for on-site visits. "SaaSification" of business models that bring the simplicity of the cloud to edge computing use cases. This includes not only resources dedicated to specific end users but also multi-tenant edge infrastructure that multiple end users share, as is currently the case for public cloud resources. Increased standardization and no-code tooling for developing AI models, including those developed at the edge. TinyML will continue to accelerate, further underscoring that the edge is a continuum spanning highly constrained devices in the physical world to regional data centers. Meanwhile, the reality will set in that many solutions that providers market as "AI" today are really just rules engines. Collaboration on the concept of trust fabrics that deliver data across heterogeneous networks will continue to grow. Data trust is critical to drive new business models and customer experiences in addition to helping businesses comply with privacy regulations and protect themselves and consumers from fake data automated by AI. An example effort here is the Linux Foundation's new [Project Alvarium](#). – Jason Shepherd, VP of Ecosystem at [ZEDEDA](#)

Harnessing time and space data will be a major market opportunity. [Projections from Deloitte suggest that](#) 40% of connected IoT devices will be capable of sharing their location by 2025, up from 10% in 2020 – making geospatial data the fastest growing space in the data landscape and creating the potential for crisis within unprepared organizations. This acceleration of geospatial data will be driven by the declining cost of sensors, more satellites gathering time/space data, and 5G roll outs. This will open up new ways of using geospatial information. But managing fast-moving, high-volume location data in a reasonable time frame has always been a challenge, and these new devices will make it even worse. IoT data has always had a time dimension, i.e. logs from smart devices about their interactions and changes in state, but now the space dimension is taking off, and many organizations don't have the skills or resources to cope with the onslaught. This will force them to explore new approaches and technologies to get the full value of time and space data. Early adopters will have a huge market opportunity within their respective industries, while slower organizations will risk getting left behind. – Kinetica's co-founder Amit Vij

Over the next few years we will see more and more companies shifting scalable data operations to the edge. For the first time ever it is becoming economically and practically feasible to run Big Data software on-prem, due to Edge-as-a-Service solutions that can be operated at a fraction of the cost of traditional server infrastructure – and require significantly less maintenance. Enterprises will be able to run Kubernetes clusters and other powerful platforms at 1,000s of remote locations, enabling them to make complex, real-time decisions. – Dominik W. Pilat, Field CTO, [Hivecell](#)

Edge Compute and Edge AI Deployment Grow As We Strive for True Smart Manufacturing – Edge compute and edge AI/ analytics deployment will continue to grow significantly through 2022 as we continue to strive for true smart manufacturing that supports mass customization, 'lights out' production, and much improved KPIs. This naturally relies on a robust infrastructure that is more deterministic by nature, as well as providing the ability to move workloads seamlessly to adapt to the plethora of vectors such as raw material availability, production capacity, grouping, location, subsystems, and beyond. In 2022, it will be interesting to see how safety will impact this trend as machinery and robots become collaborative, mobile and 'uncaged'. This isn't a challenge only being faced within the boundaries of the factory itself. This will put additional constraints on the infrastructure design and deployment of such systems, but is a fundamental issue that must be solved. – [CoreAVI](#)'s head of industrial markets, Neil Stroud

### **Kubernetes**

The establishment of Kubernetes as the platform of choice for the deployment and efficient management of most enterprise workloads – at least directionally – is no longer a question. However, the variety of Kubernetes offerings available (across cloud providers and private or so called "hybrid" offerings) has only increased. With the ease of deployment of Kubernetes clusters also decreasing, the stage is well set for "sprawl" problems of the same variety that we have been seeing for multiple cloud accounts, and prior to that for VMs in the virtualized infrastructure world. This will make the ability to manage the lifecycle and applications on these differing Kubernetes providers in a consistent manner, as well as the ability to consolidate them easily, more and more critical. – Reza Shafii, VP of Product, [Kong](#)

In 2022, automation will be key to the evolution of Kubernetes. Enterprise adoption of Kubernetes is continuing to rise, but teams are often finding themselves hindered by the skill set required to manage and operate Kubernetes. The overwhelming majority (98%) of IT executives report they are facing challenges with Kubernetes and coordinating the related ecosystem, according to our 2021 Annual Kubernetes Adoption Report findings, a report we sponsored with research conducted by Dimensional Research. This obstacle with adoption creates a need for a more simplified approach to Kubernetes management from both a user interface (UI) and automation perspective. This will also help companies remain competitive amidst the backdrop of the IT talent gap, as more user-friendly and automated Kubernetes management will not require the sophistication or experience once needed by niche and very technical IT professionals for deployment and oversight. – Terry Fu, CEO & Co-Founder at [Spectro Cloud](#)

In 2022, there will be a rise in Kubernetes usage that will lead to more multi-cluster environments as businesses try to manage heavier traffic loads. According to a 2021 survey of over 1,000 software engineers, DevOps employees and IT architects, 70% of respondents reported using Kubernetes for a business project. The survey also found that 60% of Kubernetes users are running two or more clusters at a time, while 58% of companies are running less than half their applications in Kubernetes. 2022 will be marked by an increase in Kubernetes adoption that will lead to the use of multiple clusters. These multi-cluster

environments will result in great benefits including improved availability of service (even if there's an outage somewhere in the system), enhanced resiliency, and an increase in automation of processes in the cloud, reducing workload and minimizing room for human error. While the move toward a multi-cluster world will come with growing pains, the pros greatly outweigh the cons. — Emile Vauge, CEO, [Traefik Labs](#)

More and more companies are shifting from on-prem to the cloud. The first wave of this was to simply run their existing stack on Kubernetes on the cloud. This was already a great win because they didn't need to think about provisioning and updating hardware. However, managing Kubernetes clusters is still quite a challenge, even for mature IT departments. Instead, the real virtue of the cloud comes when other people are managing the software itself. We're seeing a growth of companies like Snowflake that offer SaaS directly to customers. This ends up providing a simpler and lower-cost solution, especially when you factor in the engineering time it takes to maintain these systems. Companies, especially larger companies, are becoming more and more comfortable trusting the cloud and cloud SaaS companies. As their trust increases, so too does their willingness to give up owning their own technology stack. Owning software has turned into a liability. The factors that will result in the pullback of Kubernetes toward more SaaS products include: (i) Increased trust in the cloud and cloud-based companies; (ii) Improved integration technologies like federated auth and Single-Sign-On (SSO) allow several disparately managed SaaS products to integrate smoothly; (iii) The rapid pace of innovation among software and an increased movement of decision-making from the top of the organization to the bottom. – Matthew Rocklin, CEO of [Coiled](#)

Customers want the flexibility of multi-cloud, while the cloud providers want to make their offerings as "sticky" as possible. For organizations looking for portability, Kubernetes is becoming ubiquitous. It abstracts the underlying cloud infrastructure and simplifies running applications and CI/CD pipelines at scale. As a result, all major cloud providers are either offering or promising to offer Kubernetes options that run on-premises and in multiple clouds. While Kubernetes is making the cloud more open, cloud providers are trying to become "stickier" with more vertical integration. From database-as-a-service (DBaaS) to AI/ML services, the cloud providers are offering options that make it easier and faster to code. Organizations should not take a "one size fits all" approach to the cloud. For applications and environments that can scale quickly, Kubernetes may be the right option. For stable applications, leveraging DBaaS and built-in AI/ML could be the perfect solution. For infrastructure services, SaaS offerings may be the optimal approach. The number of options will increase, so create basic business guidelines for your teams. – Stephen Manley, CTO, [Druva](#)

The harmony of Kubernetes, GPUs, and scale-out storage will be a long-term trend to support AI applications. New workloads like AI require massive datasets, a high degree of parallelization, and high-performance compute and storage. The global enterprise Artificial Intelligence (AI) market is anticipated to grow at a Compound Annual Growth Rate (CAGR) of 39.7% to \$USD 309.6 billion by 2026. This requires scale-out computing and storage performance, GPUs, and much more storage capacity. No single vendor provides all the pieces today. – iXsystems

Businesses will embrace development platforms to increase developer productivity – With tech giants winning the race for scarce developer talent, businesses outside of the tech elite will embrace new ways to stay innovative and competitive with their own teams. Businesses are waking up to the realization that they need technology that works hard to allow their development teams to focus on creativity and innovation, instead of the tedious aspects of software development. This includes technology that handles the critical but undifferentiated tasks of development, constantly updates with the latest cloud technologies, automatically scales, and leverages containers and Kubernetes to make sure development teams deliver world class application architectures, move fast to meet changing business needs, with low risk. All this while being unencumbered from toil, unnecessary maintenance and technical debt drudgery. – Patrick Jean, CTO of OutSystems

Container support along with Kubernetes orchestration will become mandatory data warehouse features in 2022 – Containers are key to enabling an organization to meet the resource demands associated with artificial intelligence, machine learning, streaming analytics, and other resource-intensive decision intelligence processing. These workloads strain legacy data warehouse architectures. Kubernetes orchestration automates the provisioning, deployment, networking, scaling, availability, and lifecycle management of containers. – Teresa Wingfield, Director of Product Marketing at Actian

Kubernetes will develop a greater position of dominance. Kubernetes will gather mainstream acceptance to support serverless workloads and virtual machines. As such, hosting and edge platforms built to support Kubernetes will have a competitive advantage in being able to flexibly support modern DevOps teams' requirements. Edge platform providers who can ease integration with Kubernetes-aware environments will attract attention from the growing cloud-native community. For example, leveraging Helm charts-as-a service (the emerging standard in application development language), where application builders can hand over their application manifest and rely on an intelligent edge orchestration system that just works. – Daniel Bartholomew, chief technology officer for [Section](#)

### **Machine Learning**

Organizations will focus on AI initiatives that augment human performance, not replace humans with machines. Up until now, the goal of machine learning for most applications has been the replacement of the human effort with machine effort. 2022 will see machines performing tedious, tactical tasks such as information retrieval, etc. which will enable humans to focus on higher-level, strategic tasks and decision making. Single use case machine learning models will give way to the centralization of institutional knowledge for use in business processes across multiple subject domains as well. Supervised learning require training with a large set of examples for a specific use case. The time and effort required to do so is both time- and cost-prohibitive for the average user. In the coming year, expect AI providers to focus on delivering platforms that centralize data/content for use across multiple business processes. For example, a sales account executive, product manager, and customer support representative might all draw upon centralized intelligence repositories to solve their business problems; i.e., using the same institutional knowledge for different purposes. – Ryan Welsh, Founder and CEO of [Kyndi](#)

AutoML – This allows non-technical users to make use of models and techniques without requiring them to become experts in machine learning. Automating the process of applying machine learning end-to-end additionally offers the advantages of producing simpler and faster solutions to “everyday” business decision challenges. – [NWO.ai](#)

Despite the industry hype around machine learning, there has long been confusion about what value it can bring to an organization. It certainly can help get to an answer faster due to its automated nature, but – it's not necessarily better – especially when the data its models rely on is inaccurate, outdated or incomplete, which is too often the case. Organizations are now realizing the need for gap-less intelligence to deliver true personalization, where various roles and departments need a certain level of information to customize the experience for the consumer. Whether it's a bank teller, a hotel receptionist or a business leader – intelligent insights driven by machine learning are essential for delivering the next best action for each consumer. But as marketers recognize these gaps in intelligence, because of gaps in their data, they'll begin to lose trust in machine learning. Without a focus on data quality, machine learning will be inefficient, gaps in intelligence will persist and the customer experience will suffer. – John Nash, CMO, [Redpoint Global](#)

As more organizations recognize the importance of AI/ML operations in their DevOps Platform, we'll see an increase in these practices in industries that you wouldn't typically expect, such as energy, shipping and manufacturing. We've already started the transition where every company is becoming a software company, and we're now seeing these software companies adopt AI and ML. Especially with the labor and supply chain shortages and dramatic shifts in climate related events, we're seeing that companies across the globe are having to learn to do more with less in even more dynamic environments. AI/ML is well suited to solve some of these complex problems in industries we may not have expected this early. – Taylor McCaslin, Principal Product Manager, Artificial Intelligence & Machine Learning, GitLab Inc.

With advances in machine learning, the proliferation of unstructured data will yield a whole new level of business intelligence, with endless new opportunities for collaboration, knowledge sharing, innovation, and better decision making. In fact, in the future, we won't distinguish between structured and unstructured data, we'll just focus on turning information into knowledge. – CEO of Onna, Salim Elkhoul

Despite the industry hype around machine learning, there has long been confusion about what value it can bring to an organization. It certainly can help get to an answer faster due to its automated nature, but – it's not necessarily better – especially when the data its models rely on is inaccurate, outdated or incomplete, which is too often the case. Organizations are now realizing the need for gap-less intelligence to deliver true personalization, where various roles and departments need a certain level of information to customize the experience for



the consumer. Whether it's a bank teller, a hotel receptionist or a business leader – intelligent insights driven by machine learning are essential for delivering the next best action for each consumer. But as marketers recognize these gaps in intelligence, because of gaps in their data, they'll begin to lose trust in machine learning. Without a focus on data quality, machine learning will be inefficient, gaps in intelligence will persist and the customer experience will suffer. – John Nash, CMO, [Redpoint Global](#)

#### **No-code/Low-code**

Adoption of no-code, low-code AI may prove to be the biggest strength for marketers and advertisers against the backdrop of all the privacy and cookieless web debates. Giving advertisers effortless, simple and quick access to tools that can aggregate first as well as third party data, will allow them to nimbly navigate publisher and platform-specific restrictions. – Shubham A. Mishra, CEO and Co-Founder of [Pyxis One](#)

Low-code adoption will surge — but among professional developers, not business users. Demand for applications has skyrocketed and the supply of professionals remains low. Experiments with having power users build business applications are yielding fragile solutions with limited scope. Since you can't add pro developers and you can't add amateur developers, and you already believe in new methods, the only thing left is to adopt higher-productivity tools. Some low-code tools and platforms are stepping up to meet enterprise requirements for security, scaling, continuity, auditing, monitoring, deployment, and change management. They'll get adopted if they get noticed. – Mike Fitzmaurice, [WEBCON's](#) VP of North America and Chief Evangelist

IT teams will turn to low-code automation to ease their increasing workloads. We're beginning to see more automation and low-code tools designed to alleviate the pressure on IT teams. Managing solutions that utilize machine learning and AI to recognize threats, providing additional security, upholding policies and compliance are all responsibilities of IT departments. With the addition of navigating a distributed workforce, increased LOB requests and concerns around infosecurity, IT teams are completely overburdened. Low-code automation takes a bulk of the manual labor off of the plates of IT and engineering teams and allows them to focus on the core functions of their departments to support the growth of an organization. As more repetitive, manual processes are automated, skilled IT professionals can focus on revenue-generating tasks. This renewed focus on core products will result in a happier customer base and set a foundation for more tech innovation that can help the organization grow faster without IT teams being a bottleneck. – Rich Waldron, co-founder and CEO of [Tray.io](#)

Companies will continue to move away from Cloud-First to Cloud-Only – Infrastructure-as-Code low-code and even no-code platforms will make it increasingly simple (and a smart business move) for organizations without cloud-savvy DevOps engineers to migrate to the cloud and unlock new opportunities for innovation and productivity. – Venkat Thiruvengadam, Founder and CEO, [DuploCloud](#)

In 2022, automation will grow beyond the Security Operations Center (SOC) to serve as a system of record for the entire security organization. As companies struggle to adequately staff security teams—and fallout from 'The Great Resignation' adds additional stress across the organization—automation will help employees overcome process and data fatigue. Companies will seek to use low-code automation to harness the collective knowledge of their entire security organization and form a centralized system of record for operational data. – Cody Cornell, Co-Founder and Chief Strategy Officer at [Swimlane](#)

As no code and low code platforms and tools become more pervasive, methodologies and tools from the software development and DevOps worlds, such as automation, version control and declarative languages will be applied and added to these environments. Why? As business applications become more complex, the frequency of changes in them accelerates, and the teams that support them grow – so the need for structure and agility becomes more pressing. Code brings structure and enables agility in growing teams. The introduction of complementary DevOps tools and methodologies is critical to the success of no-code and low-code environments as they currently lack the much-needed structure—holding many business applications development projects back. – Gil Hoffer, Co-Founder and CTO of [Salto](#)

Democratization of ML through up-skilling will make more analysts comfortable with code. For over 20 years, different products have promised to enable advanced analytics with “no code” or “drag and drop” user interfaces. The latest wave of this trend will lose enthusiasm in favor of companies investing to upskill their workforce. Analytical programming languages like Python and R will become more table stakes (especially with the rise of data science

degree programs in secondary education), just as Excel and SQL became a decade ago. – [Domino Data Lab](#) CEO Nick Elp

2022 will continue to see growth in low-code/no-code solutions and the rise of the citizen developer movement. Access to enterprise-grade technology has been democratized, allowing companies of all sizes to modernize. Similarly, access to purpose-built solutions designed for the unique use cases of a business will empower employees to both identify issues and to implement and iterate ways to solve data and process challenges. – Joe Hagan, chief product officer, at [LumenVox](#)

Low-code/no-code technologies will continue to grow in 2022 – Low-code enables companies to operate at the scale and velocity that is required to win in today's digital agile-first market. Interest in low-code development is skyrocketing. Annual market growth is predicted to exceed 25%, growing from \$13B in 2020 to \$65B in 2027, according to research from [Brandessence Market Research](#). And for organizations looking to win in the digital-first agile world, low-code is quickly becoming a critical component of a modern enterprise technology stack. Low-code offers companies all three attributes of faster, better, cheaper in the same package. However, to help accelerate utilization of low-code and scale it across the enterprise, process intelligence is a key enabler. You cannot improve how you operate tomorrow if you don't fully understand how you work today. And, most companies truly don't understand how they operate on a daily basis, especially at a granular user activity level required to automate a process or streamline a workflow. According to [Gartner](#), more than 65% of application development in 2024 will be performed by low code platforms making it impossible to argue against the demand for low-code in the enterprise. Which is a remarkable shift for a software category that did not exist a decade ago. And the number of digital applications and services being built is exploding as well. Between 2018 and 2023, more than 500 million apps will be created according to [IDC](#) which exceeds all previous development to date. – Jon Knisley, Principal, Automation & Process Excellence, [FortressIQ](#)

As the number of organizations and products in the data management space continues to grow, the automation, self-service and no-code capabilities they are able to provide is making it easier than ever to get started building a modern data stack. However, composing all of these separate components together and knowing what to run and when is difficult to scale. Data orchestration tools are making that both possible and easy, which will result in huge growth in data orchestration in 2022. It will be exciting to see where open source projects like Airflow, Prefect, and Dagster take it. – Nick Acosta, Developer Advocate, [Fivetran](#)

Codeless reporting tools and predictive analytics will take the data analytics world by storm. Companies like ThoughtSpot, Sisu Data, and Canvas are making it more accessible and easier than ever for less-SQL savvy analysts to work with data, while simultaneously freeing data scientists and engineers up from routine ad-hoc requests and dashboard maintenance. – Barr Moses, CEO & Co-founder of [Monte Carlo](#)

Low-code will gain the respect of professional developers: Experienced developers have traditionally had a cynical view of low-code / no-code programming, and with good reason. It hindered collaboration, prevented complex problem-solving, and reduced access to underlying code, rendering them very limited. However, as the developer shortage escalates and new approaches to low-code emerge, in 2022 the resistance will start to wash away, and CIOs will accelerate adoption of low code, not merely to complement other tools, but to increase productivity of developer and business users enterprise-wide. – Asanka Abeysinghe, Chief Technology Evangelist, WSO2

Business workflows will become self-learning: Rather than needing an advanced data science degree or a Ph.D. in statistics to take advantage of AI and machine learning, professional and citizen developers will be able to use the visual tools low-code and no-code platforms provide to build dynamic business processes that are self-learning. – [CLEVR](#) CEO Angelique Schroufen

Low Code/No Code for Analytics – Right after talking about the value of APIs, wouldn't it be great to be able to execute effective analytics without coding at all? In today's world of continuous integration and continuous delivery (CI/CD), taking a week or less to create a meaningful data application is becoming increasingly common. But shrinking the development pipeline for a larger population requires a lower bar to entry into the application creation process. The growing popularity of low-code/no-code platforms in the coming year promises to accelerate that trend. Because low-code/no-code platforms enable non-coders to build their own data apps organically, the logical next step for such platforms is to tackle human curiosity and creativity as it applies to analytics. Fundamental to making this happen is the creation of modular workflows for data management and data pipelines,

and a greater adoption of robotic process automation (RPA) for data services. This is the next wave of automation for analytics, machine learning and AI. – Kyligence CEO and co-founder Luke Han

### **Observability**

Insightful data through Observability – Companies are all operating with and have access to the same data – all from the same systems. Companies that will demonstrate a leadership position in 2022 will use that data to better inform customers of what is happening in their infrastructure. How can companies better utilize the data and apply intelligence to help customers make decisions and sift through all the noise? End users expect deep insight into their data and expect vendors to offer the best experience they can to identify and resolve issues and give them observability of their systems. With true observability, we can give customers back time to focus on what really matters, which is managing the digital experiences for their customers, and empowering them to meet their customers' needs. – Frank Reno, Principal PM and Open Source Ambassador, Sumo Logic

In 2022, data will grow increasingly vital to organizational success. System uptime and application performance will demand incremental improvements as organizations work to edge one another out and claim market share. New waves of cybersecurity attacks will force novel approaches, and piecemeal data will no longer suffice. IT leaders will seek data observability solutions that can provide a holistic picture of their distributed infrastructure in real time while allowing for speed and scalability. Traditional solutions unable to log all an organization's data due to cost or technical limitations will lose traction. Furthermore, the cloud will cement its place as a prerequisite for collaboration and speed. Solutions that empower complete observability over the cloud and on-premises environments in real time will be the major winners of 2022, especially when it comes to log management. – Geeta Schmidt, VP and [Humio](#) Business Unit Lead at CrowdStrike.

### **Quantum**

With recent advances in Quantum Computing, in 2022, we will start to see the convergence of Quantum Computing with Artificial Intelligence, Knowledge Graphs and Programming Languages. These distinct technologies will start to morph into a single computing environment operating in one memory space as a fully integrated solution. The separation between programming and AI/Analytics will begin to blur as developers use Quantum-based computer languages to generate incredibly complex, next generation AI algorithms and applications that result in new discoveries based on the quantum acceleration of machine learning and deep learning. – Dr. Jans Aasman, CEO of [Franz Inc](#)

### **Robotic Process Automation**

Democratization of Digital Automation with Robotic Assistants – We've seen a great deal of innovation in the area of Enterprise RPA with repetitive tasks being efficiently automated by software robots. The term *robot* was coined by the Czech playwright and novelist Karel Capek in 1920 in his play "*Rossum's Universal Robots*." He had adopted an old Slavic term *rabota* that meant forced labor to tell the story of human-like automated machines that, until they revolted, catered to the whims of the people of Earth. The next step in the evolution of digital automation will be the democratization with the emergence of intelligent software assistants that seek to augment and not replace human work. Unlike naive *hyperautomation* intended to be used by RPA specialists with a focus on efficiency, the focus for digital automation will shift to appropriately augmenting human capacity with closed loop learning focused on business outcomes. – Rajeev Kozhikkattuthodi, VP of Product Management at [TIBCO](#)

The growth of machine identities will create an even larger identity sprawl challenge for organizations. Due to the convergence of AI innovation, digitization, and the asynchronous workforce accelerated by the pandemic—enterprises are increasingly deploying solutions like RPA to automate tasks, boost productivity, and enhance customer service. However, there's one big issue that's commonly overlooked when it comes to AI innovation – security. Today, 94% of organizations who have deployed bots or RPA report challenges securing them. What's causing this challenge is that security professionals don't realize that bots have identities just like humans. Since RPA requires access to data they ultimately need to be secured just like its human counterparts. So as enterprises exponentially deploy AI solutions like RPA, we should expect to see a string of bot-based breaches because security professionals aren't equipped to handle the identity sprawl linked to the growth of machines. – Larry Chinski, VP of Global IAM Strategy at [One Identity](#)

Machine learning and human-in-the-loop approaches to automation will displace RPA. Digital transformation efforts in a number of industries have driven massive adoption of robotic

process automation (RPA) during the past decade. The hard truth is that RPA is a decades-old technology that is brittle and has real limits to its capabilities – leaving a trail of broken bots which can be expensive and time-consuming to fix. RPA will always have some value in automating work that is simple, discrete, and linear. However, automation efforts often fall short of aspirations because so much of life is complex and constantly evolving – too much work falls outside of the capabilities of RPA. Emerging machine-learning-based technology platforms combined with human-in-the-loop approaches to automation are already redefining what it is possible to automate across a number of industries where complexity, exceptions, and outliers train the AI to work smarter, making automation stronger. – Varun Ganapathi, Ph.D., co-founder and CTO at [AKASA](#)

## Security

[According to a recent report](#), e-commerce retailers now experience an average of 206,000 web attacks per month, with 42% of businesses saying that digital fraud hampers innovation and expansion into new channels. Yet, despite this, only 34% of companies are investing in fraud prevention and mitigation. With e-commerce booming and no signs of slowing down, machine learning to defend against fraud will be on the rise. This will help online retailers keep up with fraudsters tactics and can spot patterns that might be missed by manual checks and analyze historical data and compare it to current transactions. This will be especially beneficial during the busier peak shopping seasons. – Jimmy Fong, Chief Commercial Officer at [SEON](#)

Your attack surface includes all the possible ways an attacker can get into your company's devices and networks and lock up or exfiltrate your data. So, it's essential to keep your attack surface to a minimum. The problem is that your attack surface is continually growing as more people work remotely on multiple devices and create more entry points for cybercriminals to carry out cyberattacks. Worse still, the attack surface is constantly changing. It isn't a single surface but many disparate fragments. Furthermore, control of endpoints is becoming increasingly complex as employees leave organizations and retrieval of equipment becomes harder. The bottom line is that breaches will inevitably happen. And in the coming year, companies will have to do a better job of recognizing breaches so they can extricate themselves as quickly as possible. Security and recovery strategies must be more thorough. As the attack surface expands, those strategies must cover not only your on-premises data but data in the cloud, at the edge, and everywhere in between. – Shridar Subramanian, CMO at Arcserve

Data Debt Will be a Primary Culprit of Security Breaches – Organizations have data stored everywhere, from their latest SaaS application to their oldest desktop and everything in between. And while organizations have worked tirelessly to secure their perimeters and lock down rights and access, sensitive data remains unfound and unprotected. Minimizing this data debt's security impact, begins by viewing data as a threat surface and methodically mitigating that threat based on its relative value, volume and vulnerability. In 2022, there will be many organizations, with millions of undiscovered and undetected risks across their data landscape, exposing their enterprises and their partners to significant damage. – Kevin Coppins, President & CEO, [Spirion](#)

AI will help solve security issues in government. AI will play an important role here. Federal government is embracing AI and machine learning to help understand where inefficiencies exist in systems and identify ways to save money. For example, the Health and Human Services (HHS) is using a tool called "BUYSMARTER" that trawls through the government's contract space, analyzes the contracts, and comes back with recommendations about how to save money. – George Sellner, Senior Director, Public Sector/Federal Industry Solutions at [Icertis](#)

The transition from DevOps to DevSecOps will harness the combination of AI and automation, redefining software development in 2022. Supply chain attacks, data mishandlings and not addressed known vulnerabilities over the last year made it clear that DevSecOps is the next stage of DevOps and the driving force that adds value, speed, and security to all stages of the SDLC. As we shift to that next stage, the combination of AI and automation to manage laborious security and CI/CD tasks inherent to cloud-native software development will save teams time while empowering them to proactively address any issues in the SDLC – enabling them to become an even more essential piece of business strategies. – Dynatrace's Andi Grabner, Director of Strategic Partnerships

Innovative Attack Methods Using Artificial Intelligence Will Expand the Threat Landscape – In 2022, the use of Artificial Intelligence (AI) will expand the cybersecurity threat landscape, bringing new dangers and altering the typical characteristics of threats. Attackers

will employ new and highly innovative methods, notably Machine Learning (ML), which will enable cybercriminals to use AI to carry out more cyber and ransomware strikes. AI/ML techniques will generate more sophisticated phishing intrusions, pervasive ML email attacks and zero-day attacks on top of other well-known ransomware deployments. In the hands of cybercriminals, AI/ML can create significant harm as machine-learning and deep-learning techniques will make cyberattacks more accessible. The result? Faster, better-targeted, and more destructive assaults. – Philip Chan, Ph.D., Adjunct Professor, School of Cybersecurity & Information Technology, University of Maryland

The Future of Security will be Tied to AI. – Here's a saying in the security industry: "Organizations have to be right 100% the time — the bad guys only once." The challenge in 2022 will continue to be staying ahead of attackers, even as attacks increase exponentially. The increase is partly due to organizations moving away from data centers with a single ingress and egress point to not just a cloud, but MANY clouds, some through Shadow IT which may be invisible to security. Security today can be compared to adjusting the antennae on your 1960s television set. If you held the "rabbit ears" just right, you could bring in the picture. If organizations are looking at the noise and thinking that traditional security tools can give them a clear picture, they are setting themselves up for an unfortunate experience. The best way to stay ahead of the attackers is to improve the signal-to-noise ratio. Attackers take advantage of the noise to sneak into networks or data sources. This is where AI can be most useful by picking signals from the noise. AI is effective in this situation for one very important reason: It doesn't get tired of looking at the same patterns over and over, and it can see a very big picture all at once. AI is a necessary component of a security strategy, but it is not a sufficient component. We recommend a well-rounded security strategy that includes the following: Well-considered set of traditional security tools updated for the modern environment, AI applied judiciously, human experts. Remember, AI does not yet deal well with nuance. – Theresa Kushner, Sr. Director, Data as an Asset, [NTT DATA Services](#), and Brandon Swain, Security Advisor, [NTT DATA Services](#)

Biggest data breach ever happening in 2022. We will experience the biggest breach ever coming either from a social network or a huge vendor in 2022. This data breach will have temporary consequences in the organization's market value but not a real massive impact. The breach will likely be so big that it will be sold in collections or seasons. – Ramsés Gallego, International Chief Technology Officer at CyberRes, A [Micro Focus](#) line of business

## **Storage**

I expect a broader adoption of object storage by enterprises in 2022. With the explosion of useful data, object store is becoming the standard for mass capacity because it offers advantages over traditional file stores including prescriptive metadata, scalability, and no hierarchical data structure. Plus, storage systems benefit from greater intelligence incorporated in data sets and object stores provide this intelligence. – Seagate's CTO, John Morris

As HPC deployments have become highly distributed and begun to exceed exabyte scale, it's become clear that the storage component of HPC infrastructure needs greater focus. To continue making advances in supercomputing, organizations will require highly scalable, software-defined storage that can accommodate massive data sets while easily leveraging any hardware innovations on the computing side. Parallel file storage alone cannot provide this scalability and flexibility. As a result, more organizations will use object storage as the primary storage for supercomputing deployments. – Gary Ogasawara, CTO of [Cloudian](#)

Increasingly intelligent storage – storage can no longer just be storage, vendors are realizing the enterprise desires more and more delivered natively. Better insights into data usage, better automation of scale and performance, native security capabilities such as anti-ransomware protections will all become increasingly desirable. Leading vendors are already starting to do this and more will follow. – Paul Stringfellow, analyst, GigaOm

Edge storage becomes a container play as it moves further and further out to accommodate the requirements of billions of sensors, 5G POPs and cameras. While the edge has two primary topologies, edge cache and edge storage – the growth will come in the latter. Edge storage increasingly demands containerization because it does not look like a mini data-center but rather a distributed system of endpoints. To be successful in this model, everything needs to go into the container: application code, databases, even persistent storage. To fit in a container at the edge requires lightweight, powerful, resilient, secure software components. This is why MinIO is the object store of choice – it can run in a stateless container while ensuring the data retains state. If the container fails, the data

remains safe in all but the most extreme examples of total loss. – MinIO co-founder and CEO Anand Babu Periasamy

The Era of Big Data Centralization and Consolidation is Over – The importance of centralized or consolidated data storage will also come to the forefront in 2022. To be clear this trend isn't the end of storage, but is the end of centrally consolidated approaches to data storage particularly for analytics and app dev. In 2022, we will see the continuation of the big fight that's brewing in the data analytics space as old ways of managing enterprise data, focusing on patterns of consolidation and centralization, reach a peak and then start to trend downward. Part of what we're about to see unfold in the big fight between Snowflake and Databricks in 2022 and beyond is a function of their differing approaches to centralized consolidation. But it's not just technical pressures. The economics of unavoidable data movement in a hybrid multicloud world are not good and don't look to be improving. Customers and investors are pushing back against the kind of lock-in that accompanies centralization approaches so anticipate the pendulum swinging in the direction of decentralization and disintermediation of the data analytics stack in the coming year. – Kendall Clark, Founder and CEO at [Stardog](#)

For several years, the data storage industry has recognized a need for increased automation in storage systems management. This need is amplified by data growth, and by predicted shortages in skilled human resources needed to manage these mountains of data. IDC has published reports for "the Future of Work" that provide ominous predictions that [a lack of IT skills will affect over 90% of enterprises and will cost them over \\$6.5 Trillion by 2025](#). Previous reports have predicted that storage administrators will have to manage 50 times more data in the next decade, but with only a 1.5X increase in the number of skilled personnel. The integration of AI/MLOps into large-scale data storage offerings will increasingly emerge to help administrators offload and automate processes – and to find and reduce waste and increase overall storage management efficiency. MLOps can monitor and provide predictive analytics on common manual tasks such as capacity utilization, pending component failures and storage inefficiencies. These innovations wouldn't be possible without the application of ML techniques, and their ability to consume and "train" from extremely granular system logs and event data during real-time operations. – Paul Speciale, CMO, Scality

### Verticals

Machine learning applications enable the processing of large amounts of data sets and reaching valuable conclusions which, by using its algorithms, can drive effectiveness and provide efficiencies including time saving opportunities. It analyzes patterns in real-time enabling quick decisioning. A range of financial services applications already use AI/ML today for everything from fraud detection, lending approvals, and AML screening, to risk monitoring and investment predictions. Machine Learning is constantly evolving, and Fintech will continue to be one of main industries to benefit from the power of AI/ML. – Abdul Naushad, President and CEO, [Buckzy](#)

AI will become more engrained in transportation and road safety. This trend will escalate even more now with the passing of the Infrastructure Bill, which includes measures around incorporating anti-drunk driving technology in all new cars. We will also continue to see a rise in the technology used in commercial vehicles to make trucking a safer and more appealing career, which is crucial amid the current driver shortage further disrupting supply chain operations. AI will create even more connected data to create a safer world, from continued development of connected commercial vehicles to in-car sensors that measure the levels of elements in a driver's exhale to detect health issues. These innovations through AI will ultimately create safer roads for all drivers. – Ryan Wilkinson, Chief Technology officer at [IntelliShift](#)

*Sign up for the free [insideBIGDATA newsletter](#).*

Join us on Twitter: @InsideBigData1 – <https://twitter.com/InsideBigData1>

tweet

share

share

email