CBINSIGHTS

2022

AI trends to watch in 2022

CBINSIGHTS

Where is all this data from?

The CB Insights platform has the underlying data included in this report

Click here to sign up for free

Table of Contents

Intro

The era of machine learning has arrived.

AI companies raised a record $66.8B in funding in 2021 — over double the global funding total in 2020. And while quarterly AI funding slipped 12% from Q4'21 to Q1'22, this decline was less dramatic than venture as a whole, which saw funding fall 19% during the same period.

In our 2021 enterprise AI trends report, we highlighted themes that continue to remain relevant: natural language processing models are getting bigger and better, AIOps vendors are raising multi-million-dollar funding rounds, and more vendors are integrating no-code features into their platforms.

Now, new trends — as well as new solutions to old problems — that are relevant to consumers and enterprises alike are appearing across industries.

The "metaverse" has emerged as a major theme in the last year, as online worlds and games grew in popularity and Facebook made its highly publicized name change to Meta. This has in turn fueled the need for new AI-powered content moderation tools to monitor for toxic and hateful behavior online.

From the war in Ukraine to presidential campaigns in South Korea, deepfakes have become more prominent in mainstream media and politics — putting deepfake detection and media provenance tech back in the spotlight.

Finally, to comply with data privacy laws, corporations across healthcare, telecommunications, and finance are turning to synthetic data: fake data that closely mimics real-world data.

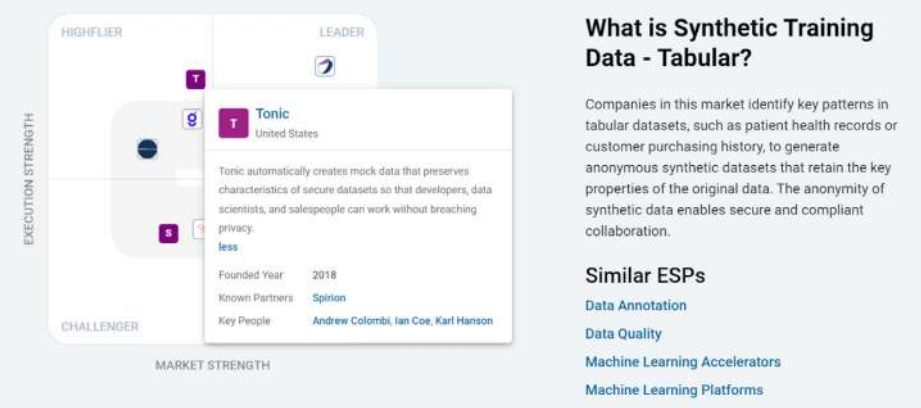
From familiar technology to cutting-edge advancements, read on for 7 AI trends to watch closely this year.

Synthetic data for privacy

From healthcare to finance, corporations are experimenting with synthetic datasets to enable data sharing and collaboration — while complying with GDPR and other privacy laws.

In industries where sufficient real-world data to train AI is unavailable or sparse, or where compliance and privacy are major concerns, enterprises are turning to synthetic datasets: fake images, videos, or tabular data that resemble real-world datasets.

Despite the challenges in accurately mimicking real-world data, big corporations are betting on the tech.



SOURCE: CB INSIGHTS ESP VENDOR MATRIX FOR SYNTHETIC TABULAR DATA PROVIDERS

Illumina is using synthetic genomics data developed by startup Gretel for medical research. In a joint case study, the 2 companies highlighted issues like restrictive legislation and patient consent requirements that limit the speed and scale of medical research that relies on sensitive patient data. In response, Gretel used real genotype and phenotype data to train an AI algorithm to generate artificial genomic data.

> **"Along with privacy enhancing technologies such as differential privacy, synthetic data has the potential to address the deep privacy concerns working with genomic data, enabling faster sharing of data and unlocking innovation."**
>
> **— Gretel-Illumina joint case study**

In the financial sector, J.P. Morgan is training financial AI models with fake data, while in the telecom industry, Telefónica has worked with Mostly AI to create GDPR-compliant synthetic customer profiles that mirror statistical patterns of real customer data. It's estimated that up to 85% of real customer data in the telecom industry is unusable due to lack of customer consent, making behavior analysis and prediction difficult.

In March 2022, MDClone raised $63M from Lightspeed Venture Partners, OrbiMed Advisors, and others for its synthetic health data platform. The company claims that synthetic data is a viable alternative to other methods of data de-identification (i.e., removing personally identifiable information from patient records) where patients may potentially be re-identified through cross-referencing.

Another vendor in the space, Tonic, has introduced AI features on its platform to improve the fidelity of fake data. It raised $35M last year from Insight Partners, Bloomberg Beta, and others, and works with clients including Flexport, eBay, and Oscar.

While synthetic image and video generation has taken off — thanks to GANs, a now popular approach that uses 2 neural networks to generate hyper-realistic synthetic videos and images — synthetic tabular data generation is still in its early stages. But the upsides of data anonymization, privacy compliance, and potential bias reduction make it an appealing option for corporations across industries.

The $67B chip race

As companies scramble to dominate in AI chips, new innovations from 3D chips to in-memory computing are poised to shake up the space.

The rapid commercial success of AI across industries is creating demand for specialized hardware that can support compute-intensive AI workflows, both in cloud data centers and edge devices like cameras.

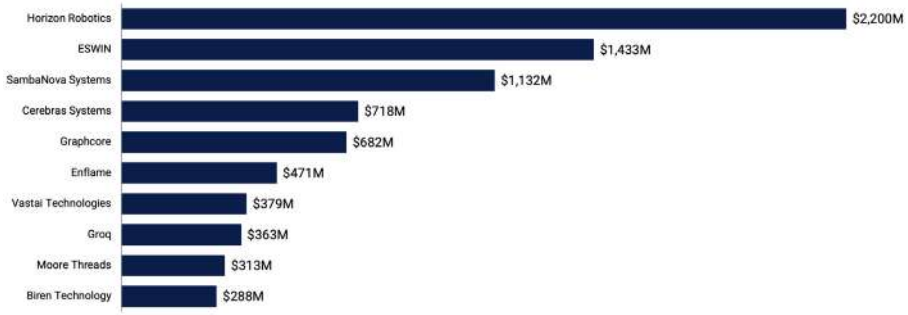
Once dominated by Nvidia's GPUs, the $67B AI chip market has been blown open by a number of players.

Google's latest Pixel phones are powered by the tech giant's own Tensor processors to support on-device AI applications, while Amazon launched Graviton3, its custom chip for AI inference, in Q4'21.

Beyond the incumbents, startups are also entering the fray. Cerebras Systems offers what it claims to be "the largest chip ever built," with 2.6T transistors and 850,000 AI cores. The company raised $250M in Q4'21 at a $4B valuation.

Top-funded chip providers for AI workloads

Total disclosed equity funding (as of 05/13/2022)



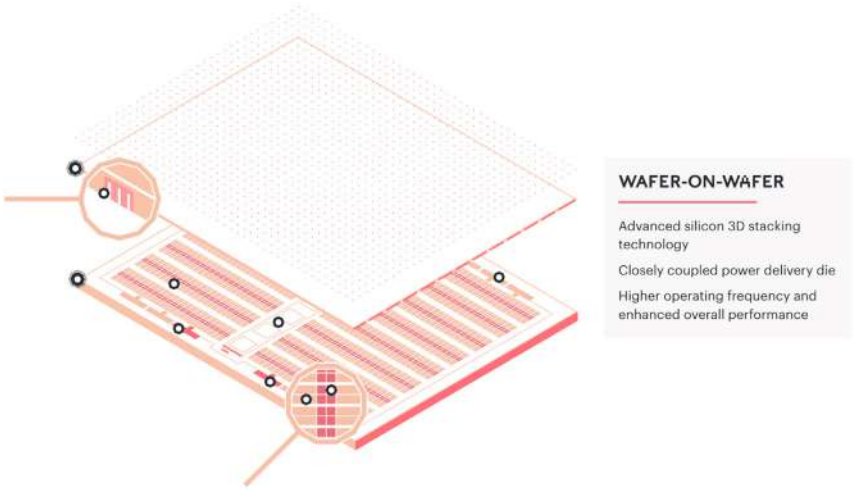
CBINSIGHTS

However, mammoth chips aren't practical for many everyday AI applications, due to space and energy limitations. Seeing this opportunity, a growing number of companies are offering AI chips that can be used with low-power devices like car sensors, store cameras, automated factory robots, and more.

On-device AI processing can power low-latency — and potentially privacy-preserving — experiences. Startups including Mythic, Syntiant, and Kneron have each raised more than $100M to develop this technology.

SOURCE: SYNTIANT

Meanwhile, startups like Untether AI and HOUMO.AI are working on "in-memory computing." This approach brings AI processors and memory together onto a single chip. The close proximity and high level of integration in these systems can deliver dramatic performance gains compared to conventional methods: Samsung reports it used the approach to more than double the speed of a speech recognition neural net while cutting its energy consumption in half.



SOURCE: GRAPHCORE

AI chip unicorn Graphcore is also playing with chip structure to boost performance gains using an approach called "3D chips." The tech involves bonding multiple chipsets together to create a fused stack of intricately aligned components. Graphcore used this method to place a power delivery chip directly on top of an AI processor, allowing the chip to train a neural net 40% faster than its previous generation while using less energy.

Other companies are ditching the physics of conventional AI chips altogether by turning to photonic processors, which shuttle data around using light as opposed to electrical signals. The big advantage of photonics is speed — light can move information faster, with greater bandwidth, and using less energy than relying purely on electrons flowing through cables. With the demands of AI tools expected to continue growing exponentially, photonics promises a way to avoid hitting a looming hardware barrier.

Researchers have struggled for decades to make practical, general-purpose chips using photonics. However, startups including Lightmatter and Luminous Computing are betting that they can use photonic chips optimized for tasks like deep learning to build "AI supercomputers" capable of handling even particularly taxing algorithms.

Specialized AI chips are here to stay — and incumbents like Nvidia and Intel are facing fierce competition from startups and big tech leaders alike. New technologies are emerging that will make these chips perform better and use less energy.

Looking further ahead, cutting-edge technology like quantum machine learning — which uses quantum computers to help run certain AI tasks much more efficiently — will complicate this landscape even further.

Protecting virtual worlds

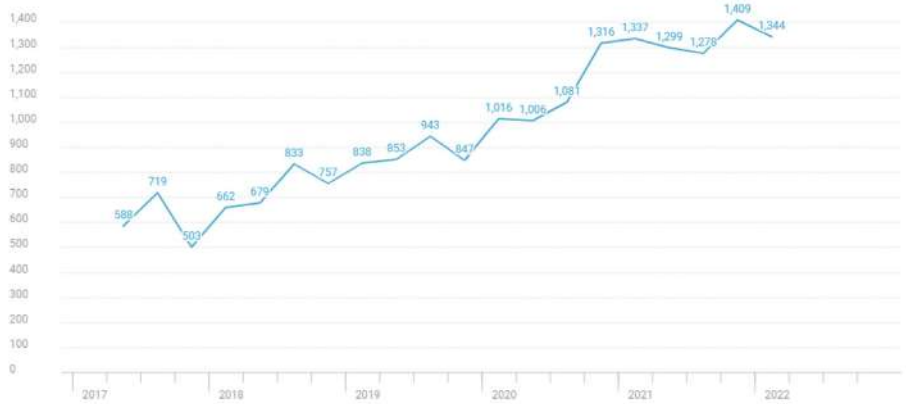
The threat of toxic content and behavior has spilled over from social media to a new front: online worlds. Companies are using AI to detect harmful behavior in gaming and other virtual spaces.

In the US, the number of people playing video games is at an all-time high. Additionally, 76% of Americans under 18 play games, raising concerns that children will be exposed to inappropriate or hateful content.

Detecting hate speech online is not a new problem. Meta reportedly spent $13B on content moderation between 2016 and 2021, while TikTok claimed in 2020 to have 10,000 people working on content moderation. Both TikTok and Meta have been sued by content moderators who suffered psychological damage on the job.

Now, given the increased hype around the metaverse and a rapidly evolving online gaming ecosystem, the threat of toxic content and behavior has spilled over from social media sites to a new front: online gaming and virtual worlds.

Mentions of gaming and e-sports trend up on earnings calls



CBINSIGHTS

Games can quickly turn toxic in the form of hate speech, cyberbullying, griefing (intentionally helping the other team win to frustrate your teammates), and rage quitting (refusing to play a game to ruin the experience for everyone else). A study from the Anti-Defamation League found that up to 80% of players have experienced harassment in some of the most popular multiplayer games.

Startups are both pivoting and emerging to address this market with the belief that AI can combat this behavior at scale. Spectrum Labs claims its natural language processing (NLP) platform reduces audio- and text-based content moderation efforts by 50% while improving toxic behavior detection by 10x.

Meanwhile, in March 2022, GGWP emerged from stealth and announced its intention to combat toxic behavior in games by using AI to monitor chat logs and gameplay data. The startup had previously raised a $12M seed round backed by investors like BITKRAFT Esports Ventures, Sony Innovation Fund, and Riot Games.

In April 2021, Hive raised a $50M Series D at a $2B valuation.The startup, which was founded in 2013 as a platform to crowdsource AI training data, is providing automated content moderation to help metaverse companies like HighRise and Naver Z (which runs the Zepeto platform) stay on top of player hate speech, cyberbullying, and more.

SOURCE: HIVE

Big tech players are also snapping up AI startups to bring moderation capabilities in house. In October 2021, for instance, Microsoft acquired Two Hat, an AI content moderation service that lists Roblox, Epic Games, and Microsoft-owned Minecraft as its customers.

Ultimately, perfect content moderation may be impossible. Online communities constantly adapt to escape platform censorship. However, breakthroughs in key areas like NLP and image-based classification, as well as recent funding rounds to AI-powered content moderation startups, indicate that AI will be at the forefront of the content moderation war.

Weeding out deepfakes

As deepfake candidates campaign for votes and AI videos spread wartime misinformation, deepfake detection tech competes with the AI creating it.

Deepfakes have expanded from hyper-realistic images to voice and video spoofs — including face reenactment, where one face in a video is swapped out with another.

Powered by self-learning algorithms, these fakes have gotten better and better over time. The sheer number of publicly available videos and voice recordings makes it easy to train AI algorithms and generate deepfakes — especially of celebrities. Researchers say it will only get harder to distinguish AI-generated faces, objects, and videos from real ones.

Deepfakes have proliferated in the media in 2022, especially in the political sphere. In March 2022, The Wall Street Journal reported how Yoon Suk-yeol, a candidate in the South Korean presidential elections, was using deepfakes of himself to improve his public image with young voters. The "AI Yoon," as the deepfake version was called, came off as more humorous than its real-life counterpart.

YOON SUK-YEOL. SOURCE: THE WALL STREET JOURNAL

In Europe, viral deepfakes of Volodymyr Zelenskyy and Vladimir Putin have added to wartime misinformation in Ukraine and Russia in recent months. And in 2021, a wave of politicians around Europe were tricked during video calls by live deepfakes of Russian political actors.

Beyond fake news and political disinformation, deepfakes are also targeting enterprise customers — and have the potential to become a mainstream tool in consumer phishing and extortion scams.

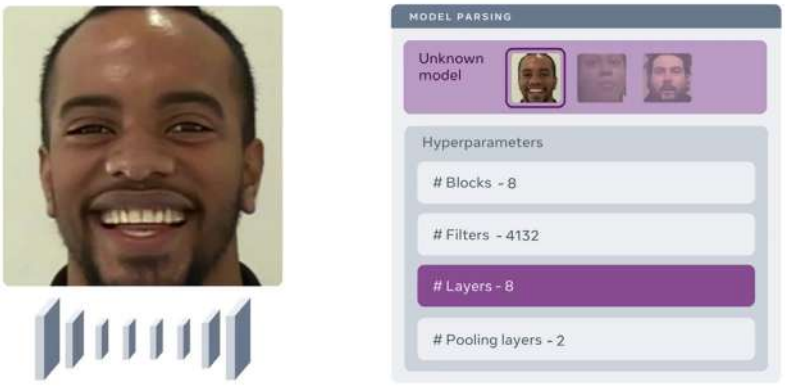
> **"What we observe is that fake media is getting better and better and better, and the ability for computers to tell what's real and what's fake, or users to tell what's real and what's fake, is rapidly tending to zero."**
>
> **— Paul England, engineer at Microsoft Research Lab**

Arguing that AI-powered defense cannot fully combat AI-generated fakes, Microsoft (alongside partners like Adobe and the BBC) launched Project Origin, a service that allows publishers to authenticate media using tamper-proof metadata.

Microsoft has also separately partnered with startup AI Foundation, which operates both commercial and non-profit ventures. AI Foundation's nonprofit is headed by Twitter co-founder Biz Stone, Trion Worlds creator Lars Buttler, and others.

In 2020, AI Foundation raised $17M to build deepfake avatars, while also announcing the launch of a deepfake detection platform, Reality Defender, with Microsoft. Reality Defender has partnerships with ABC, Homeland Security, and the Department of Defense, among others.

Taking a different approach, Truepic uses cryptography and blockchain-based tech for photo and video identification. Truepic has been working with Qualcomm to add a cryptographic seal around pictures captured by Qualcomm chipset-powered smartphones. It raised $27M last year from Adobe, Microsoft's M12, Sony Innovation Fund, and others.



Each generative model leaves its imprint on the deepfakes it creates...

SOURCE: META AI

Last year, researchers at Meta announced that they can now "reverse engineer" a deepfake image — that is, not just determine whether an image is fake or not, but also dissect the attributes of the AI model used to create the deepfake. This makes it possible to determine whether multiple fake media originated from a single source.

As a first wave of response to this growing cybersecurity threat, tech companies have experimented with approaches ranging from on-device authentication software and APIs to blockchain-based solutions. However, deepfakes will continue to become more advanced and ubiquitous, forcing tech companies to find new ways to detect and disrupt them.

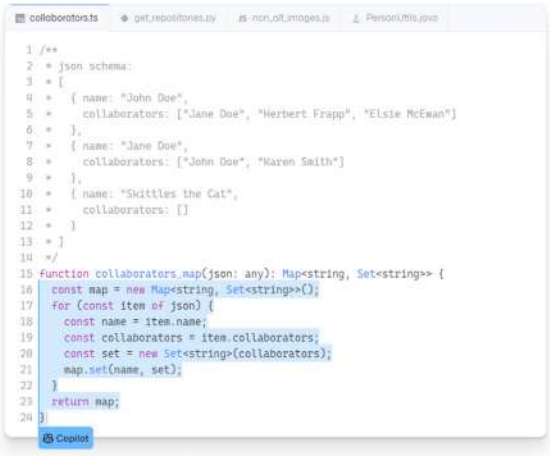
Augmented coding

AI is making coding easier and more efficient by assisting with code completion and automating software testing.

Algorithms are taking natural language commands and translating them into computer code, signaling a new era for software development — especially among citizen developers.

In June 2021, GitHub (acquired by Microsoft in 2018) and OpenAI (in which Microsoft owns a $1B minority stake) tag-teamed to launch GitHub Copilot. Trained on publicly available GitHub data, Copilot converts English-language comments into code. Copilot is particularly useful in mapping smaller programming tasks to existing code.

Convert comments to code. Write a comment describing the logic you want, and let GitHub Copilot assemble the code for you.



SOURCE: GITHUB COPILOT

Microsoft isn't the only big tech player making moves in the space. In February 2022, Google's DeepMind released AlphaCode, AI-generated programs that were tested on more complex problems in competitive programming. After evaluating its AI on recent competitions hosted on Codeforces, DeepMind claimed that its AI performs "roughly at the level of the median competitor."

On the private market side, startups are more focused on software testing: automating quality checks and unit tests of code. Key vendors in the space include Mabl (backed by CRV and Google Ventures), Autify (backed by Salesforce Ventures), and University of Oxford spinout Diffblue (which raised a $10M Series B in Q1'22).

Automated programming is still in its infancy. But recent advances favor a growing enterprise trend toward no-code/low-code solutions that enable non-technical users to participate in data science projects, close skills gaps, and speed up production cycles.

Multimodal AI

Tech companies are making strides in multimodal AI — a single AI model that can understand concepts from multiple modalities, like video, text, and 2D images — to improve search and content generation.

Today, an AI model trained on video data can be used for predicting video content, a model trained on text can be used for text predictions, and so on.

To go beyond this, multimodal AI research aims to be more holistic, using a single AI model to conceptualize information across multiple types of data like text, 2D images, and videos to make a prediction.

For example, in early 2021, OpenAI trained an AI model called DALL-E to generate images based on a text prompt. In the image below, the AI generates avocado-shaped armchairs following a prompt for the same.

SOURCE: OPENAI

In January 2022, OpenAI released DALLE-2, which improves the original model's output image resolution by 4x.

In May 2022, Google launched Imagen, a text-to-image project that reportedly outperforms OpenAI's model in terms of the quality of images generated, as well as the alignment between the input (text) and output (AI-generated image).

SOURCE: GOOGLE

Earlier this year, Meta published a paper called "Omnivore: A Single Model for Many Visual Modalities." The paper details an AI model that, when trained to recognize 2D images of pumpkins, can also recognize pumpkins in videos or 3D images without requiring additional training for the latter 2 media types.

Multimodal AI is growing beyond academic research labs to find practical applications.

Google, for instance, is using multimodal AI to improve search. In the future, a user could take a photo of their hiking boots and ask a query like, "Can I use these to hike Mt. Fuji?" The search engine would recognize the image, mine information on the web about Mt. Fuji from text, image, and video data, and connect the dots to provide a relevant answer.

SOURCE: GOOGLE

On the startup side, Twelve Labs raised $5M in seed funding from Index Ventures, AI researcher Fei-Fei Li, and other investors in March 2022. The company is developing AI to understand context from both visuals and audio data to index videos and make them searchable.

Multimodal AI research is poised to go beyond corporate research labs to power the next era of search and content generation, among other applications.

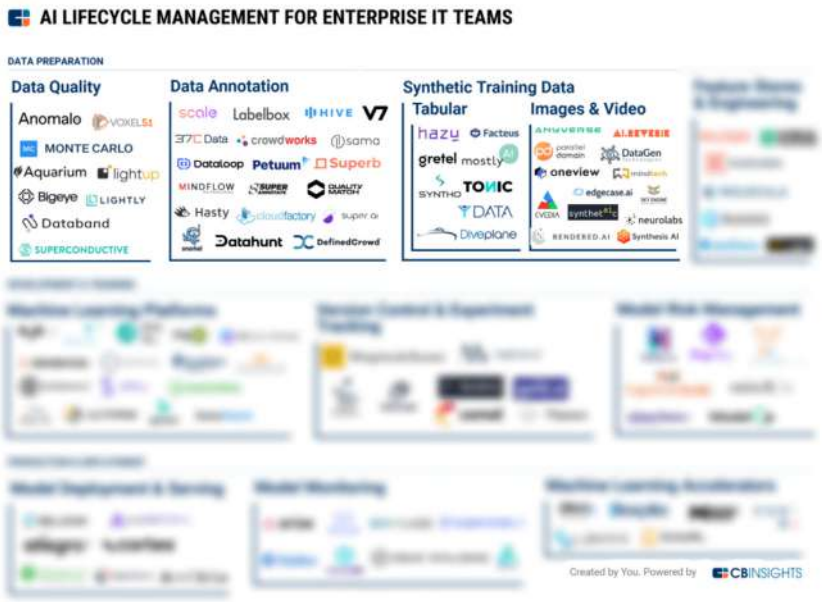
End-to-end machine learning

In the fragmented market for AI development, end-to-end machine learning platforms that take a project from start to finish will emerge as a compelling option for enterprises struggling with AI implementation.

As commercial applications of AI scale rapidly, enterprises are looking to overhaul existing data management practices to become "AI-first."

This strategy is easier said than done. Taking a project from raw data to production-ready AI is a multi-step process, from sourcing data and running data quality checks to developing models and monitoring post-production performance.

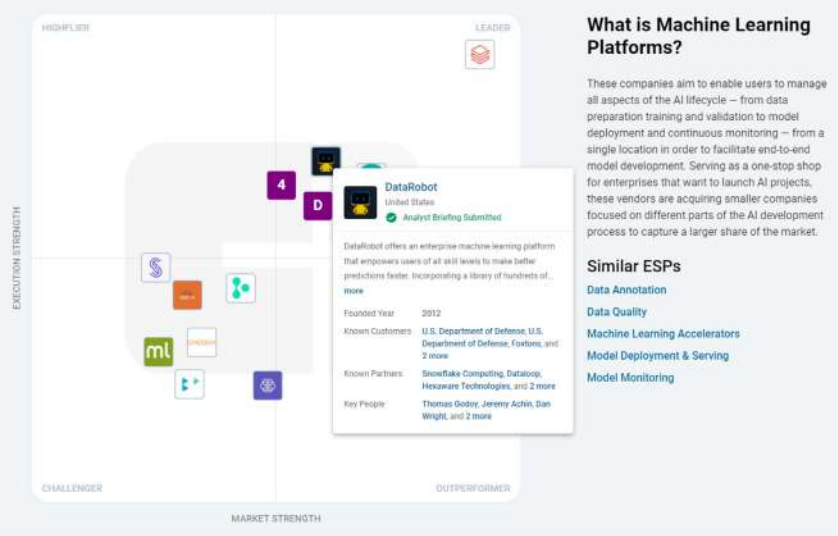
Hundreds of vendors have emerged in the MLOPs (machine learning operations) market to tackle different parts of the process. (Clients can dig deeper in our AI lifecycle management Tech Market Map, which maps vendors across 11 different steps in the AI development process.)



End-to-end machine learning vendors — who combine multiple steps in the AI lifecycle management process into a single SaaS product — will emerge as a more scalable option for enterprises that want to build AI systems quickly and efficiently.

Google released an end-to-end AI development platform, Vertex AI, in May 2021, touting it as a one-stop shop for data science teams that don't have prior machine learning experience.

DataRobot, a company valued at $6.3B and a leader in our ESP Vendor Matrix for machine learning platforms, has been actively scaling its product through M&A for several years. Last year, DataRobot made 3 acquisitions to capture a larger share of the enterprise AI market.



SOURCE: CB INSIGHTS ESP VENDOR MATRIX FOR MACHINE LEARNING PLATFORMS

ML platforms are also building no-code and low-code features to bridge the talent gap in AI, using plug-and-play features to encourage non-experts to participate in AI projects. To this end, Databricks, the second-most highly valued AI unicorn in the world at $38B, acquired 8080 Labs, developer of a no-code Python data analysis tool, in 2021.

Meanwhile, "AI for AI" — using AI to automate aspects of the AI development process itself, such as data quality checks or parts of model development — is another growing space, with most established vendors now offering AutoML (automated machine learning) features.

In the increasingly crowded AI development space, big tech companies and unicorn startups will continue to compete for a greater share of the market by offering an ever-wider range of AI services.