AI 2.0: Upgrade Your Enterprise With Five Next-Generation AI Advances

Advanced Level: Technology Practices For Insights-Driven Businesses

by Kjell Carlsson, PhD, Brandon Purcell, and Mike Gualtieri

February 9, 2021

Why Read This Report

Just when you thought you knew what AI could and couldn't do, it quietly gets a radical upgrade. Five key advances redefine AI use cases and break the chains that have shackled enterprise AI adoption. Though you've likely never heard of them, these AI 2.0 advances are already entering commercial products, and forward-looking enterprises need to start preparing if they want to reap their competitive advantages. AD&D pros and other tech leaders should use this report to understand this next generation of AI capabilities and how to start on their AI 2.0 journey.

Key Takeaways

Five Key Advances Upgrade AI To Version 2.0
Hidden behind the terms transformer networks, synthetic data, reinforcement learning, federated learning, and causal inference are capabilities that redefine the art of the possible with AI.

AI 2.0 Is A Radical Enterprise Upgrade
In addition to tackling the data, accuracy, speed, and security constraints that have held back broad implementation, AI 2.0 unlocks AI's creativity, its multipurpose capabilities, and brings it to the edge.

Get Started On Your AI 2.0 Journey
Most of the tools and services you need to start developing AI 2.0 solutions are already available. Build your knowledge, team, and strategy to acquire these capabilities and exploit them.

AI 2.0: Upgrade Your Enterprise With Five Next-Generation AI Advances

Advanced Level: Technology Practices For Insights-Driven Businesses

by Kjell Carlsson, PhD, Brandon Purcell, and Mike Gualtieri
with Srividya Sridharan, Martha Bennett, Charlie Dai, Diego Lo Giudice, Michele Goetz, and Jeremy Vale
February 9, 2021

Table Of Contents

Related Research Documents

The Forrester Tech Tide™: Artificial Intelligence For Business Insights, Q3 2018

The Future Of Machine Learning Is Unstoppable

FORRESTER®

Forrester Research, Inc., 60 Acorn Park Drive, Cambridge, MA 02140 USA
+1 617-613-6000 | Fax: +1 617-613-5000 | forrester.com

The Upgrade To AI 2.0 Will Redefine Enterprise AI

AI has just undergone a step-change in its evolution, but you would be forgiven for having missed it. It's obfuscated by a deluge of impressive, but incremental, advances in AI-accelerated hardware, platforms, and AI-embedded enterprise solutions. However, hidden in the latest, cutting-edge offerings are AI advances that radically extend the art of the possible and unshackle AI from the data, accuracy, speed, and security constraints that have limited enterprise use cases and adoption. Together, these advances are a discontinuous change in the capabilities of AI. It's a movement from AI 1.0 — characterized by pattern recognition, task-specific models, and centralized training and deployment — to 2.0 — characterized by language, vision, and other general data generation models, trained and embedded everywhere.[1] Best of all, these AI advances are already here. Enterprises should now take advantage of AI 2.0's ability to be:

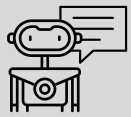
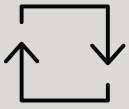
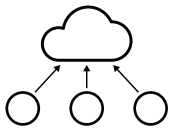
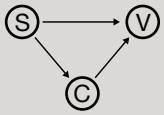
- **Creative.** AI 1.0 revolutionized our ability to detect patterns in unstructured data, e.g., object detection in images and intent recognition in chats. With AI 2.0, enterprises can automatically generate marketing content and software code, propose new drug molecules, and synthesize artificial training data to make other AI models more robust and accurate, both faster and cheaper.[2] Further, enterprises can use AI 2.0 technologies to propose and test never-before-seen solutions to complex optimization tasks, whether they be manufacturing processes or marketing campaigns.
- **General.** In AI 1.0, you could use transfer learning to customize models trained to recognize one set of patterns on another (e.g., customize a model trained on dogs to recognize cats). AI 2.0 takes this to a whole other level with giant models that span a host of tasks like summarization, question answering, topic mining, and query generation. These models even span different types of data (e.g., generating captions from images and vice versa). Why does this matter? Because these multitasking models require less training for each new domain and task.[3] Indeed, they are starting to solve tasks that they were never trained for (just don't hope for self-thinking machines anytime soon).
- **Anywhere.** AI 1.0 was throttled by an organization's ability to transfer data to a central location to train and run its models. In AI 2.0, data never has to move. AI models are deployed and even trained at the edge, enabling new AI applications that need to be cheaper, faster, and more secure. Your smart speaker learning your speech patterns isn't nearly as creepy when your voice isn't sent back to HQ, especially when it is more responsive, capable, and understands you better because your model has been combined with everyone else's.

Five Breakthrough AI Advances Drive AI 2.0

Much like there are many technologies under the AI umbrella, AI 2.0 is a collection of new technologies and techniques that are maturing simultaneously in terms of their technical feasibility and business applications. Five key technological advances are most important for enterprises, both because of the radical new capabilities they unlock and because they are already in production and driving outcomes

today (see Figure 1). However, unless you are a data scientist, you probably haven't heard about them, as most have developed rapidly in the last one to three years, have names that mask their importance, and would never be featured in a science fiction film.

FIGURE 1 The Five Enterprise-Changing AI Advances That You've (Probably) Never Heard Of

Technology	What is it?	What is it used for?	Who is using it?
Transformer networks 	Giant pretrained, customizable, hyperaccurate, multitasking deep learning models	Any hard problem with a significant time or context dimension (e.g., understanding and generating text, software code, etc.)	Hyperscalers (Amazon Web Services, Google, IBM, and Microsoft), the advanced guard of speech and text analytics vendors, and many startups
Synthetic data 	Generative models and simulated virtual environments used to create or augment existing training data	Accelerating the development of new AI solutions, improving the accuracy and robustness of existing AI models, and protecting sensitive data	Autonomous vehicles, financial services, insurance and pharmaceutical firms, and every computer vision vendor
Reinforcement learning 	Machine learning approaches that test their way to optimal actions via simulated environments or a large number of micro-experiments	Constructing models that optimize many objectives/constraints or deciding on action based on positive and negative environmental feedback	Firms targeting particular B2C marketing tasks, optimizing repeatable manufacturing processes, and robotic learning
Federated learning 	A managed process for combining models trained separately on separate data sets	Sharing intelligence between devices, systems, or firms to overcome privacy, bandwidth, or computational limits	Hyperscalers, AI-enabled application vendors, and consumer electronics companies
Causal inference 	Approaches such as structured equation modeling and causal Bayesian networks that help determine cause-and-effect relationships in data	Business insights (e.g., attribution analysis) and bias prevention where insights and explainability are as important as prediction accuracy	Innovation teams at leading organizations (e.g., determining treatment efficacy for a given disease at healthcare providers)

1. **Transformer networks unlock language (and more).** Heard of the op-eds and blogs that AI can write or the images it can draw from captions?[4] They were generated by transformer networks, such as OpenAI's GPT-3, which tackle many hitherto human-only tasks, especially those with a time or context dimension like natural language processing and generation.[5] Transformers make it possible to train giant models that tackle these and multiple other tasks at once (e.g., machine translation, natural language generation, and code generation simultaneously) with greater accuracy and less data than separate models. They are also easier to maintain and customize across many use cases.

   Hyperscalers are already embedding transformer networks in their business applications. For example, Microsoft uses them in its offerings for tasks like natural language search, autocaptioning of images, moderating inappropriate gamer language, and automated customer support. Salesforce Research developed Photon, which uses transformer networks to turn business users' questions into automatically generated SQL queries.[6] In addition to a deluge of startups and cutting-edge vendors — particularly in text and speech analytics — that are incorporating these models, you can soon expect enterprises to leverage pretrained transformer networks to create armies of more capable, hyperaccurate virtual agents.

2. **Synthetic data improves your data for every AI solution.** Obtaining the data you need is slow, costly, and sometimes impossible — even with unlimited resources. Insufficient, incomplete, or just plain bad data not only dramatically slows projects and limits the number of use cases an enterprise can tackle; it also reduces the accuracy of models and introduces bias and risk. Enter synthetic data. Typically created using generative models or simulated virtual environments, it augments your data, improving the accuracy, robustness, and generalizability of your models, plus it can be used instead of sensitive data to enable internal and external data scientists to experiment with the data, without the risk. Synthetic data can also jump-start projects where you have little or no training data to begin with. With synthetic data, don't fake it 'til you make it; do both.[7]

   Enterprise data science teams can generate their own synthetic data, but it is not always feasible or practical to create it in-house. Thankfully, an ecosystem of synthetic data vendors — most specializing in particular methods or use cases — is developing. For example, Virtusa offers a range of generative models that create digital twins of patient populations in 35 US states. In contrast, Israeli startup OneView synthetically generates images using video game engines. In one case, it improved oil leak detection for pumps and pipelines by a factor of three using synthetically generated aerial images.

3. **Reinforcement learning helps you adapt to changing dynamics.** In 2016, DeepMind's AlphaGo bested world Go champion Lee Sedol in what was a watershed moment for AI. That system used reinforcement learning to learn how to play the game. Well, it's not all fun and games for reinforcement learning anymore. 2020 showed us the need to be able to react to changes in data, as customer behavior changed almost overnight due to the pandemic. Because reinforcement learning learns from interacting with its real or simulated environment through trial and error rather than historical data, it can adapt to similar shifts, helping to circumvent data and model drift.[8]

Today, some companies are using reinforcement learning in autonomous systems. For example, an oil and gas exploration company is leveraging Microsoft's Project Bonsai for geosteering — finding the optimal path to drill horizontally underground. And Microsoft is using reinforcement learning for its own marketing — to contextually personalize products displayed on its homepage.[9] In the future, companies will turn to reinforcement learning to identify and deliver the next best experience to customers, optimizing journeys to maximize customer lifetime value.[10]

4. **Federated learning shares intelligence within and with others.** Getting data from different sensors, systems, and parts of the organization, let alone from other organizations, is a big challenge that can prohibit a multitude of highly valuable AI use cases. Transferring the data is costly, difficult, and often risky from a security, privacy, or competitiveness perspective. Federated learning overcomes this by developing AI models separately and sharing the model, or features of the model, instead of the underlying data. The information is a fraction of the size and hard to extract sensitive data from, which then is used to train a "federated model" that combines the intelligence across the different sources. The result? Intelligence is shared quickly, cheaply, and more securely within and across organizations.

   There is a good chance that you are using federated learning already. Google's Android 11 uses it to generate smart replies (and suggest emojis), and many machine-learning-enabled offerings from other vendors use federated learning to improve, without sharing customer data.[11] IBM's AutoAI and Federated Learning — developed by IBM Research — use it to train models on data across disparate systems, while ACI Worldwide, a payment solution software company, enables its members to share features to improve fraud detection.[12]

5. **Causal inference improves insight and AI robustness.** Machine learning techniques make it easier than ever to discover correlations in data, but since most do not establish causality, they also make it easier to find reverse or spurious correlations.[13] This can lead to faulty business decisions based on incorrect analyses or poorly performing models, and in some cases may harm customers through discriminatory or unfair practices. Causal inference techniques identify cause-and-effect relationships between variables. While they don't usually prove causality on their own, they suggest which relationships are supported by the data and which to test for causality and counterfactuals.[14]

   Relative to the other AI 2.0 advances, causal inference offerings are less mature, and many companies will need their own or third-party AI research teams to use them. Refinitiv is working with the MIT-IBM AI Watson Lab to apply causal inference techniques to investment decisions, while Ericsson is researching the use of causal inference to determine the causes of network outages.[15] Expect more companies to use causal inference techniques to determine optimal strategies for engaging customers, allocating marketing budget, improving operational efficiency, and detecting and preventing algorithmic or human bias.

Recommendations

Get Your Enterprise Ready For AI 2.0

The opportunity to get in on the ground floor of a transformative set of technologies doesn't come along often. When one does, it is usually inaccessible to all but a select group of specialists. For now, AI 2.0 has leveled the playing field by eliminating many barriers to entry built on years of expertise in AI domains like natural language processing, computer vision, and data advantages painstakingly built over years. Newcomers are outperforming veterans, and startups are building new applications that used to take years or were infeasible. Could you wait and take advantage of AI 2.0 solutions once they are mature? Yes, but you would forgo the opportunity to outperform your industry. To develop your AI-enabled competitive advantage:

- **Accelerate your AI 1.0 journey in tandem.** No enterprise is done democratizing and operationalizing the current generation of AI technologies like AutoML, ModelOps, explainability, chatbots, and robotic process automation.[16] Don't put those efforts on hold. In many cases, these capabilities are prerequisites to driving business value with AI 2.0, and it extends and improves on them. In instances where there is overlap (e.g., reinforcement learning in marketing or transformer networks in voice-of-the-customer analysis), the AI 2.0 technologies will not replace all the 1.0 functionality for the foreseeable future. That being said, prioritize new implementations that are already incorporating or laying the groundwork for AI 2.0 to future-proof your investments.
- **Build and empower an AI-forward team.** Individuals with proven experience driving business value with AI 2.0 technologies aren't just rare; they don't exist yet. Instead, you will need to upskill a hybrid strategy and innovation team to understand and assess the landscape of AI 2.0 technologies, monitor their development and opportunities, strategize, undertake pilots, and chaperone the successful adoption of these technologies.
- **Design your AI 2.0 strategy.** It's time to have an intelligent conversation about conversational intelligence and the host of new use cases for your enterprise that AI 2.0 can enable. Evaluate and prioritize use cases that score highly from both a business value and feasibility point of view, and seek out partners with the AI 2.0 capabilities to accelerate your journey.
- **Stand on the shoulders of (tech) giants.** Amazon, Google, IBM, and Microsoft are investing heavily in AI 2.0 for themselves, but they already offer these technologies as services that their customers can use to develop their own AI 2.0 solutions.[17] Expect them not just to support the full suite of AI 2.0 technologies but to improve on their range of capabilities, ease of use, and cost. Of course, you could turn directly to OpenAI to leverage the latest GPT-3 models, but it could take the vendor a while to get through the 10,000-plus organizations on its waitlist.[18]
- **Scan the AI horizon regularly.** Will AI-enabled quantum computing platforms transform the enterprise? Not anytime soon. Independent, self-evolving AI models on a blockchain? Don't even think about it. However, an ever-growing host of other AI technologies are developing rapidly and

just waiting for someone to stumble on the right killer use case or technological breakthrough to go mainstream. It's too early for most enterprises to worry about adversarial defense, neurosymbolic AI, neuromorphic computing, machine olfaction, and a plethora of other promising AI technologies in development, but that could change in the blink of an android's eye.

Supplemental Material

Companies Interviewed For This Report

We would like to thank the individuals from the following companies who generously gave their time during the research for this report.

Amazon Web Services
C3.ai
Google
IBM

Microsoft
NVIDIA
OneView.ai
OpenAI
Salesforce
Virtusa

Endnotes

[1] A brief history of the (commercially feasible) AI revolution:

AI 0.0 (pre-2010): business rules, expert systems, ontologies, OCR, machine vision and graph-based systems mimicking human-like interaction with little or no use of machine learning.

AI 1.0 (post-2010): commercially feasible, large ("deep") neural networks (e.g., convolutional neural networks and recurrent neural networks) for extracting insights at scale from unstructured and semistructured data (e.g., images, text, and speech, logs) and combinations of data.

AI 2.0 (2020 onward): Advances in deep learning, reinforcement learning, and other nontraditional machine learning methods dramatically changing the capabilities and use cases of commercially feasible AI solutions

[2] Source: Cade Metz, "Meet GPT-3. It Has Learned to Code (and Blog and Argue)." The New York Times, November 24, 2020 (https://www.nytimes.com/2020/11/24/science/artificial-intelligence-ai-gpt3.html) and David H. Freedman, "Hunting for New Drugs with AI," Scientific American, February 1, 2020 (https://www.scientificamerican.com/article/hunting-for-new-drugs-with-ai/).

[3] For example, Salesforce's decaNLP tackles a variety of natural language processing tasks all in one model. Source: "Salesforce Research Introduces the Swiss Army Knife of Natural Language Processing," Salesforce, June 20, 2018 (https://www.salesforce.com/news/stories/salesforce-research-introduces-the-swiss-army-knife-of-natural-language-processing/).

[4] Source: "A robot wrote this entire article. Are you scared yet, human?" The Guardian, September 8, 2020 (https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3); Karen Hao, "A college kid's fake, AI-generated blog fooled tens of thousands. This is how he made it." MIT Technology Review, August 14, 2020 (https://www.technologyreview.com/2020/08/14/1006780/ai-gpt-3-fake-blog-reached-top-of-hacker-news/); and "DALL·E: Creating Images from Text," OpenAI Blog, January 5, 2021 (https://openai.com/blog/dall-e/).

[5] For example, Google has developed transformer networks for better text summarization, while OpenAI has developed transformer networks that can generate new images from text descriptions and automatically label images using natural language. Source: Peter J. Liu and Yao Zhao, "PEGASUS: A State-of-the-Art Model for Abstractive Text Summarization," Google AI Blog, June 9, 2020 (https://ai.googleblog.com/2020/06/pegasus-state-of-art-model-for.html); "DALL·E: Creating Images from Text," OpenAI Blog, January 5, 2021 (https://openai.com/blog/dall-e/); and "CLIP: Connecting Text and Images," OpenAI Blog, January 5, 2021 (https://openai.com/blog/clip/).

[6] Source: Victoria Lin, "Talk to Your Data: One Model, Any Relational Database," Salesforce Research blog (https://blog.einstein.ai/talk-to-your-data-one-model-any-database/).

[7] Synthetic data should not be confused with other, more nefarious forms of artificial data, such as the fake data used in adversarial attacks and the creation of deepfakes. Source: "Predictions 2021: Artificial Intelligence," Forrester (https://www.forrester.com/fn/5FQLFysVYlhIZLOTHA0O9Z).

[8] For a vivid example of a reinforcement learning system adapting to situations it hasn't been trained for, check the following. Source: Matt Simon, "Watch a Robot Dog Learn How to Deftly Fend Off a Human," Wired, January 5, 2021 (https://www.wired.com/story/watch-a-robot-dog-learn-how-to-deftly-fend-off-a-human/).

[9] Source: Jennifer Langston, "With reinforcement learning, Microsoft brings a new class of AI solutions to customers," The AI Blog, December 7, 2020 (https://blogs.microsoft.com/ai/reinforcement-learning/).

[10] See the Forrester report "Come Together (Right Now) To Deliver The Next Best Experience." Source: "Predictions 2021: Customer Insights," Forrester (https://www.forrester.com/fn/4wfoTYqbJb7cxeEd4JyEvN).

[11] Source: Rahul Roy-Chowdhury, "How we keep you safe online every day," The Keyword, October 7, 2020 (https://blog.google/technology/safety-security/how-we-keep-you-safe-online-every-day/).

[12] IBM's AutoAI and Federated Learning are available as part of Watson Studio on Cloud Pak for Data.

[13] For example, the correlation between the growth in cancer rates and the subsequent growth in cell phone usage does not imply that cancer caused cell phones. Source: "Cell Phones," xkcd (https://xkcd.com/925/).

[14] Counterfactuals: what would have been true under different circumstances.

[15] Source: Hanna Helin and Hossein Adeli Jelodar, "Applying causal inference to machine learning problems," Refinitiv, June 15, 2020 (https://www.refinitiv.com/perspectives/ai-digitalization/applying-causal-inference-to-machine-learning-problems/) and Ravi Pandya, "From how to why: An overview of causal inference in machine learning," Ericsson Blog, February 6, 2020 (https://www.ericsson.com/en/blog/2020/2/causal-inference-machine-learning).

[16] AutoML: automated machine learning; ModelOps: model plus operations.

[17] For example, currently pretrained BERT models can be fine-tuned on AWS, Azure, and Google Cloud Platform. GPT-2 and Turing-NLG models can be developed on Azure. IBM Research has made its Causal Inference 360 Open Source Toolkit available online. The open source federated learning framework TensorFlow Federated is also available. Source: IBM Research Trusted AI (https://cif360-dev.mybluemix.net/) and "TensorFlow Federated," TensorFlow (https://www.tensorflow.org/federated).

[18] Source: Tiernan Ray, "What is GPT-3? Everything your business needs to know about OpenAI's breakthrough AI language program," ZDNet, August 25, 2020 (https://www.zdnet.com/article/what-is-gpt-3-everything-business-needs-to-know-about-openais-breakthrough-ai-language-program/).

FORRESTER®

We work with business and technology leaders to develop customer-obsessed strategies that drive growth.

PRODUCTS AND SERVICES

- Core research and tools
- Data and analytics
- Peer collaboration
- Analyst engagement
- Consulting
- Events

Forrester's research and insights are tailored to your role and critical business initiatives.

ROLES WE SERVE

Marketing & Strategy Professionals

CMO
B2B Marketing
B2C Marketing
Customer Experience
Customer Insights
eBusiness & Channel Strategy

Technology Management Professionals

CIO
- Application Development & Delivery

Enterprise Architecture
Infrastructure & Operations
Security & Risk
Sourcing & Vendor Management

Technology Industry Professionals

Analyst Relations

CLIENT SUPPORT

For information on hard-copy or electronic reprints, please contact Client Support at +1 866-367-7378, +1 617-613-5730, or clientsupport@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.

163520