STATE OF AI IN 2021

Two tech investors, Ian Hogarth and Nathan Benaich, have been detailing the State of AI every year since 2018. They released their 2020 version in early October this year.

Hogarth and Benaich take a macroscopic look at the global artificial intelligence playing field, which they've summed up in a 177-slide report! (It's a long but fascinating read for AI followers.) They define AI as:

> *" A multidisciplinary field of science and engineering whose goal is to create intelligence machines."*

Before we jump into their 2020 assessment, let's look at how they fared with their 2019 predictions.

2019 predictions

In their 2019 State of AI report, Hogarth and Benaich made six predictions—4.5 of which were correct. The team accurately predicted:

- **New natural NLP companies will raise $100 million in 12 months.** Many companies achieved this, including io at $200 million, K Health at $48 million, and Chorus.ai at $45 million.
- **No autonomous driving company will drive more than 15 million miles in a year.** Waymo was the closest, but still significantly short of the milestone, driving 1.45 million miles. Cruise and Baidu didn't crack the million-mile mark.
- **A Fortune 2000 company outside the Big Tech 5 will adopt privacy-preserving ML.** A research consortium including GlaxoSmithKline, Merck, and Novartis used machine learning

ledger orchestration for drug discovery—a mouthful known as MELLODY.
- **Universities will build *de novo* undergraduate degrees in AI.** Carnegie Mellon graduated its first class of AI undergrads in 2020 and Singapore University of Technology and Design (SUTD) created an undergrad degree in design and AI.

The team was partially correct on one prediction: that Google will have a major quantum breakthrough and that five new startups will form with a focus on quantum ML. Of course, Google arguably achieved "quantum supremacy" in October 2019—that's the correct half. Many quantum startups were launched in 2019, but fewer than five are focused specifically on quantum ML.

One prediction the team got wrong? AI governance. In the 2019 report, they predicted (hoped?) that governance for artificial intelligence would become a key issue, with at least one significant AI company making a big change to their governance model. Unfortunately, this did not happen.

State of AI in 2020-2021

This year's report considers a few dimensions:

- **Research.** Breakthroughs, capabilities, and innovation.
- **Talent.** The makeup of AI talent and its supply/demand.
- **Industry.** Commercial applications of AI and the business impact.
- **Politics.** Regulation and economic implications of AI, as a new AI geopolitics emerges.
- **Predictions.** What they believe will happen.

Here's a summary of key findings in each area.

Research

The buzziest area of AI today is natural language processing, or NLP. Next gen transformer language models, like GPT-3, will continue to unlock new NLP use cases, some of which we may even go on to use practically. But the concern here is the sustainability of such research. These systems are huge; GPT-3 uses hundreds of billions of parameters to run. Most companies can't afford to live in this world of boundless compute power and endless associated costs.

This lack of access can significantly reduce competition—which isn't good for research. Machine learning researches believe that progress in ML has stalled, leaving few organizations that can experiment significantly.

This concern plays out in the real world: many companies say they haven't adopted big pre-trained AI Software like IBM Watson because they don't have a true understanding of how it's trained, how it will perform. This could easily introduce problems around ethics, reputation, and compliance—as we'll see shortly.

Other research-related findings:

- Biology is getting in on AI, with universities and companies using it for medical imaging, genetics, chemistry, and even drug discovery. Healthcare-related AI has grown by 50% every year for four years running.
- With 85% of research projects in AI closed sources, the team is considered that overall AI research could suffer from accountability and reproducibility issues that are usually avoided when projects are open source instead.

Talent

The U.S. maintains its reign of AI dominance, measured in acceptances of major academic papers. But if you look at the makeup of these researchers, more than two-thirds are born outside the U.S. Of AI researchers in America:

- 27% are from China
- 11% are from India
- Another 11% are from Europe

This indicates that AI talent is still drawn primarily to the U.S.—at least for now. Beyond the U.S., several new institutions around the world are dedicated to AI are formed (one of the team's accurate predictions from last year).

Still, despite these findings, demand continues to outpace supply for AI talent. In 20202, academic brain drain—researchers leaving academic research—is acute, driven by increases in corporate recruiting. At least for now, this brain drain is having a negative impact on entrepreneurship.

Industry

Industry is the opposite of research—it's what Hogarth and Beniath call the "task-specific" domain of AI. Despite more research breakthroughs seeming to come more often, it's the on-the-ground AI that companies need. Of the fastest growing GitHub projects since July 2020, 25% focus on MLOps, short for machine learning operations. The investor team see this as a move away from R&D to operations—how to actually run models

Most businesses today that need an AI system actually need a not an overarching system, but one that can perform a very specific task and do so incredibly well.

For example, several companies have developed chatbots that outperform perform the research behemoth models like GPT-3 and BERT. These smaller models are important: they use much less compute power, making this type of AI adoption financially feasible for companies. NLP applications are being used widely, with implementation in Google Search and Microsoft Bing.

Other takeaways here:

- Self-driving cars are hanging out in the trough of disillusionment, but open sourcing of data is growing in order to crowdsource new ideas.
- NVIDIA, Google, and Graphcore continue to advance their AI hardware platforms

Politics

Globally, we've reached the tipping point of technology. Internet users are becoming savvier and we can all point to a few instances where technology has hurt, not helped.

In artificial intelligence, we are beginning to realize ethical risks. Researchers have long warned about this black box problem of bias and accuracy, but it wasn't until the Clearview AI scandal that facial recognition and data scraping became larger public concerns. Globally, some nations have begun passing laws that allow them to scrutinize when foreign parties takeover AI companies.

In response, some nations and U.S. states are starting to pass laws that limit certain technologies, including facial recognition. The next area ripe for regulation might be how algorithms work in other

fields, like financial or insurance decisions.

Increasingly, companies are also becoming cogs of geopolitics. Semiconductor companies, crucial to the development of AI hardware, are unwittingly becoming geopolitical pawns. Some governments are even scrutinizing when foreign parties takeover AI companies.

The U.S. military, too, is experimenting with AI, likely adding capabilities to its strategies and techniques. The DoD—and thus the entire field of established defense contractors—is getting in on the AI action, working on projects ranging from intelligence analysis software to systems that could auto detect and disrupt electronic communications. The implications of this could be significant. (This ongoing experimentation might also explain why the DoD just released new regulations for data governance.)

AI predictions for 2020-2021

With this background, we may increasingly see a new form of machine learning, one that can preserve and protect privacy. Known as federated learning, ML could offer a way for different parties to learn from the same data—without compromising privacy.

So, what does the team predict for the next 12 months? Of eight predictions, we think these are the most interesting:

- Some party will build an AI language model that runs on 10 trillion parameters.
- A big corporate AI lab shuts down due to changing strategy from its parent company.
- Chinese and European defense-focused AI startups will collectively raise $100 million in a year—a response to the U.S. military activity.
- Facebook has a big breakthrough in augments and VR with 3D computer vision.

With a strong track record, some of these predictions will certainly come true.

Additional resources

For related reading, explore these resources:

- BMC Big Data & Machine Learning Blog
- AIOps: A Beginner's Guide
- 4 Types of Artificial Intelligence
- Machine Learning: Hype vs Reality
- Machine Learning with TensorFlow & Keras, a tutorial series