
The Overview of AI in Drug Discovery in 2019: The “Proof-of-concept Year”

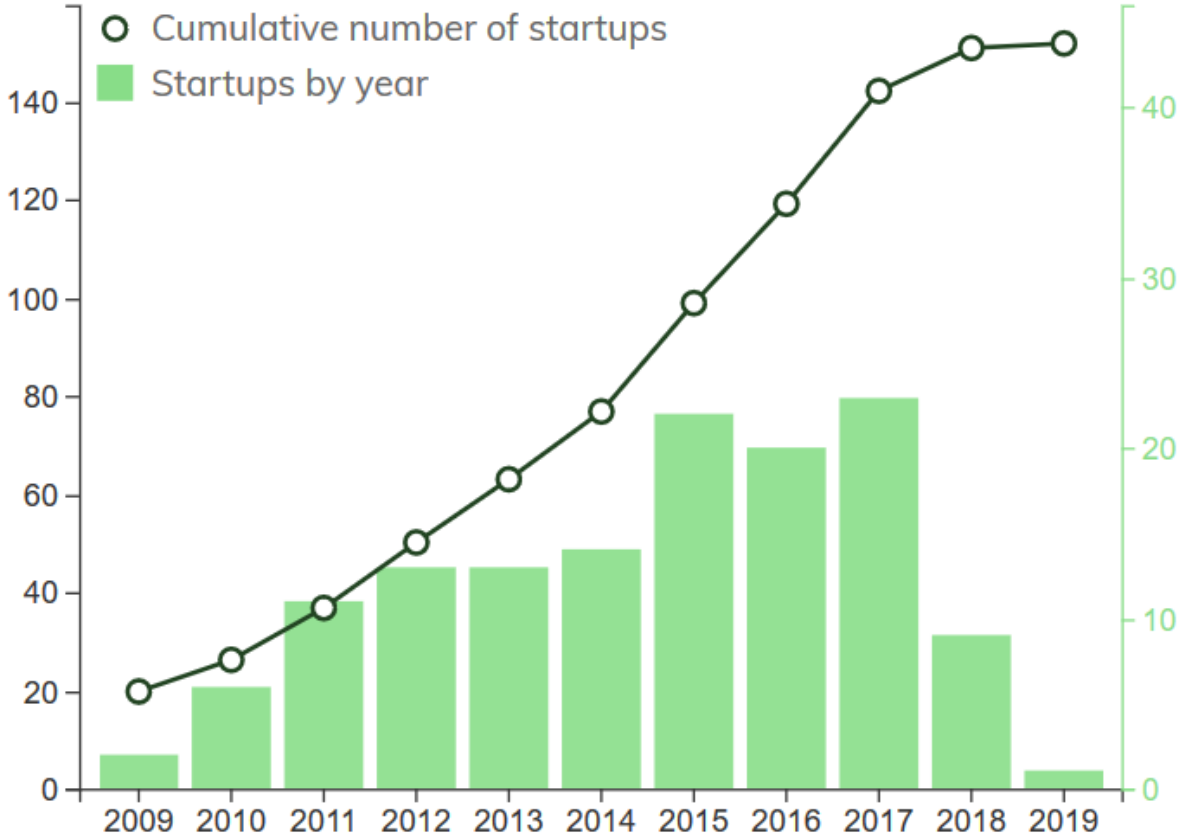
Oct. 2, 2019 by Andrii Buvailo

The race for adopting new machine learning (ML), deep learning (DL) and related technologies (for simplicity -- “artificial intelligence”/“AI”) keeps rapidly unfolding in the pharmaceutical industry, albeit with varying rate of progress across different use cases.

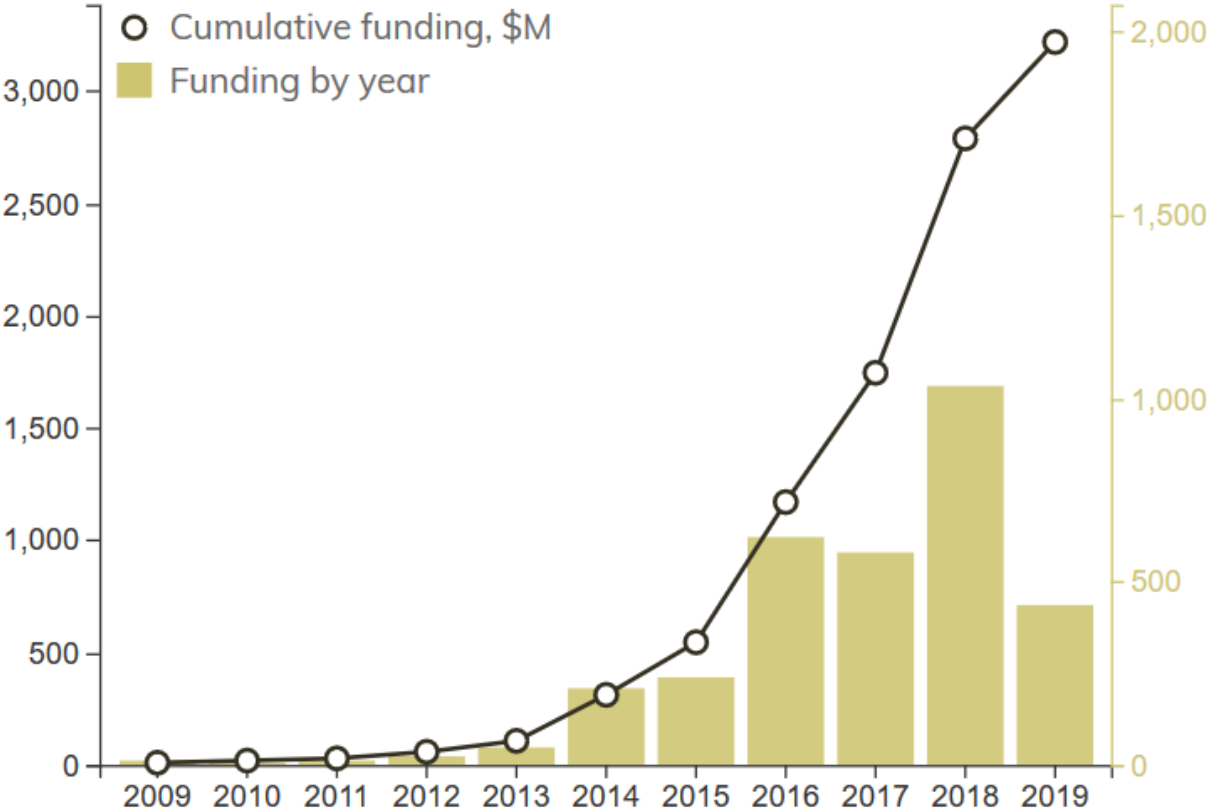
Let’s review retrospectively some of the key developments in the drug discovery area in 2019 and see how they characterize the current state of AI in the pharmaceutical industry (“hype vs reality”). Note, that I do not cover the healthcare sector in this post (diagnostics, medical applications of AI, digital health etc) -- those will be discussed in one of the future posts.

1. Notable Venture Rounds, New AI Startups

According to BPT Analytics report, the overall number of new AI-startups per year has been declining since 2017, with only 9 newcomers in 2018, and only one notable new AI startup focusing on drug discovery in 2019, so far -- Acorn AI, a “big data”-driven AI-company founded by Medidata to accelerate drug discovery via optimizing clinical trials.



The total amount of disclosed venture capital raised for AI-driven drug discovery startups in 2019 is around \$0.5B so far, the number likely increasing by the end of the year.



Below is a summary of some of the most important venture rounds in the “AI for drug discovery space” in 2019:

BlackThorn Therapeutics

San Francisco based BlackThorn Therapeutics, a clinical-stage biotech company focusing on developing first-in-class drugs for neurobehavioral disorders, closed a \$76 million Series B financing from multiple investors. The money will be used to further advance its drug discovery research efforts, including the application of AI-driven cloud-based platform PathFinder™ to generate objective neuromarkers, which support drug target identification, and improving the success of its drug candidates in clinical trials.

Exscientia

This British drug discovery company focusing on an “end-to-end” drug candidate development using its AI-driven platform Centaur Chemist™ raised \$26 million in round B in January, to scale up collaborations with big pharma, as well as build own portfolio of small molecules.

Insilico Medicine

The Hong Kong-based company Insilico Medicine, a developer of a comprehensive drug discovery and biomarker development platform GENTRL, and a pioneer in the application of generative adversarial networks (GANs) to drug discovery use cases, raised a \$37 million round B to commercialize its AI-driven “end-to-end” drug design platform GENTRL, and expand its business development team.

Insitro

Lead by a tech “star” Daphne Koller, Insitro is a San Francisco-based AI-driven company, focused on drug discovery, which raised its \$100 million Series A round earlier this year from a number of high-profile investors. Not much is disclosed publicly about the company’s platform and research strategy, but from its recent research collaboration with Gilead it is known the company will focus on disease modeling and lead molecules discovery to potentially treat nonalcoholic steatohepatitis (NASH).

Notable Labs

Another San Francisco-based company Notable Labs, which is an AI-driven personalized testing service for predicting which types of cancer patients are most likely to respond to a drug, closed \$40 million Series B round to scale up its capabilities and expand its platform globally.

Standigm

This South Korean drug discovery company focused on small molecule drug discovery using three of its proprietary AI-driven platforms -- Standigm BEST®, Standigm Insight®, and Standigm Expander® -- raised \$11.5 million Series B funding round to advance its drug candidates pipeline toward license-out.

Recursion Pharmaceuticals

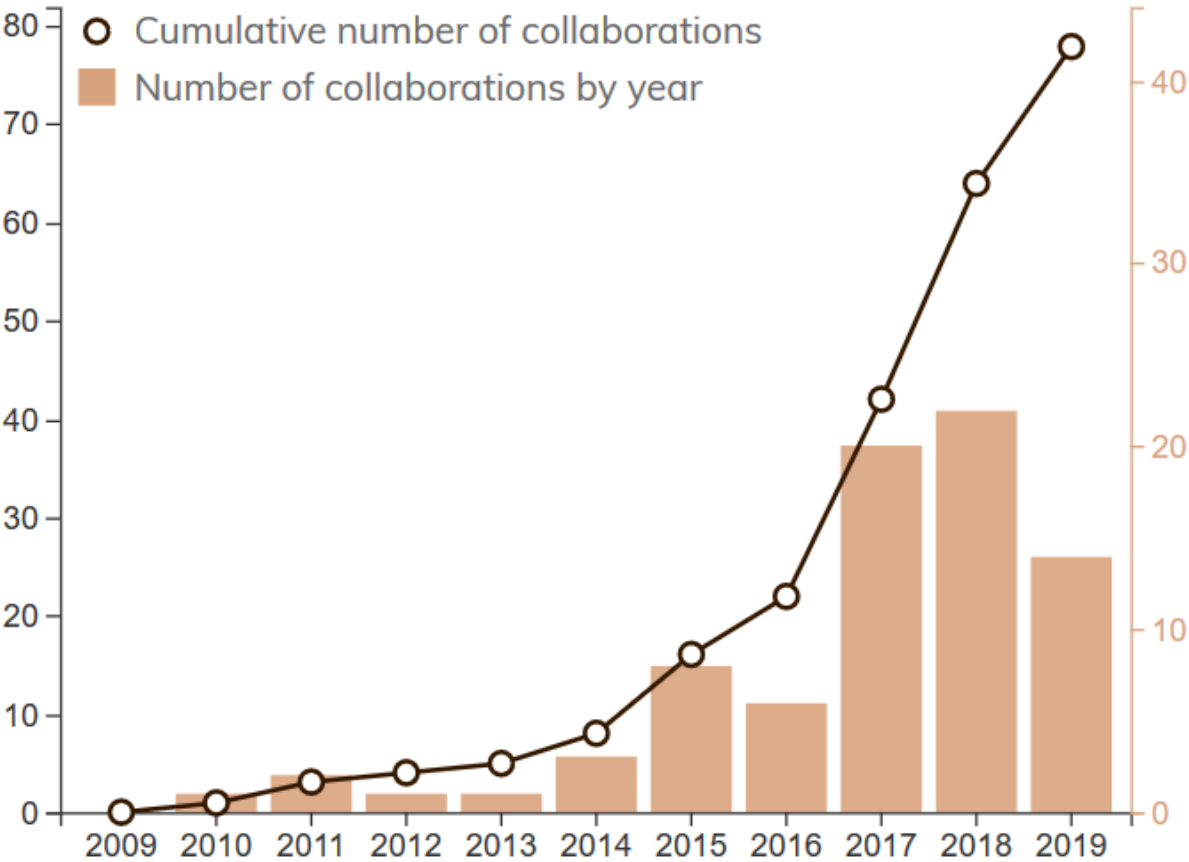
The Salt Lake City-based biotech company Recursion Pharmaceuticals, which uses automated, experimental biology with artificial intelligence to reveal new drug targets and develop drug candidates, raised \$121 million in a Series C round to further develop its in-house pipeline of small molecules and the technology.

Read also: [194 Companies Offering AI Solutions For Drug Discovery and Development](#)

2. Notable Research Collaborations

AI-startups with big pharma

While 2018 was a record year by the number of publicly announced research partnerships between big pharma and AI-driven startups to run pilot programs with new technologies (22), the year 2019 is characterized by a smaller number of “big pharma -- AI startup” deals so far (14).



Some of the most active AI-driven startups in deal-making with large pharma players in 2019 are Atomwise (on a deal-making spree, having partnered with Charles River Laboratories, Eli Lilly, Hansoh Pharma, and smaller players); BenevolentAI (partnered with AstraZeneca and Novartis); Exscientia (partnered with Roche and Celgene); and Iktos (partnered with Merck and Janssen).

Notably, the average deal valuations reported in 2019 are generally larger compared to the previous year, with a number of deals exceeding a billion dollars in potential milestone payments, as a recent Insitro deal with Gilead, worth up to \$1.05B in milestone payments. Two of 2019 deals had “record-breaking” status in this area:

- Atomwise and Hansoh Pharma (September 2019), reportedly, the “Largest China-US Collaboration for AI Drug Discovery”, potentially worth up to \$1.5B;

- Exscientia and Celgene, reported as “the largest deal in AI for drug discovery” by the time of the news communication in March 2019.

“Big pharma” with “big tech”

Big news just came out days ago about a large-scale five-year collaboration between Novartis and Microsoft aiming at bringing the latter’s advanced data analytics and AI capabilities to the far front of the drug maker’s ongoing research plans. Three main directions are in scope: personalized therapeutics for macular degeneration and irreversible blindness; cell and gene therapy manufacturing; and overall accelerating Novartis’ drug discovery and development processes.

It is notable that Microsoft earlier announced another AI collaboration with AstraZeneca to form AI Factory for Health -- an accelerating program for healthcare startups.

3. The “Proof-of-concept Year” for AI in Drug Discovery

It would take too long to list all the interesting announcements involving AI in drug discovery in 2019, but what is clear is that big pharma is now after AI, and the appetite of C-suite executives has grown dramatically since the last couple of years.

Let’s review several important developments that illustrate a sort of “technical” proof-of-concept for the AI in pharma research.

Deep Genomics revealed new drug target and a drug candidate ready for IND via AI

The company claims its AI-driven discovery platform AI Workbench, including more than 20 machine learning models, managed to identify a new genetic target for Wilson disease, and consequently -- a new oligonucleotide therapy DG12P1, which the company intends to advance into clinical studies within two years. The whole process took less than 18 months, compared to 3-6 years for a “traditional” case scenario.

Exscientia reaches a milestone in AI collaboration with GSK

Exscientia announced that it has managed to reach its first major milestone within its AI-drug discovery collaboration with GSK, which started in 2017. The AI company delivered a highly potent in vivo active lead molecule targeting a novel pathway for the treatment of chronic obstructive pulmonary disease (COPD). The in vivo active molecules were discovered with 5 cycles and only 85 compounds tested.

Insilico Medicine validates its AI-drug discovery platform with in-vivo tests

A recent article in Nature Biotechnology “Deep learning enables rapid identification of potent DDR1 kinase inhibitors” published by Insilico Medicine made quite a noise in the pharmaceutical media space after the company managed to confirm in-vitro and in-vivo activity for some of the lead molecules, “imagined” from scratch by their AI-driven system GENTRL (Generative Tensorial Reinforcement Learning) -- in a matter of weeks. This is an important conceptual milestone for the progress of generative adversarial networks (GANs) in drug discovery -- a type of deep learning that Insilico has been pioneering in pharma since 2016 (a historic timeline is included in the referenced post).

Recursion Pharmaceuticals reaches a milestone in AI collaboration with Takeda

Recursion recently announced that it discovered a number of novel preclinical and clinical small molecules in over 60 unique indications, with new therapeutic candidates identified for over a half a dozen diseases -- within its AI-driven collaboration with Takeda Pharmaceuticals, started in 2017. This year Takeda has exercised its option for drug candidates in two rare diseases, and the companies extended collaboration. Clearly, the pharma giant is happy with what Recursion offered via its AI-driven effort.

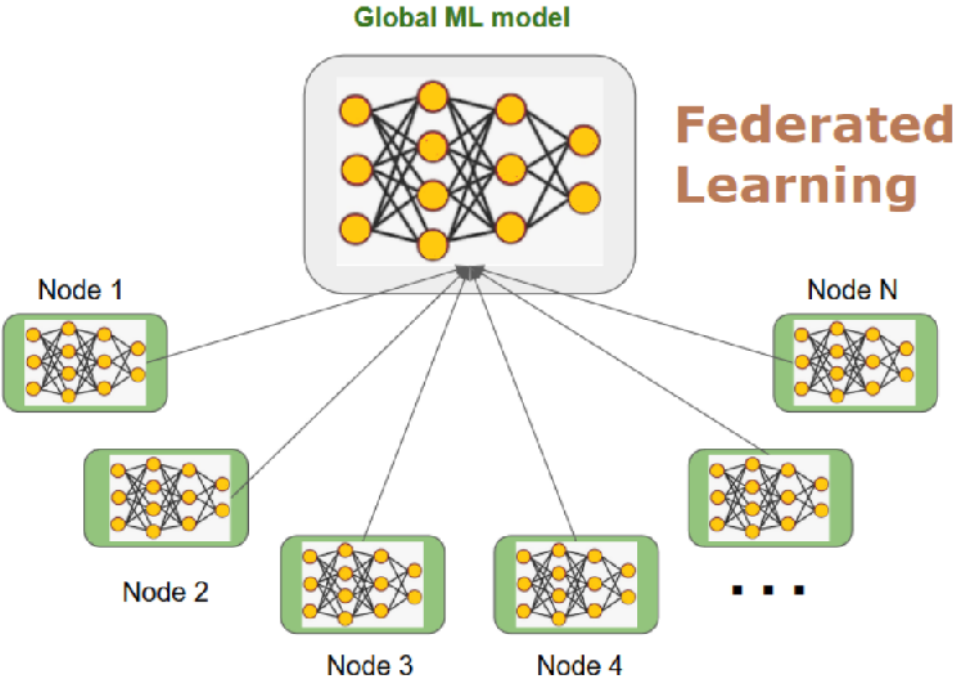
4. Pharma Looks for New Collaboration Models to Fit AI

Deep learning technologies, the cornerstone of “AI-revolution”, require a lot of data that is not always available, or easily accessible in the pharmaceutical industry for the AI-developers to properly train their systems.

Another problem is intrinsic biases in the available data -- all because the scope of data sources is often limited, meaning all available data might have a limited representation of features, not reflecting the true nature of what is modeled. There is an insightful article illustrating this issue on the example of hidden biases in a well-known DUD-E dataset, widely used for training convolutional neural networks: DUD-E, Where's My Accuracy?

Generally, data availability in drug discovery is a very different situation from what we see in more "classical" applications for AI -- natural language processing and image recognition, where data is much more abundant and diverse (millions of text samples in many languages; millions of images with diverse contents etc). So the impressive progress of AI in those areas is not necessarily possible as quickly and comprehensively in the pharmaceutical industry.

One important step to tackle data availability and data bias in pharma have been undertaken in 2019 by a group of pharmaceutical companies and research institutions with a project MELLODDY (Machine Learning Ledger Orchestration for Drug Discovery), which uses such advanced technologies as blockchain and federated learning. Conceptually, the project will allow training machine learning models on numerous private datasets owned by various companies, without compromising their intellectual property, or confidentiality of sensitive data -- here is a more in-depth explanation of how it works.



It is now hard to say how far MELLODDY project will go, but the concept is powerful and I am sure it will be chased by numerous other life science organizations. Undoubtedly, being able to use everyone's research data (currently hidden under companies' firewalls) for the common benefit of training AI models will be a game-changing step towards AI-driven transformation of the pharmaceutical industry.