MIT-IBM Watson AI Lab

Annual Report

2021
2022

Table of Contents

Message from the Directors

With five years under our belt, the MIT-IBM Watson AI Lab has found its stride and is continuing to pick up the pace. Our work — to advance AI science and discovery with hardware, software, and algorithms; examine how AI impacts society on an economic and ethical basis; and translate our research into real-world applications — spans from the foundational to exploratory. This coming year, we've added 30 new projects to our roster, bringing our total portfolio to more than 80 collaborations. We're maintaining our broad coverage of fundamental and emerging research topics, and with insights from our corporate members, we're working to evolve and accelerate our "lab to market" pipeline, which taps into areas of finance, manufacturing, telecommunications, heavy industries, and more along the way.

In the last year, we've been able to further facilitate industry-academia research in our beautiful space, allowing for team meetings, executive discussions, educational webinars, student mentorship and internship opportunities, events, and much more. Many of you joined us for our in-person open house and industry showcase, when we discussed and demonstrated current and future research and directions, such as a focus on impactful work with our corporate members and our benefit offerings.

This year, we look forward to further ramping up our face-to-face collaborations. Lab-supported graduate students have joined our ranks to help solve challenges our industry members are facing, and appreciably innovate beyond the problem at hand; invaluable guidance from you has helped to frame their work and their education. Another cohort will join us this year, in addition to undergraduates and students from around the world through our internship program.

We're thrilled with what the Lab has accomplished to date, as well as the direction it's moving. In addition to engagement with corporate-directed work and young researchers, the Lab has continued to submit patents, and present and publish in top-tier journals and conferences. The following is a highlight from where we stand, and we look forward to what's on the horizon.

David Cox and Aude Oliva
IBM Research MIT

Lab by the Numbers

Over 700 peer-reviewed papers

Over 500 news and media mentions

Over 420 project proposals submitted from MIT and IBM

$240M 10-year investment to found a joint lab

159 researchers across MIT and IBM

Over 100 principal investigators across 5 schools and 1 college

Over 80 active projects

40 patent disclosures

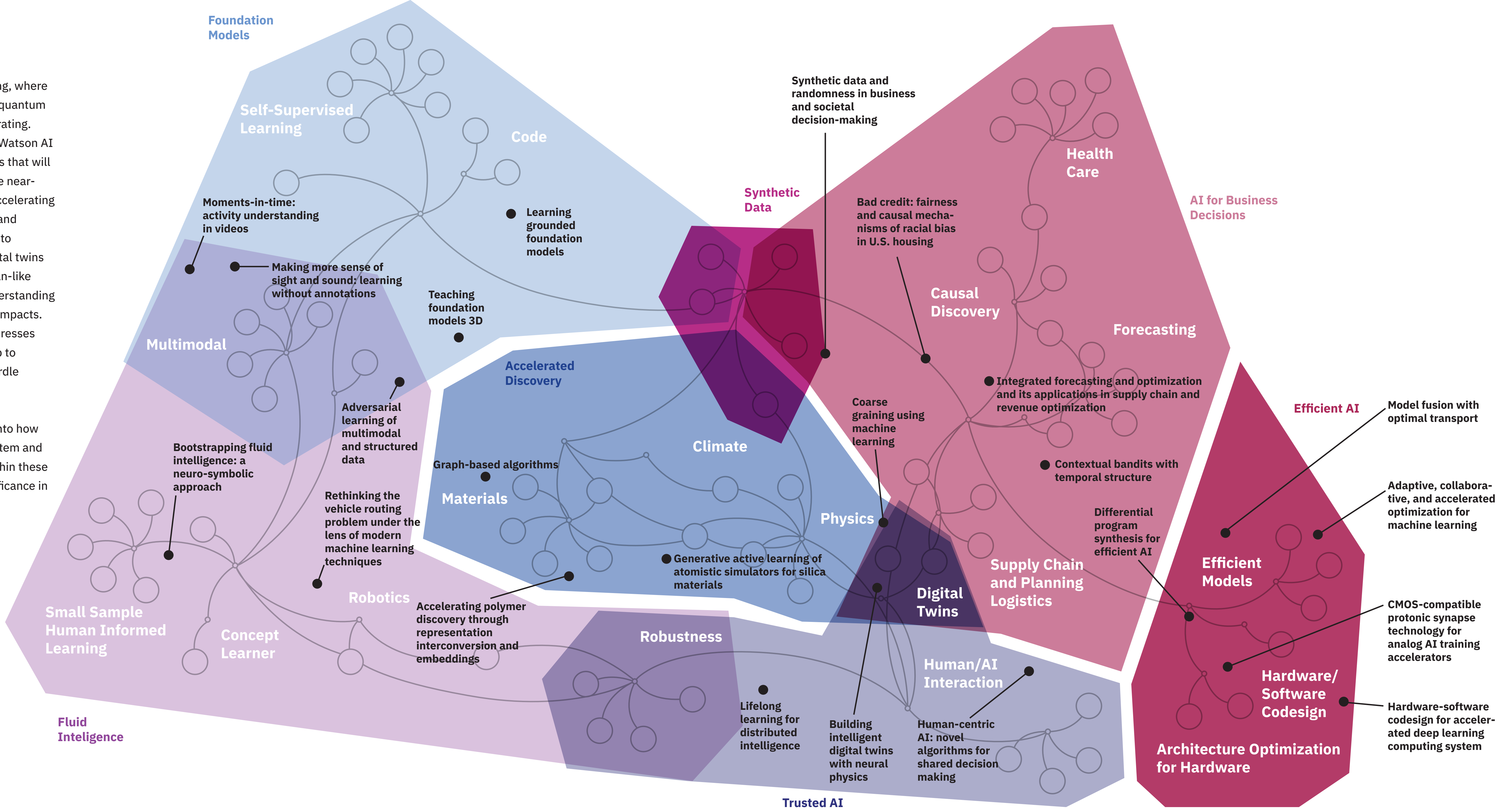
25 industry use cases

5 current member companies

Selected Projects

We're in the next wave of computing, where classical computing meets AI and quantum computing, and progress is accelerating. Holding this in mind, the MIT-IBM Watson AI Lab invests in a portfolio of projects that will develop innovations for gains in the near- and long-term. This ranges from accelerating discovery within the medical field and building AI models that are robust to adversarial attacks to creating digital twins of physical systems, infusing human-like reasoning into AI models, and understanding supply chains and socioeconomic impacts. Each of these tailored projects addresses an urgent need and sets the Lab up to understand and tackle the next hurdle presented by industry and society.

The following provides a glimpse into how the Lab thinks about the AI ecosystem and highlights some of our projects within these overlapping thrusts and their significance in the real world.



Moments-in-time: activity understanding in videos

Aude Oliva, SouYoung Jin,
Mathew Monfort (MIT)
Rogerio Feris,
Dan Gutfreund,
Rameswar Panda (IBM)

Animals and humans are bombarded with visual information as we move about the world. But the brain manages to extract just enough detail to make sense of it all and not get overwhelmed. The virtual world of video, by contrast, contains far more information, much of it redundant. Each second of a clip may contain more than a dozen frames that are unnecessary to understanding the events that are unfolding. In this project, researchers are experimenting with ways to cut as many frames as possible to build video-recognition models that are substantially smaller and more efficient than those used today. The goal is that the systems will be able to capture semantic relationships between events in dynamic settings, recognize particular actions in other domains or without seeing it before, and synthesize new videos based on language prompts.

"Sixty percent of the brain is dedicated to visual processing, reasoning, and imagination, capabilities that are at the core of human cognition. So, artificial systems that incorporate this are and will become even more ubiquitous in everyday life."

Model fusion with optimal transport

Justin Solomon (MIT)
Mikhail Yurochkin,
Kristjan Greenewald (IBM)

Modern AI models excel at specialized tasks, but fail when asked to learn too much, especially when different types of data are thrown at them. In this project, researchers will use optimal transport, a set of principles borrowed from geometry and probability theory, to fuse together a variety of specialized models to build an AI system with greater versatility. The work has applications in privacy-preserving federated learning, transfer learning, multi-modality inference, and natural language processing.

"Geometry provides a natural viewpoint to understand data and AI models. Our group developed geometric tools to understand text, image, and biological data to anticipate failures of deep models and to enhance reliability of machine learning."

> “Our project seeks to develop deep-learning models that can **learn high-level semantic correspondences** between visual and auditory, especially spoken, inputs that could improve the performance of multimodal AI systems, shed some light on how humans learn audio and visual representations, and assist humans in low-resource domains, like in customer care.”

James Glass,
Antonio Torralba (MIT)
Samuel Thomas,
Dhiraj Joshi,
Rogerio Feris,
Yang Zhang,
Leonid Karlinsky,
Hilde Kuehne (IBM)

Making more sense of sight and sound: learning without annotations

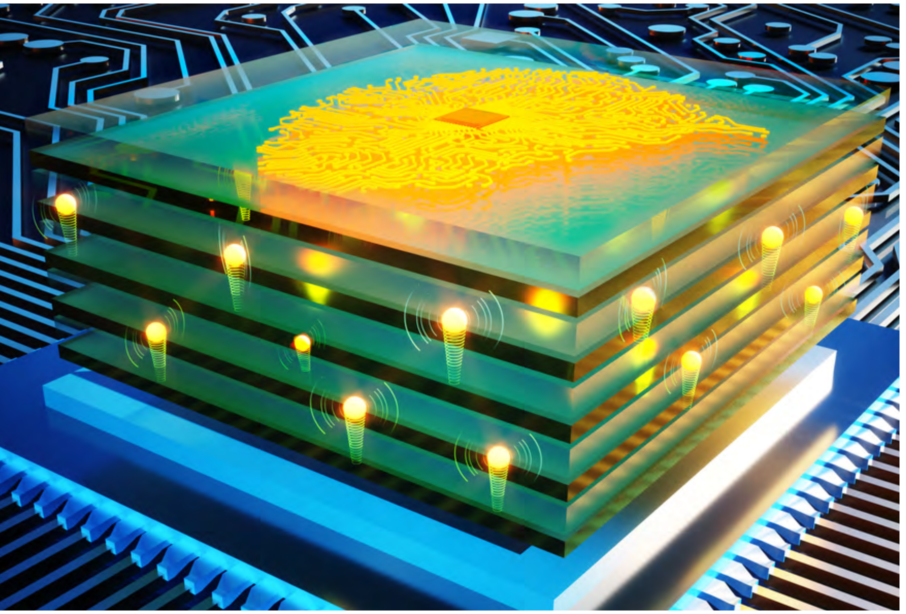
YouTube’s massive trove of videos gives researchers a rich source of data to leverage in training AI models to recognize how objects and activities shown in videos correspond to their spoken word representations. In this project, researchers will use cooking shows and other instructional videos to train a deep learning model to learn to associate objects and actions, like “hot dog” and “cutting,” with their raw speech representations. Applications for the work include indexing the mountains of audio and video surfacing online daily.

Jonathan How (MIT)
Gerald Tesauro,
Matthew Riemer,
Miao Liu (IBM)

Lifelong learning for distributed intelligence

In an increasingly automated future, humans and robots will likely work together in fast-paced, changing environments. In this project, researchers will address some of the key challenges of representing new skills and new models of the environment in a robot agent, sharing and combining this information with other robots on the team, and enabling robots to respond rapidly to new scenarios. The challenge includes creating principled methods for robots to communicate with humans and other robots in ways that are easily understood by humans, and quickly adapt their behaviors and learn to solve tasks in new environments using minimal data.

> “Our research has created new distributed learning algorithms for solving important, practical, and challenging problems in which there are multiple **independent decision-makers interacting in real-world environments** (e.g., supply chain management or finding an efficient economic policy) with possibly conflicting or changing interests over time.”

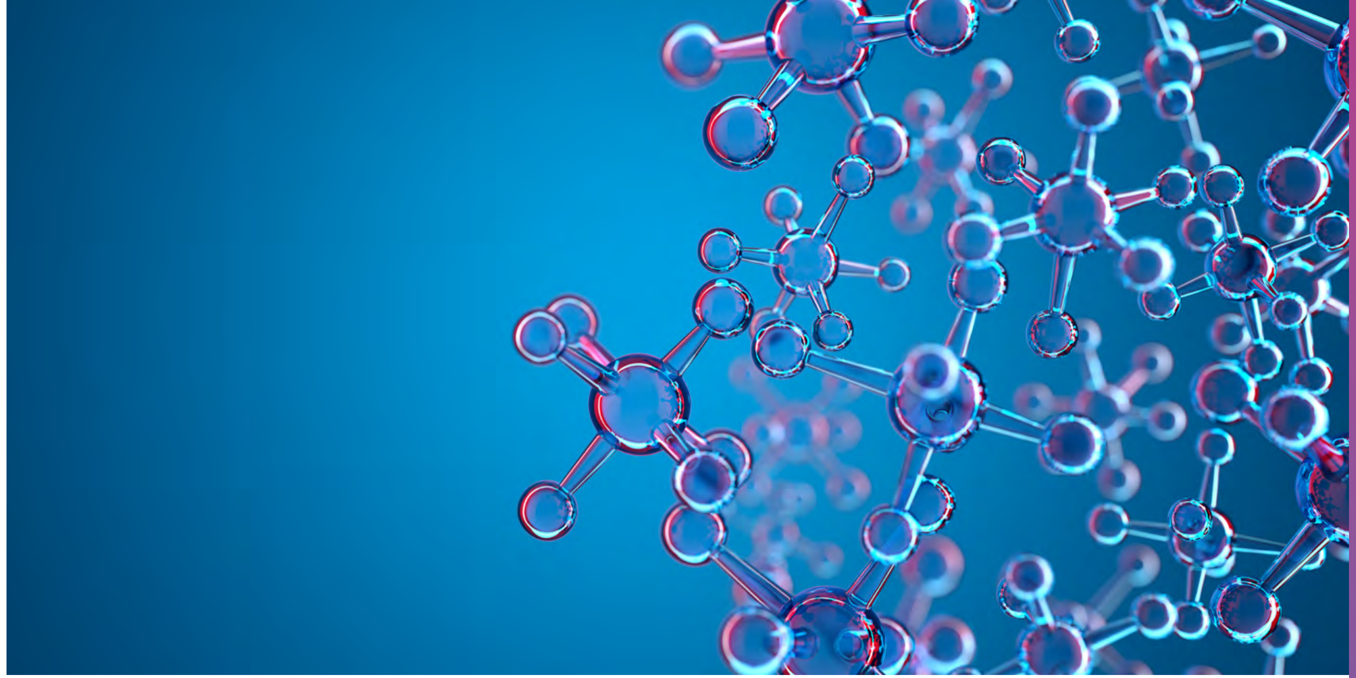


CMOS-compatible protonic synapse technology for analog AI training accelerators

Jesus del Alamo,
Ju Li, Bilge Yildiz (MIT)
John Rozen,
Dirk Pfeiffer,
Takashi Ando,
Paul Solomon,
Malte Rasch,
Teodor Todorov (IBM)

As deep learning models become exceedingly more complex, they require more energy, time, and money for training and inference. To keep up, researchers are designing and fabricating biologically inspired hardware arrays of programmable resistors — i.e. analog “neurons” and “synapses” — that can execute computation significantly faster and more efficiently than the human brain. In this project, researchers are working to optimize processing circuitry architecture with small memory devices and manufacturing techniques called complementary metal–oxide–semiconductors (CMOS), which store input and output settings and minimize energy loss. Using this process, the team is developing a programmable, nonvolatile resistor, where a proton is pushed from a palladium solid hydrogen reservoir through a protonic electrolyte (nanoporous phosphosilicate glass) and into an insulating oxide, like a tungsten trioxide channel. Voltage pulses modulate the device reversibly, symmetrically, and reproducibly. Such proton intercalation devices allow for computation within the device’s memory and parallel operations with the promise of being scalable.

“Our project aims to scale up **a new class of nonvolatile protonic programmable resistor technology** for analog deep learning pioneered at the MIT-IBM Watson AI Lab. We will develop an integrated technology suitable for large-scale arrays and on-chip training demonstrators.”



Graph-based algorithms

Wojciech Matusik (MIT)
Jie Chen, Payel Das,
Veronika Thost (IBM)

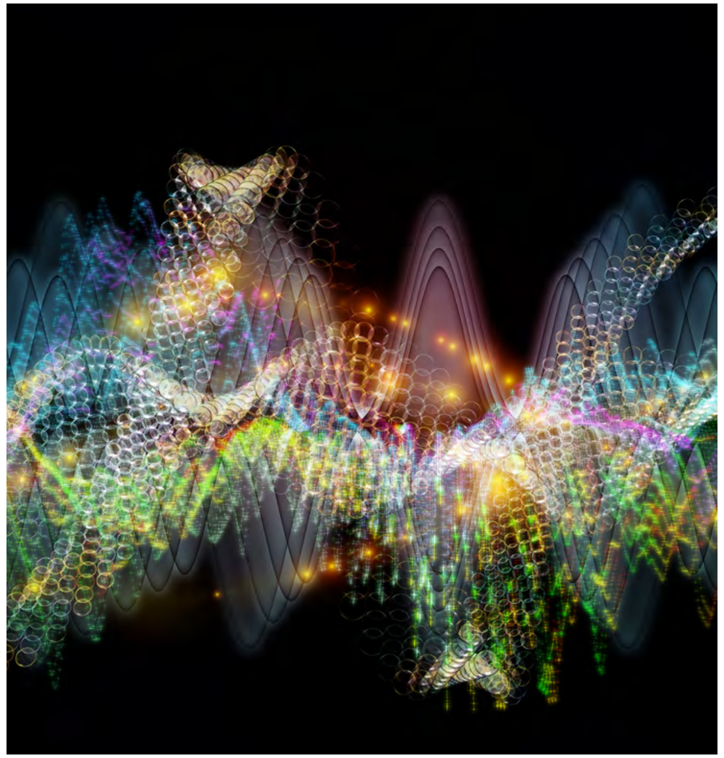
“This project exploits the learning of symbolic representation — grammar — over the intractable space of graphs, to deliver a data-efficient solution for graph-based designs that **accelerate molecular discovery** to meet growing societal challenges and build a sustainable future.”

A graph structure can represent rich data and relationships between them. Graphing algorithms that visualize how everything from computers to atoms connect with each other have become powerful design tools. In this project, researchers will explore Bayesian optimization methods and develop a grammar-like set of constraints to guide neural networks as they design new ways of arranging computers and atoms to create new computer systems and never-before imagined molecules. The constraints will allow the neural networks to efficiently search outside training datasets for valid and synthesizable structures that fit a desired set of properties. This pipeline will ultimately serve to design novel and optimal cyber-physical systems and chemical compounds.

Synthetic data and randomness in business and societal decision-making

Dean Eckles (MIT)
Akash Srivastava,
Naoise Holohan,
Kristjan Greenewald,
Sara Magliacane (IBM)

Methods for privacy-preserving sharing of data are seeing rapid development and adoption in industry and government. This adoption is motivated by both new regulations, but also recognition that high-dimensional models often memorize data. Through this work, researchers will increase the usefulness of privacy-preserving data sharing (especially via synthetic data generation) and study fruitful connections with randomized experimentation, fairness, and causal inference. First, firms and government agencies are increasingly turning to generating synthetic datasets, rather than only specific, privacy-preserving statistics. This work will advance methods for creating synthetic datasets that can be used for multiple purposes, which often cannot be precisely articulated ex ante. Second, privacy-preserving methods, including synthetic data, can affect downstream decisions, including by implicitly creating randomness in decisions. This research studies this randomness, including developing ways to make it useful (e.g., it creates "free" randomized controlled trials) and ways it may result in systematic unfairness. Third, the work develops privacy-preserving methods for estimating causal effects. Not only is causal inference increasingly a core part of business analytics (and so synthetic data will be used for causal inference), but it is also uniquely enabled by the randomness studied here.



"Privacy considerations can present barriers to data sharing and use within and between organizations. In this project, we explore how **statistically private synthetic data,** which preserves key features of the real data while providing formal privacy guarantees, can mitigate these barriers."

"Machine learning promises to automate tasks currently performed by humans, but what happens when the algorithm's prediction is off target? Our research enables machine-learning algorithms to know when to predict and when to allow humans to take over and furthermore teaches human users to better understand the algorithms' failures, to **enable more accurate decisions** in domains such as healthcare and content moderation."

David Sontag,
Arvind Satyanarayan (MIT)
Subhro Das,
Dennis Wei,
Prasanna Sattigeri (IBM)

Human-centric AI: novel algorithms for shared decision making

Humans make around 35,000 decisions in a day, ranging from simple to complex and stressful ones, which can lead to decision fatigue; AI can help relieve some of this. As the capabilities of machine-learning agents expand, particularly in the areas of healthcare, content moderation, collaborative writing and autonomous driving, humans will increasingly be able to defer to these AI systems for decision making. However, most models cannot predict with 100 percent certainty, so researchers are designing AI systems to collaboratively interact with humans to recommend decisions; for expert deferral, when the AI system lacks confidence in its decision; and for multi-round communication to teach the human user when to accept or override the AI's recommendation and refine the model. Researchers will test this work with radiology data and investigate how to mitigate bias within the model. Such processes will augment the tasks that human-AI systems will be able to complete with optimal decision outcomes.

People

Aude Oliva

MIT Director, MIT-IBM Watson AI Lab
Director of Strategic Industry Engagement, MIT Stephen A. Schwarzman College of Computing

David Cox

IBM Director, MIT-IBM Watson AI Lab
Director of Exploratory AI Research, IBM

Anantha Chandrakasan

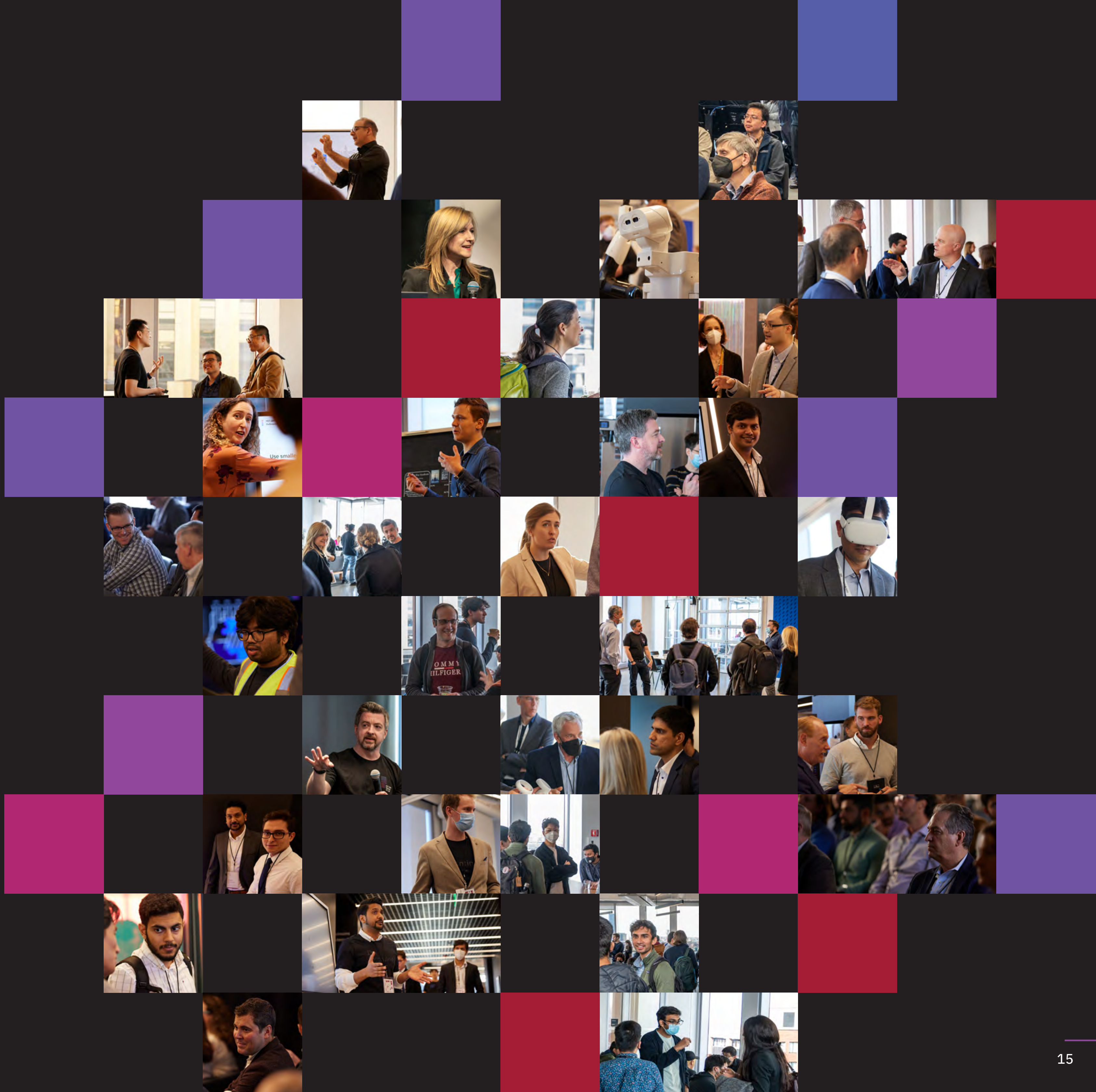
MIT Chair, MIT-IBM Watson AI Lab
Dean, MIT School of Engineering

Dario Gil

IBM Chair, MIT-IBM Watson AI Lab
Senior Vice President and Director of Research, IBM

Daniel Huttenlocher

MIT Co-Chair, MIT-IBM Watson AI Lab
Dean, MIT Stephen A. Schwarzman College of Computing



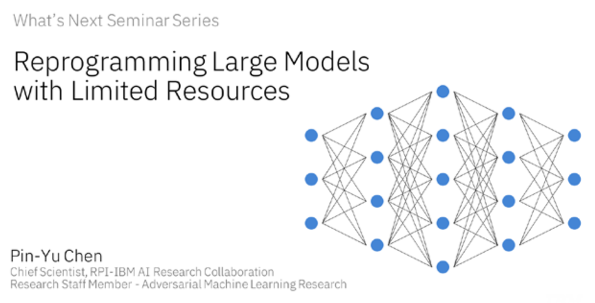
Events

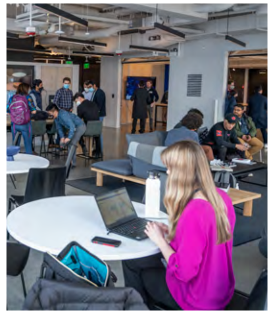
Climate Implications of Computing and Communications

During a two-day, virtual workshop in the spring of 2022, the MIT-IBM Watson AI Lab, the MIT Stephen A. Schwarzman College of Computing and the MIT Climate and Sustainability Consortium brought together an interdisciplinary group of MIT researchers and industry leaders across sectors to discuss and explore collaborative approaches to reduce energy consumption for the future of computing.

What's Next? Seminar Series

In order to be on the cutting edge of industry, corporations strive to be knowledgeable about and implement the latest technologies, while planning for the future innovations. To help them do this, the MIT-IBM Watson AI Lab invited corporate members and convened leading Lab researchers for seven, virtual seminars throughout the year to discuss technical advances in the fields of AI, quantum computing, and accelerated discovery, followed by a moderated discussion for members' employees.





MIT-IBM Watson AI Lab Open House

Community is central to innovation. In April 2022, the MIT-IBM Watson AI Lab invited researchers and students into the Lab, providing a behind the scenes look at current and upcoming research projects with member companies, and MIT and IBM researchers. Lab co-directors Aude Oliva and David Cox opened the event, followed by a poster session, videos about projects, and networking via live-chat.

MIT-IBM Watson AI Lab Industry Showcase

In May 2022, industry members, researchers, and collaborators gathered in the Lab for advisory meetings, keynote and panel discussions, and exhibits concentrating on AI deployment. Industry attendees met with research teams for assessments of projects and tailored conversations, such as AI educational overviews. Lab co-directors Aude Oliva and David Cox provided executives with headlights into AI's future and the Lab's corresponding research strategy.

Focusing on company-wide implementation and real-world impacts, MIT economist Neil Thompson examined financial, compute, and human infrastructure needed for AI utilization, with insights from industry members. Dario Gil PhD '03, IBM chair, senior vice president and director of research at IBM, shared how classical, cloud, and quantum computing will accelerate discoveries across sectors.

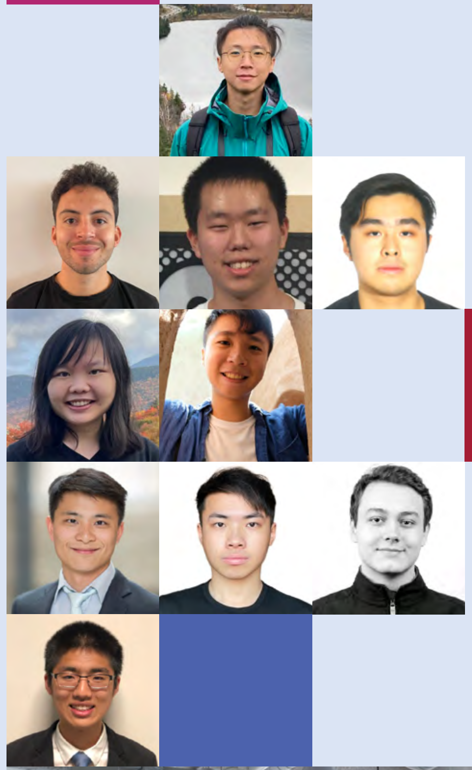
Campus Engagement

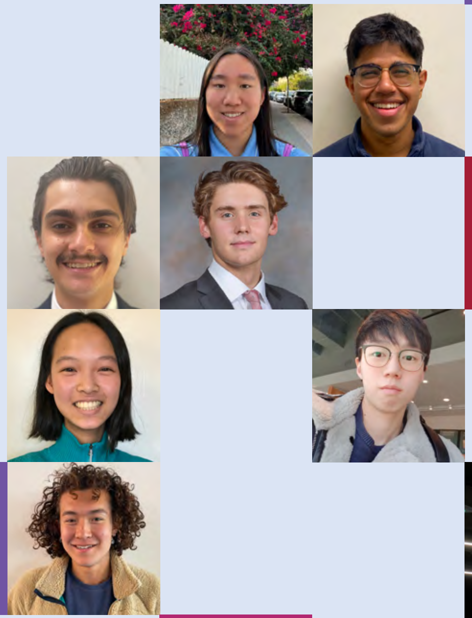
To keep up with the extraordinary rate of progress on AI, the MIT-IBM Watson AI Lab supports the development of the next generation of thinkers, beginning with undergraduate students up through early career researchers.

Working with the Lab, MIT undergraduates can participate in their own education for a semester/summer, up to a year or more, by experiencing what it’s like to contribute to a real AI project and seeing how premier researchers are thinking about challenges and potential applications. Here, students apply their knowledge to research in their area of interest, guided by Lab advisers. Over its lifetime, the Lab has funded research for roughly 100 undergraduate students.

Graduate students, largely from programs in the MIT Department of Electrical Engineering and Computer Science, work alongside MIT and IBM Research experts to immerse themselves in applications of their current field of study, while completing their thesis. Some graduate students can further opt to collaborate on industry-led projects, receive mentorship, and provide solutions to frontline challenges in AI, enhancing their understanding of AI use in the real world and insights for use toward their graduate degree.

In addition, the Lab hosts summer interns and is working with the Institute to further increase the equity of opportunities for young researchers and diversity of people and ideas entering the fields of computing and AI, as they will make critical impacts in these areas and for society.





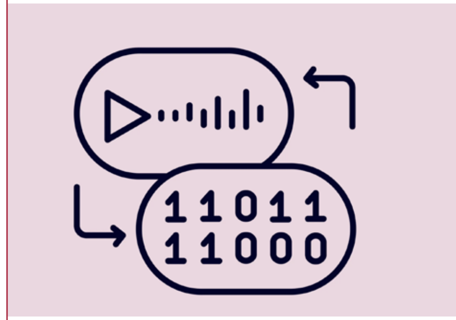
Since the summer of 2021, the Lab has seen 20 masters and 9 undergraduate students pass through our doors.

In the Media

The MIT-IBM Watson AI Lab and its researchers were cited or featured by numerous media outlets, including *Wired, TechCrunch, Nature, Popular Science, National Public Radio, The Conversation, IEEE Spectrum, Quanta Magazine,* and *Smithsonian Magazine*.

IEEE Spectrum

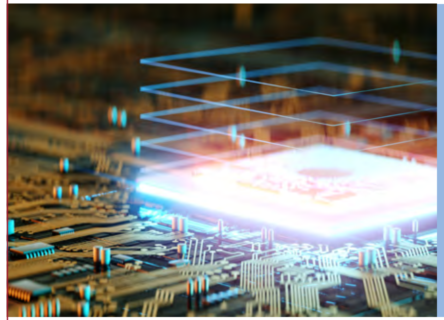
Clever compression of some neural nets improves performance



An AI method from the MIT-IBM Watson AI Lab may help democratize automatic speech recognition, particularly for uncommon spoken languages. The technique, PARP (Prune, Adjust, and Re-Prune), reduces the complexity of a powerful pretrained speech recognition, machine-learning model, improving the model's efficiency and boosting its performance.

New Scientist

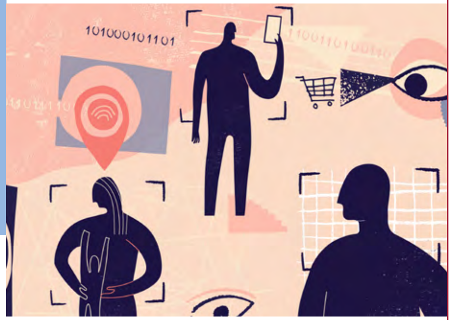
Researchers from the MIT-IBM Watson AI Lab have created an "artificial synapse" device, specifically, a network of nanoscale resistors, that can transmit protons across terminals at speeds 10,000 times faster than biological synapses. Further, this analog machine-learning technique requires little energy and is efficient possibly offering advantages over digital machine learning.



'Artificial synapse' could make neural networks work more like brains

Venture Beat

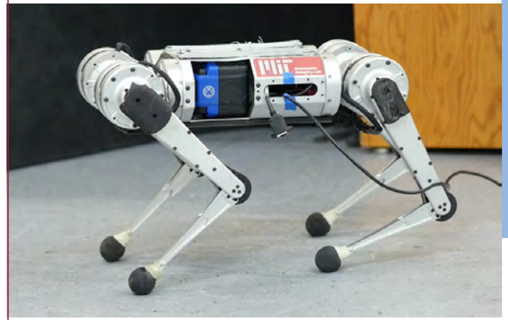
Humans excel at abductive reasoning: piecing together snapshots from our experiences and environment to predict what is happening around us, like a doctor or juror examining partial evidence and drawing a conclusion. A dataset from the MIT-IBM Watson AI Lab, called Sherlock, helps to assess the same ability in AI systems through a collection of paired visual scenes and text clues.



This new dataset shows that AI still lacks commonsense reasoning

Popular Science

Machine learning helped MIT's cheetah robot break its own speed record



MIT-IBM Watson AI Lab-supported research uses a model-free reinforcement learning system, which allows the robotic mini cheetah to learn entirely by trial and error in simulation, and maximize its speed based on the terrain. With this capability, it has broken the record for the fastest run recorded.

Quanta Magazine

AI researchers fight noise by turning to biology

Humans are adept at picking out important information for predictions and aren't distracted by artificial noise, which can fool a neural network. To understand and combat this problem in visual and auditory prediction, MIT-IBM Watson AI Lab researchers are leaning on neuroscience and employing noise generator filters that mimic the brain's noisy neurons, with success.

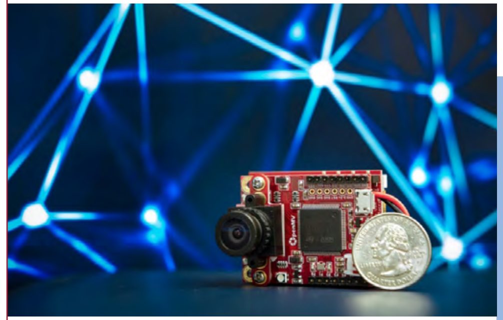
TechTalks

AI is a great tool for understanding patterns; however, researchers are trying to take it a step further, to plan actions, understand goals of other agents, and solve problems, like humans can. In an IBM workshop, MIT professor and MIT-IBM Watson AI Lab researcher Joshua Tenenbaum described how he's thinking about intuitive physics and psychology.

What happens when you combine neural networks and rule-based AI?

Tech Monitor

Putting AI in IoT chips? It's a question of memory



MIT-IBM Watson AI Lab researchers are working to shrink deep learning down to run on internet-of-things (IoT) microcontrollers without the need for cloud computing — a technology called tiny machine learning (tinyML). In this way, the technology can preserve privacy, reduce costs to train and run machine-learning models, reduce latency, and make the IoT device more reliable for households.

National Public Radio

MIT-IBM Watson AI Lab researchers developed a machine-learning method that can learn to control a fleet of autonomous vehicles as they approach and travel through a signalized intersection in a way that keeps traffic flowing smoothly and avoids idling at red lights. The technique also reduces fuel consumption and emissions, while improving average vehicle speed.

Green driving

Member Impacts

Case Studies

As a course of practice, the MIT-IBM Watson AI Lab tackles fundamental questions in the AI and applications space, as well as real business challenges encountered by our industry members. Through our symbiotic relationship, we're able to develop new, customized research and solutions that are then implemented with demonstrated impacts.

Reducing the dependence on data for manufacturing inspections

A member company sought to leverage computer vision to automate quality inspection processes across their various product lines. However, traditional data-hungry techniques would have required training hundreds of new models, one for each product line. Any cost savings from the automation would not have justified the total model training costs. The Lab developed a new transfer learning technique that leverages correlations to take a model, trained in one domain (e.g., a product line), and applied it to accurately make predictions in a new domain (e.g., a different product line). When prototyped and piloted at the member company's production line, this technique was both more accurate and required less data than traditional methods.

Key Metrics:
- 70% reduction in training data requirements for research prototype
- 16 months to onsite pilot

Separating correlation from causation in customer behavior

When analyzing customer behavior, it is critical to separate correlation from causation. For example, customer attrition might be highly, positively correlated with sales support. That does not mean, however, that the extra sales support causes attrition. Rather, there is a spurious correlation in the data that the customers most at risk of leaving are given the most extra support. Traditionally, the only way to define causal effects is through costly experimental interventions. The Lab's researchers have developed new methods that leverage AI to search out and identify causal relationships within large historical datasets, avoiding the need to run large, expensive experiments. After applying causal inference methods to a member company's customer attrition data, the Lab provided the corporate member with a prioritized list of potential marketing interventions, ranked by their ability to reduce customer attrition.

Key Metrics:
- 175 marketing interventions prioritized by causal impact
- 13 months to prototype

Member Benefits

A new way to engage for R&D progress

The MIT-IBM Watson AI Lab is the world's leading academic-industry alliance for advanced AI research, empowering scientists from MIT and IBM Research to co-conceive, co-publish, and co-patent groundbreaking work. While it's not unusual for industry to fund academic research, this unique commitment is designed to foster long-term relationships, diverse research activity, and cross-pollination of themes.

Our work is bolstered by the Lab's membership program, which invites companies to co-invest in the Lab in exchange for intellectual property and other membership benefits. All of our members are market and technology leaders in their respective industries. They bring a breadth of perspective and strategic influence by providing real-world challenges where innovation is needed, helping us achieve our mission to push the envelope of what AI can do and to make a positive impact on business and society.

To achieve this, the Lab provides access to core research and engagement teams, who identify pathways to solutions within the prescribed domain space, develop novel application-focused research, and ensure the scope and goals of the project align with those of our stakeholders. The Lab does this by providing use case workshops and project updates; collaborating through regular office hours, code-sharing, and communication; and coordinating project reviews and prototyping. Together, the Lab and our member companies can move in lockstep to solve our members' pressing problems while improving AI's capabilities.

Member Benefits:

Cutting edge research

- Collaborate with researchers to invent novel AI methods
- Secure license to selected intellectual property

Applied project engagement

- Drive project direction around your use cases and data
- Build prototypes with our engineers to showcase your research applications

External eminence

- Co-author pieces on AI innovation for conferences, news articles, and other opportunities

Talent development

- Skill-up top talent by embedding them on project teams with MIT and IBM researchers
- Recruit student talent at MIT through targeted on-campus recruiting programs

MIT-IBM
Watson
AI Lab

MIT Schwarzman College of Computing

MIT School of Engineering

IBM Research