

Science AMA Series: We see an opportunity to achieve a deeper understanding of intelligence. We are MIT faculty members Anantha Chandrakasan, Daniela Rus, and James DiCarlo. AMA!

MIT *official*¹ *and* / *ScienceAMAs*¹

¹Affiliation not available

April 17, 2023

Abstract

Unfortunately, that's all the time we have to answer your questions today. Thanks, everyone for your engaging questions! Follow: @MIT, @MITEngineering, @MIT_CSAIL, and @mitbrainandcog to continue to get news around our work. ++++++ At MIT, we are on a quest to answer two big questions. How does human intelligence work, in engineering terms? And how can we use that deep grasp of human intelligence to build wiser and more useful machines, to the benefit of society? We aspire for our new knowledge and newly built tools to serve the public good. Read this MIT news article to learn more: <http://mitsha.re/5k6D30i80qQ> About us Anantha Chandrakasan: I am the dean of the School of Engineering at MIT. Before being named Dean, I was the Vannevar Bush Professor and head of the Department of Electrical Engineering and Computer Science (EECS). During my tenure at EECS I spearheaded a number of initiatives that opened opportunities for students, postdocs, and faculty to conduct research, explore entrepreneurial projects, and engage with EECS. Daniela Rus: I am the Erna Viterbi Professor of Electrical Engineering and Computer Science and Director of the Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT. I imagine a future where robots are so integrated in the fabric of human life that they become as common as smart phones are today. James DiCarlo: I am the head of MIT's Department of Brain and Cognitive Sciences and the Peter de Florez Professor of Neuroscience. My research goal is to reverse engineer the brain mechanisms that underlie human visual intelligence, such as our ability to recognize objects on a desk, words on a page, or the faces of loved ones. This knowledge could inspire novel machine vision systems, illuminate new ways to repair or augment lost senses and potentially create new methods to treat disorders of the mind.

[REDDIT](#)

Science AMA Series: We see an opportunity to achieve a deeper understanding of intelligence. We are MIT faculty members Anantha Chandrakasan, Daniela Rus, and James DiCarlo. AMA!

MIT_OFFICIAL [R/SCIENCE](#)

Unfortunately, that's all the time we have to answer your questions today. Thanks, everyone for your engaging questions! Follow: @MIT, @MITEngineering, @MIT_CSAIL, and @mitbrainandcog to continue to get news around our work.

+++++

At MIT, we are on a quest to answer two big questions.

How does human intelligence work, in engineering terms?

And how can we use that deep grasp of human intelligence to build wiser and more useful machines, to the benefit of society?

We aspire for our new knowledge and newly built tools to serve the public good. Read this MIT news article to learn more:

<http://mitsha.re/5k6D30i80qQ>

About us

Anantha Chandrakasan: I am the dean of the School of Engineering at MIT. Before being named Dean, I was the Vannevar Bush Professor and head of the Department of Electrical Engineering and Computer Science (EECS). During my tenure at EECS I spearheaded a number of initiatives that opened opportunities for students, postdocs, and faculty to conduct research, explore entrepreneurial projects, and engage with EECS.

Daniela Rus: I am the Erna Viterbi Professor of Electrical Engineering and Computer Science and Director of the Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT. I imagine a future where robots are so integrated in the fabric of human life that they become as common as smart phones are today.

James DiCarlo: I am the head of MIT's Department of Brain and Cognitive Sciences and the Peter de Florez Professor of Neuroscience. My research goal is to reverse engineer the brain mechanisms that underlie human visual intelligence, such as our ability to recognize objects on a desk, words on a page, or the faces of loved ones. This knowledge could inspire novel machine vision systems, illuminate new ways to repair or augment lost senses and potentially create new methods to treat disorders of the mind.

[◊ READ REVIEWS](#)

[✍ WRITE A REVIEW](#)

CORRESPONDENCE:

DATE RECEIVED:
February 03, 2018

DOI:
10.15200/winn.151757.79486

ARCHIVED:
February 02, 2018

CITATION:

Everyone not living underneath a rock is aware of the massive progress in modern deep learning over the past decade. Since deep learning is largely based neural networks and convolutional architectures, which were originally inspired by biological neurons, one might hope to learn how the brain might work by understanding how these DL neural nets learn. With some exceptions, I haven't seen this pan out, and neuroscience and deep learning seem to be moving further apart.

The main differences I see are,

Biological neurons spike, DL neurons don't.

The vanilla backprop algorithm seems to be biologically implausible.

MIT_official , r/Science , Science AMA Series: We see an opportunity to achieve a deeper understanding of intelligence. We are MIT faculty members Anantha Chandrakasan, Daniela Rus, and James DiCarlo. AMA!, *The Winnower* 5:e151757.79486 , 2018 , DOI: [10.15200/winn.151757.79486](https://doi.org/10.15200/winn.151757.79486)

© et al. This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



New architectures in DL (such as ResNet) are engineered instead of inspired by biology.

Given these differences, do you think neural nets have anything to teach us about the brain?

Is it plausible to think that new (DL) neural network architectures could be discovered by studying the brain?

If we do assume a similar learning algorithm to DL at play in the brain, how does the brain generate it's "cost functions" it needs to optimize?

Thanks!

[MeissnerEffect](#)

This is Jim DiCarlo -- Hi Everyone!

Great points and questions [u/MeissnerEffect](#) !

First off -- I do think that it is not only possible -- it is guaranteed that new neural network architectures will be discovered by studying the brain! The only question is in what ways will these brain architectures surpass the currently engineered architectures and by how much. Key ways of interest are: performance in tasks, efficiency of learning, and power efficiency. In many task domains, the brain surpasses AI in all of these ways, and those are some of the most active areas of research in our community here at MIT.

You are also absolutely correct to point out that deep learning grew out of the intersection of brain and cognitive sciences and engineering. The key breakthrough of deep artificial neural networks (now "deep learning") came when researchers used a combination of science and engineering. Specifically, some researchers -- Hinton and LeCun are two of the best known -- began to build algorithms out of brain-like, multi-level artificial neural networks so that they had neural responses like those that neuroscientists had measured in the brain. They also used mathematical models proposed by scientists to teach these deep neural networks to perform visual tasks that cognitive scientists had found humans to be especially good at — like recognizing objects from many perspectives. This combined approach rocketed to prominence in 2012, when computer hardware had advanced enough for engineers to build these networks and teach them using millions of visual images.

In 2013, we found that some deep feedforward ANNs were remarkably good matches to the neural responses in the brain that are part of the brain's deep neural network for processing visual images at the center of gaze (called the ventral visual processing stream). This was -- and still is! -- very exciting because it suggested that ANNs can be used to gain a new understanding of previously mysterious visual processing.

My lab and others here at MIT and beyond are hard at work on that. In that regard, I agree with you that, with a few exceptions, as deep learning has progressed, it has deviated from the actual brain. And you have mentioned some of those deviations. One of the missions of the MIT IQ Core is to use results from the brain and the mind (science) to pull the development of some ANNs (engineering) back toward the actual NN at work in the human brain. This tight intersection of science and engineering -- which we call "reverse engineering" -- will certainly lead to an improved understanding of how human intelligence works (and how it can break down), and we believe that this approach with be a safer, more efficient approach to artificial intelligence. Given the past history (above), this is one of the best ways that we (esp. [MIT IQ](#)) can replenish the well of AI algorithms.

Edit: added a link

Hello! Just wanted to ask, what should one study to one day maybe have a chance to work with you?

Obviously I understand that a set path doesn't exist. I am truly interested in the fundamental nature of intelligence and I'd like to tackle it from both the AI and human intelligence standpoint. Are degrees like Computer Science and Cognitive Science a good place to start? How about Neuroscience?

[Gazparo](#)

Welcome everyone. This is Anantha. As part of the new MIT IQ initiative (<https://iq.mit.edu/>), we need collaboration between cognitive science, computer science and neuroscience. Brain scientists learn from machine learning experts and vice versa. At MIT, one of the most popular undergraduate and graduate courses is machine learning. Machine learning is the new literacy. I think starting where you've said sounds like a great place.

My own research is focused on the design of low-power integrated circuits. There is also tremendous opportunities for application of new devices and circuits in the field of human and artificial intelligence. Getting background in hardware can also bring a important perspective to this field.

Kind of a weird question that I suppose there isn't really an answer to, but I'll ask it anyway.

I can't really think of a good way to phrase this, but how close are we to designing a system that can infer an object's purpose based on its context and surrounds? Sort of how if we see a square thing with a whole lot of buttons next to a tv, we can immediately infer that's it's a remote and not a cellphone and that the arrows will change channels and volume and the numbers are for choosing a certain channel, even if it looks completely different to any remote we have seen before.

Obviously this is not an exact science, but I was just wondering how close we are to emulating this "human perception" of objects, where we can pick up something we've never seen before and realistically guess what it's purpose is.

Sorry, I'm pretty tired and this comment probably makes no sense, so please let me know if you need any clarification.

Also, less important side question: what courses / classes would you recommend to an MIT freshman going into the new year?

Thanks!

E: 2105730

[ChrisWalley](#)

This is Daniela: This kind of common sense reasoning is a very hard question for machines. Most learning algorithms today require millions of examples and can identify pixel patterns rather than purpose. Situational awareness from images is a big challenge. Researchers like my colleague Prof. Josh Tenenbaum are working towards new algorithms that will learn more like humans from much smaller datasets (maybe n=1), and even exhibit simplified forms of common sense reasoning.

Hello! Are there any interesting trends in the progression/evolution of human intelligence that you've observed?

[raygemage](#)

This is Anantha. One important direction is the amazing new hardware capabilities that are allowing us to probe at finer granularity the brain so that we can deeply understand how brain activities give rise to human activities. This is enabled by low-power electronics. Such tools will be critical in understanding

and reverse engineering human intelligence.

Daniela Rus: What level of intelligence do you one day hope that these "common place" robots will possess? Do you see them as being tools, or are we all going to be starting in "iRobot?"

James DiCarlo: Are you trying to build a mechanical replica of the human brain? If you are, is that the level of intelligence that you are attempting to give to robots? That, combined with Daniela's dream of robots being common place, is starting to feel a lot like "iRobot," or maybe a robot version of "Planet of the Apes."

[Spider-Man-2099](#)

This is Daniela: Robots are not that great at figuring things out today, but we are striving to make them much better at understanding the world. Think of robots as tools and assistants, and over time these assistants will be able to do increasingly more for you. Unfortunately, I don't think it is in our future to all be movie stars in iRobot.

Daniela Rus: What level of intelligence do you one day hope that these "common place" robots will possess? Do you see them as being tools, or are we all going to be starting in "iRobot?"

James DiCarlo: Are you trying to build a mechanical replica of the human brain? If you are, is that the level of intelligence that you are attempting to give to robots? That, combined with Daniela's dream of robots being common place, is starting to feel a lot like "iRobot," or maybe a robot version of "Planet of the Apes."

[Spider-Man-2099](#)

[u/Spider-Man-2099](#) thanks for the question!

This is Jim: My lab is not currently trying to build an exact mechanical replica of the brain. But we are trying to use computers to simulate how different types of replicas (aka models) would match up with the brain in terms of its behavior (e.g. visual comprehension) and its internal neural responses at different stages of visual processing.

Such models could be the basis of hardware that replicates those functions -- and to help people that have (e.g.) lost vision.

There's often great confidence (at MIT, particularly) placed on the idea that a better understanding of the brain and the mind that runs on it could very well lead to better computer algorithms and to general, strong AI. Two instances where this idea does seem to have had moderate success are perhaps reinforcement learning (somewhat inspired by VTA error signaling neurons) and deep neural nets (somewhat inspired by visual cortex).

And yet so many other computational systems that exhibit astounding amounts of problem-solving capabilities seem to draw very little inspiration from nervous systems -- Wolfram Alpha, for example, or Boston Dynamics' robots, or a lot of the expert systems AI work done a few decades ago.

What is the argument for going about attempting to figure out human intelligence such that we can use it on machines, besides the couple examples I listed above? Wouldn't it perhaps be easier to just focus on building general AI? In other words, why the focus on human intelligence?

[dopu](#)

This is Jim. Great question!

You correctly point out that not all areas of progress in AI-related systems have been driven by detailed knowledge of the brain and the mind (Although many of these are at least brain-inspired).

So one way to phrase your question is this: what is the most efficient path to discover human-level AI? Path 1: Have engineers work on their own to see how far they can get. Path 2: Have engineers work with guidance from the brain and the mind. No one knows the answer to the question of which path is faster.

However, the recent successes (esp. reinforcement learning and deep CNNs/deep learning) have shown that Path 2 can deliver a very impressive return. The human brain has had millions of years to develop its capabilities -- while our engineers could probably work faster than evolution, it might take many many years to find processing strategies that work as well as the brain in some aspects of intelligence. Thus, Path 1 may be very, very long. So why not take a huge shortcut and look to the brain and the mind (Path 2)?

Also note that the above considerations are only about the question of paths to AI. But Path 2 has additional human benefits to it as well. An engineering description of the brain will not only allow us to build better machines. It will also allow us see new ways to repair, educate, and perhaps even augment our own minds!

Hello,

Thanks for giving us this chance for open questions.

Why do we continue to push AI & Deep Learning when many prominent scientists and technical people have voiced their concerns over making machines smarter than us? Stephen Hawking, Elon Musk, Nick Bostrom, etc. have all said that it could be our greatest threat as humans to invent ourselves into obsolescence.

<https://www.livescience.com/49419-artificial-intelligence-dangers-letter.html>

I wonder why it's really a benefit to have our data tracked by smart machines and compiled to give us robo-butlers, when so much potential for misuse is inherent in those services. I do see how smart machines that can identify diseases and help spot patterns in new research are useful and beneficial, but the amount of machine intelligence that has already become integral to normal, daily life seems to point more and more towards automating our lives so much that we hardly have to do anything for ourselves. Not to mention all the data breaches and theft of secure, personal records that have already occurred.

Thanks for you responses!

[Cynghen](#)

This is Daniela: Machines and people do not have to compete, especially when they can do so much more working together. Machines are better than people at things like remembering and crunching data, while people are better than machines when it comes to creative thinking and strategic reasoning. While we cannot stop technology from advancing and changing the world, it is important to think about the consequences and put in place provisions that ensure changes are for the greater good. We still have a lot to do before we can fully understand the science and engineering of intelligence, and I believe that work will result in breakthroughs that will provide us with AI and robotics systems that will help make a better world. I also agree with what President Reif said when we announced MIT IQ, namely that AI can benefit all research fields. It is important to understand that AI is an advanced tool developed by people for people. It is an incredibly powerful tool, but like all tools it is

not inherently good or bad; it is what we choose to do with it and I believe we can do incredible things.

Hello,

Thanks for giving us this chance for open questions.

Why do we continue to push AI & Deep Learning when many prominent scientists and technical people have voiced their concerns over making machines smarter than us? Stephen Hawking, Elon Musk, Nick Bostrom, etc. have all said that it could be our greatest threat as humans to invent ourselves into obsolescence.

<https://www.livescience.com/49419-artificial-intelligence-dangers-letter.html>

I wonder why it's really a benefit to have our data tracked by smart machines and compiled to give us robo-butlers, when so much potential for misuse is inherent in those services. I do see how smart machines that can identify diseases and help spot patterns in new research are useful and beneficial, but the amount of machine intelligence that has already become integral to normal, daily life seems to point more and more towards automating our lives so much that we hardly have to do anything for ourselves. Not to mention all the data breaches and theft of secure, personal records that have already occurred.

Thanks for you responses!

[Cynghen](#)

Why do we continue to push AI & Deep Learning when many prominent scientists and technical people have voiced their concerns over making machines smarter than us? Stephen Hawking, Elon Musk, Nick Bostrom, etc. have all said that it could be our greatest threat as humans to invent ourselves into obsolescence.

This is Anantha. AI can have tremendous positive impact on society. For example, we can prevent deaths from cancer by using deep learning for early detection and personalized treatment. It is critical to think about the ethical implications as we develop and deploy AI technologies. This applies across a range of applications from the decisions that self-driving cars to the use of data. It's also important to consider the policies that govern new technologies. We see that AI can have truly a positive impact on virtually every application area, from material design and drug discovery to transportation and energy and the environment.

Is it possible to gain intelligence or lose it?

[datboihasnain](#)

This is Daniela: Your IQ will drop if you don't sleep enough.

One of them, "The Core," will advance the science and engineering of both human and machine intelligence. A key output of this work will be machine-learning algorithms.

Algorithms only for machine learning? I appreciate the power of machine learning, but can we expect algorithms from other branches of AI as well?

[drsjsmith](#)

This is Daniela again: There is a multitude of subfields of AI and each one of them is moving forward

with new ideas and new algorithms. I actually have a nice chart that has about 50 fields, including speech understanding, robotics, symbolic reasoning, knowledge representation, etc. But while we are making progress in all fields to understand the science of autonomy and the science of intelligence, and we are building machines with increased autonomy, these systems remain limited. We need new insights and new algorithms, and in the MIT IQ we are looking toward biology and neuroscience for inspiration.

How many years in the future could AI actually capable to replace lawyer?

[xipha](#)

This is Anantha. We imagine artificial intelligence as complementing our own intelligence. As MIT economist Daron Acemoglu says, (<http://news.mit.edu/2018/3q-daron-acemoglu-technology-and-future-work-0201>) let's think about tax professional. AI removes the need for seasoned accountants to perform numerically-related tasks. But we need tax professional to inform clients about their choices and options in some sort of empathetic human way. They will have to become the interface between machines and customers. It's very important to think about the future of work as we conceive and deploy AI technologies.

How many years in the future could AI actually capable to replace lawyer?

[xipha](#)

Daniela, again: I talked about how machines and humans can be much better when they work together. Machines and humans can also be better lawyers. Word processing, Internet and email have revolutionized document drafting, access to information and sharing information. These innovations change law practice in fundamental ways. The next wave of technologies, natural language processing, promises similarly far-reaching effects. By interpreting text, natural language processing systems could predict decisions to support the work of lawyers. But while software is replacing some tasks, computers can't counsel clients, write compelling briefs or persuade judges. They can't be lawyers, but they can change the type of work that lawyers do by finding patterns and correlations, predicting outcomes and improving analysis accuracy.

My question is for Jim. Can you expand on the phrase 'reverse engineering the brain mechanisms'? Does that mean fitting a model (like a CNN) to the neural data, and then trying to understand how the CNN works? Or does it mean optimizing a CNN according to some criteria and observing that it works *and* that it has something in common with the neural data?

P.S It's Leo. Thanks for the wonderful class last semester!

[volvo_physics](#)

Hi Leo! When I say "reverse engineering", I mean using results from the brain (e.g. neural activity patterns) and the mind (e.g. behavior) to both inspire and concretely guide the development of neural-network models of brain subsystems.

The basic intuition is this: Because many possible neural network algorithms might explain any given layer of human intelligence, engineers working alone are searching for the proverbial needle in a haystack. However, when we guide algorithm-building and testing efforts with discoveries and measurements from brain and cognitive science, we get advances like deep vision networks and deep learning.

So I think of it more like your second answer: optimizing a model (not necessarily a CNN) to explain/predict all currently available data. Then more data are collected. Then more optimization occurs. The convergence of that process is a model that I would call a form of "understanding" -- it certainly enables a very wide range of applications ranging from AI, to brain-machine interfaces, to understanding how the brain processing can go awry (brain disorders), to new ways to better educate.