

# Edge

*To arrive at the edge of the world's knowledge, seek out the most complex and sophisticated minds, put them in a room together, and have them ask each other the questions they are asking themselves.*

<https://www.edge.org/response-detail/26197>

Printed On Sun June 11th 2023



## 2015 : WHAT DO YOU THINK ABOUT MACHINES THAT THINK?

[In the News \[ 27 \]](#)

| [Contributors \[ 192 \]](#) | [View All Responses \[ 192 \]](#)



**[John Tooby](#)**

*Founder of field of Evolutionary Psychology; Co-director, Center for Evolutionary Psychology, Professor of Anthropology, UC Santa Barbara*

### The Iron Law Of Intelligence

As luck would have it, I am myself a machine that thinks, so I will share the special insight this gives me with those of you who don't share my good fortune. To dispense with vestigial metaphysical objections, we know that machines that think like humans are possible, because they have been overrunning the landscape for millenia. If we now want human-like intelligences that are made, not begotten, then it will be extraordinarily useful to achieve an understanding of the human-like intelligences that already exist—that is, we need to characterize the evolved programs that constitute the computational architecture of the brain.

Not only has evolution packed the human architecture full of immensely powerful tricks, hacks, and heuristics, but studying this architecture has made us aware of an implacable, invisible barrier that has stalled progress toward true AI: the iron law of intelligence. Previously, when we considered (say) a parent and child, it seemed self-evident that intelligence was a unitary substance that beings had more or less of, and the more intelligent being knows everything that the less intelligent knows, and more besides. This delusion led researchers to think that the royal road to amplified intelligence was to just keep adding more and more of this clearly homogeneous (but hard to pin down) intelligence stuff—more neurons, transistors, neuromorphic chips, whatever. As Stalin (perhaps) said, Quantity has a quality all its own.

In contrast, the struggle to map really existing intelligence has painfully dislodged this compelling intuition from our minds. In contrast, the iron law of intelligence states that a program that makes you intelligent about one thing makes you stupid about others. The bad news the iron law delivers is that there can be no master algorithm for general intelligence, just waiting to be discovered—or that intelligence will just appear, when transistor counts, neuromorphic chips, or networked Bayesian servers get sufficiently numerous. The good news is that it tells us how intelligence is actually engineered: with idiot savants. Intelligence grows by adding qualitatively different programs together to form an ever greater neural biodiversity.

Each program brings its own distinctive gift of insight about its own proprietary domain (spatial relations, emotional expressions, contagion, object mechanics, time series analysis). By bundling different idiot savants together in a semi-complementary fashion, the region of collective savantry expands, while the region of collective idiocy declines (but never disappears).

The universe is vast and full of illimitable layers of rich structure; brains (or computers) in comparison are infinitesimal. To reconcile this size difference, evolution sifted for hacks that were small enough to fit the brain, but that generated huge inferential payoffs—superefficient compression algorithms (inevitably lossy, because one key to effective compression is to throw nearly everything away).

Iron law approaches to artificial and biological intelligence reveal a different set of engineering problems. For example, the architecture needs to pool the savantry, not the idiocy; so for each idiot (and each combination of idiots) the architecture needs to identify the scope of problems for which activating the program (or combination) leaves you better off, not worse. Because different programs often have their own proprietary data structures, integrating information from different idiots requires constructing common formats, interfaces, and translation protocols.

Moreover, mutually consistent rules of program pre-emption are not always easy to engineer, as anyone knows who (like me) has been stupid enough to climb halfway up a Sierra cliff, only to experience the conflicting demands of the vision-induced terror of falling, and the need to make it to a safe destination.

Evolution cracked these hard problems, because neural programs were endlessly evaluated by natural selection as cybernetic systems—as the mathematician Kolmogorov put it, "systems which are capable of receiving, storing and processing information so as to use it for control." That natural intelligences emerged for the control of action is essential to understanding their nature, and their differences from artificial intelligences. That is, neural programs evolved for specific ends, in specific task environments; were evaluated as integrated bundles, and were incorporated to the extent they regulated behavior to produce descendants. (To exist, they did not have to evolve methods capable of solving the general class of all hypothetically possible computational problems—the alluring but impossible siren call that still shipwrecks AI labs.)

This means that evolution has only explored a tiny and special subset out of all possible programs; beyond beckons a limitless wealth of new idiot savants, waiting to be conceived of and built. These intelligences would operate on different principles, capable of capturing previously unperceived relationships in the world. (There is no limit to how strange their thinking could become).

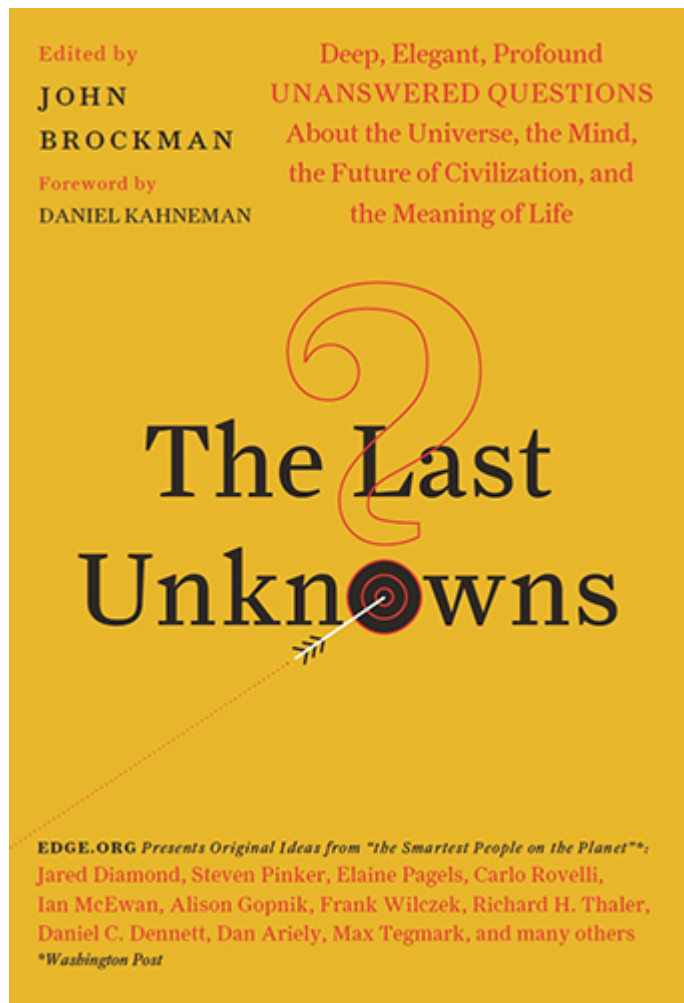
We are living in a pivotal era, at the beginning of an expanding wave front of deliberately engineered intelligences—should we put effort into growing the repertoire of specialized intelligences, and networking them into functioning, mutually intelligible collectives. It will be exhilarating to do with nonhuman idiot savant collectives what we are doing here now with our human colleagues—chewing over intellectual problems using minds equipped interwoven with threads of evolved genius and blindness.

What will AIs want? Are they dangerous? Animals like us are motivated intelligences capable of taking action (*MICTAs*). Fortunately, AIs are currently not *MICTAs*. At most, they are only trivially motivated; their motivations are not linked to a comprehensive world picture; and they are only capable of taking a constrained set of actions (running refineries, turning the furnace off and on, shunting packets, futilely attempting to find

wifi). Because we evolved with certain adaptive problems, our imaginations project primate dominance dramas onto AIs, dramas that are alien to their nature.

We could transform them from Buddhas—brilliant teachers passively contemplating without desire, free from suffering—into MICTAs, seething with desire, and able to act. That would be insane—we are already bowed under the conflicting demands of people. The foreseeable danger comes not from AIs but from those humans in which predatory programs for dominance have been triggered, and who are deploying ever-growing arsenals of technological (including computational) tools for winning conflicts by inflicting destruction.

[Return to Table of Contents](#)

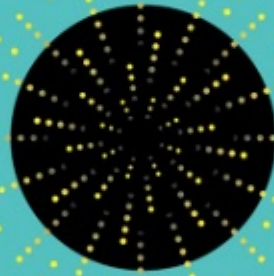


2018 : [WHAT IS THE LAST QUESTION?](#)

2017 : [WHAT SCIENTIFIC TERM OR CONCEPT OUGHT TO BE MORE WIDELY KNOWN?](#)

Lost, Overlooked, and Underappreciated  
Scientific Concepts Everyone  
Should Know

# This Idea Is Brilliant



"An epicenter of  
bleeding-edge insight."  
*-The Atlantic, on Edge.org*

Edited by  
**JOHN  
BROCKMAN**  
*Editor of Know This*

**EDGE.ORG** Presents *Original Ideas From:*

Jared Diamond, Steven Pinker, Richard Dawkins, Lisa Randall,  
Lawrence Krauss, Carlo Rovelli, Ian McEwan, Alison Gopnik,  
Richard Thaler, Helen Fisher, Frank Wilczek, and many others

2016 : [WHAT DO YOU CONSIDER THE MOST INTERESTING RECENT \[SCIENTIFIC\] NEWS? WHAT MAKES IT IMPORTANT?](#)

Edited by  
**JOHN BROCKMAN**

Editor of *This Idea Must Die*

"A lavish cerebral feast."  
-*The Atlantic*

# Know This

Today's Most Interesting and Important  
Scientific Ideas, Discoveries, and Developments

EDGE.ORG Presents Original Ideas from "the World's Finest Minds"\*

Jared Diamond, Steven Pinker, Carlo Rovelli, Frank Wilczek,  
Alison Gopnik, Freeman Dyson, Martin Rees, Jonathan Haidt,  
Rebecca Newberger Goldstein, and more "Very, VERY smart people"†

\**The Guardian* †Boing Boing

2015 : [WHAT DO YOU THINK ABOUT MACHINES THAT THINK?](#)

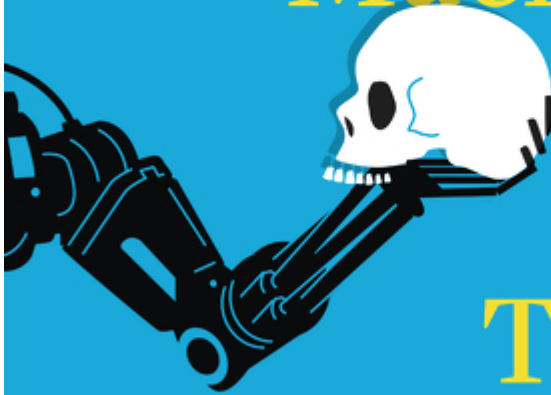
Edited by  
JOHN  
BROCKMAN

*Editor of This Explains Everything*

Today's Leading  
Thinkers on the Age of  
Machine Intelligence

"An epicenter of  
bleeding-edge insight."  
—*The Atlantic*,  
on Edge.org

# What to Think About Machines That Think



**EDGE.ORG** *Presents Original Ideas from:*

Steven Pinker, Tim O'Reilly, Freeman Dyson, Daniel C. Dennett,  
Kevin Kelly, Richard Thaler, Helen Fisher, Martin Rees, Esther Dyson,  
Frank Wilczek, Richard E. Nisbett, Nick Bostrom, *and more*

2014 : [WHAT SCIENTIFIC IDEA IS READY FOR RETIREMENT?](#)

# This Idea Must Die

Scientific  
Theories That  
Are Blocking  
Progress



**EDGE.ORG** *Presents Today's  
Leading Thinkers*

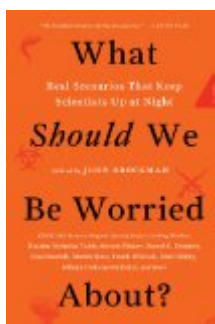
Jared Diamond, Richard Thaler,  
Steven Pinker, Richard Dawkins,  
Nassim Nicholas Taleb,  
Ian McEwan, Rebecca Newberger  
Goldstein, Freeman Dyson,  
Daniel Golman, Alan Guth,  
*and more*

Edited by

**JOHN  
BROCKMAN**

Author of

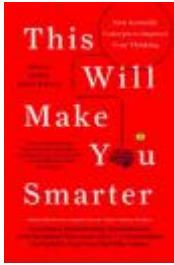
*This Explains Everything*



2013 : [WHAT \\*SHOULD\\* WE BE WORRIED ABOUT?](#)



2012 : [WHAT IS YOUR FAVORITE DEEP, ELEGANT, OR BEAUTIFUL EXPLANATION?](#)



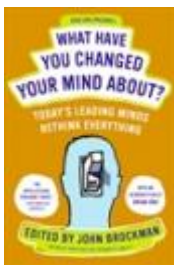
2011 : [WHAT SCIENTIFIC CONCEPT WOULD IMPROVE EVERYBODY'S COGNITIVE TOOLKIT?](#)



2010 : [HOW IS THE INTERNET CHANGING THE WAY YOU THINK?](#)



2009 : [WHAT WILL CHANGE EVERYTHING?](#)



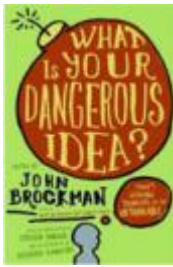
2008 : [WHAT HAVE YOU CHANGED YOUR MIND ABOUT? WHY?](#)



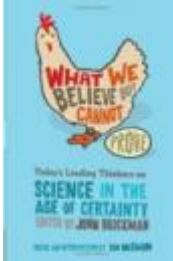
2007 : [WHAT ARE YOU OPTIMISTIC ABOUT?](#)

2006 : [WHAT IS YOUR DANGEROUS IDEA?](#)





2005 : [WHAT DO YOU BELIEVE IS TRUE EVEN THOUGH YOU CANNOT PROVE IT?](#)



2004 : [WHAT'S YOUR LAW?](#)

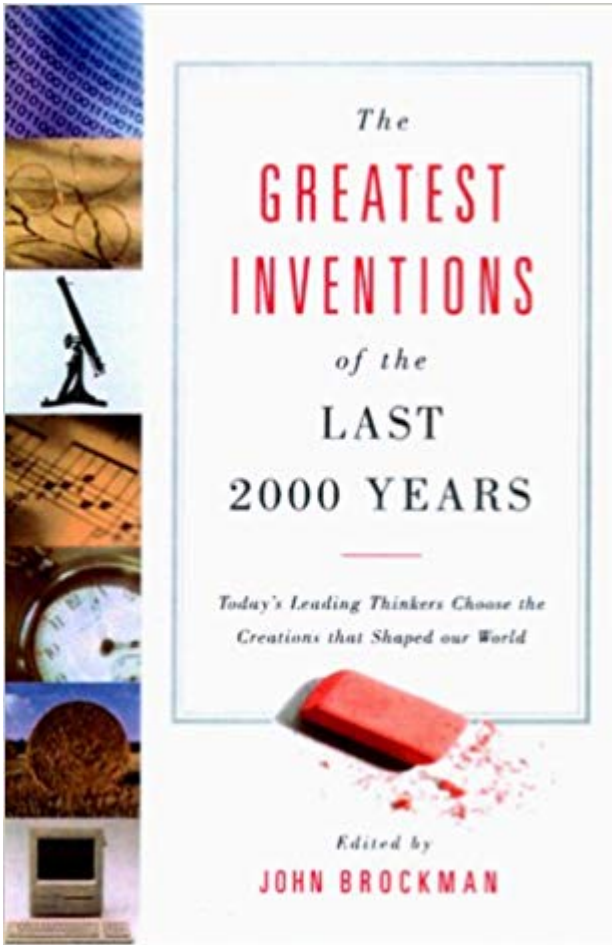
2003 : [WHAT ARE THE PRESSING SCIENTIFIC ISSUES FOR THE NATION AND THE WORLD, AND WHAT IS YOUR ADVICE ON HOW I CAN BEGIN TO DEAL WITH THEM? - GWB](#)

2002 : [WHAT IS YOUR QUESTION? ... WHY?](#)

2001 : [WHAT NOW?](#)

2001 : [WHAT QUESTIONS HAVE DISAPPEARED?](#)

2000 : [WHAT IS TODAY'S MOST IMPORTANT UNREPORTED STORY?](#)



1999 : [WHAT IS THE MOST IMPORTANT INVENTION IN THE PAST TWO THOUSAND YEARS?](#)

1998 : [WHAT QUESTIONS ARE YOU ASKING YOURSELF?](#)

---

Contact Info:editor@edge.org

Edge.org is a nonprofit private operating foundation under Section 501(c)(3) of the Internal Revenue Code.

Copyright © 2023 By Edge Foundation, Inc All Rights Reserved.