

The 3 Laws of Ed-Tech Robotics #TEDxNYED

I gave my first (and probably last) TEDx talk this weekend at TEDxNYED on the topic of ed-tech, science fiction, and ethics. Unfortunately (or fortunately -- depending on how you view things), the livestream wasn't working. I'll post a video if and when it becomes available (although I'm not sure my talk will past TED muster). But I've posted below a rough transcript of my talk, along with a couple of the slides I showed:

The drive to create a machine that duplicates the human mind, that duplicates the human body has ancient roots. It has roots in mythology and in literature far older than humans' ability to build this sort of machine through science, engineering, and technology.

But here we are. 2013. And there are almost daily headlines to this end — robots — designed to replicate, enhance, and automate some function of our lives, of our world.

Robots on the factory floor. Robot game show contestants. There are robots that offer medical advice, that perform surgery Robots that play heavy metal music. Lunar robots. Robots on Mars. Robot cheetahs. Robot cockroaches. Robot lawyers. Robot sparrows. Robot spies. Yes, robots in our classrooms. Robots to defend us and watch us and teach us and attack. And drones. My God, the drones.

There are robots that can debone chickens.

This is the development that's been stuck in my head, I confess. And it is actually quite a remarkable one, as one of the things that has prevented robot chicken de-boning in the past is, as the sub-header reads here, that all chickens — even factory-farmed chickens — are slightly different. As such it's been a challenge to engineer a robot that can make the necessary allowances and exceptions as it plucks and carves and de-bones.

The de-boning and the butchering, however, standardizes the meat.

And perhaps tying this chicken deboning robot to standardization and teaching and learning and technology requires a bit of a conceptual leap — but this is a TED Talk, so what the hell.

Here's a link: we have reached the level where we can train robots to beat Ken Jennings at *Jeopardy* and to recognize all the anatomical differences in chickens' bone structures, but are we able to recognize and cultivate differences elsewhere — oh say, the cognitive differences and intellectual capacities in humans, in students?

In other words, can we build and train robots to help us — in all of our uniqueness — learn? What sort of standardization, what sort of differentiation would that entail? More testing, more data-gathering?

And this is the more important question: even if we can build automated instructional software, do we want to? Why?



Gratuitous chicken de-boning image here, my apologies, let's agree that there might just be cause for concern when it comes to robots. Take one look at this image and you have to shout, "Oh my God. Do we really want to train robots with precise knife-wielding skills?!"

No doubt our fears about what will happen if we build automatons are as old as our drive to build them. But the word "robot" is itself quite new. It was coined in 1920. It comes, not from engineering but from theatre — "Rossum's Universal Robots" by the Czech playwright Karel Capek. Capek credited his brother for the invention of the word, but in Czech "robota" means "forced labor." In the play, robots take over factory work and eventually the world, leading to the extinction of the human race.

Much like Frankenstein's monster sought to destroy his creator, many literary robots have aimed to overthrow their human masters.

But there is another possibility as science fiction fans who've read their Isaac Asimov will reassure us, "Of course, we can train robot chicken de-boners *and* remain safe as humans, as long as we maintain the 3 laws of robotics."

Law 1: A robot may not injure a human being, or through inaction, allow a human being to come to harm.

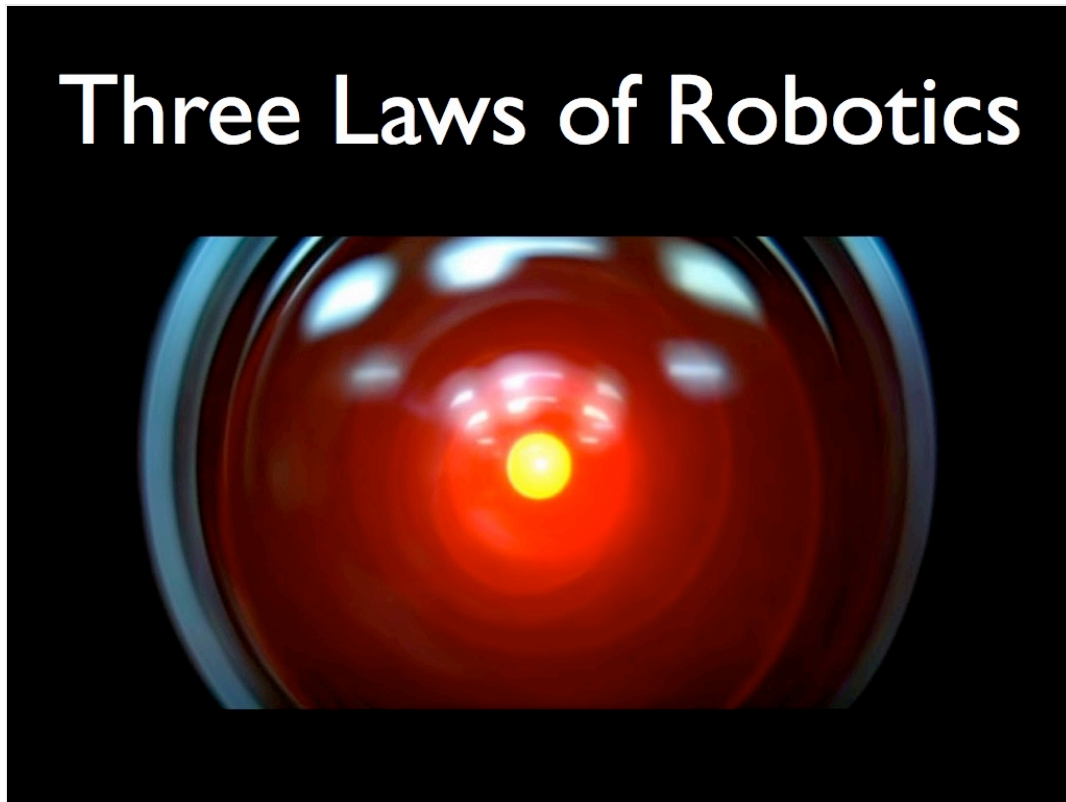
Law 2: A robot must obey the orders given to it by human beings, except where such orders would conflict

with the First Law.

Law 3: A robot must protect its own existence as long as such protection does not conflict with the First and Second Laws.

The Zeroeth Law: A robot may not harm humanity, or by inaction, allow humanity to come to harm.

In other words — in Asimov’s framework, at least — robots are programmatically and thus necessarily protective of human safety, of human life. They are here to help us not harm us.



But not all robots know these rules, of course. And let’s be clear — the 3 Laws do not resolve *all* ethical questions regarding robots. They just reassure us that there will be fewer knife-wielding robot rebels.

Nor do all robots appear as threatening as robot chicken de-boners. The robots in education technology certainly don’t. Sure, there are the battle-bots in robotics competitions. But most of what the field of artificial intelligence has brought into the classroom is software: adaptive learning software, intelligent tutoring systems, automated assessment tools.

We have authored no “three laws of ed-tech robotics” in response — neither in science fiction nor in our district tech procurement guidelines. We rarely ask, “what are ethical implications of educational technologies?” (Mostly: “will this raise test scores?”) We rarely ask, “Are we building and adopting tools that might harm us?”

I frame this as a science fiction question in part because science fiction helps us tease out these things: the possibilities, horrors, opportunities of technology and science. I turn to science fiction because novelists and researchers and engineers and entrepreneurs construct stories using similar models and metaphors. I refer to science fiction because what our culture produces in the film studio or at the writer's desk is never entirely separable from what happens in the scientist's lab, what happens at the programmer's terminal. We are bound by culture. And there are some profoundly important — and I would add, terribly problematic — views on teaching and learning that permeate both science fiction *and* technological pursuits.

The view of education as a “content delivery system” for example appears as readily in ed-tech companies' press releases as it does on the big screen. Take *The Matrix*, for example, where Keanu Reeves utters one of his finest lines as a computer injects directly into his brainstem all the knowledge he needs:

“Whoa, I know kung fu.”

Or take the behaviorist B. F. Skinner's “teaching machine” — a mechanical device developed in the 1950s to offer “programmed instruction” — it's a vision that appears in many SF novels (*Enders Game*, for example) and one that isn't that different from many ed-tech products on the market today.

Or take the latest craze, these massive open online courses or MOOCs — many of which have their origins in the university AI lab — Udacity and Coursera in Stanford's AI department and the head of edX from MIT's AI lab. Many of these rely on automated grading mechanisms alongside their video-taped lecture lessons. The leap — as a metaphor and as a model — from Google's robotic, self-driving car to the self-driving online course is not that far — literally — it's built by the same fellow.

The automation of teaching and learning — the development of teaching machines, robots in the classroom — is a long-running plot line in science fiction and a long-standing goal of AI itself: the creation of artificial intelligence-backed machines that claim to automate and “personalize” lesson delivery and assessment. Many remain crude, using mostly multiple choice for the latter.

But last year, the Hewlett Foundation sponsored a contest to get some of the best minds in the world of machine learning — a revealingly named subsection of AI — to design a programmatic way to automate, not the grading of multiple choice tests, but the grading of essays. They put a \$100,000 bounty out for an algorithm that grades as well as humans. And it was a success; the bounty was claimed. Not only is it possible to automate essay grading, researchers contended, robots do this just as well as human essay-graders do.

Investors, big testing companies, and all those unemployed robots with degrees in rhetoric and composition, were thrilled.

Robots grade more efficiently. Robots — unlike those pesky human English instructors and graduate students — remain non-unionized. They do not demand a living wage or health care benefits.

A computer can grade 16,000 SAT essays in 20 seconds, I've heard cited, and if teachers don't have to worry about reading students' work they can assign more of it. They can also have more — massively more — students in their classes. Hundreds of thousands more.

Not everyone is sold on the idea that robot graders are the future. Many critics argue that it's actually pretty easy to fool the software. That's because robots don't "read" the way we do. They do things like assess word frequency, word placement, word pairing, word length, sentence length, use of conjunctions, and punctuation.

But robot essay graders might reveal something else — as writing professor Alex Reid argues, "If computers can read like people it's because we have trained people to read like computers."

Let's look closely at the human graders that these robots in the Hewlett competition were compared to. The major testing companies hire a range of people to grade essays for them. They don't have a personal relationship with the students they're grading, obviously. These graders are given a fairly strict rubric to follow — a rubric that computers, no surprise, follow with more precision. Human graders are discouraged from recognizing "creative" expression.

Robot graders, of course, have no idea what "creative expression" even means.

What does it mean to tell our students that we're actually not going to read their papers, but we're going to scan them and a computer will analyze them instead? What happens when we encourage students to express themselves in such a way that appeases the robot essay graders rather than a human audience? What happens when we discourage creative expression?

Robot graders raise all sorts of questions about thinking machines and thinking humans, about self-expression and creativity, about the purpose of writing, the purpose of writing assignments, the purpose of education, and the ethics of education technology and of robots in the classroom.

And what are the implications of automating the teaching and learning process? Why does efficiency matter so much? Why do we want the messy and open-ended process of inquiry standardized, scaled, or automated? What will all of this artificial intelligence drive us to do about human intelligence?

Three Laws of Robotics

0.

A robot may not harm humanity, or by inaction, allow humanity to come to harm.

Do robot essay graders violate the Laws of Robotics?

The promise from the proponents of these technologies — from adaptive learning companies, and AI professors, and MOOC startups and the like — is that someday we'll have efficiency and order and standardization and scale. I find that chilling.

As the French philosopher Jacques Ellul argued in the 1960s — not too long after the creation of B. F. Skinner's "teaching machine" or Isaac Asimov's *I, Robot* in fact, the drive for efficiency is all-encompassing and a dominant force in our technological age,

"The human brain must be made to conform to the much more advanced brain of the machine. And education will no longer be an unpredictable and exciting adventure in human enlightenment but an exercise in conformity and an apprenticeship to whatever gadgetry is useful in a technical world."

What "use" will our technical world have for us, for our students?

And I think this is why I like the Laws of Robotics very much — not for answering these questions in form of a simple plot device or character restraint, but for prompting us to ask questions. The Zeroeth Law in particular reminds us about the potential of harm to humanity.

This isn't simply about "the survival of the human race" with or versus the machine — it's about our humanity. Indeed humanity and learning are deeply intertwined — intertwined with love, not with algorithms.

I'm not sure we need to devise laws of ed-tech robotics — not exactly. But I do think we need to deliberate far more vigorously about the ethics and the consequences of the technologies we are adopting.

It's clear that building teaching machines has been a goal in the past. But that doesn't mean that doing so is "progress." What is *our* vision for the future? Either we decide what these new technologies mean for us — what our ethical approach to technology will be — or someone else will.

Icon credit: The Noun Project

Audrey Watters, The 3 laws of Ed-Tech Robotics #TEDxNYED

29 April 2013

<http://hackededucation.com/2013/04/29/the-3-laws-of-ed-tech-robotics-tedxnyed/>