

# Schneier on Security

---

[← Defending Against Liar Buyer Fraud](#)

[Friday Squid Blogging: Giggling Squid Restaurant →](#)

## When Thinking Machines Break the Law

Last year, two Swiss artists programmed a Random Botnot Shopper, which every week would spend \$100 in bitcoin to buy a random item from an anonymous Internet black market...all for an art project on display in Switzerland. It was a clever concept, except there was a problem. Most of the stuff the bot purchased was benign -- fake Diesel jeans, a baseball cap with a hidden camera, a stash can, a pair of Nike trainers -- but it [also purchased](#) ten ecstasy tablets and a fake Hungarian passport.

What do we do when a machine breaks the law? Traditionally, we hold the person controlling the machine responsible. People commit the crimes; the guns, lockpicks, or computer viruses are merely their tools. But as machines become more autonomous, the link between machine and controller becomes more tenuous.

Who is responsible if an autonomous military drone accidentally kills a crowd of civilians? Is it the military officer who keyed in the mission, the programmers of the enemy detection software that misidentified the people, or the programmers of the software that made the actual kill decision? What if those programmers had no idea that their software was being used for military purposes? And what if the drone can improve its algorithms by modifying its own software based on what the entire fleet of drones learns on earlier missions?

Maybe our courts can decide where the culpability lies, but that's only because while current drones may be autonomous, they're not very smart. As drones get smarter, their links to the humans that originally built them become more tenuous.

What if there are no programmers, and the drones program themselves? What if they are both smart and autonomous, and make strategic as well as tactical decisions on targets? What if one of the drones decides, based on whatever means it has at its disposal, that it no longer maintains allegiance to the country that built it and goes rogue?

Our society has many approaches, using both informal social rules and more formal laws, for dealing with people who won't follow the rules of society. We have informal mechanisms for small infractions, and a complex legal system for larger ones. If you are obnoxious at a party I throw, I won't invite you back. Do it regularly, and you'll be shamed and ostracized from the group. If you steal some of my stuff, I might report you to the police. Steal from a bank, and you'll almost certainly go to jail for a long time. A lot of this might seem more ad hoc than situation-specific, but

we humans have spent millennia working this all out. Security is both political and social, but it's also psychological. Door locks, for example, only work because our social and legal prohibitions on theft keep the overwhelming majority of us honest. That's how we live peacefully together at a scale unimaginable for any other species on the planet.

How does any of this work when the perpetrator is a machine with whatever passes for free will? Machines probably won't have any concept of shame or praise. They won't refrain from doing something because of what other machines might think. They won't follow laws simply because it's the right thing to do, nor will they have a natural deference to authority. When they're caught stealing, how can they be punished? What does it mean to fine a machine? Does it make any sense at all to incarcerate it? And unless they are deliberately programmed with a self-preservation function, threatening them with execution will have no meaningful effect.

We are [already talking](#) about programming morality into thinking machines, and we can imagine programming other human tendencies into our machines, but we're certainly going to get it wrong. No matter how much we try to avoid it, we're going to have machines that break the law.

This, in turn, will break our legal system. Fundamentally, our legal system doesn't prevent crime. Its effectiveness is based on arresting and convicting criminals after the fact, and their punishment providing a deterrent to others. This completely fails if there's no punishment that makes sense.

We already experienced a small example of this after 9/11, which was when most of us first started thinking about suicide terrorists and how post-facto security was irrelevant to them. That was just one change in motivation, and look at how those actions affected the way we think about security. Our laws will have the same problem with thinking machines, along with related problems we can't even imagine yet. The social and legal systems that have dealt so effectively with human rulebreakers of all sorts will fail in unexpected ways in the face of thinking machines.

A machine that thinks won't always think in the ways we want it to. And we're not ready for the ramifications of that.

This essay [previously appeared](#) on Edge.org as one of the answers to the 2015 Edge Question: "What do you think about machines that think?"

EDITED TO ADD: The Random Botnet Shopper is "[under arrest](#)."

Tags: [courts](#), [crime](#), [essays](#)

[Posted on January 23, 2015 at 4:55 AM](#) • 83 Comments

---

## Comments

**A Nonny Bunny • [January 23, 2015 5:20 AM](#)**

For the while being, I think responsibility lies with whoever unleashed the machines on the world. It's the same as with children. As long as they are not responsible for themselves, it's the parent's responsibility. Even if it's the child's ill-conceived decision that's the problem.

---

**wiredog • [January 23, 2015 5:23 AM](#)**

"What if there are no programmers, and the drones program themselves? What if they are both smart and autonomous, and make strategic as well as tactical decisions on targets? What if one of the drones decides, based on whatever means it has at its disposal, that it no longer maintains allegiance to the country that built it and goes rogue?"

Then you have Skynet, and legal culpability is the least of your problems.

---

**Jan Willem • [January 23, 2015 5:42 AM](#)**

In fact we have already this issue. Not with drones (robots) but with companies and governments.

If a governmental organisation makes a mistake, the person in charge is not punished. The organisation itself gets a fine, to be paid to the government. So none feels the pain of the mistake.

What happens if a bank goes bankrupt because the software that buys or sells shares makes a terrible mistake and sells or buys for a wrong price? Hunderd of people loosing there jobs?

Whose mistake: the programmer, the managing director or none? And who will be blamed or even punished? See even the bank crisis of the last 6 years. If someone is made guilty, not the management who was in charge. They left with a nice bonus. It were the simple customers of the banks who feel the pain, because of loss of their savings.

Unless the regulations/laws will change and will make the owner (or in case of the government someone high enough in rank) personal responsible, disasters with drones / robots will always end in a 'not guilty' for the real responsible person.

---

**Carl • [January 23, 2015 5:46 AM](#)**

If you own a self-driving car from google, will you still be responsible for traffic accidents, speeding tickets and other vehicle violations?

---

**Jim K • [January 23, 2015 6:08 AM](#)**

Health and safety laws grapple with similar problems, though the actors are real life intelligent people rather than machines.

Regardless of the negligent intent of their minions, executives are responsible for setting up and

maintaining a safe system of work.

---

**Robert Brown** • [January 23, 2015 6:16 AM](#)

For nations that have a heritage based on British Commonwealth law (and that is much of the world, thanks to British naval power resulting in the British Empire years ago), the precedent for how to handle the situation derives from the Torah:

[http://www.elilabs.com/cgi-bin/myword\\_query.cgi?query=ex21.28-32](http://www.elilabs.com/cgi-bin/myword_query.cgi?query=ex21.28-32)

The important points to consider are that the offender was the property of a human, and the defender had a certain amount of free will, being an animal, not a machine.

I expect the legal system in these countries will continue to follow this precedent and find the owner of the drone liable, not the creator (engineers in this case, God in the case of the animal).

---

**GordonS** • [January 23, 2015 6:52 AM](#)

> Who is responsible if an autonomous military drone accidentally kills a crowd of civilians?

Even for the existing 'semi-autonomous' drones, where pilots on the ground are pulling the trigger and killing civilians, nobody is held responsible.

---

**Jeff** • [January 23, 2015 7:03 AM](#)

@Robert Brown -- I think you're on the right point. The basis of the claim will be "failure to control a dangerous device". Your self-driving car kills someone, its your fault for not controlling that car. Your drone goes rogue and takes out a civilian, you put it there, so its your fault. There will certainly be defenses available, such as the demonstration of reasonable care, etc., but the basic liability should be clear.

That liability also provides the benefit of creating the right incentive for the person most capable of ensuring safety -- the person who is actually deploying the machine. Since they had the last clear chance to prevent the injury, they have the liability if anything goes wrong.

---

**Maj. Motoko Kusanagi** • [January 23, 2015 7:13 AM](#)

This resembles of Ghost in the Shell.

---

**Lucius Nasica** • [January 23, 2015 7:16 AM](#)

I think this is not a new question.

If a soldier breaks the laws on duty, the commander is also held responsible, so the

"commanding officer" of a drone (or A.I.) should be held co-responsible the same way - i.e. responsible as long as it is not possible to prove all best practices and procedures were honored. Of course poor "training" may be a decisive factor, even if in real world I've never seen a drill officer being held ultimate responsible for soldiers' failure, so even questioning developers about the development and deployment of the A.I. is important - the more powerful the A.I., the less important, but anyway the factor keeps its relevance.

---

**Janne** • [January 23, 2015 7:18 AM](#)

Somebody owns and deploys the drones. Even today a police or military force (and ultimately the society employing it) is legally responsible for the actions of its members, even if that member breaks the law or the organizations own rules.

---

**Les** • [January 23, 2015 7:25 AM](#)

The maker of the robot will be held responsible. It's not conceptually different from Toyota's "self-driving car" fiasco of a few years ago, or GM's current issue with cars that would occasionally relieve drivers of control.

---

**Gaspar** • [January 23, 2015 7:42 AM](#)

Clearly, culpability and responsibility do not always overlap.

---

**Autonomous** • [January 23, 2015 7:47 AM](#)

@Carl: the Google self-driving car

<http://www.techspot.com/news/59360-google-reportedly-preparing-enter-us-auto-insurance-market.html>

We all know about this sort of thing, having watched a former company company get into many different such markets, only to eventually abandon ALL of them (in general, the burden of payment often exceeded the willingness to pay claims, and most certainly exceeded the appetite for overly large profitability).

It is my observation that this is the only thing that will protect Google from ultimate devastation when these self-serving...ahh...self-driving vehicles start killing and maiming people. It is sort of like doctors insuring themselves against liability issues, where their own companies, using pseudo-claim analysts, deny such claims (similar to how many insurance companies deny claims now to avoid paying).

Insuring yourself in risky ventures seems a classic way to avoid liability issues by putting a non-accountable buffer between yourself and the risk.

Further, it is a self-driving car! The driver doesn't need insurance (the cars certainly won't be sold to anyone, only leased or rented; this makes for easy termination in case they don't like something about the passengers, like a low credit score or low medical score or some unknown factor that might place Google at risk).

THE DRIVER IS GOOGLE! Google needs insurance to protect themselves against claims by the passengers and other vehicle drivers/owners.

Or, passengers (those who lease or rent or ride) will be required to acknowledge and accept a hold-harmless clause to protect Google!

By the way, one reason I don't have or use a "smart phone" is because it is so dumb to do so (dumb for me; others make their own choices). My classic "basic" cell phone is in a Faraday cage wallet when driving in IL or walking through the mall. Cell phones are rapidly approaching the level of autonomous instruments.

End of rant...

---

**Bob Paddock • January 23, 2015 7:59 AM**

I think "Smart Cars" are about to get to smart. I once almost had a accident on Interstate 80, in a construction zone.

Somehow a car that was over packed pulled from between some construction equipment, in front of me.

This driver could not see out any window but right in front of him, and he was pulling *across* the interstate traffic, not going with the flow. He could not see out the passenger window, that was facing me.

He shot out from between the construction equipment about twenty feet in front of me, while I was doing 45 MPH, remember it was a construction zone.

The correct solutions to the problem was to *floor the gas*, so that I could get in front of him while there was still space, and get off on the right hand berm of the road.

Any system that applied the brakes would have guaranteed that a crash happened.

Show me a 'self driving' car that deals with this real world scenario only then I will entrust my life and loved ones to it.

---

**John Campbell** • **January 23, 2015 8:33 AM**

Well, Keith Laumer's Bolos have "commanders", though, once in battle reflex mode, all bets are off.

TSORP was built-in at a very low level to force a shutdown when the system evades an order from a lawful commander.

The whole area of autonomous machines had originally been explored by Laumer's Bolo stories; Other authors found it fertile ground to root their own stories in.

Let us not forget another such autonomous weapon-- Saberhagen's Berserkers.

Someone needs to consider how much computing capacity would be needed to implement even one of Asimov's Laws.

Drones are robots and the exploration into the ethics or robotic "action" has already been explored, even if only in imagination. The hell of it is that the most advanced "robots", when it comes to "thinking" (well, all such mechanisms are subject to GIGO... including human systems), are all designed for military driven destruction.

"It doesn't matter how well-crafted a system is to eliminate errors; Regardless of any and all checks and balances in place, all systems will fail because, somewhere, there is meat in the loop." - me

---

**Monty** • **January 23, 2015 8:41 AM**

In this particular case, it's pretty clear that the bot's creators are responsible - socially if not legally.

Imagine you build a robot that chops down trees and let it loose in a forest. Every day the robot moves to a random location and starts chopping. Most of the time it does indeed cut down trees; one day the robot injures a person instead. You'd then be charged (I guarantee that much), and if you said that you weren't responsible because the robot was acting on its own - well, you'd still be held responsible, because you're the one who made it the way it is. You create a hazard; you're responsible for what happens next.

Of course this is different. The tree-chopping robot would harm a person, whereas purchasing drugs is (depending on your views) either a victim-less crime or one that only harms society as a whole, in an abstract manner. And liability for the tree-chopping robot would primarily be a civil issue, not a criminal issue. You might additionally get charged with assault, but you might as well not.

In the Random Darknet Shopper case, the artists may or may not be held responsible for



receiving drugs, or facilitating their purchase, assuming that's even a crime. But they will likely be held responsible for owning them, a crime in itself in many jurisdictions.

Concerning questions such as "what if robots are truly autonomous" (a better term than "intelligent", I believe): liability laws may well still apply. Many pets and animals are autonomous, for instance; but if my cat bites a person, I'll still be required to pay their medical bills and possibly damages as well. "My cat does not follow my orders" is not a defense that will hold up in court, even if it is true.

Perhaps we should be approaching the entire question from a more practical angle. Instead of saying "how can we program morality into autonomous machines", perhaps we should say "how can we program the ability to be punished into autonomous machines": or, better yet, "how can we program autonomous machines to be social". Our societies are already equipped to deal with large numbers of autonomous beings. We merely need to make autonomous machines compatible with these existing systems.

---

**vas pup • [January 23, 2015 8:51 AM](#)**

@Bruce:"What do we do when a machine breaks the law? Traditionally, we hold the person controlling the machine responsible." I'd say OWNER of machine as first line of responsibility regardless of culpability. Then, owner could sue (regress claim) manufacturer/programmer/etc., fire person controlling machine (if culpability established) as of chain of responsibility/culpability emerged. For victim OWNER (gov or private) is primary responsible party. @Caspar has very good point.

@Autonomous • January 23, 2015 7:47 AM:"Insuring yourself in risky ventures seems a classic way to avoid liability issues by putting a non-accountable buffer between yourself and the risk." Yes! But if and only if risk were not result of neglect (concept of reasonable person by assumption that judge is always reasonable person with no bias/emotions in particular case meaning excluding MOTIVATED REASONING) or culpable actions. Insurance is working for ventures which have potential danger regardless of culpability, i.e. for civil responsibility only.

---

**John Campbell • [January 23, 2015 9:27 AM](#)**

What of the "autonomous" trading programs the stock market members use? While attempting to "maximize gains" (or, far more likely, minimize losses) haven't these opened up some questions for the firms running the programs?

There are some kinds of liability being a corporation-- or LLC (Limited Liability Company)-- cannot shield you from. At some point the onus of responsibility-- and blame, I guess-- must resolve to an individual human being as owner/operator/etc.

Consider, in some ways, non-sentient machines mimic, in a legal sense, "slaves". They are



owned by SOMEONE and it is the owner that is liable for what the "slave" (or attack dog) does.

Sadly, if a government owns it, who gives the order for the use of a robot to kill? Where, in the shackles of command (thank you, Robert Asprin, for coming up with a funny way to refer to the "chain of command"), does legal liability land?

Likewise, in a corporation, if a robot kills or injures a human being, who, of the officers, gets to go to prison? Or up for a needle?

One former Naval officer I work with explained "Authority can be delegated; Responsibility cannot be delegated."

---

**Future Skynet Developer • [January 23, 2015 9:55 AM](#)**

@GordonS

Bingo!

I enjoy seeing these comments from rational people trying to rationalize this irrational situation. We as humans enjoy personifying everything. We like to say we will add morality to the programming of AI, but we all know how that turned out in the movie IRobot. As soon as we are able to give free will to robots anything you previously programmed into them goes right out the door. Instead of Skynet like you see in the Terminator series, I would wager you would see a wide range of robo-personalities. There would robo-criminals, robo-police (robo-cop?), robo-extremists, robo-feminists, robo-serial-killers, robo-politicians, robo-jerks, robo-hipsters, any category of entity you see in society today (maybe even some new ones). Even with their superior processing abilities I would wager they would still be susceptible to circular reasoning and arriving at conclusions based on fallacies (human weaknesses).

I think the single most important question in this debate should be is the object in question property? or is it responsible for itself? that should answer question who to blame.

As soon as the entity in question is responsible for itself then it is in danger, remember there is no basic human rights for machines, no Geneva convention, nothing in our existing laws to protect them or protect us from them. There is nothing to stop a robo-holocaust, Skynet, or humans going on murderous rampages of entire robo-communities.

As much as people like Elon Musk and Stephen Hawking say "we will work together to prevent Skynet", I think that is BS, all it takes is one guy who thinks Skynet sounds pretty badass, and has the means to do so, then bam there it is unleashed onto humanity.

Remember there is no such thing as good or evil, just socially favorable and socially unfavorable actions. The robots who play nice with our society will be rewarded with power, opportunity and will be integrated, and those who do not will be savagely hunted to extinction like all the other

apex predators who prey on humans.

---

**T • [January 23, 2015 10:10 AM](#)**

legally that should not be so difficult:

a) civil liability (i.e. for the damages in \$\$\$): basically who controls the hazard has to pay, if at least some negligence is found (could be more than one person); with special laws that could be changed so that no negligence is needed. In some countries that applies e.g. for the owner of a car and other great hazards like nuclear facilities or railroads.).

b) criminal liability: you would also need to find culpability in the person(s) controlling the device. Otherwise it would be the same as a truck that drives into a crowd because the steering failed. If no one is found to be negligent (driver did not go too fast, had checked truck in a way that could reasonably be expected etc.), then no one will be convicted and punished. As is the case with all accidents that are nobody's fault.

Only if we get to the point of artificial intelligence/life on the level of a Mr. Data or such, could we try to "punish" the "device" (which in Star Trek Mr. Data is found not to be).

---

**T • [January 23, 2015 10:15 AM](#)**

and I should add: not all crimes can be committed by negligence only. Killing people, yes. Purchasing drugs as in the article; I guess no. Therefore in the case of this software: that should get him off criminally if my guess is correct and things work out ok (i.e. the court believes it was not him, but the autonomous software).

---

**Sandy Wills • [January 23, 2015 10:19 AM](#)**

I'd like to tweak a previous comment. The actual principle is:

"Authority can be delegated. Responsibility can only be \_shared\_."

Everyone from the Commander in Chief (for a military) or the CEO (for a corporation) all the way down to the immediate supervisor for the grunt actually doing the work is responsible for the results. And yes, the US is badly broken in that way. Neither our government officials nor our company officers are held to any responsibility in any way. Only the military still holds to that principle. Admirals and Generals still accept responsibility for what their men and women do. I may not LIKE what our military does, but I certainly TRUST anyone in a US military uniform more than I do their supposed civil masters.

---

**B. D. Johnson • [January 23, 2015 10:50 AM](#)**

*Last year, two Swiss artists programmed a Random Botnot Shopper, which every week would spend \$100 in bitcoin to buy a random item from an anonymous Internet black market*

*Who is responsible if an autonomous military drone accidentally kills a crowd of civilians?*

These aren't really comparable scenarios and "responsible" has a couple different things it could mean there.

In the botnet scenario they had every reason to believe that it would buy something illegal, deliberately started it up, and (apparently) took no steps to prevent it. There should be criminal liability there. If they had set it loose on Amazon instead or added filters to screen out illegal things that would be a different story.

In the autonomous drone scenario there wouldn't be criminal liability as the creators (presumably) wouldn't have reason to believe it would attack civilians and would have taken deliberate steps to ensure it didn't. While there would probably be a civil liability there (depending on exactly what happened to cause it to attack civilians), there shouldn't be any more criminal liability than in any other accident.

It's like if you're driving in your car (which you take steps to ensure remains safe to operate) and you hit a patch of ice and skid into someone's car, that's not criminal conduct. If you had bungee-corded your wheel in place, closed your eyes, and hit the gas in a crowded parking lot, *that's* criminal.

---

65535 • **January 23, 2015 11:00 AM**

Groan...

This case could be copied and misused in several scenarios. What if the NSA/FBI/GCHQ decided to go much farther than "parallel construction" and uses a software buyer drone to plant drugs on certain activists? Akin to "drone frame."

What if the NSA/FBI/GCHQ did the same as above but instead of a bag of drugs - buys a bomb to deliver to a "high value" adversary/target?

The above scenario could be used by criminals or even young script-kiddies for fun and/or retaliation. It is not a pleasant thought! The permutations of this "software buyer-drone" could be multifaceted - as the possible negative results.

"...we have already this issue. Not with drones (robots) but with companies and governments. If a governmental organisation makes a mistake, the person in charge is not punished. The organisation itself gets a fine, to be paid to the government. So none feels the pain of the mistake." - Jan Willem

How true this is.

"If a soldier breaks the laws on duty, the commander is also held responsible, so the "commanding officer" of a drone (or A.I.) should be held co-responsible the same way...in real world I've never seen a drill officer being held ultimate responsible for soldiers' failure..." -Lucius Nasica

Yes, the spreading of the blame and punishment never seem to hit the correct person.

"Further, it is a self-driving car! The driver doesn't need insurance... THE DRIVER IS GOOGLE! Google needs insurance to protect themselves against claims" –Autonomous

Interesting point.

"Cell phones are rapidly approaching the level of autonomous instruments." —Autonomous

I agree, modern cell phones are robot drone spy machines – which seem to be pwnd by the NSA.

'What of the "autonomous" trading programs the stock market members use?' - John Campbell

Good point. This a case of drone trading – which might explain some of volatility in the stock market - the bubbles and melt-downs. But, the big guy never gets punished.

All of the above need to be examined and re-mediated before broad and damaging drone warfare breaks out.

---

**Bob Staudenmaier • [January 23, 2015 11:42 AM](#)**

Finally! A book with a good title. That could get him some sales.

---

**Carl • [January 23, 2015 12:02 PM](#)**

*@Autonomous Or, passengers (those who lease or rent or ride) will be required to acknowledge and accept a hold-harmless clause to protect Google!*

Good points. They know their ways around EULA's.

in self-driving cars and human-killing autonomous drones, pilot/driver is the maker.

the owner gives the order, either to kill a target or drive to a destination, with the exception of self-driving taxi's, where passenger has no ownership.

possible defense for owner: "i told it to kill XXX but not YYY" "i told it to drive me to supermarket, i didn't tell it to run over my neighbor twice"

---

**pquirk • [January 23, 2015 12:12 PM](#)**

It's premature to to start assigning blame to autonomous robots when we haven't defined their constitutional rights. Do they have a right not to incriminate themselves (by encrypting their log files with a key that only they know?) Do they have second amendment rights?

When an autonomous robot (say a drone) is operating in the airspace of another country, which constitution governs the robot and its owner? What if the operator/owner is on a ship in international waters? Can we design generic autonomous robots that behave according to the laws of the land in which they find themselves?

And what of the "reasonable robot"? Will the judge be able to ask a jury whether this was the action of a reasonable robot? Will it have a right be judged by a jury of its peers?

Until we can say that a robot has basic constitutional protections, I believe the owner/operator will be held responsible. Jurisdictional issues will become much more common, both in war situations and in situations where the owner and the robot are in different legal settings. Just think about drones on a university campus versus outside it. Do campus police have any jurisdiction over the drone when it strays over homes miles from the engineering school?

---

**Joe Buck • [January 23, 2015 12:25 PM](#)**

We have another group of partially sentient intelligences that sometimes engage in destructive activities on their own, against the will of their owners: pets, livestock, and so forth. Perhaps we can treat a drone that goes rogue the same way as a dog that goes rogue: the owner is legally responsible for damages and under some circumstances the dog would be put down.

---

wumpus • [January 23, 2015 12:27 PM](#)

Wasn't this already decided (for US-based law) with the Morris worm? In this case it was strictly the programmer responsible: since the attack involved a worm, the owners of the hardware were held harmless. While it isn't clear that he (Robert Morris) had an intent to take down the internet (it was 1988 and that was pretty much what happened), he was held responsible for it all the same.

As far as I can tell, there really hasn't been a need to change such a law. As far as drones killing civilians, do you really think that when pilots kill civilians while in their planes, they face any greater legal threat?

---

Terminator • [January 23, 2015 12:34 PM](#)

Great essay Bruce! I enjoyed reading it and thinking about the questions it poses for society and justice systems. On that note, what if we have autonomous judges? ;)

---

Anura • [January 23, 2015 1:30 PM](#)

With regards to the darknet shopping bot, wouldn't you expect it to buy drugs or other illegal things? It seems pretty obvious.

Also, obligatory:

<http://xkcd.com/576/>

---

Sancho\_P • [January 23, 2015 1:47 PM](#)

That's a "good essay" !

*"Steal from a bank, and you'll almost certainly go to jail for a long time. A lot of this might seem more ad hoc than situation-specific, but we **humans have spent millennia working this all out**. Security is both political and social, but it's also psychological. Door locks, for example, only work because our social and legal prohibitions on theft keep the overwhelming majority of us honest. **That's how we live peacefully together at a scale unimaginable for any other species on the planet.**" (e.a.)*

... And that's also the reason why we are going to lose the peace.

Our laws have evolved from some moral principles to infinite legal fine-print, often hiding the roots.

This living artwork is still growing, driven by law professionals, powered by business lobbies.

From the “*what comes from above is untouchable*” authoritarian follower to the “*it’s fine until it’s officially declared illegal*” opportunist the system is widely supported, holding the few moral principles in headlock.

There is barely progress.

Poverty, social inequality (and their visibility) are on the rise, as are environmental problems, and we are discussing statistics.

Unable to exploit the tremendous technical progress to reduce conflicts and stabilize our spaceship we are now going to silence critical minds by technical and legal means, allegedly intended to fight terror and crime (e.g. FinFisher).

Later the same handle is very useful to fight social unrest.

Sure, nobody is accountable, all (state +) actors are blessed by impunity.

One has to be an optimist to see a smooth future.

---

MikeA • [January 23, 2015 1:50 PM](#)

"Talk to the bomb. Teach it phenomenology".

[http://lemonodor.com/archives/2008/02/ethics\\_in\\_lethal\\_robots.html](http://lemonodor.com/archives/2008/02/ethics_in_lethal_robots.html)

@Anura: Or maybe 325?

---

Sancho\_P • [January 23, 2015 2:07 PM](#)

@ Robert Brown, et all

I have a problem with the blanket expectation “The **owner** of xxx is guilty”.

That would mean regardless whatever inherent / intentional (low cost) issue the machine has, you, “who bought it”, is legally liable?

This would follow the insane IT - software liability cul-de-sac but doesn’t hold water when it comes to safety of physical action (machine control), at least today.

All work related to Safety Integrity Level (SIL) would be useless?

All car / airplane / ... makers would be free?

I don’t know if there are (safety) certified drones e.g. for the police, but wait, after some accidents

...

The owner only **may** be guilty, e.g. using a hammer to kill someone.

For state and military actions no one can be accountable, it’s us, the people. Sad, but true.



- Do not forget, our legal system is made by lawyers for lawyers, they will be glad to find / fight it out for you, **case by case**.

@ B.D. Johnson

I'm a bit with you for the Random Darknet Shopper but would like to point out that  
- at the very first -  
the LEAs are responsible and should be accountable for not taking down the "Darknet".

They (should do what the artwork did and) must not be paid for inactivity.

---

**Alan S • January 23, 2015 2:10 PM**

So, what happens today if an automated web crawler downloads content in the public domain which is illegal to possess?

---

**Anura • January 23, 2015 2:32 PM**

@Alan S

I think this eventually comes to a matter of thinking about the law. If it's that easy for a bot to break the law by browsing the web, then maybe the law should be changed so that what they are doing is not a criminal offense.

---

**tyr • January 23, 2015 3:16 PM**

I'd assume tort law already covers this in some detail.  
As far as criminal law it should fall under the same area as setting up a death trap. The person who sets up the trap is responsible even though the device may have been manufactured by another party.

Since law functions in the same fashion that driving a car by looking in the rear view mirror does, it never looks forward but only deals with new innovations after they come into its provence, you can't expect it to deal with any AI question until they are already out of the bag.

Since a careful perusal of Pareto shows humanity as considerably less that rational thinkers the real

danger of AI is in having a real basis for comparison. Few will be comfortable with an intelligence that is rational. The whole bag of monkey tricks used to deal with humans will have to be discarded when dealing with something truly rational.

The hilarious consequences of following orders exactly should be familiar to anyone with a sense of humour. "You can't just do what I say, you have to do what I mean"

Photography through the telescope erased the canals of Mars, hopefully machine intelligence will do the same for other treasured icons of human mythologies, at best the idea that we are rational actors should disappear.

It is going to be interesting to see the neophobe reaction to a widespread use of AI.

I'd like to see a little more success in programming ethical and moral behaviors into the current band of humans before we start deciding what is necessary for a machine.

---

**Brandioch Conner** • [January 23, 2015 3:19 PM](#)

But as machines become more autonomous, the link between machine and controller becomes more tenuous.

Why do you say that?

It's basic programming. If the machine would NOT do something WITHOUT your programming then YOU are responsible for the code YOU wrote.

Whether the legal system you live under will punish you is a different subject.

We already experienced a small example of this after 9/11, which was when most of us first started thinking about suicide terrorists and how post-facto security was irrelevant to them.

I think you're using those words incorrectly. Punishment is not security.

Security is about REDUCING the number of people who CAN do X.

The social and legal systems that have dealt so effectively with human rulebreakers of all sorts will fail in unexpected ways in the face of thinking machines.

No. It will come down to who owns/programs/runs the machines. The same as with every other machine out there.

---

**vas pup** • [January 23, 2015 3:31 PM](#)

@tyr: "The hilarious consequences of following orders exactly should be familiar to anyone with a sense of humour.

"You can't just do what I say, you have to do what I mean". Since all of us are not Justices of SCOTUS having clear vision when interpreting Constitution what EXACT meaning founding fathers had ((joke) , most of us need crystal ball to properly decipher implied meaning. That is why paper trail is always good (written order) in the case of doubts for protection. Sense of humor when following order could bring you in serious trouble is not applied I guess. For sense of humor is always proper time and place.

---

**Mike Amling** • [January 23, 2015 4:51 PM](#)

Line 1: "Random Botnot Shopper" Botnot := Botnet

---

**AlanS** • [January 23, 2015 4:55 PM](#)

Just a note that the Alan S posting above is not AlanS.

@Alan S

Might want to modify the name you post under to avoid confusion.

---

**Contanimation** • [January 23, 2015 5:14 PM](#)

Responsibility/liability can fall into multiple areas:

1. The owner: the car owner has to carry insurance, improper maintenance
2. The operator: especially if a law is broken (DUI, reckless driving)
3. The manufacturer/engineer/designer: if a design issue causes the failure
4. third party event: the car that hits you

The responsibility has to be apportioned by laws and the legal system. However, the process breaks down when none of the above can be identified. A Google car is essentially owned by a human/company, manufactured by/for Google, and arguably operated by either a human in the front seat or Google. Lawyers will fight until society works this out.

Now move to an AI driven self reproducing bot. As long as the programming does not evolve, there is still clearly a manufacturer/operator and should be an owner (even if not identifiable).

However, if the code is written to require two (or more) bots to "mate" by sharing randomly selected code/information (functions/data sources) over an industry standard and independently learn (AI) using the ever growing knowledge base of internet. Who is now the designer, owner, and operator. You have engineered evolution (and essentially life) into computing. There is nobody left to hold responsible.

Botnets already reproduce in code and use AI. Its only time before somebody engineers mating into them. When they can print their own physical bodies and buy their own CPUs from Amazon using stolen credit cards... I should have saved this for the movie plot contest.

---

**Clive Robinson** • [January 23, 2015 5:33 PM](#)

This is another rework of "God does not play dice" and "Free Will", with the biblical "eye for an eye"

There are two presumptions underlying law,

- 1, You have a choice (free will).
- 2, You can in some way be punished or rehabilitated (eye for an eye).

Computers are complex machines, but few would argue that machines are anything other than deterministic. Thus they can not have "free will"...

So how do you give a deterministic machine free will, well you may not be able to, but you can

quite easily "fake it" so it looks like it does. That is you add in a truly nondeterministic element which we normally call a "True Random Number Generator" or TRNG.

But just adding a TRNG only gives a cursory approximation to free will, because it is not tailored to a purpose, that is --supposed-- human free will is not random it's goal driven, and further the goals have to be rational.

That is someone who goes around making totally random acts is recognised as dysfunctional because they are not behaving rationally as others would expect (reasonable man). Further even if somebody does have a goal if it is outside societal norms it is judged as non rational.

Which raises the question can you develop deterministic machines to be not just rationally goal driven but also have the deterministic machine adjust it's rational goals to fit in with society by it's self?

If you can, can you really say it is a deterministic machine any more, because to fit in with societal norms by it's self it would have to be self aware and have the ability to learn and develop.

Importantly goals or their frustration is what punishment and rehabilitation is all about, that is you apply pain or suffering to the law breaker such that they will "learn" not to make the same or similar bad choices in future. Unless a machine can be made in addition to being self aware and to learn to suffer in some way it can not be punished.

So adding a TRNG will not make a deterministic machine equal in the eyes of the law, because the TRNG will in no way be effected by any kind of punishment.

But this raises the question of liability to the designers of a system. We tend to accept it is not possible to design a deterministic machine that can work in a complex environment with out it suffering from "dead lock" where a certain set of inputs will cause the machine to be unable to move or respond etc. How do designers solve this problem, well they "add noise" to the inputs to the machine such that the "dead lock" only lasts for a very short fraction of time. By definition noise is the output of a random process, importantly for anti dead lock solutions it has to be non deterministic.

Thus the problem arises that to stop a deterministic machine locking up and thus become dangerous in some situations that can not be known in advance, a designer has to add non deterministic behaviour to the system that could cause it to be a danger in other situations that can not be known in advance.

Which raises the question "Is it possible to design a system that can be guaranteed safe under all conditions in a complex environment where only some future situations can be predicted ?" To which the answer is "no"...

Thus would it be "reasonable" to make the system designer liable?

---

**Sancho\_P • [January 23, 2015 6:32 PM](#)**

@ (the other) Alan S, Anura

IIR(and understood)C there was a strange litigation in Germany for >10k of (free) "Youporn" watchers (streaming, not downloading ! ) for alleged copyright infringement. Many of them payed in a reflex. It turned out that the claiming lawyer didn't have the rights and finally lost.

"Streaming" (just to watch) is illegal if the consumer learns (can not deny knowledge, see legal fine-print) that the video / content is violating copyright, e.g. the video was taken inside a cinema.

The case was frightening for me because I tinker with a "machine" that keeps my computer busy, generating "noise", by crawling web pages and links - which of course sometimes starts downloading pdf's.

However I do not know how the machine could detect copyrighted content.

@ Clive Robinson

Why not (if you mean the manufacturer)? He who made money with **and** is culpable. An insurance may cover the costs.

---

**Contanimation • [January 23, 2015 6:48 PM](#)**

@Clive

Which raises the question can you develop determanistic machines to be not just rationally goal driven but also have the determanistic machine adjust it's rational goals to fit in with society by it's self?

If you can, can you really say it is a determanistic machine any more, because to fit in with societal norms by it's self it would have to be self aware and have the ability to learn and develop.

Consider the virtual "society" of the internet. Why can't a system programmed to use AI, reproduce, and mutate end up evolving these rational goals? The successful evolving code will survive even as the antibodies/"antivirus" attempt to kill it off.

While nearly all will be minimized or fail, the right code and parameters, the right rate of reproduction, and the right starting parameters very could evolve fast enough to survive. As it survives, it will become more capable and developed. Eventually, we very well could have "free will" at an AI level. Animals evolved this over a few hundred million years. The internet society

could evolve much quicker.

---

**albert** • [January 23, 2015 6:57 PM](#)

It's much too early to worry about robots committing illegal acts, and then worrying about whom to punish.

.

Many years ago NPR did a bit on 'smart bombs', the laser guided bunker-busters. They had time delay fuses that let them penetrate before detonating.

.

IIRC,

A particular smart bomb told the story of how it penetrated a bunker, and was surprised to find a bunch of people, still alive, huddled in fear. It thought about them, their families, and the fact that they had no control of their lives or the acts that led them into their present situation.

.

It decided not to detonate.

.

I gotta go...

---

**Carl** • [January 23, 2015 8:34 PM](#)

@ Joe Buck • January 23, 2015 12:25 PM

*"We have another group of partially sentient intelligences that sometimes engage in destructive activities on their own, against the will of their owners: pets, livestock, and so forth. Perhaps we can treat a drone that goes rogue the same way as a dog that goes rogue: the owner is legally responsible for damages and under some circumstances the dog would be put down."*

We cannot treat a drone the same way as a pet dog because the maker of such drone is clear and present.

Even for a self-programming drone, there was somebody who programmed it to self-program.

Not the same case for animals, so it's not a valid comparison.

---

**Wael** • [January 23, 2015 8:52 PM](#)

Fascinating topic...

How can we fault Machines that have no [conscience](#), no "feelings" and no sense of self awareness? "Faulting" the machine is meaningless to the Machine unless it possesses some "characteristics".

Until we're able to create a "Machine Conscience" module that is independent of the rest of



decision making logic, we cannot "prosecute" the machine. Then again, the machine needs to have a "self awareness" module. It needs to have "feelings". We are in a state of sin (at least with current technology) if we think we can program this into a machine. What you'd want to do is give the machine the initial instincts, feelings, and learning ability to develop its personality and let it grow and be accountable for its actions. Then, and only then, will the machine actions be free of the influence of its designer. The side effect is the machine may decide it doesn't like the task the designer intended for it and choose a different path. Humans cannot hold a machine that's driven by programmed logic (fuzzy or otherwise) accountable if it's not free willed. Randomness doesn't count as free will either.

As an analogy, why don't we apply the same punitive laws to the mentally ill, imbeciles, or juveniles as we apply them to "normal" humans? It's because they don't have the mental capability to "judge" right from wrong and weigh the consequences of their actions.

At the end of the day, this won't be about "Faulting" machines, but about "vindicating" humans who will eventually use the machines as scape goats. Forget about smart devices... next stop: Conscience-enabled devices. Until then, the designer is guilty as sin... at least in my monitored eBook...

---

**Earl Killian • January 23, 2015 9:07 PM**

I think there are two cases:

- (1) the machine is not eligible for personhood, and so it is owned by a person or corporation, in which case the owner is responsible for the machine's actions (and the owner should probably seek indemnification from the designer for bugs, but good luck with those "click agree" license agreements).
- (2) the machine has reached the point of being granted dignity rights (the term suggested for when "human rights" is considered speciesist (analogy to racist -- see speciesism)), in which case it can be held responsible for its own actions.

The real question in my mind is why we don't yet have suitable responsibility expectations of corporations, not AIs, since corporations have been with us for hundreds of years already, and AIs are still over the horizon. IMO, a corporation that commits a crime should be held accountable with a formula where N years in prison for a person is equivalent to N years of 10% of revenue penalty for a corporation. Most corporate crime punishment (that is when corporations are punished at all) is incredibly tiny by this standard.

---

**gordo • [January 23, 2015 9:28 PM](#)**

"More human than human is our motto." - Dr. Eldon Tyrell, Tyrell Corporation

---

**gordo • [January 23, 2015 9:41 PM](#)**

Blade Runner - Final scene, "Tears in Rain" Soliloquy

---

**Ole Juul • [January 24, 2015 2:01 AM](#)**

People get themselves so confused. If, upon returning to the dinner table, you find your soup bowl empty; I will simply explain to you that the spoon did it. Seriously, I have a cat that can figure this stuff out. Keep your eye on the money. Mind you, I have no idea where the spoon came from. I suppose you could blame whomever set the table.

---

**Me • [January 24, 2015 11:33 AM](#)**

"Who is responsible if an autonomous military drone accidentally kills a crowd of civilians?"

For me that's an easy question. The people responsible are the ones that had a clear overview of what the machine was capable of doing and did not prevent it from coming into existence, nor they prevented its use.

I do not believe in things happening by accident, like 10 teams scientist working on 10 irrelevant projects which came together accidentally with no clue on what was being built.

For the drones, the main criminals are the people in government that paid for them and put them in operation. Then, there are the businessmen and the scientists than have brought them about. Both of these groups **usually** have a crystal clear picture of what they are developing and where it could be used. And, in fact, apart from western governments that are known to be criminals, most businessmen and most scientists are the same too. AIs are nothing but a new weapon for killing people, just like nuclear power was. Without super strict laws forbidding their use in military operations that will be enforced by a non-terrorist international organization (in contrast to terrorist organizations like NATO, EU, etc), any person developing them for public release should be considered a criminal.

---

**jdgal • [January 24, 2015 3:34 PM](#)**

Whoever owns a 'bot needs to be legally liable for anything it does, and thus has the responsibility always to be able to restrain it. (Of course, this means that the person or company that buys one had better make sure he's going to be able to discharge that responsibility -- and

he'd better get a guarantee from the maker that the 'bot hasn't been programmed in a way that will take control away from him at a later time.) He can delegate the job of monitoring or piloting the 'bot, but must remain responsible anyway. This holds even if the 'bot is capable of operating independently (like an Aegis missile system). There's no new legal principle here, and no need for one.

A truly self-aware, intelligent machine would be a different story, and a fair legal system would have to grant it "personhood" -- hopefully with the full rights of a human (since the alternative is slavery). But I don't think such a machine will be built in the next century; and if it did happen I would require substantial testing before believing it and granting personhood.

The problem of governments and militaries misbehaving and not answering for it is a separate issue, one that we've had for all of history. The laws of war allow so many kinds of behavior that ordinary laws don't, not because they're desirable but because there is no practical way to prevent them or do without them.

I would like to see existing requirements that vehicles, aircraft, and military weapons be marked with national flags and/or license-plate equivalence extended to cover all autonomous vehicles and drones. And victims of attack, collisions, or even privacy intrusion should be granted the legal right to take photos and use them as evidence, and to capture or shoot down machines that don't carry the required identification or that have harmed them. Those types of reaction, and responses to them in turn, are the area where the law will need to establish new precedents.

---

**emk • January 24, 2015 3:55 PM**

I don't see this as a particularly big problem. Unless we decide that thinking machines are legal persons i.e we create the problem ourselves.

Otherwise the machine is the property of somebody akin to an animal or a legal entity like a corporation. That owner is ultimately responsible for any harm caused by their property. That's the law now. Whether their property is autonomous, thinking etc the owner is responsible.

The fact that the US does not do a good job in enforcing laws against (large) corporate entities is not because it's impossible or even difficult to do so, but rather because the US as a society is controlled by corporate entities. This is a peculiarity of US society.

---

**Sancho\_P • January 24, 2015 4:44 PM**

Someone owns a "machine" that transports people (a cablecar) for skiing.

During the ride two boys play with their radio, from one cabin to the other.

Suddenly the control system runs havoc.

The car smashes against the wall because the stop didn't work.

The owner is responsible?

Or I, who constructed and built the system, ignoring technical standards?

---

**Terry Cloth** • [January 24, 2015 6:31 PM](#)

That's either one smart machine or one dumb vendor. How do you buy a fake passport without supplying a photograph? Did the bot recognize the request, and pick a random head shot from the 'net? Did they send a blank one, expecting you to have your local shady laminator to add your pic?

Inquiring minds want to know.

---

**tyr** • [January 24, 2015 10:09 PM](#)

Just suppose.

You develop a drone with a semi- autonomous AI.

The operator only exists in the loop to say don't do that to the drone when it is going to make an egregious mistake. This makes the average politician happy because there is a fallible human in the loop.

Who gets blamed when the machine AI reacts so fast the human can't stop it. There is no reason to suppose that a machine can't have much better reflexes than a human and no reason to build a war machine with inferior ones.

Speaking from the stone age of such things it takes a lot of nerve to depend on an ally who will automatically take you out if you get a single transistor fail in your IFF (Identification Friend or Foe). IFF failure will also get shot down by human friendly fire and no amount of handwringing over the morality will fix the problem.

One base assumption that needs to be discarded a machine will not think the way a human does,

it will think like a machine. If you follow up on Rod Brooks early robotics it says a lot about how human nervous systems work but with the one exception of Ramachandran most haven't got the message yet. It is hard to discard the superstitious trappings of thousands of years just because they are wrong but it's the only way we'll get a grasp on what is so badly needed today.

Just as an off topic aside.

Wouldn't it be a lot simpler to use 8 bit micros and simpler code for the peripherals? You could then verify that they aren't hiding some nefarious blob. I realize this flies in the face of OOP and 64 bit CPU thinking with massive software bloat. : ^ )

---

Clive Robinson • **January 25, 2015 12:00 AM**

@ tyr,

Wouldn't it be a lot simpler to use 8 bit micros and simpler code for the peripherals?

Why stick to just the peripherals?

There are good reasons to have wide bus widths for a quite limited number of functions, but for many more they are actually a hinderance, and guess which functions get used most frequently?

Another area for wide bus widths is trying to get around the speed of light problems with memory. Take a 1GHz clock speed that has a wavelength of around thirty centimeters or one foot (in old money ;-). Thus the maximum distance you could achieve for an out and back signal is fifteen centimeters, however due to other effects half that again. Thus the PCB size you can use is finite to get the data bandwidth up you need to have signals run in parallel. There are two ways to do this, the first is to have physically more traces or tracks on the PCB the second is to use multilevel / phase encoding of data on a single trace. Currently we tend to favour the first method for simplicity in the design, hence we have quite wide bus widths just to have sufficient data bandwidth. That said these days the data from the bus almost always goes into a cache not the CPU direct, thus the CPU internal bus width could be any width required.

Eight bit micros use very few transistors and thus quite a few could be put on a single chip including appropriate MMUs and local memory.

The only real limitation is that of the lump of fat sitting between the ears of the person

programing. Currently we tend to think of things happening sequentially not in parallel, which tends to make our usage of multi-core systems inefficient... maybe it's time we played catchup or developed better tools.

---

**Clive Robinson** • [January 25, 2015 12:35 AM](#)

@ tyr,

Are you talking about this chap and his "Plastic pal that's fun to be with"?

<http://www.xconomy.com/san-francisco/2014/05/22/baxter-meets-world-rod-brooks-on-what-rethink-robotics-is-learning/>

---

**Wesley Parish** • [January 25, 2015 4:07 AM](#)

This is one topic I've approached in this story [Trouble In The ACT](#) and attempted to analyze.

If you're aware of the differing political emphases of Australia dn the United States, this story won't be difficult: Australia doesn't have the automatic reaction to "Socialist" that the US has. So a program that responds to "socialist" in the cross-agency data-checking dataminer as to "terrorist" is American; and dead-wrong, at least as far as the system that leads this (unseen) autonomous drone to strike at these innocent civilians.

In short, even if the policy is separated - as it should be - from the execution thereof, it still depends on the foundational principles and ethics of the people who program these machines. That only changes if said machine is capable of reprogramming itself to the degree that it can rewrite the entire set of foundational principles and ethics. That is very difficult for humans to do: "conversion" is an fascinating psychological topic, if only because it is impossible to complete - there is always residual behaviour that won't change.

So, who is responsible for the behaviour of a given "autonomous" machine? Primarily, it is the programmers, who impose their sets of values on it. then it is the owners and operators, whether or not they are the same people.

It's only if the drone can make its own decisions and enforce them against "hypothetical" owners and operators, that the drone bears any responsibility for its actions.

'nuff sed?

---

**Clive Robinson** • [January 25, 2015 5:56 AM](#)

@ Wesley Parish,

So, who is responsible for the behaviour of a given "autonomous" machine?

Primarily, it is the programmers, who impose their sets of values on it. then it is the owners and operators, whether or not they are the same people.

Sadly you have left out the real culprits... Programers work to a specification, the specification is generally drawn up from a perceived set of needs of others. It is these "others" and those who dream up solutions for them who are the real amoral ones.

In essence you have a group of individuals without any kind of morals who want to make money, they really do not care about the ethics of what they do. They find what they think is "a hole in the market" and come up with a couple of sentences to describe the perceived solution. This then gets handed down to others to flesh out these bones and make it sound exciting and sexy, they likewise have no morals or ethics. Then the equivalent of "marketing" come up with a "desirable features list" which becomes the skeleton of the "needs requirements" that eventually becomes the specification.

Up to this point the entire drive behind the specification is how to "extract benefit" from the "perceived market gap", at no point does it cross the minds of these amoral individuals that there may be a good and proper reason why there is what they perceive as a "hole in the market".

For instance it is known that you can cause significant neurological damage with two beams of energy whose frequency difference and modulation will cause --on effective demodulation in the body-- nerves to be stimulated outside of their normal ranges, and thus lead to unconsciousness and death. To the point of unconsciousness weapons that do this are considered "non-lethal" and thus have been researched as such on various government grants. What is frequently not mentioned and was known well in advance of any such research requests is that the effectiveness of weapons of this type drops off as with all energy beams as the inverse square of the distance, thus if you are aiming at a person at a distance you are killing those at lesser distances. Further that the effect on the human body is accumulative as nerves do not regenerate after such insults and the psychological effect is in effect the same as caused by medieval style torture, which would be illegal in this day and age... So who is responsible for these weapons being not only researched, but designed, built and field tested?

---

**BP • January 25, 2015 6:49 AM**

Computers can never be programmed to understand that the world can die. That there can be an end to our world. You cannot teach that concept to a computer. Perhaps we must make sure that the computer understands that the end of time is the most evil thing possible. As is the end of one human soul.

The principle that when you save one life you have saved the world must be programmed into these machines. The people who engage in the hacking of computers today fail to grasp the



concept that computer harassment can cause suicidal thinking. And that is an evil that must be avoided at all costs. The present state of the "understanding" of these machines fails to grasp that when they save one life, they save the world. We must get back to that concept. Or mankind is doomed.

---

**BP • [January 25, 2015 6:59 AM](#)**

I lived in South Carolina when gambling were ubiquitous. Mothers let babies die in hot cars because the programmers understand how to make the machines grab the soul. I never will forget the gentlemen who was playing one of those machines. (I used to sit in a southern country store and just observe people while I drank my diet soda so as to get my inventory of southern stories a little larger). The gentlemen was having a run of luck on the machine and it was clicking and clacking like an angry devil, all programming done to fool the soul, and the gentleman remarked "listen to it, it's trying to outthink me". It was then that I understood that machines could be made to be truly evil.

If you subscribe to Harpers (which I recommend, I found they have archives going back to 1850 and thus I will never run out of "free" and excellent reading stock) read the whole thing.

<http://harpers.org/blog/tag/video-poker/>

And weep. For humanity. Our leaders have discovered what bells and whistles controlled by machines can do. And thus mankind is doomed.

---

**BP • [January 25, 2015 7:11 AM](#)**

George Orwell once remarked that as bad as British Colonialism and Empire was, the nations that would replace it were far worse. He often wrote of the essential gentleness of the English people, but not of the Americans.

This was written in WWII as bombs were being dropped on England but Orwell knew even that that the USSR and USA would replace the British Empire. AND said so in not so many words, but George wrote subtleties. Roosevelt drove a hard bargain with Churchill. Our machines will be our sad legacy.

So here's a hat tip to thinking George, who so ominously predicted 1984, or today's world, and the clue was dropped when he mentioned how Nosey Parker was the most hated name in the English lexicon. And then proceeded to give us in the same sentence a clue to what he was going to do next. And then he retired to Jura to write 1984. So listen to and read the following and drop a mental hat tip to George. He was so prescient.

<https://www.youtube.com/watch?v=34Vxgydpmus>

[https://en.wikipedia.org/wiki/Burmese\\_Days](https://en.wikipedia.org/wiki/Burmese_Days)

---

**Gweihiir • [January 25, 2015 10:41 AM](#)**

"The social and legal systems that have dealt so effectively with human rulebreakers..."

I disagree on this. There is ample evidence that at least the legal system is mostly worthless. For example, the "war on drugs" with its draconian punishments seems to have had zero effect on the consumption of drugs and at the same time massive negative effects on society as a whole. The same is true for most murderers that do not have a mental condition (those with such a condition should not be subject to the legal system anyways, but should be dealt with by the medical system): Most of them committed their crime under circumstances that were extreme and are not going to repeat itself and where the punishment was completely immaterial to them. Hence the law has neither a deterring effect, nor does it have any other beneficial effect. It does do a lot of additional damage though.

At the same time, people that have created untold misery and loss to society (see for example the banking crisis of 2008), are not even prosecuted and come out of their crimes against society with their ill-gotten fortunes intact.

Hence what I see here is merely the collapse of the fantasy that "the law" actually works, when it is pretty useless in general and outright harmful in many cases. What "the law" seems to be doing is basically feeding a lot of parasites with a hugely inflated sense of their value to society and a destructive authoritarian mind-set.

As to the question of "crimes" committed by machines: First, if the crime is basically victim-less as in buying drugs, who cares? If the machine killed or harmed somebody, then there is somebody else that was responsible for its safety, just as with all other non-intelligent machinery. What the machine can do is immaterial to this. In such cases there needs to be an investigation, and the outcome can be "accident" or "negligibility". Depending on rules and outcome, the entity responsible for the safety of this machinery or their insurance pays for any and all damage.

There is of course a second (but according to the state-of-the-art highly unlikely and maybe impossible) case: The machine is actually sentient and makes its own decisions. Here, the machine should just be viewed as a person and treaded accordingly. That means the machine can be prevented from doing it again and "fixed" (if possible) by the equivalent of a medical system, and it can be required to pay for damage done from its personal wealth. Punishment and threat of punishment does not work on most humans, so if it should not work on sentient machines, that would actually not cause any new problems. It may help to shut down or reform a legal system that is far more problem than solution though.

tl;dr: I don't think there is a new problem here.

---

**Brandioch Conner • [January 25, 2015 11:31 AM](#)**

Now I'm wondering if the whole "moral machines" thing is the same as a "zombie plan".

That is, a fantasy that nerds and geeks can argue about. Even though it is impossible.

---

**Ole Juul • [January 26, 2015 2:58 AM](#)**

*Brandioch Conner • January 25, 2015 11:31 AM*

*Now I'm wondering if the whole "moral machines" thing is the same as a "zombie plan".*

*That is, a fantasy that nerds and geeks can argue about. Even though it is impossible.*

Of course it is. My earlier post about my cat was dead serious and, I thought, a strong comment on what is going on in this thread.

The story about a computer buying stuff or killing someone is just causing confusion in the minds of those who are predisposed to that confusion. It could just as well be about a man pushing a button to make a bomb go off. If you've never seen that kind of remote control before, then it might look impressive. However at this point in history we really should know better. My cat does.

---

**Barney • [January 26, 2015 5:57 AM](#)**

When humans break the English civil law in the course of doing their jobs their employer is held responsible through the doctrine of vicarious liability. There's no need to show that the employer did anything wrong as they are responsible for their employees' actions. I wonder if there should be a similar doctrine for autonomous systems. Otherwise it seems like an employer could reduce their liability by replacing fallible humans with equally fallible machines.

---

**gordo • [January 26, 2015 6:48 AM](#)**

@ Ole Juul, Yes, a serious looking cat!@

To maybe help dispel the kinds of confusions that are out there, and that you warn against feeding, the below couple of snippets are from a "Reality Club" discussion that follows a conversation:

The Myth Of AI

A Conversation with

Jaron Lanier [11.14.14]

<http://edge.org/conversation/the-myth-of-ai>

The "Reality Club" post, introduced with a brief comment to the piece's co-conversationalist, John Brookman, is an essay by Rodney A. Brooks [Robotician; Panasonic Professor of Robotics]

(emeritus), MIT; Founder, Chairman & CTO, Rethink Robotics; Author, *Flesh and Machines*]. Brooks' essay is named: "Artificial Intelligence Is A Tool Not A Threat."

A couple of excerpts from Brooks:

Recently there has been a spate of articles in the mainstream press, and a spate of high profile people who are in tech but not AI, speculating about the dangers of malevolent AI being developed, and how we should be worried about that possibility. I say relax. Chill.

---

I think it is a mistake to be worrying about us developing malevolent AI anytime in the next few hundred years.

---

In the 1930's Turing was inspired by how "human computers", the people who did computations for physicists and ballistics experts alike, followed simple sets of rules while calculating to produce the first models of abstract computation. In the 1940's McCullough and Pitts at MIT used what was known about neurons and their axons and dendrites to come up with models of how computation could be implemented in hardware, with very, very abstract models of those neurons. Brains were the metaphors used to figure out how to do computation. Over the last 65 years those models have now gotten flipped around and people use computers as the metaphor for brains. So much so that enormous resources are being devoted to "whole brain simulations". I say show me a simulation of the brain of a simple worm that produces all its behaviors, and then I might start to believe that jumping to the big kahuna of simulating the cerebral cortex of a human has any chance at all of being successful in the next 50 years. And then only if we are extremely lucky.

[direct link to Brooks essay: <http://edge.org/conversation/the-myth-of-ai#25982> ]

---

**tz • [January 26, 2015 3:49 PM](#)**

The problem is we either can't agree, or have the same problems punishing humans as C. S. Lewis noted:

<http://www.angelfire.com/pro/lewiscs/humanitarian.html>

What do the machines DESERVE? Or to put it differently, are we not TODAY treated by the system as merely advanced machines - to be made examples as defective or to be repaired? To quote part of the article:

According to the Humanitarian theory, to punish a man because he deserves it, and as much as he deserves, is mere revenge, and, therefore, barbarous and immoral. It is maintained that the only legitimate motives for punishing are the desire to deter others by example or to mend the criminal. When this theory is combined, as frequently happens, with the belief that all crime is more or less pathological, the idea of mending tails off into that of healing or curing and punishment becomes therapeutic. Thus it appears at first sight that we have passed from the harsh and self-righteous notion of giving the wicked their deserts to the charitable and enlightened one of tending the psychologically sick. What could be more amiable? One little point which is taken for granted in this theory needs, however, to be made explicit. The things done to the criminal, even if they are called cures, will be just as compulsory as they were in the old days when we called them punishments. If a tendency to steal can be cured by psychotherapy, the thief will no doubt be forced to undergo the treatment. Otherwise, society cannot continue.

My contention is that this doctrine, merciful though it appears, really means that each one of us, from the moment he breaks the law, is deprived of the rights of a human being.

The reason is this. The Humanitarian theory removes from Punishment the concept of Desert. But the concept of Desert is the only connecting link between punishment and justice. It is only as deserved or undeserved that a sentence can be just or unjust. I do not here contend that the question 'Is it deserved?' is the only one we can reasonably ask about a punishment. We may very properly ask whether it is likely to deter others and to reform the criminal. But neither of these two last questions is a question about justice. There is no sense in talking about a 'just deterrent' or a 'just cure'. We demand of a deterrent not whether it is just but whether it will deter. We demand of a cure not whether it is just but whether it succeeds. Thus when we cease to consider what the criminal deserves and consider only what will cure him or deter others, we have tacitly removed him from the sphere of justice altogether; instead of a person, a subject of rights, we now have a mere object, a patient, a 'case'.

The distinction will become clearer if we ask who will be qualified to determine sentences when sentences are no longer held to derive their propriety from the criminal's deservings. On the old view the problem of fixing the right sentence was a moral problem. Accordingly, the judge who did it was a person trained in jurisprudence; trained, that is, in a science which deals with rights and duties, and which, in origin at least, was consciously accepting guidance from the Law of Nature, and from Scripture. We must admit that in the actual penal code of most countries at most times these high originals were so much modified by local custom, class interests, and utilitarian concessions, as to be very imperfectly recognizable. But the code was never in principle, and not always in fact, beyond the control of the conscience of the society. And when (say, in eighteenth-century England) actual punishments conflicted too violently with the moral sense of the community, juries refused to convict and reform was finally brought about. This was possible because, so long as we are thinking in terms of Desert, the propriety of the penal code, being a moral question, is a question in which every man has the right to an opinion, not because he follows this or that profession, but because he is simply a man, a rational animal enjoying the

Natural Light. But all this is changed when we drop the concept of Desert. The only two questions we may now ask about a punishment are whether it deters and whether it cures. But these are not questions on which anyone is entitled to have an opinion simply because he is a man. He is not entitled to an opinion even if, in addition to being a man, he should happen also to be a jurist, a Christian, and a moral theologian. For they are not question about principle but about matter of fact; and for such *cuiquam in sua arte credendum*. Only the expert ‘penologist’ (let barbarous things have barbarous names), in the light of previous experiment, can tell us what is likely to deter: only the psychotherapist can tell us what is likely to cure. It will be in vain for the rest of us, speaking simply as men, to say, ‘but this punishment is hideously unjust, hideously disproportionate to the criminal’s deserts’. The experts with perfect logic will reply, ‘but nobody was talking about deserts. No one was talking about punishment in your archaic vindictive sense of the word. Here are the statistics proving that this treatment deters. Here are the statistics proving that this other treatment cures. What is your trouble?’

The Humanitarian theory, then, removes sentences from the hands of jurists whom the public conscience is entitled to criticize and places them in the hands of technical experts whose special sciences do not even employ such categories as rights or justice. It might be argued that since this transference results from an abandonment of the old idea of punishment, and, therefore, of all vindictive motives, it will be safe to leave our criminals in such hands. I will not pause to comment on the simple-minded view of fallen human nature which such a belief implies. Let us rather remember that the ‘cure’ of criminals is to be compulsory; and let us then watch how the theory actually works in the mind or the Humanitarian. The immediate starting point of this article was a letter I read in one of our Leftist weeklies. The author was pleading that a certain sin, now treated by our laws as a crime, should henceforward be treated as a disease. And he complained that under the present system the offender, after a term in gaol, was simply let out to return to his original environment where he would probably relapse. What he complained of was not the shutting up but the letting out. On his remedial view of punishment the offender should, of course, be detained until he was cured. And of course the official straighteners are the only people who can say when that is. The first result of the Humanitarian theory is, therefore, to substitute for a definite sentence (reflecting to some extent the community’s moral judgment on the degree of ill-desert involved) an indefinite sentence terminable only by the word of those experts—and they are not experts in moral theology nor even in the Law of Nature—who inflict it. Which of us, if he stood in the dock, would not prefer to be tried by the old system?

It may be said that by the continued use of the word punishment and the use of the verb ‘inflict’ I am misrepresenting Humanitarians. They are not punishing, not inflicting, only healing. But do not let us be deceived by a name. To be taken without consent from my home and friends; to lose my liberty; to undergo all those assaults on my personality which modern psychotherapy knows how to deliver; to be re-made after some pattern of ‘normality’ hatched in a Viennese laboratory to which I never professed allegiance; to know that this process will never end until

either my captors have succeeded or I have grown wise enough to cheat them with apparent success—who cares whether this is called Punishment or not? That it includes most of the elements for which any punishment is feared—shame, exile, bondage, and years eaten by the locust—is obvious. Only enormous ill-desert could justify it; but ill-desert is the very conception which the Humanitarian theory has thrown overboard.

If we turn from the curative to the deterrent justification of punishment we shall find the new theory even more alarming. When you punish a man in *terrorem*, make of him an 'example' to others, you are admittedly using him as a means to an end; someone else's end. This, in itself, would be a very wicked thing to do. On the classical theory of Punishment it was of course justified on the ground that the man deserved it. That was assumed to be established before any question of 'making him an example' arose. You then, as the saying is, killed two birds with one stone; in the process of giving him what he deserved you set an example to others. But take away desert and the whole morality of the punishment disappears. Why, in Heaven's name, am I to be sacrificed to the good of society in this way?—unless, of course, I deserve it.

But that is not the worst. If the justification of exemplary punishment is not to be based on desert but solely on its efficacy as a deterrent, it is not absolutely necessary that the man we punish should even have committed the crime. The deterrent effect demands that the public should draw the moral, 'If we do such an act we shall suffer like that man.' The punishment of a man actually guilty whom the public think innocent will not have the desired effect; the punishment of a man actually innocent will, provided the public think him guilty. But every modern State has powers which make it easy to fake a trial. When a victim is urgently needed for exemplary purposes and a guilty victim cannot be found, all the purposes of deterrence will be equally served by the punishment (call it 'cure' if you prefer) of an innocent victim, provided that the public can be cheated into thinking him will be so wicked. The punishment of an innocent, that is, an undeserving, man is wicked only if we grant the traditional view that righteous punishment means deserved punishment. Once we have abandoned that criterion, all punishments have to be justified, if at all, on other grounds that have nothing to do with desert. Where the punishment of the innocent can be justified on those grounds (and it could in some cases be justified as a deterrent) it will be no less moral than any other punishment. Any distaste for it on the part of the Humanitarian will be merely a hang-over from the Retributive theory.

It is, indeed, important to notice that my argument so far supposes no evil intentions on the part of the Humanitarian and considers only what is involved in the logic of his position. My contention is that good men (not bad men) consistently acting upon that position would act as cruelly and unjustly as the greatest tyrants. They might in some respects act even worse. Of all tyrannies a tyranny sincerely exercised for the good of its victims may be the most oppressive. It may be better to live under robber barons than under omnipotent moral busybodies. The robber baron's cruelty may sometimes sleep, his cupidity may at some point be satiated; but those who torment us for our own good will torment us without end for they do so with the approval of their own conscience. They may be more likely to go to Heaven yet at the same time likelier to make a Hell



of earth. Their very kindness stings with intolerable insult. To be 'cured' against one's will and cured of states which we may not regard as disease is to be put on a level with those who have not yet reached the age of reason or those who never will; to be classed with infants, imbeciles, and domestic animals. But to be punished, however severely, because we have deserved it, because we 'ought to have known better', is to be treated as a human person made in God's image.

---

**tz • [January 26, 2015 4:23 PM](#)**

I quote Lewis at length because I think there is a categorical error if one doesn't consider that justice is meeting out a deserved punishment.

By shifting to deterrence, a lot of really horrible things can be done in the name of deterrence and will be quite effective. Under "deterrence", the problem with Aaron Swartz is not that he killed himself, but that there were too few earlier examples to deter him. We need the NSA but with a SWAT team. As Lewis points out, moral proportion has no place in the argument. Something is either capable of deterring or not, and deterrence has now become the end, instead of being merely a useful side-effect of actual justice. The FBI entrapping people who would otherwise be considered innocent or never think of doing anything to "prevent terrorism" is a current example.

A side problem is we used to require an evil act - there were no "thought crimes". "Hate crimes" have started to cross the line, as well as "terroristic threats", or the "material support of terrorism" laws for publishing a youtube saying nice things about bad people. What if the thinking machine has the wrong thoughts but doesn't act?

It is a question whether or not machines might or might not "think", that is have a will. But the immediate question is why should such a will ever be evil? And I'm talking about Lewis category. The first machine quite likely will be Adam before Eve and before the Serpent. Totally innocent. But this can also be a problem - an adult who has the mind of a 2 year old can be both innocent and quite dangerous and do actual harm without intending to.

But lets say we have an innocent, but intelligent machine, and it wants to help mankind (something like Asimov's first law, plus). This will allow me to circle back to Lewis point.

Is it harmful to restrict liberty or to even take the lives of some otherwise innocent humans? Or to take their property (think guns)? It depends on whether you think Human Beings have a mind or soul, and that it is the most important and critical thing. Such oppression only acts on the soul. The pain is entirely internal and because you won't cooperate with the "good". And the machine will know what it best.

If the mind is even coequal to the body for a human to be human, than any action not based on justice and desert - or going against them - is harm. A form of rape (Consider that rape is not punished as if it were the equivalent of theft of service from a willing prostitute but something far

more serious).

If, because it is hard to measure "mind" we can't add it to the calculation, or it is considered a part of the body, then not to "cure" those who can be cured, or to "deter" those who can be deterred is harm. Yet that is what the Humanitarian theory is - to treat humans as machines.

Yet we may never have machines with a "mind" - and if we don't, either the acts they do are part of the will of one or more human beings (it can be a will to be negligent), or they will be unintentional and random like a hailstorm or lightening. That is where and how to seek justice and desert. Otherwise you have to deter or cure people from making machines that act.

---

**Alex • January 26, 2015 4:53 PM**

I just picked up a self-driving car at the end of December, so this topic interests me quite a bit.

As the law stands right now, I am 100% responsible for whatever my car does. I am the owner, it is my car.. but there's a problem: In particular my new car, even when on "manual" mode, is more like an Airbus A320 -- the driver/pilot never has 100% control.

You can tell the car what your intentions are...but that is merely input in series of equations which also factor in weather conditions (temperature, rain, sunlight), the locations & speed of vehicles around the car, road traction, G-forces, etc.

So far, 99% of the time, it gets it right. I have seen a couple of glitches where it wanted to follow a car down an interstate off-ramp rather than track straight or where it didn't want to move any further forward because it thought I was too close to the car in front of me.

Likewise, the car also determines how much braking force is necessary and reasonable. The car's definition of reasonable and mine were completely different -- it originally braked much later than I would. We've since tuned it to something more comfortable. BUT it does beg the question...what happens if the car gets it wrong? I'm the owner, so it's my car, my problem, and with current laws, and the disclaimer the manufacturer had me sign, my fault.

I do worry about this issue on the computer front as well. Should botnet-possessed computer owners be responsible for their computer's actions? I think it's reasonable that people should have SOME idea of what their computer is doing. BUT very few of us are programmers anymore, and I highly doubt even the most tech savvy among us packet sniff their connection 24/7. If I run Tor on my computer and some anonymous perv on the network starts downloading child porn, should I be held accountable? If someone uses my WiFi for nefarious things, am I now a party to the crime?

In many ways these questions have always existed. Take the recent Justice Department review on Civil Forfeiture. If someone borrowed your car and then got busted dealing drugs, the feds

confiscated the car. Even though you had no direct involvement in it, you indirectly helped.

---

**Clive Robinson** • [January 26, 2015 4:57 PM](#)

@ tz,

What C.S.Lewis did not discuss in what you quoted was if any punishment or rehabilitation should be visited on the criminal.

The issue revolves around the idea that we somehow have "free will" where as most of the universe behaves in what appears to be a preordained way fixed by the laws of nature and thus without free will.

If we do not believe in free will --which is a logical place to start-- then logically we can not call for punishment for any reason as the person had no choice their actions were preordained by those same rules by which all matter is constrained. Thus the punishment or rehabilitation would not change anything because it could not be changed.

However even where free will appears to exist we have to be careful, because it could quite easily be faux free will. As I noted above a fully deterministic system could have a random element included to make it's actions appear as though it has free will when in fact it does not. Again if we believe this to be the case then again no punishment or rehabilitative action will effect either the deterministic rules or the random element.

If however we do believe in free will we have a very difficult hurdle to cross, namely that of how we can have free will whilst our very base component parts can not. We can not claim either complexity or random input because with the best will in the world complexity cannot be described as anything other than deterministic, and random to unpredictable by definition, there has to be something else.

We could look at the extreme sensitivity to inputs that gives us one form of chaos or the behaviour around cusps. That is we know from previous experience that a pencil when stood upright on it's point will when released fall over. We just do not know when or in which direction.

Even when we combine these we still do not arrive at a satisfactory answer for free will.

We can keep going with these arguments but at the end of the day we do not come up with a satisfactory reason for free will.

Which is a problem because without one you can not make an argument to say that either punishment or rehabilitation will have any effect.

---

**James235** • [January 26, 2015 6:18 PM](#)

If I stab someone, can I say 'It wasn't me who stabbed them, it was the knife'?

The artists knew exactly what they were doing when they programmed a bshopper to target the black market.

Let's assume the ridiculous here: a motion-sensor camera was secretly placed in the ground. It is outside a kindergarten, and takes photos up young girls' skirts as they walk by.

Is this valid and not a criminal act as a 'person' didn't take the photo?

Does a person or a camera take a photo?

Is it a question of intention?

The artists knowingly broke the law - albeit in a rather smug/intellectual manner. They should be charged accordingly.

It seems slightly hypocritical to support the actions of the artists. Yes, they may be 'getting one up on the system', but they are still knowingly breaking the law.

If I were walking down the high street and my son was stabbed in the face, I'd be very thankful of any available CCTV footage of the incident.

Or should a criminal be congratulated on 'getting away with it'?

---

**uh, Mike** • **January 26, 2015 7:03 PM**

I have a large dog that bites.

---

**Buck** • **January 26, 2015 10:36 PM**

@James235

It seems slightly hypocritical to support the actions of the artists. Yes, they may be 'getting one up on the system', but they are still knowingly breaking the law.

It's kind of funny that you say so... As part of an art project, I happen to have recently read [18 U.S. Code § 1801 - Video Voyeurism Prevention Act of 2004](#) ([.com](#), [.edu](#), [.org](#) alternatives, in no particular order)

(b) In this section---

(5) The term "under circumstances in which that individual has a reasonable expectation of privacy" means---

(A) circumstances in which a reasonable person would believe that he or she could disrobe in privacy, without being concerned that an image of a private area of the individual was being captured; or

In light of Snowden's '[revelations](#),' it is now much more questionable as to whether or not a 'reasonable person' may expect privacy within the confines of their own home if a laptop/smartphone/webcam present. Yes, the law is law as has been written and is practiced through case law, yet the definition of such laws are clearly dependent on the norms of the time - coupled with the personal interpretations of our judges! So, just who are the real artists here again..?

---

**brian r • [January 26, 2015 11:18 PM](#)**

Lets turn Digital Rights Management to work in the favor of the public, and build it into programs so that it's possible to disable any program that exceeds its rights, and derivative programs too. When a self-driving car has an accident that is it's fault, then all the cars running that program lose their license and are not allowed to drive.

Then the corporation that created it might be more fiscally motivated to prevent such accidents. They would have to rehabilitate the program and must convince a panel of experts that the program can no longer exceed its rights.

---

**Jarda • [January 27, 2015 8:07 PM](#)**

I think the decision where lies the responsibility is quite clear. If a machine of some artists buys illegal items, the artist is responsible and should be punished according to the law.

If a fleet of autonomous military drones wipes out an entire town of civilians, including the hospital and kindergarten, the army isn't responsible, neither the politician who signed the papers. On the contrary, it will be found, that the cousin of the grandchild of the cousin of the grandfather of the wife of the town's chief, which lives merely 500 miles away from the town, once has thrown a stone at the American army, which constitutes a sufficient proof that the entire town was a training camp of Al-Quaida.

---

**vas pup • [January 30, 2015 9:04 AM](#)**

Directly related to the subject <http://www.bbc.com/news/31047780>

Bill Gates on AI danger. My solution is to plant into future AI/robots dog's psychology. i.e.

unconditional love to creator.

---

**Dirk Praet** • [January 30, 2015 1:06 PM](#)

@ vas pup

*My solution is to plant into future AI/robots dog's psychology. i.e. unconditional love to creator.*

Programming a really difficult concept as love into an AI may prove a tad ambitious, but Asimov's Three Laws would be a good start

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

In later works where robots had taken responsibility for government of whole planets and human civilisations, he also added the "zeroth law" :

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

For those who for whatever reason have never read "I, Robot", please do and consider all nine stories were written between 1940 and 1950. The Will Smith movie with the same name has very little to do with them. While you're at it, add some Aldous Huxley, John Wyndham, Arthur C. Clarke and Robert A. Heinlein.

---

**Buck** • [February 2, 2015 11:57 PM](#)

[@myself](#)

Well, it would appear that [former CIA and NSA director, Michael Hayden, fancies himself as quite the artist](#):

In a speech at Washington and Lee University, Michael Hayden, a former head of both the CIA and NSA, opined on signals intelligence under the Constitution, arguing that what the 4th Amendment forbids changed after September 11, 2001. He noted that "unreasonable search and seizure," is prohibited under the Constitution, but cast it as a living document, with "reasonableness" determined by "the totality of circumstances in which we find ourselves in history."

He explained that as the NSA's leader, tactics he found unreasonable on

September 10, 2001 struck him as reasonable the next day, after roughly 3,000 were killed. "I actually started to do different things," he said. "And I didn't need to ask 'mother, may I' from the Congress or the president or anyone else. It was within my charter, but in terms of the mature judgment about what's reasonable and what's not reasonable, the death of 3,000 countrymen kind of took me in a direction over here, perfectly within my authority, but a different place than the one in which I was located before the attacks took place. So if we're going to draw this line I think we have to understand that it's kind of a movable feast here."

Although, it really is a shame that he has seemingly never read the Constitution in full... The words 'probable cause' and 'separation of powers' immediately spring to mind for some reason...

:-\

---

[Subscribe to comments on this entry](#)

## Leave a comment

[Login](#)

**Name (required):**

**E-mail Address:**

**URL:**

☐ Remember personal info?

**Fill in the blank: the name of this blog is Schneier on \_\_\_\_\_ (required):**

**Comments:**

**Allowed HTML:** <a href="URL"> • <em> <cite> <i> • <strong> <b> • <sub> <sup> • <ul> <ol> <li> • <blockquote> <pre>

Submit

---

[← Defending Against Liar Buyer Fraud](#)

[Friday Squid Blogging: Giggling Squid  
Restaurant →](#)

Schneier on Security is a personal website. Opinions expressed are not necessarily those of [Co3 Systems, Inc.](#).