Demetri Kofinas:	00:00	Today's episode of Hidden Forces is made possible by listeners like you. For more information about this week's episode, or for easy access to related programming, visit our website at hiddenforces.io and subscribe to our free email list. If you listen to this show on Your Apple podcast app, remember, you can give us a review. Each review helps more people find the show and join our amazing community. And with that, please enjoy this week's episode.
Demetri Kofinas:	00:32	Since the turn of the 20th century, algorithms have assumed the power previously associated with pontiffs or the divine rite of kings. In an instance of late 20th century lore, the great chess champion Gary Kasparov, reflecting upon his historic loss to IBM's Deep Blue, described the algorithm that defeated him in less than 20 moves as, "having suddenly played like a god for one moment."
Demetri Kofinas:	01:04	Kasparov's experience, that of having been unnerved by the intelligence and obstinate posturing of an otherwise lifeless machine, has not remained confined to the narrow dimensions of his chessboard.
Demetri Kofinas:	01:18	In the 20 years since his loss, increasingly intelligent algorithms seem to be overtaking our world and making humanity obsolete in the process. But in the age of the algorithm, there are those who believe that humanity's place has never been more important. That we should stop seeing machines as objective masters, and start treating them as we would any other source of power, questioning their decisions, scrutinizing their motives and holding them accountable for their mistakes.
Demetri Kofinas:	01:54	This week, on Hidden Forces, Hannah Fry: Power, Complacency and Humanity in the Age of Algorithms.
Demetri Kofinas:	02:21	Hannah Fry, welcome to Hidden Forces.
Hannah Fry:	02:22	Thank you for having me.
Demetri Kofinas:	02:23	It's so wonderful having you in the United States.
Hannah Fry:	02:25	Yeah, it's nice to be here. A bit rainy for me today.
Demetri Kofinas:	02:28	It's rainier than it is in the UK.
Hannah Fry:	02:30	Yeah.
Demetri Kofinas:	02:30	There, it drizzles.

Hannah Fry:	02:32	There, it's just gray sky that reaches down and touches the ground, whereas here it actually, you get wet here.
Demetri Kofinas:	02:37	This is not normal though. It's also UN week and the traffic's insane. When did you get here?
Hannah Fry:	02:42	On Sunday.
Demetri Kofinas:	02:42	Sunday. And you leave?
Hannah Fry:	02:44	On Friday. I'm a little confused about which day's which at the moment. I'm a bit-
Demetri Kofinas:	02:48	Today is, I think today is Tuesday.
Hannah Fry:	02:49	Yeah.
Demetri Kofinas:	02:51	Well, it's wonderful having you here. I told you I read your book and for our video viewers, this is the US copy. This is what the US copy looks. I told you I was not sure I have a picture of what the ... this is a picture from the run down ... this is what the UK copy looks like.
Demetri Kofinas:	03:05	I couldn't decide which one I liked more, but I think I was going with the American version. I think you said you like the UK version more.
Hannah Fry:	03:11	Yeah, it's close. But yeah, I think I slightly prefer the British one. See the thing is actually, it was my friend who made the British design, so I'm emotionally invested in it.
Demetri Kofinas:	03:19	Oh really? I like it. Like I told you, we both agree it's the nerdy, geekier one.
Hannah Fry:	03:23	Maximum nerd.
Demetri Kofinas:	03:24	So the name of the book is, Hello World. The subtitle is, Being Human in and Age of Algorithms.
Demetri Kofinas:	03:30	I told you one of the things I really like about the book is that it takes an even keeled perspective. Oftentimes you find that people take one or the other. And I think because it's hard to do. Also, it sells better to ... particularly...
Hannah Fry:	03:42	Don't tell me that now.

Demetri Kofinas:	03:44	...the doom and gloom perspective, but I like this because it does actually walk that line and I think part of that's captured in your title. So tell us why you chose the title, Hello World?
Hannah Fry:	03:55	Yeah, I think you're right. I think there is a lot of nuance in these discussions. I love nuance in these big questions. I think it's really important to capture. And the reason why I chose the title, Hello World, there's a couple of reasons really.
Hannah Fry:	04:06	The first is that anybody who's ever learned to program is going to be familiar already with this phrase. It's almost like a rite of passage. If you learn how to code, the very first thing that you do is you program the machine to flash out the words, 'hello world' onto the screen.
Hannah Fry:	04:21	And the reason for that is a tradition that dates back about the 1970s when a chap called Brian Kernigan wrote it in his C programming book. And he said that he did that because he'd seen this cartoon where a little chick was hatching out of an egg and it was chirping, "hello world," as it was born, and it was just something that stuck in his mind.
Hannah Fry:	04:43	But the thing that I really like about that, really the reason why I chose that for the title of this book is that he was never totally clear about who the chick was in that analogy. Whether it was the human who was learning to program for the first time, learning how to create algorithms for the first time, or whether it was the machine itself that was kind of being awoken.
Hannah Fry:	05:02	And I just like that idea that there's this moment where you can't tell which is which. It's a moment of dialogue really, between human and machine, where both are amateurs and it's like a shared journey of possibilities that you're both embarking on together. And I just thought that was exactly the right sentiment for where we're at now in time, in really the age of algorithms.
Demetri Kofinas:	05:23	I agree. That was what struck me about it as well though, use of the word dialogue, because I think that often times the conversation is about what is technology doing to us? As opposed to, I'm gonna mention a previous guest who I've mentioned before on this show, Tim O'Reilly, for the exact same reason.
Demetri Kofinas:	05:41	He has this great analogy of this symbiotic organism. And I think that's helpful. Do you find that you see the world as on where we are increasingly interacting with machines, and computers,

and algorithms, as opposed to these siloed communities. One is being built here and it's eventually gonna take over the other?

Hannah Fry: 05:58 I'd like to, but I think that's the world that we want to live in. Or certainly the world that's the future I'm hoping for. One where it is the symbiotic relationship, this partnership. But I do think actually, that more and more, we're not seeing that. I think that we're seeing a question, a conversation about human versus machine.

Hannah Fry: 06:15 Machines are coming to take all of our jobs. Machines, the ones with the authority, and the ones with the power. And I think that actually that's a lot about what this book is about is kind of just taking that version of the future and questioning whether that's really what we want to end up with.

Demetri Kofinas: 06:31 How do you think that narrative came to be? Where do you think that comes from? Is that just a general displaced anxiety about change? Or is it very specific to technology? Or is it both?

Hannah Fry: 06:42 It's a really tough question. I think in part, it's because of the way that we are deeply flawed as humans. So I sort of think we've got this really strange relationship with machines because on the one hand, we are really trusting of technology. There are stories of people quite literally driving their cars off the edge of a cliff because their sat map told them to do it.

Hannah Fry: 07:05 But there are other examples of, how many times when you go on sky Scanner do you bother to double check the websites of all of the airlines to make sure that you're being given the cheapest airline. I certainly don't I trust the technology-

Demetri Kofinas: 07:16 I don't do that, but I've heard you should do it.

Hannah Fry: 07:18 Should you?

Demetri Kofinas: 07:18 I've heard there are all sorts of tricks to getting a better airline.

Hannah Fry: 07:21 Oh, that's true. That's true and, yeah. Getting cheaper tickets.

Demetri Kofinas: 07:25 But no. I don't do it. I mean, we're probably too busy to do it.

Hannah Fry: 07:27 Well, it's true, but I also think that across the board really, we're seeing lots of occasions where people are quite happy to just take a bit of a cognitive shortcut, right? Algorithms and computers give you an easy way out sometimes. They give you an easy source of authority.

Hannah Fry: 07:42 But then on the other hand, we also have this real habit of just dismissing technology completely if it ever is shown to make any mistakes. So you've got people shouting at Siri for being stupid, or just throwing away algorithms that have got any kind of flaws in them.

Hannah Fry: 08:00 And I kind of think that that conversation about human versus machine that's so prevalent in the way that we think about them really stems from that. The way that we think of machines as either totally perfect and we wholeheartedly trust them completely, or complete junk that we can never use at all. When in reality, we should be looking at this as something slightly more in the middle.

Demetri Kofinas: 08:20 One of the things that came up when you were talking, is this relationship of master/slave. And I think it's a term in computer science as well, but also, I think that we see computers either as our slaves, or if they're not, then they might be our masters. And that makes me think of the story of Kasparov and Deep Blue.

Demetri Kofinas: 08:39 You start the book, I think it's the first chapter or the second chapter, with this story of this famous chess player, Kasparov, and his famous match against IBM Deep Blue which he lost. I'm gonna give it away to those who don't know. Tell us about that story. And I should say also, the book is full of this.

Demetri Kofinas: 08:57 One of the other things that I liked about the book is, it's all of these anecdotes and it personalizes things and through those stories, you're able to kind of connect with the particular issues, whatever it is that we're talking about. Tell us that story and why you chose to put it in the book.

Hannah Fry: 09:12 Yeah, it was a tough call, that one actually because I know that the story of Kasparov and Deep Blue is one that's been retold thousands and thousands of times. I mean there's no chess match in history that's been pored over more than that one.

Hannah Fry: 09:23 But for me, I have a slightly different take on what happened during that match than the normal stories tend to cover, because I think for me it wasn't so much a story about how remarkable the achievement of IBM's Deep Blue had been, despite the fact that this was a remarkable achievement. Technically it was a remarkable achievement.

Hannah Fry: 09:42 I think for me, this was the story about how Kasparov let his human flaws lose him the match. So Kasparov, you have to

understand, this man is unbelievably good, like unbelievably good. He was so intimidating.

Hannah Fry: 09:58 I spoke to several chess grand masters when I was researching this book. And they described him as being like this tornado when he entered the room. Everyone was kind of pinned to the sides watching this great man walk through. And he had these tricks where he would totally out psyche his opponents before he'd even really sat down at the board.

Hannah Fry: 10:15 And one of the things that he used to do is he would take off his watch as he would start to play and he would put it on the table next to him. And he would kind of toy with his opponent for a bit. Let them think that they were having some kind of game, and then when he was bored of toying with them, he would pick up his watch and he return it to his wrist. And that was the signal that everyone in the room recognized, which is that Kasparov is done with you now. He was finished toying with you.

Demetri Kofinas: 10:38 He's like, "I'm bored."

Hannah Fry: 10:39 I'm bored. You need to resign this game now because otherwise I'm just gonna take you down. Either way, defeat is coming. But when it's against a machine, you just can't do that. You can't use any of those kind of tactics. And yet the machine could use those tactics on Kasparov.

Hannah Fry: 10:54 So one of the lesser known things about this machine was that, it was programmed to be able to play chess of course, but it was also programmed to try and psyche Kasparov out. So it was programmed, when it came up with an answer, to sometimes it would just sit on that answer for a little while and just count down the timer, just let the timer-

Demetri Kofinas: 11:13 It would crank the wheels and the gears.

Hannah Fry: 11:13 That's right. Steam coming off it, and the idea behind that is, it wanted Kasparov to start thinking about what was going on inside the machine. To start second guessing the machine. So Kasparov got caught in this thing of, "Oh well I must have pulled it into this position where it's struggling with this calculation. I've done something that it's finding really tricky." When in reality, the machine was just sitting back and letting the timer tick down.

Hannah Fry:	11:37	And Kasparov, in his own book, has written about his emotional status during that process. Because you know ultimately, I think that widely, in the chess community, they regard that Kasparov was still better than the machine. At that moment in time, Deep Blue was not a better chess player than Kasparov, but it Kasparov's attitude towards the machine that was the thing that lost him the match.
Hannah Fry:	11:59	And I think that's something that we see across the board, well outside of the world of chess. It's the way that we react to technology and allow technology to control us that really is where the questions lie.
Demetri Kofinas:	12:12	And that brings up the fear component. The fact that there's this relationship of fear oftentimes with computers and machines and the case of Kasparov is a classic example.
Hannah Fry:	12:21	Yeah. Sometimes it's fear. Sometimes it's faith. Sometimes it's ... We've got this really, we're just not good at having these kind of clear, and logical, rational relationships with machines.
Demetri Kofinas:	12:31	So one of the things you also do in the book and it's called, Being Human in the Age of Algorithms, you focus in on algorithms and I think most people don't actually know what an algorithm is, so I think it would be helpful for you to define that, and tell us what that is and maybe break down the different types of algorithms and how you do that in the book.
Hannah Fry:	12:46	Yeah, completely. So I think it's completely understandable really, why people don't totally know what the word algorithm means. Because, it doesn't really mean very much, to be honest. It's just this really broad umbrella term that doesn't convey much meaning.
Hannah Fry:	13:02	But very, very simply, all an algorithm is, is a series of instructions. It's something that takes you from some kind of input via some logical steps, through to some kind of output. So, in theory, a cake recipe counts as an algorithm, Your input is the ingredients and your output is the cake at the end. And your logical steps is the recipe itself.
Hannah Fry:	13:23	But the way that people tend to use the word, is the instructions that you give to a computer. And there are lots of examples of algorithms that we're really familiar with. Google Search, for example is an algorithm. Or Facebook's newsfeed, or Amazon's recommendation engine. These are all kind of algorithms that we're just interacting with all the time.

Hannah Fry:	13:41	But we're now seeing algorithms in all kinds of different places. So they're in our cars. They're in our courtrooms, in our hospitals, in our schools. They're there kind of behind the scenes making just little decisions about who gets to do what. And about how we're living our lives.
Demetri Kofinas:	13:59	You break down four different categories for how these are used. We have prioritization. Is that what's used in Google, I think? In Google search engines? Thing like that. What are those for, can you tell us?
Hannah Fry:	14:10	Oh crikey. You're testing me now.
Demetri Kofinas:	14:11	I've got them for you here. Prioritization, classification, association and filtering. I think the reason I bring those up is because, I wonder for people, if it would be helpful to think about some of the places where they encounter these different algorithms and how they differ. And I also think at some point, I'd like to talk about the difference between deterministic algorithms and these machine learning algorithms that use inductive reasoning.
Hannah Fry:	14:36	Yeah absolutely. These different categories, there's not like this official categorization of algorithms, so I think there are a lot of computer scientists who would argue with the way that I've chosen to group these here.
Demetri Kofinas:	14:47	Is that right?
Hannah Fry:	14:47	But you know, there isn't an official list, so you've got to pick the line somewhere, but I think they do a good job of explaining the different types of things that algorithms can do. For example, prioritization algorithms. All these do is they give you a list that's in order of priority. Google search, that's exactly what it's doing. You put in a search term, and it prioritizes the entire internet for you and spits it back at you in an ordered list.
Hannah Fry:	15:12	But you could also have a prioritization algorithm that was used by the police to tell them where in a city they should target their forces on a particular evening, if they want to crack down on the most crime.
Demetri Kofinas:	15:23	You had a great chapter on that, by the way. What was that decay, what was it called, distance decay or buffer zone or something like that? Interesting things-

Hannah Fry:	15:31	Oh yeah. God, you have properly read this book. I'm so impressed.
Demetri Kofinas:	15:33	I enjoyed it. I wasn't aware of that. Anyway, I interrupted you. Please continue. But we'll get into that when we deal with crime and algorithms.
Hannah Fry:	15:40	So you also have things like the classification algorithms. Okay, so classification algorithms will be like, if you are female in your early 30s and you've been with your partner for a long time on Facebook. On Facebook, they can tell that. Then they would classify you as someone who is probably going to get engaged, or probably gonna want a baby and serve you up some adverts for diamond rings or pregnancy tests. Those kind of things.
Hannah Fry:	16:04	But classification algorithms could also be classifying people in courtrooms as to whether they're high risk of going on to offend or not in the future. So these are the very, very simple basics. The type of tasks that algorithms can do.
Demetri Kofinas:	16:16	So let's just take classification, for example. You can see with that one right there, how beneficial it could be, but at the same time, the dangers of it because it's stereotyping on steroids, right? You can get kind of nuts with the stereotyping and segmenting people, and of course, we've seen that in so many areas.
Demetri Kofinas:	16:33	One obvious case is on Facebook, where people live in the eco-chambers of very particular types of news because they've been stereotyped as hardcore wing Republican. Well, good luck changing because you're getting dished increasingly radical content. Or vice-versa on the left.
Demetri Kofinas:	16:49	Anyway, I mean the power of algorithms is strong and you make that point also with this great example of Robert Moses, the great New York landscape architect. I don't know what the official term for that-
Hannah Fry:	16:59	Is great the best word to use? I guess he did a lot of work.
Demetri Kofinas:	17:02	Great as in huge. His impact was enormous.
Hannah Fry:	17:06	Enormous.
Demetri Kofinas:	17:07	Enormous. You gave the example of the racist bridges. Give that example because I think it will be helpful to the audience to understand the omnipotence of these algorithms.

Hannah Fry:	17:15	Yeah, this is something that I think people in the UK haven't heard of at all, but I guess if you're American, you're a lot more aware of the work that Robert Moses did. Robert Moses was an urban planner, a town architect. Is that fair to say? A town architect around the 1930s.
Demetri Kofinas:	17:24	Yeah, an urban planner, town architect. I called him a landscape .... what did I call him? A landscape engineer?
Hannah Fry:	17:34	That works.
Demetri Kofinas:	17:34	I think urban architect probably makes a lot of sense.
Hannah Fry:	17:37	Okay, we'll go for that. In the 1930s, he built a state park in Long Island. Jones Beach, it's called. Beautiful state park. And he was basically a massive racist. He was incredibly racist. And he didn't want anyone who wasn't white to frequent his new beautiful state park.
Hannah Fry:	17:56	So what he did to prevent that from happening because he had no political power. What he did to prevent other types of people from visiting his park was that in the highway that approached this beach, he deliberately built these bridges that hung very, very low over the traffic. So they are extremely low. I think sometimes they only leave about nine feet of clearance from the tarmac below.
Hannah Fry:	18:16	And the idea behind that, of course, was that white and wealthy Americans would be visiting the beach in their private cars and could easily slip underneath these narrowed bridges, or these low hung bridges. But people who were traveling in buses, which they were more likely to be if they were from the poorer black neighborhoods, wouldn't be able to pass under these bridges, and therefore wouldn't be able to go to the park.
Hannah Fry:	18:36	Now, the point that I was really making in using that example, because there's no real algorithms that are involved in that, but it's this idea that you can have objects that aren't attached to humans, these inanimate objects that end up having a kind of clandestine control over people, that end up being oppressive almost, just because of the way that they're designed.
Hannah Fry:	18:58	Sometimes, when they're deliberately designed to oppress people like in the case of those bridges, but sometimes because people are just a little bit thoughtless when they omit being able to include people in the way they're designing things.

Demetri Kofinas:	19:10	And also because algorithms don't come out of thin air. We create them and we imbue them with our biases, and some of those biases, to your point in the book, are necessary. Because if you don't have those biases, then the algorithm doesn't work. And yet, having those biases can create a feedback loop where you have cases like these decision trees in parole boards and criminal hearings.
Demetri Kofinas:	19:33	Judges decide whether to let someone out early or not based on a test which is heavily influenced by the neighborhood in which they live. Inadvertently, their skin color. And give that example. That's a great example of ... what are they called?
Hannah Fry:	19:49	Recidivism algorithms.
Demetri Kofinas:	19:50	There you go.
Hannah Fry:	19:51	Yeah, exactly. So these have actually been around a really long time.
Demetri Kofinas:	19:54	Which I was not aware of. I'll tell you something, that was very educational for me. I wasn't aware of the fact that people use these types of decision trees to make these decisions, and I actually want to mention one more thing that I thought about with that and that feeds into the medicine aspect as well.
Demetri Kofinas:	20:09	I think there's something attractive to people in such professions like the law and medicine, where liability is very high for making bad decisions.
Demetri Kofinas:	20:19	I think they welcome the opportunity to have an algorithm where they can offload responsibility to a computer, even though that may not be the ideal outcome for society, or for the people that they are either seeing as patients, or paroling.
Hannah Fry:	20:31	Yeah, I totally agree with you. I think if you're faced with a tricky decision, I mean I don't know if you ever do this. I don't want to make a decision, can someone else just be in charge? I don't care-
Demetri Kofinas:	20:40	I feel like that all the time. I feel like that all the time and I will also say, increasingly, to the prior point I was making, I also worry about being blamed for things.
Hannah Fry:	20:50	Exactly.
Demetri Kofinas:	20:51	You know what I mean?

Hannah Fry:	20:51	Yeah, it's an easy way to hand over responsibility. I entirely agree with you. I think, especially when you're talking about judges, for example, where there's really, really, difficult, difficult decisions that they have to make about people's freedom, robbing people of their freedom.
Hannah Fry:	21:06	I think that when you have an algorithm there telling you an answer, it's very, very hard not to over-trust that. It's very hard not to take the cognitive short-cut of that very easy sort of authority, because you can rely on it later to say, "Well, the algorithm told me."
Demetri Kofinas:	21:24	That's hugely important in medicine. If you're a radiologist and you miss, let's say somebody's cancer, and they sue you, wouldn't you be better off just saying, "I was just following a computer's algorithm."?
Hannah Fry:	21:33	Yeah, you would. You definitely would. But I think actually in medicine, is one area where I think that they really have understood these pitfalls and these big problems that are on the horizon by having these algorithms telling you what to do, or telling you their opinion.
Hannah Fry:	21:50	And the way that they've done it in medicine is that they have much more of a partnership because ultimately, the things that humans are really bad at, is we're really sloppy. We get tired really easily. We make lots of mistakes. We miss little tumors, that kind of thing. We're really bad at that stuff.
Hannah Fry:	22:07	Algorithms, on the flip side, don't have any of those problems. They never get tired. They're very consistent, they're very precise. But they don't understand context and they don't understand nuance. They make little mistakes. So the idea in medicine that they do is they create this partnership between human and machine.
Hannah Fry:	22:25	So rather than having the machine make the diagnosis and then have the human come in and check whether the machine is correct or not, the machine is just designed to flag areas that it thinks are suspicious and then the human comes along afterwards and actually does the diagnosis, looks through all the areas that the machine will have flagged as suspicious, and saying yes, no, malignant, benign, malignant. It's really hard to say those in a row-
Demetri Kofinas:	22:48	Malignant, benign.

Hannah Fry:	22:49	But by doing that you avoid this problem of, a machine says it's 18 months in jail, what do you want to do? Instead, you're saying the machine thinks that these areas are worthy of your attention, what do you want to do? It's a different way to frame the problem.
Demetri Kofinas:	23:04	It's also coming at it from saying what are machines really great at? And what are people still good at? And how do we work together to get a better outcome. And I think you've agreed, that's a more productive path to a future where people, again to bring concerns about the fear aspect, people have a fear of being replaced, right? By machines.
Demetri Kofinas:	23:23	I mean we saw that during the industrial revolution and we're seeing it in terms of automation. And people are worried that that's gonna take their job, right?
Hannah Fry:	23:31	Yeah, and understandably so, I think. I think that actually the workplace is going to change on the basis of a lot of this stuff. And I think one of the things about the industrial revolution at least, is that we had a lot of ... there was time. There was a very big change but it was a big change over a number of years, whereas this stuff is happening really quickly.
Hannah Fry:	23:50	But I also think that I don't want to go into a hospital where an algorithm is in charge. I want that human to be part of that process. I don't think we're anywhere near the stage where you could hand over this stuff to machines without any kind of human input at all. We're a long way away from that.
Demetri Kofinas:	24:06	How much do you think that is, 'cause that's a common response. It's not just you, right? A lot of people feel that. In fact, I think I saw a video of you. Maybe it was the Ted talk, or maybe it was an interview I heard, but you talked about how, when someone goes in front of a judge, would they rather have a cold piece of steel cranking out some decisions, or do they want a warm beating heart, even if they are worse.
Hannah Fry:	24:30	Someone who can look you in the eye...
Demetri Kofinas:	24:31	Look you in the eyes.
Hannah Fry:	24:31	...as they send you to jail, exactly.
Demetri Kofinas:	24:33	And that brings us back to the Kasparov thing.

Hannah Fry:	24:35	So actually, you know it's funny because I've asked a few different audiences this exact question now and most people say that if they were the one in the dock, they would rather it was a human who was making a decision about their future, that it was a human making that prediction. But I think part of the reason for that is, when humans make mistakes, people imagine that those mistakes will go in their favor.
Hannah Fry:	25:00	If you explain to people that actually, the variation in the type of decisions that a human judge will make are massive, people always imagine that they're going to be one of the luckier ones, right? I kind of think actually, I'm not against algorithms in the court room. I just think they have to be designed in the right way. Because algorithms can make better, more accurate and fairer assessments of people.
Hannah Fry:	25:26	And I think that's the difference when you're thinking about yourself in the courtroom, or whether you're thinking about the whole country's courtrooms altogether. Because I think when you're thinking about the bigger system, you really want stuff to be as fair as it possibly can be, and as unbiased as it possibly can be, but I actually think that algorithms do have a role to play in that.
Demetri Kofinas:	25:44	I have another theory about that and it just came to me. I think that there's some sense of appeal. There's room for negotiation. You tell me something. You said you're a recent mother. You said about a year old, your daughter. You're not yet at that stage, but we all have met children that love negotiating, right? So we learn very early that if we don't like results, let's talk about it. Let's talk about it. Where can I meet you half way?
Hannah Fry:	26:10	Where's the wiggle room?
Demetri Kofinas:	26:10	There's no wiggle room with a computer.
Hannah Fry:	26:12	No, you're right.
Demetri Kofinas:	26:12	You know what I mean?
Hannah Fry:	26:13	No, you're right.
Demetri Kofinas:	26:14	I think there's some sense in that, even though it's illogical. It's irrational. There isn't, once a judge hands down his or her decision that's it. You don't like it? Too bad. There's appeal in the court of law, but the same thing would happen with the computer.

Demetri Kofinas:	26:27	I want to switch a little bit and talk about something that I think the public has a fascination with. And I want to ask you first, why you think it is? And that is autonomous driving, right? Because there's so many fascinating technologies in the pipeline. And some of them, I think, could capture the imagination of the public like 3D printing.
Demetri Kofinas:	26:44	But for some reason, autonomous driving vehicles have just become the talk of the town. It is such a big topic in the media. People want to know more about it and they want to understand ... and it's happened so quickly. It came out of nowhere.
Demetri Kofinas:	26:58	You brought up the great story of DARPA's race in the Mojave desert of 2004 with those cars. Man, were those cars ... one was falling over the other. It was a giant ... I think the winner went seven miles before flipping over, you know what I mean?
Hannah Fry:	27:12	Hilarious, I know.
Demetri Kofinas:	27:13	And now, here we are 14 years later, and cars are driving themselves getting on and off highways. How did this happen, first of all? And that's kind of a way of getting into a conversation about these machine learning algorithms and how computers learn using neural nets, but also, I'm curious. Why do you think that people have such a fascination with autonomous cars?
Hannah Fry:	27:35	Well, I think we've always had a fascination with autonomous ... or we've had them for a very long time anyway. Surprisingly, autonomous cars is not a new fantasy. This is something that actually belongs to the era of jet packs and tinfoil hats-
Demetri Kofinas:	27:47	The Jetsons.
Hannah Fry:	27:48	Yeah, totally.
Demetri Kofinas:	27:48	Did you guys have the Jetsons in the UK?
Hannah Fry:	27:49	Yes, completely.
Demetri Kofinas:	27:50	You did? Well, did they speak with a British accent?
Hannah Fry:	27:52	No, I don't think so. I think it was your...
Demetri Kofinas:	27:57	We had maids. You might have had butlers. You know, this is a slight interjection. This is a funny moment, but I'll say I had this

guest on. Henry Timms, I think it was Henry Timms and Jeremy Hymens, I think It was Henry Timms though, who had the story. Henry Timms, I believe it was. He worked for a very rich woman in England and he's British and she apparently had only retired butlers from the queen. So he has fascinating stories about what it's like to actually be serving.

Demetri Kofinas: 28:26 And he has this one great example where he's on a private plane and he had forgotten to buy a gift for his wife, and literally the moment he realized it, he didn't say anything. Literally, someone just ... a butler came and said, "Will this do sir?" He goes, "I took the liberty," And he had already bought the gift, knowing that he had forgotten it, without even telling him. He saw it in his face.

Hannah Fry: 28:48 That's absolutely incredible.

Demetri Kofinas: 28:48 Remarkable. But I interrupted you. Jetsons. We were talking about the Jetsons and about people always having this vision of a future with autonomous cars. You brought up the world's fair, I think, in the book?

Hannah Fry: 28:57 Yeah, that's true.

Demetri Kofinas: 28:58 And, tell that story and we'll keep going forward...

Hannah Fry: 29:00 Yeah, this is back in the 1950s, where people are really seriously thinking about autonomous cars. There's been plenty, plenty, plenty of attempts to do it. There's one particularly, well I find it quite amusing, attempt in the UK, of where they built this car.

Hannah Fry: 29:14 I think it was a Citroën. I can't remember what car it was. A Citroën, anyway. They built it so that it would follow a bit of copper piping that's built into the motorway. They dug up one of the freeways, laid this copper piping, and proved that it could work, but then just sort of ran out of steam. So now still, under one of the motorways in Britain, there's this piping which was supposed to be this 1960s autonomous vehicle.

Demetri Kofinas: 29:36 How far does it run?

Hannah Fry: 29:37 Oh, it's like rubbish. It's like three miles or something.

Demetri Kofinas: 29:44 That's interesting.

Hannah Fry: 29:44 A really short demonstration.

Hannah Fry:	29:46	But yeah, I think it's understandable that we have this dream. The idea of getting out of your house and sitting in a car and then telling it where you want to go and just chilling in the back while you're driven. You're chauffeured.
Demetri Kofinas:	30:00	While you check the stock price of your...
Hannah Fry:	30:01	It's the butler for everyone. It's the chauffeur for everyone. Real luxury for every individual. Of course it's a dream. I think it makes perfect sense that it is. And I think that now, after DARPA, seeing this is much more of a reality, it's understandable that people are getting excited about it.
Demetri Kofinas:	30:17	And it's also been part of the popular culture, like KIT, Knightrider. Everyone wants their own car. There was a 1970s horror movie where the car was ... I forget the name, but it was Christine, I think that was the name of the movie, where the car actually tried to kill everyone. It would just drive around trying to kill people.
Hannah Fry:	30:31	Some similarities to modern driverless cars.
Demetri Kofinas:	30:33	Exactly, yes. Exactly. But I think autonomous driving cars are cool for a number of reasons. One is because the decision making is so complicated.
Demetri Kofinas:	30:45	The trolley problem is something you bring up in the book, and you're dealing with incomplete information and you're dealing with nuanced decision, right? And there's another great quote I mentioned in the book you had. I think I have it here. Here it is. I'm gonna quote it. You know what I'm gonna say, "You're going to mug them right off. They're going to stop and you're just going to nip round..."
Hannah Fry:	31:04	This is the results from a survey, and it's very English people and it's very amusing to hear you say.
Demetri Kofinas:	31:10	I can't say it with your accent. I didn't even want to try.
Hannah Fry:	31:14	Mug them right off...
Demetri Kofinas:	31:14	But then you're just gonna nip round. But let's use that as a segue to talk about what it is that I was intending, because I do want to talk about the challenges of making autonomous driving real. And how that is a subset for a larger conversation about bringing machine learning algorithms into the world.

Hannah Fry:	31:33	So that exact quote that you mentioned there, it was in answer to a question of, how members of the public imagined it might look if we had autonomous cars on our roads. And the thing is, is that if you're creating a driverless car, the number one rule, the number one thing that that driverless car had to be able to do is to avoid a collision where possible.
Hannah Fry:	31:54	That's the number one rule. So what that means is this. If you step out in front of a driverless car in a situation where it doesn't need to run you over, it's gonna have to stop for you. And suddenly, if you have a series of cars that you can know will reliably stop if you bully them, then that kind of changes the rules of the road, so you can mug them off.
Demetri Kofinas:	32:19	So mug them off means what?
Hannah Fry:	32:21	It basically means bully them. That's what it means.
Demetri Kofinas:	32:23	All right. Got it.
Hannah Fry:	32:25	And nip round them means that you can overtake them.
Demetri Kofinas:	32:29	It's like you're in full control of the car as a pedestrian. It's a reverse. You don't have to be afraid. You know you're in control because you know how it thinks and it has to protect you at all costs.
Hannah Fry:	32:39	Exactly. And suddenly you have people who, traditionally, didn't have very much power on the roads like cyclists, for example who suddenly become much more powerful.
Demetri Kofinas:	32:47	You haven't spent enough time in New York. They've got a lot more power than you might think.
Hannah Fry:	32:50	Oh really?
Demetri Kofinas:	32:50	Yeah, yeah. They're pretty aggressive here. But you bring up the trolley problem also, which deals with some of this stuff. And I love this problem because once you begin to think about it, it shows you the enormity of the calculations required, and the value-set.
Demetri Kofinas:	33:06	So this is, what is it called in machine learning in AI? The values, when you program values into a machine?
Hannah Fry:	33:15	As in ethical values?

Demetri Kofinas:	33:16	Utility functions.
Hannah Fry:	33:17	Yeah, okay, yeah.
Demetri Kofinas:	33:18	A utility ... something like that. If there's one driver in the car and it's a 50 year old man and there's a child on the street, if you swerve to avoid the child, there's a 75% chance you might die, the computer calculates, but if it hits the child, the child will 100% die. Do you hit the child or do you swerve?
Hannah Fry:	33:38	Completely. The thing about the trolley problem, which is what you're describing. That ethical dilemma of what do you do in this situation, when it was originally devised, it was devised to demonstrate how, in some situations, there isn't a right answer.
Hannah Fry:	33:52	Where, actually it's really hard to know what the morally correct thing to do is. And the thing about driverless cars, the reason why so many people talk about this trolley problem, is that you can now imagine a situation in which you actually have to decide, regardless of whether you want to or not, whether there's a right answer or not, you have to decide what you're going to do in a situation.
Hannah Fry:	34:15	If you speak to people who are experts in the field, the people who build these driverless cars, they actually tend to be quite dismissive of the trolley problem. They kind of say, "Well, it's such a rare occasion, such an unusual thing that it won't ever happen, so we don't need to worry about it. We're certainly not going to program this in."
Demetri Kofinas:	34:31	What do they mean that it's a rare problem?
Hannah Fry:	34:32	That the idea of having to choose who to kill.
Demetri Kofinas:	34:35	Why is that rare? Why do they say that that's rare? It seems like it would be fair to common enough. Not an every day occurrence but...
Hannah Fry:	34:42	I have changed my mind on this actually because I used to kind of agree with them. I remember having a conversation with Professor Paul Newman. Who runs Oxbotica, which is a driverless car program in the UK.
Hannah Fry:	34:55	And he said to me, "Has it ever happened to you? Have you ever had to choose who to kill while driving home?" Well no I haven't. And he was like, "Do you know anyone whose had to choose someone they had to kill?" And I said, "No, well I guess I

		don't." "have you ever heard of anyone whose had to choose who to kill?" I said, "Well, no. I suppose not." He's like, "Well there you go. We're getting sidetracked by this problem, when actually there are much more interesting questions to be had about driverless cars."
Hannah Fry:	35:16	And I kind of bought into that argument, but then the trolley problem happened to my husband. Exactly the trolley problem happened to my husband.
Demetri Kofinas:	35:24	Wow. Can you tell that story?
Hannah Fry:	35:26	He was driving down the road and then coming around the corner was someone who was trying to escape from the police. And they cut into the wrong lane, so they were driving directly towards him and there was traffic going in the other direction.
Hannah Fry:	35:40	He essentially had the choice of staying where he was and have a head on collision with person who was escaping the police. Swerve to the right and go head on to the traffic in the other lane, or swerve to the left where there was a cyclist in that lane who he was gonna hit if he swerved left.
Hannah Fry:	35:55	And basically, he had our daughter in the car so he decided to save himself essentially, went for the cyclist.
Demetri Kofinas:	36:02	Oh wow.
Hannah Fry:	36:02	But thankfully, thankfully the cyclist saw the situation unfold and went up on the pavement to get out of his way, knowing that he was gonna cut him out.
Demetri Kofinas:	36:09	Wow.
Hannah Fry:	36:09	So I kind of think that actually, maybe this isn't as rare as some people imagine.
Demetri Kofinas:	36:16	Wow. Thank you for sharing that story. It highlights a few things. I thought about that when I was reading the book. We don't know how we'll react until we're in that moment.
Hannah Fry:	36:25	Of course, of course. But then, however you react, you can justify it. Because you can say, let's imagine that something really bad happened in that situation and someone did end up getting hurt, you take my husband up to the docks in the courtroom afterwards and ask him to explain his decision, and

		he can say, "Well, it was a snap decision. I didn't know. I just had to make a decision in the moment."
Hannah Fry:	36:45	You can't hold a person accountable for what they do in that high pressure situation, whereas if there was an algorithm driving the car, you can go through, look at line by line by line of code. and see exactly what it was that gave rise to the decisions that it made. And that just changes the rules slightly.
Demetri Kofinas:	37:04	But doesn't the decision it makes also depend on the values we assign to the participants in this scenario?
Hannah Fry:	37:10	Only if you decide to program it that way. Only if you decide to address the trolley problem head on and say, "How old is that person? How old is that person? What are the chances of this person surviving?" And you don't necessarily have to do that.
Demetri Kofinas:	37:23	But wouldn't the car's decision, to your point, you said and autonomous car's number one job is not to hit anyone, right?
Hannah Fry:	37:29	Mm-hmm.
Demetri Kofinas:	37:30	The way I think a lot of people mistakenly think about how these cars will be programmed is that their number one objective will be not to endanger the driver or the passengers in the car. But we haven't come to a determination about how that's gonna play out because you could imagine a scenario where, let's say every car's autonomous, or 50% of the cars, or whatever it is. That the sum total of everyone involved in driving these cars or being in them, would be better off with an algorithm that functions in a different way, that prioritizes life differently than just the person in the car.
Demetri Kofinas:	38:05	So it's straightforward if it's the people in the car, but if it's not that way, how do you make those decisions?
Hannah Fry:	38:12	Well, I know. It's really a difficult problem and not one with an easy answer. But I think that ultimately, there was a survey done, in fact going back to that comment that you mentioned earlier. I think it was the same study which asked people, "If you were making the rules for driverless cars, how would you want the rules to go?" And basically everyone agrees that the cars should save the most people possible. That's how it should work.
Hannah Fry:	38:36	But then if you ask people, "Would you buy a car that would save the most people as possible if that sometimes meant that

		you, yourself would be killed?" They're like, "Well no. Of course not. Of course I wouldn't."
Hannah Fry:	38:46	I think it comes back, we see this in the courtroom as well. There's a difference between your own personal incentives and your own personal motivations and what you want for everybody.
Hannah Fry:	38:56	What's right for the individual and what's right for the group are not always necessarily the same thing.
Demetri Kofinas:	39:00	I'm stuck on this thing about values because I thought about it before. Have you read Nick Bostrom's book, or any of his work on-
Hannah Fry:	39:05	Super Intentions.
Demetri Kofinas:	39:06	Yeah.
Hannah Fry:	39:06	Yeah.
Demetri Kofinas:	39:07	During the time that I was reading that book, I don't know where I read this other historical account of how the insurance industry was born, and how actuary tables were developed.
Demetri Kofinas:	39:17	But it was the first time that, as populations were moving, I think it was in London, as populations were moving into the city, the need for life insurance and things like that developed, and there needed to be some way to value people.
Demetri Kofinas:	39:30	And that was the first time that people had done that and I think this is like a further evolution of that which is to try and assign values to human life and to people and to instances.
Demetri Kofinas:	39:40	Another thing that you've talked about in the book, which we've mentioned it, I believe only in episode ... for sure in Episode Two with Jim Rickard's base theorem, but perhaps in one or two other episodes.
Demetri Kofinas:	39:51	It's something that I've wanted to do a full episode on. I'd like for you to at least take this opportunity to talk about it a little bit. Why you speak about it in the book and its relevance to machine learning and to algorithms.
Hannah Fry:	40:03	Yeah, it's just one of the most powerful equations that has ever existed really. And at its heart, it's a really simple idea. I think

		the way that I describe it in my book. I mean, I love frivolous example, so I use a quite silly example to illustrate.
Hannah Fry:	40:21	But the idea is that you update your knowledge based on the information you have. You throw away the idea that something is 199% absolutely true or false and you start talking about your certainty in something. You start talking about your belief almost in something.
Hannah Fry:	40:35	So it's like, let's imagine we were in a restaurant somewhere and you said to me, "Oh, I think that's Lady Gaga over there sitting at that other table."
Hannah Fry:	40:45	And before I turn around to have a look, I will have some idea as to how much I believe that your hypothesis that that's gonna be Lady Gaga, right? So maybe I'll take into account where we are in the world, how posh the restaurant is, how likely I imagine it would be that Lady gaga would be in this restaurant.
Hannah Fry:	41:01	But then as I look at the woman you're pointing out, there's gonna be different things I take into account, so maybe whether she's got bodyguards with her, maybe she's got blonde hair, all of these different kind of things. And every new piece of information that comes in, I update my belief accordingly.
Hannah Fry:	41:14	It can either go up or down based on all of these new pieces of information. And then perhaps if I notice that she's wearing a meat dress or something, which is something that you don't tend to find in people who aren't Lady Gaga, that might be enough to tip me over my threshold to conclude that I'm happy enough to believe that that's Lady gaga.
Hannah Fry:	41:32	And that's essentially the idea, the very basics of the idea behind base theorem, is that it's a systematic way to update your belief in something. It stops you having to be absolutely certain about things, and starts dealing with theories and uncertainties.
Hannah Fry:	41:50	And in the case of driverless cars, this stuff is really, really important because if you think about the blue dot on your GPS when your phone is telling you where you are. What that blue dot is signifying is there's some uncertainty in your position.
Hannah Fry:	42:06	And when you've got a mobile phone and walking around with your mobile phone, it doesn't particularly matter whether you're here or three meters to the left.

Hannah Fry:	42:13	But when you're in a driverless car, or in a car full stop, that three meters could be the difference between being in your lane and being in a lane where you're driving straight into traffic.
Hannah Fry:	42:23	So you need to start being able to deal with messy data that has errors in it with uncertainties. And that's` why you need your car to have a belief about where it is, not just a measurement about where it is.
Demetri Kofinas:	42:37	It's the difference, another way to talk about it is the difference between deductive and inductive reasoning, right?
Hannah Fry:	42:42	Yeah.
Demetri Kofinas:	42:43	Instead of saying there's certainty, all I have to do is solve this equation and I'm gonna have a clear answer. It's dealing with complexity, a world that's so complex that you can't come to a definitive solution to the problem and you need to do your best. And the other thing you mentioned is this updating and learning.
Demetri Kofinas:	42:59	And these machine algorithms are learning and they're learning in a way that produces answers which are better than anything else we've seen before and yet, because of how they're learning, we don't know exactly why they came to the answer they did, right?
Hannah Fry:	43:13	Often that is the case, yeah. Often that's the case. So the big analogy that I like to use is the difference between traditional types of programming and this machine learning and artificial intelligence that we're seeing. The newer stuff that's being implemented much more now, is like when you try to train a dog how to sit.
Hannah Fry:	43:31	If you train a dog, you don't lay out a list of instructions for the dog, right? You don't say, I want you to sit. And what I'd like you to do is to move this muscle and then that muscle, and then move your tail down and whatever.
Hannah Fry:	43:41	You don't do that. All you do to the dog is you clearly communicate with it the objective of what you'd like it to do. So maybe you push down its bum and say the word sit.
Hannah Fry:	43:52	And then you repeat that process over and over and you reward it whenever it gets something right and you ignore it whenever

it does something wrong. And over time, if you repeat that enough times, the dog works out what you want it to do.

Demetri Kofinas: 44:04 It seems like a miracle whenever it happens.

Hannah Fry: 44:05 It is a miracle when my dog manages to do it. It is an actual miracle.

Hannah Fry: 44:11 But all of the steps in between, the way that that dog decides to manage to achieve the objective, it makes all those decisions itself. It works out the process itself. And that's really what's happening with machine learning algorithms is that you clearly communicate with the computer what your objective is. You reward it when it gets it right, and you let it work out all the steps in between by itself.

Demetri Kofinas: 44:33 Now that we're on autonomous vehicles, one thing that I don't want to forget, and I want to circle back and discuss is this notion of complacency because that's something that, as these computers become better and better and better, and these algorithms become better and better and better at doing thing we increasingly suck at, we get even worse at it, right?

Hannah Fry: 44:52 Yeah.

Demetri Kofinas: 44:52 I had lived in Italy out of college and it was a good example where it was a brand new city. I was based out of Florence. And at first, I didn't know where the hell I was, but within a few weeks or a month, I would be in neighborhoods where, previously, I didn't understand where this place was. But in my head, I had a map. I knew where I was and I could get home at any point in time.

Demetri Kofinas: 45:14 New York's different story but, you lead me somewhere else now, I will get lost. And another example. I will bring this one up because it's a great story. I learned to drive stick shift and I learned to drive cars with no bells and whistles, certainly no parking alarms and things like that.

Demetri Kofinas: 45:30 And I hadn't driven a stick in a few years and I was in Europe and rented a car and realized I had gotten used to parking with a camera that created a CGI graphic of my car. I was like, "How am I gonna park?" I didn't realize how much my senses had deteriorated.

Hannah Fry: 45:48 Yeah, it's true.

Demetri Kofinas:	45:50	What do we do about this, that these machines are increasingly doing way better jobs than we can do. We're offloading these menial tasks to them so we can become more efficient at doing other things, I hope. That's the theory, supposedly, right?
Demetri Kofinas:	46:04	But we're losing the ability to do things in certain cases that we will need to do, in certain moments, in a crisis, we will need to be able to do these things, and we can't.
Hannah Fry:	46:13	Yeah, you're exactly right and I think that the case of the autonomous cars is a really important illustration at this point because ultimately the way that the driverless cart is set up at the moment is, the car is in charge. The car's doing all the driving, but you are there to monitor the car, and to step in the case of emergency.
Hannah Fry:	46:29	And that's sounds absolutely fine except that, for one thing, imagine that we get to a stage where a driverless car makes a mistake maybe once every ten thousand miles. Suddenly, you are gonna lose your skills. The moment you have to step in and you have to take over, the worst, most difficult moment of all, is when you have to be absolutely on it, when you need your highest skill in driving. And that's incredibly difficult to do.
Hannah Fry:	46:55	But even aside the loss of skill stuff, there's also the issue of, we're not very good at tirelessly paying attention and looking out for something that's going wrong.
Hannah Fry:	47:06	The really tragic example of this is the Uber crash where the human monitor, as they call it-
Demetri Kofinas:	47:13	The human monitor?
Hannah Fry:	47:15	Yeah.
Demetri Kofinas:	47:15	That's the driver?
Hannah Fry:	47:15	That's the official driver, yeah.
Demetri Kofinas:	47:17	I got it, right.
Hannah Fry:	47:18	...was looking down at their lap the moment when a person walked out in front of the car, and the car failed to break.
Demetri Kofinas:	47:24	Was the person killed?
Hannah Fry:	47:25	Yeah.

Demetri Kofinas: 47:25 Really?

Hannah Fry: 47:25 Yeah, it's really awful. It's really awful and this is a completely innocent bystander. It's just someone who was crossing the road. Really awful stuff. But ultimately, we've known about this for quite a long time. Back since the 1980s where people were beginning to automate nuclear power plants, and there were all these essays that came out at the time saying, "Hang on a second. If you want the majority of the grunt work in the hands of the machine, and you're just expecting this human to sit that and stare at the monitor and wait for something to go wrong, and then jump in and be extremely skilled at exactly that moment, that is just a recipe for disaster."

Hannah Fry: 48:02 And I do think that we have to worry a tiny bit about this with the modern autonomous cars, while we're still at this stage where people have to step in in emergency.

Demetri Kofinas: 48:11 Well, there's a really fascinating book on near nuclear catastrophes. I don't know if you're familiar with it. I forget the name of the author. But it lists a number of them and one of them was Petrov, which you talked about in the book, which this guy was ... What is it where you work at the area where it's the central command for Nuclear-

Hannah Fry: 48:31 The missile detection system.

Demetri Kofinas: 48:32 Missile detection. And he saw a missile, supposedly being launched from the United States coming inbound to Russia, and he had to make a decision about whether to launch a counter strike, right?

Hannah Fry: 48:42 Yeah, so his job was just to monitor this automatic system and if any missiles were detected, to pick up the phone immediately and inform his superiors who would then launch counter strike and nuclear war would take over the planet.

Hannah Fry: 48:55 And this happened. It was, as we're recording this in fact, it was exactly 35 years ago.

Demetri Kofinas: 48:59 Oh really?

Hannah Fry: 48:59 Yeah, it's tomorrow I think is the anniversary. And he was monitoring the system in the middle of the night. This alarm went off. Missiles were detected, and he's orders were very, very clear. Immediately make that phone call. But something gave him pause.

Hannah Fry:	49:13	He was just a bit concerned about the accuracy of this system because for one thing, it has said it only detected five missiles and he was like, hang on a second. If America are launching nuclear war, why are they only sending five missiles?
Demetri Kofinas:	49:27	Inductive reasoning. Base theorem.
Hannah Fry:	49:27	Exactly right. Why wouldn't they?-
Demetri Kofinas:	49:28	I have to update my assumptions.
Hannah Fry:	49:33	Completely, so he kind of froze in his chair and was like, "I'm not sure what to do." Because he knew that if he made that call, there'd be no one else to stop this from happening. He knew that him making that call meant nuclear war. But then at the same time, every second that he waited, was really eating into Russia's chance to launch a counter strike, if this was indeed true, it spelled the end for the USSR if indeed he was wrong.
Hannah Fry:	49:58	He just had no way of knowing for sure one way or the other whether this was right or wrong. And in the end, thank goodness for all of us, he decided to just assume the machine had made a mistake and held off. And it was 23 minutes later I think that the ground detectors failed to detect these missiles landing and he knew that he had been right.
Hannah Fry:	50:15	And actually, in the end, what had happened was the sunlight bouncing off of clouds. The algorithm had mistaken those for missiles. So thank goodness. He literally averted nuclear war by trusting his own instincts over an algorithm.
Demetri Kofinas:	50:27	That's one of my concerns. Not specifically the case of nuclear war but the issue of complacency. It's high on my list because I experience it in my own life and I can see that it's not frivolous concern, that it happens. My skills have deteriorated in all sorts of areas, and we have these hair trigger systems like our nuclear systems which, I think they still do have human intermediaries. We're not at the sky net moment yet, but I think that's the bigger concern.
Demetri Kofinas:	50:54	The popular media has made it one where the machines are gonna take over. But the more realistic scenarios are that a mistake happens.
Hannah Fry:	51:02	Of course. And mistakes do happen and to be honest, I kind of think mistakes are completely inevitable. We can worry about minimizing mistakes and I think that's an important thing to do,

		but I also think that we have to be realistic, that you can't ever eliminate all mistakes.
Hannah Fry:	51:18	I mean I struggle, algorithm or not, in a kind of computerized world or otherwise, I really struggle to think of any system in the world which is perfectly fair, perfectly unbiased, and never ever makes mistakes. I just don't think that that happens.
Hannah Fry:	51:32	So my big argument that I try and make in this book is that maybe we're thinking about this in all the wrong way. Maybe we should just accept the fact that these machines are gonna have flaws, are gonna make mistakes and we should just design them to wear their uncertainty very proudly and design them that we can appeal them when they do make mistakes.
Demetri Kofinas:	51:54	So that's the importance of having the code be available for people to be able to see what's going on. There's not being black boxes. 'Cause if you have these algorithms that have so much control over our lives, but that we can't actually ... something comes out.
Demetri Kofinas:	52:06	You have the great example of the voters in Idaho, where was it?
Hannah Fry:	52:10	Yeah.
Demetri Kofinas:	52:10	What do you think of that? And then this could bring us into a conversation about reform regulation, the law. Are things being done better in Europe with GDPR what is it called, or in the United States.
Demetri Kofinas:	52:22	Talk to us a little bit about this, about how to navigate this better.
Hannah Fry:	52:26	Yeah so, the example that you're talking about in Idaho was, it was a story about there was 16 disabled residents who got a pretty terrible news actually. The Department of Health and Welfare in Idaho had decided that they were going to introduce this new budget tool that was going to work out how much state support each of these people were entitled to.
Hannah Fry:	52:45	These were people with very severe disabilities, who qualified for residential care, but who had chosen to be cared for in their own homes instead.
Hannah Fry:	52:52	So the money that they were receiving was really important to their independence and for their care in general. And so they

each went into the state department and sat down and talked through the system and the budget tool came up with an overview of how much money they were entitled to.

Hannah Fry: 53:07 And it was just really strange because some people ended up with more money than they had in previous years. So it certainly wasn't a political decision to just slash funding. But then other people ended up with way less than they'd had. Tens of thousands of dollars down on what they needed to have.

Hannah Fry: 53:22 And they tried to appeal. They tried to ask why on earth these decisions had been made, but no one in the state department would help them so they ended up filing a class action lawsuit and getting the budget tool to be turned over for scrutiny.

Hannah Fry: 53:36 And when it was, the thing that I find quite extraordinary about this is that this algorithm that was holding so much power over them, this algorithm that so many people in the state department really put their faith in, it wasn't like some clever AI. It wasn't some really beautifully crafted mathematical model, it was an Excel spreadsheet. And actually quite a crappy Excel spreadsheet.

Hannah Fry: 53:57 So there was all these errors in the formulas. The data was ...loads of bugs in this thing, and yet just because it was wrapped up as this spancy algorithm-

Demetri Kofinas: 54:09 It's a partial machine.

Hannah Fry: 54:10 Yeah, it was given all this power and authority. I just think we've been living in the wild west. When you can collect data on whoever you want, about whatever you want. You can create an algorithm that makes any decision that you want that impacts anyone you want and there's no one to stop you from doing that.

Hannah Fry: 54:27 I think we used to be in a world where you could just put any colored liquid in a glass bottle and sell and say it was medicine and make a fortune.

Demetri Kofinas: 54:34 Yeah, snake oil salesman in the 1800s.

Hannah Fry: 54:37 Yeah, but we stopped doing that because, for one thing, it's morally repugnant, but also, it harms people. And we haven't, at the moment, got that back stuff of something like the FDA who are just checking that the things that are impacting this much

on people's lives have benefits that outweigh the harms that they're imposing on people.

Hannah Fry: 54:58 And I really think that that's what we need. We need some kind of regulatory body that approves algorithms and says, "Yes, this one is good enough to use on the public."

Demetri Kofinas: 55:09 A tall order. I mean right now; the market incentivizes and rewards people to work in the private sector. And the private sector, of course, generates the profits that then lobby the government to prevent the regulations, right?

Hannah Fry: 55:21 Yeah, it's true. It's kind of the best and worst of capitalism all in one. On the one hand, you have driving innovation forward, creating these incredible, remarkable new technologies at speeds that are just lightning compared to what any state funded groups could manage.

Hannah Fry: 55:37 But then, at the same time, without regulation and control, there's no real way to get the positive benefit out of those fully for society, or be sure that you're really seeing the positive benefits for those for society.

Demetri Kofinas: 55:49 It's interesting you bring up the industrial revolution, right? Because that was one place where we saw similar dynamics. One particular part of the economy, the industrial parts, the steel, the oil, the railroads, they generated so much profit and had so much power, that I think there's something similar happening today with respect to technology companies.

Demetri Kofinas: 56:11 They've amassed so much power. In terms of data, Facebook has, and it's not just Facebook, but Facebook's kind of the poster child, have had at it with our data and there has been an exchange, but it's not clear what the value of that exchange is, and would people be willing to give up all of this privacy for what they're getting in return, which is to be able to send a few messages to their friends,

Demetri Kofinas: 56:34 But in other cases, some really powerful and great things. You bring up 23 and Me, and the potential. And I mentioned to you before we started, I had Eric Schott on the program from the Icahn Institute and the challenge that he has and others have, which is getting the data you need to do the analytics to actually be able to save people.

Demetri Kofinas: 56:53 To make leaps and bounds in cancer research and other areas. There's so much potential that comes from being able to

		harness the power of these algorithms, but at the same time, there are risks and there are dangers and how to navigate that is not an easy task.
Hannah Fry:	57:08	No, it's not an easy task. It's not an easy task at all. But I kind of think that actually we can't be, as individuals, we can't be complacent about this stuff. I think we have to recognize the traits that we are making. I think we have to be slightly more switched on about what is going on around us. We kind of let it all happen.
Hannah Fry:	57:27	Data has been the new oil, right?
Demetri Kofinas:	57:29	Indeed.
Hannah Fry:	57:29	And we've been living in the Wild West and I know that's mixing up my oil and gold analogies.
Demetri Kofinas:	57:33	No, it's good although both ... Well, I guess they weren't all in the West, but there was a lot of oil exploration in the West.
Hannah Fry:	57:38	Okay, fine. All right.
Demetri Kofinas:	57:39	Texas was a big part of that. And the Midwest maybe.
Hannah Fry:	57:42	Okay, okay, okay. We've been living in the Midwest.
Demetri Kofinas:	57:44	You're excused. You're coming in from the Transatlantic. Hannah, it was wonderful having you on. I want you to stick around because I want to ask you some other questions for our more nerdy viewers that we're gonna make available to our Patreon, but I appreciate you coming on the program.
Hannah Fry:	57:58	Well, thank you very much for having me.
Demetri Kofinas:	57:59	Thank you.
Demetri Kofinas:	58:01	And that was my episode with Hannah Fry. I want to thank Hannah for being on my program. Today's episode of Hidden Forces was recorded at Edge Studio in New York City.
Demetri Kofinas:	58:12	For more information about today's episode, or if you want easy access to related programming, visit our website at: hiddenforces.io and subscribe to our free email list.

Demetri Kofinas:	58:24	If you're a regular listener of this show, take a moment to review us on Apple podcasts. Each review helps more people find the show and join our amazing community.
Demetri Kofinas:	58:36	Today's episode was produced by me, and edited by Stylianos Nicolaou. For more episodes, you can check out our website at, hiddenforces.io. Join the conversation at Facebook, twitter and Instagram at @HiddenForcesdPod or send me an email.
Demetri Kofinas:	58:57	As always, thanks for listening. We'll see you next week.