

You're listening to Imaginary Worlds – a show about how we create them and why we suspend our disbelief. I'm Eric Molinsky.

And this is Joana Bryson.

JOANNA: In America, I'm professor JB, in the UK I'm called Dr. Joanna Bryson because you're only professor when professor so I'm a reader which is a really cool title because nobody knows what it means.

She does the same work on both sides of the pond – teaching and designing artificial intelligence.

In the 1990s, she was working with scientists at MIT who believed that robots should have human characteristics like big eyes, because that will encourage people to interact with the robots. But the robot they were working on was no C3PO.

JOANNA: In fact it was just a torso, it didn't even have arms. But it had a head and it had two cameras, in fact four cameras where the eyes should be.

And the software wasn't working. But:

JOANNA: People coming by would say, it would be unethical to unplug that, and I was like but it's not unplugged, but if you did it plug it in, it would be unethical to unplug it, and I was like, well it doesn't work. And I was mystified because this is a piece of scrap – well, nice scrap, but people immediately thought they owed ethical obligation to it.

So it worked. People attributed human traits to this robot because it looked like a head and a torso. But she thought this is working too well. People are imagining this heap of metal and wires had consciousness and unplugging it would be the same as killing it.

So she wrote an academic paper about her experience called "Just an Artifact," but it didn't get any traction. She tried to publish it again with a different title – nah. Then, she got one more shot at publishing it.

JOANNA: And so I thought okay, this is my chance to really get this message across, but I thought okay this is going to be the third time lucky, I'm going to call it robots should be slaves.

Robots Should Be Slaves. If there ever was click-bait for an academic paper – this was it.

JOANNA: Now, I regret this now because I've realized there is nothing you can do to try and to break the idea that slaves are humans you own because of this horrible legacy we have, but people took it to be it's okay for them to be human but we should treat them badly and it's like no, no, no.

Some of her fiercest critics were science fiction fans.

JOANNA: I love it when people tell me, I have a PHD in Artificial Intelligence, and people tell me you don't understand AI because you didn't watch AI the movie!

CLIP: AI

And when she reminded them that most of the robots in that movie were abused or abandoned.

JOANNA: They say oh no, I don't want to own them; these are going to be our children. And whatever, it's a natural drive but it's a drive that we have to over come, and people are right when they argue to me you don't understand, because I'm not a parent, you're not a parent you don't understand pass along mantle,

This really frustrates to her. The human body may be an effective biological machine, but it's a clumsy, inefficient design for a metal machine. She wishes sci-fi depicted more robots doing what they do best – impossibly hard tasks over and over again, really efficiently – which should free us up to have more leisure time -- which I guess is another way of saying robots should be slaves – even though that still sounds wrong.

JOANNA: I mean I can entertain possibility, and in some of my papers I say, we should look at this, could it be that we could build something we owe obligation to?

But she thinks that's a mental exercise at best. She worries sci-if is leading us astray, filling our heads with fantasies of self-conscious robots that we want to adopt, liberate or kill before they kill us.

How much does the past haunt our vision of the future?

That is just after the break.

>> UNDERWRITING

The first modern robot story was a play from Czechoslovakia in 1920 called R.U.R. -- Rossum's Universal Robots.

GH: In Czech, robot does mean slave. That is quite literal.

Gregory Hampton teaches literature at Howard University.

GH: One of my mantras about literature is that literature is a direct reflection of people who produce it. And so if you want to learn about people their aesthetic, their value system, just read literature, they're going to be put things in that they may not be conscious of.

So when he reads RUR, he sees European-style Marxism. And then when he looks at American robots stories, he sees Uncle Tom's Cabin and Nat Turner's Rebellion.

GH: I teach narrative in seven moments. There's the I was born section, the intro to robot, there's the description of suffering in slave and robot, there's the description of the family that brought robot into household, there's this moment where robot or slave becomes enlightened, and after there's this moment robot or slave wants to be free, then there's a plot to escape or in some instances destroy the master.

How does this play out? Take the movie Bicentennial Man, based on a book by Isaac Asimov. The robot Andrew is played by Robin Williams. shortly after he meets his new family:

CLIP: BICENTENNIAL MAN, JUMP OUT WINDOW

GH: The film uses that as comic relief but horrific.

CLIP: BICENTENNIAL MAN

GH: When Andrew or Robin Williams comes through front door, father has house meeting, and says to the girls Andrew is piece of property I'm aware of that, but for the purposes of making household stable and happy, I'm going to demand you treat him as though he were person, and that's where real problems started, slave household in

antebellum America, this blurring of the line, crossing of the line consistently, you say they're not human but depend on their humanity.

CLIP: BICENTENNIAL MAN

Eventually, Andrew becomes self-aware, buys his freedom, and tries to gain human rights.

GH: He is changing himself having surgeries, having replacements, having skin grafts done, replacing mechanical organs to the point where he looks human, and he goes to court, and he goes to the human Supreme Court or something.

CLIP: BICENTENNIAL MAN

GH: We've transcended antebellum America and now we're into Civil Rights, this time period what won't the African American do to be included?

The first time I made this connection, I was listening to public radio story about slavery and the Civil War. They decided to follow that serious subject with a lighter piece about a new fangled "cleaning robot" called the Roomba, which does all your vacuuming for you. And I thought – has anyone else noticed these parallels?

That's how I found Joanna and Gregory, and came across articles by Erik Sofge. He covers robotics for Popular Mechanics. He wanted to cover this beat so he could dispel myths people get from sci-fi. But he's also a fan who gets sucked into these stories. Like when he watches Star Wars, his heart goes out to the droids because they're bought and sold callously.

ES: Most of the characters, especially early movies are so awful to the droids, will just threaten to destroy them for anything and C3PO has clearly been affected by this to a huge degree, he's paranoid that he's going to be deactivated all the time.

CLIP: C3PO

And that kind of uneasiness over how we treat robots, leads to the other slave narrative -- an uprising against the master.

ES: its basically invaded research even, there's a robotics papers where they talk about the Terminator scenario.

The Terminator Scenario was all over the news recently when the robotics company Boston Dynamics put out a video where an engineer kept poking a barrel-chested humanoid robot with a stick while it was trying to lift boxes, to show how adept the robot was at getting back up again. The video went viral because it looked like the robot was being tormented while it was doing manual labor, as if it kept doing its job out of dignity or fear, even though it was just following its programming.

That video even creeped out Joanna Bryson, the author of “Robots Should Be Slaves.”

JOANNA: My conscious intellectual thought was being impressed they got that much further, but my gut response same as yours I’m sure.

Shortly afterward, Boston Dynamics was dropped by their parent company, Google. There were business reasons but leaked internal memos showed the Terminator jokes didn’t help – like here’s Trevor Noah on the Daily Show.

CLIP: DAILY SHOW

Gregory Hampton wasn’t laughing.

GH: Does that say something about what this society thinks about marginalized bodies? I think it is. I think we can only imagine marginalized in particular way and the most handy reference is the slave. For a lot of engineers proudly involved in humanoid robots, these images are what’s leading them and not conscious of what does that entails, they’re not conscious of what does master slave relationship, owner servant and how we treat these things, what does that do to our psyche?

EM: But some robotocists have argued that robots are never going to be self aware, they’re very useful, let’s stop being afraid of this and embrace they’re our servants because they don’t have a consciousness.

GH: Yeah, and this is the same argument pro-slavery people in antebellum American, they’re not human, they’re not intelligent.

EM: But they're talking about things that literally are not human and they design robots, they are not human, they do not have the consciences that a human being will have.

GH: I guess I want to suggest even if that's the case, even if consciousness is not developed. The AI not as advanced as some would say doesn't take away from the idea that there are going to be side effects. If you treat a thing like a slave, you're going to develop certain symptoms. If you embark upon this relationship in particular way in the way you've done in the past with humans, there's going to be a side effect you had when participating in slavery.

In other words, it doesn't matter whether robots develop feelings. The question is how it changes us, and what we consider acceptable behavior.

Erik Sofge says if you want to see the real future of robots and people interacting, look at the other project Google is heavily invested in: self-driving cars.

ES: When there is coverage in advances there aren't this talk of uprising and what these things could do to us. It's interesting and I think a lot of it has to do with the fact that there isn't anything anthropomorphic about a robot car, and it's about the car and about people despising the business of commuting, but the car as a chore.

The programming of these self-driving cars does reflect science fiction, or at least the three laws of robotics that run throughout Isaac Asimov's stories, like Bicentennial Man.

CLIP: ASIMOV

But on the road, a robot car won't have clear moral choices.

ES: If a robot has to choose who to kill, a driver or another driver or a bystander, who should it kill? If there's school bus, you hitting a streetlight or hitting a school bus, what should it do?

Erik worries that living alongside robot servants, even hyper intelligent cars, could bring out the worst in us.

ES: I'm positive the human drivers will treat those cars like crap, because they know they can push the around, they can cut them off because they know that robot car will do everything it can to be completely safe.

Interestingly, Joanna Bryson decided to rewrite Isaac Asimov's laws of robotics because in sci-fi, we keep imagining that the robots is making the moral choice.

JOANNA: The idea that the robot is the moral agent is broken.

Her five principles of robotics reiterate that people manufacture robots: therefore the burden of responsibly is on us. We shouldn't worry about treating them badly -- we should worry about why we want to treat them badly. We shouldn't worry about them trying to kill us either. Their programming will always reflect whatever human desires we had to build them.

Well, that's it for this week's episode, thanks for listening. Special thanks to Joanna Bryson, Erik Sofge and Gregory Hampton, with music by Alexis Cuadrado and Sono Sanctus.

Imaginary Worlds is part of the Panoply network. You can like the show on Facebook. I tweet at emolinsky. The show's website is imaginary worlds podcast dot org.