EXAMINING ETHICS
TRANSCRIPT — JEFF SEBO: THE MORAL CIRCLE
(01.31.25)

I'm Alex Richardson, and this is Examining Ethics, the show designed
to bring insights from the cutting edge of moral philosophy and ethics
education to the rest of us. We've probably all spent some time
thinking about what we owe to each other as members of a human moral
community. But we don't quite as often think carefully about whether
we may have obligations to other kinds of beings as well. Our guest
today works at the cutting edge of philosophical debates on moral
standing and was recently recognized as part of Vox Media's Future
Perfect 50, a list of innovators, thinkers, and change-makers in 2024.
In his new book, The Moral Circle, Who Matters, What Matters, and
Why?, he argues that we ought to be prepared to expand or, in his
words, explode our circle of moral consideration to septillions of
beings, not only including animals, but a potentially vast class of
other kinds of beings as well. Jeff Sebo, welcome to the show.

Thanks so much for having me.

Of course. Glad to have you.

Let's start with you just telling us a little bit about your work and
background generally.

So I am currently associate professor of environmental studies and
affiliated professor of bioethics, medical ethics, philosophy, and law
at New York University. I also direct the Center For Environmental and
Animal Protection, which is a research center that conducts and
supports research about important issues at the intersection of
environmental and animal protection. And I direct the center for mind
ethics and policy, which is a research and outreach center that
conducts and supports different kinds of work about the nature and
intrinsic value of nonhuman minds, especially insects, other
invertebrates, and AI systems. And I co direct the wild animal welfare
program, which, along with my colleague Becca Franks, I should say is
a research and outreach program that examines what wild animals are
like and how humans and wild animals interact and how we can improve
our interactions with them. I have always had an interest in how we
can apply these philosophical arguments, these philosophical tools to
questions about the nature and intrinsic value of nonhuman minds and
what we owe nonhuman beings. And a lot of my own work, especially in
recent years, has focused on the moral status of different nonhumans,
and you and I will discuss that a lot today. That includes animals, it
includes parts of us, it includes collective devotes, it includes now
new technology like AI systems, but then also the relationship between
different pressing global concerns like animal welfare and global
health and the environment and AI safety and and other other issues

that are are quite concerning to many people.

Your new book, just freshly off the press, is all about the evolving concept of moral standing. Can you say a little bit more about how this concept evolved over time and why it's now particularly important for the average person to think carefully about it?

Yeah. Absolutely. And maybe we can start by defining our terms like any good philosopher. When I talk about moral standing, like many philosophers, what I mean is a certain kind of intrinsic value. A a lot of beings, of course, have a lot of different kinds of value. For example, our possessions or environments, they have a kind of instrumental value. That means they matter for us because of the benefits that they provide for us. But I am talking about a a further kind of value that some beings have when they matter intrinsically, they matter for their own sakes. They matter for or to themselves. So, you know, my dog, for example, he does have moral standing because my dog has consciousness and emotionality and bonds of care and interdependence, and so it matters to him what happens to him. And for that reason, I have responsibilities to him, not just responsibilities about him to other humans. And so the question here is how has our conception of moral standing, of who matters for their own sake in that kind of way? How has that conception evolved over time? And, of course, the answer varies from region to region, from community to community. But broadly speaking in the west, we can say that the history of thinking about the moral circle has been a history of moral circle expansion. Generally starting from a very exclusionary and hierarchical understanding of who matters and how much they matter. And then gradually, sometimes reluctantly expanding it out to include more beings and give them more weight in our deliberations. And, of course, this is still a work in progress even within our own species, even with our fellow humans. Some humans are sometimes excluded entirely or at least not given as much weight, as much concern as they should be given. But it especially tends to neglect or exclude or discount the interests of various nonhumans. Even nonhumans were very similar to us, like, other primates or other mammals. And the interesting recent development in our thinking about the moral circle is that after a long period of mostly entirely excluding animals and then grudgingly including a few of them, we're now contemplating including many more and and reckoning with what that might mean. And that is what this book addresses.

The book begins with a really fascinating point about a kind of forward looking uncertainty. And this can be both an empirical uncertainty and a moral uncertainty about what kinds of things and what kinds of developments we may see in the future when it comes to what kind of beings deserve moral consideration. So given this framing associated with uncertainty, how should we approach decision making in the space?

Yeah. I think the uncertainty framing is very important because when people ask questions about moral standing or the moral circle, they tend to ask them in yes or no, all or nothing ways. What does it take to matter for your own sake and who has what it takes? But the reality is we have ongoing substantial disagreement and uncertainty about both the values, what it takes to matter and the facts. Who has that? And my claim is that these issues are so important and so difficult and so contested that it would be arrogant; it would be hubristic for us to just assume that our own current views are correct and make life or death decisions about how to treat other beings based on that presumption. Especially when we look into the even recent past and think about how arrogantly wrong our predecessors were or even our own past selves were about these exact same questions. It it would be surprising if we were the 1st generation in the history of the world to get this exactly right. And so for me, when we make these life or death decisions about how to treat other beings, we should do that not by asking, do they matter, but rather by asking, might they matter? Given the best information and the best arguments currently available to us, is there a realistic non negligible chance that they matter? That in the fullness of time, it will turn out that they really do possess some capacity or some relationship that gives them an intrinsic value. And if there is a realistic non negligible chance that they matter, then we should, in the spirit of caution and humility, give them at least a little bit of consideration now when making decisions that affect them.

I do really like the framing here, particularly when you associate it with our sort of ability, if not propensity, to cause harm. So you frame it as a kind of risk perspective. Right? We ought to make descriptions of moral standing from a perspective of risk with respect to what we might discover in the future. But philosophers kind of like their moral certitude. Right? I kind of wonder what kind of pushback you get when you share this view with people, particularly people who are also thinking about moral status, but maybe do it in this sort of, like, you cross the line and you possess the property sort of way.

Yeah. Absolutely. And I do get different kinds of pushback. And to be clear, I think a lot of the pushback is completely reasonable and it merits further conversation. So I can give you two examples of forms of pushback that I think are reasonable and that we should take seriously. There are, of course, silly forms of pushback too, but we can set those aside. So one form of pushback that I think is reasonable concerned the nature of moral uncertainty. As you and I both noted a few minutes ago, there are 2 kinds of uncertainty that matter here. One is uncertainty about what it takes to have moral standing. Do you need to be sentient, able to experience happiness and suffering? Do you need to be conscious, able to have subjective experiences at all? Do you need to have agency, be able to set and pursue goals? People disagree about the answer to those moral questions. And then, of course, on the scientific side, we have

disagreement on uncertainty about which beings have those features. Can AI systems, for example, suffer? Can they have conscious experiences? Can they set and pursue their own goals? So some people push back experiences? Can they set and pursue their own goals? So some people push back by saying, hey. Look. I get that we need to think in terms of risk when we have disagreement and uncertainty about science. Like, if I am not sure how bad a pandemic will be or how bad climate change would be, maybe I should on the side of caution to mitigate harm. But people feel unsure whether we should apply those same tools to disagreement and uncertainty about ethics in part because people disagree about what ethics even is in the first place. Like, some people think ethics is trying to describe objective truths in the same way that science is trying to describe objective truths. On that view, maybe it does make sense to think in terms of risk. But other people think ethics is an expression of our own individual most deeply held beliefs and values. And so they see less of a connection with how we think about risk in the case of science. But I think that thinking about disagreement and uncertainty in terms of risk makes sense either way. Even if I am what philosophers call an anti realist, and I think that ethics is nothing more than the expression of my own most deeply held beliefs and values. I can still be unsure about what my own most deeply held beliefs and values are or will reveal themselves to be in the fullness of time as I reckon with these issues more. And so thinking about ethics in terms of risk is kind of like placing a bet on what my own beliefs and values about ethics will be as I get more information and resolve contradictions in my own values. So no matter what, I think it makes sense. Another form of pushback that I think makes sense is we have to be careful about how we think about risk because it might seem to make sense that we should give weight to non negligible risks and we should factor them into our decision making. But then if we take that seriously and we follow it to its logical conclusion, we realize that we should all of a sudden be considering many possible impacts for many possible beings. I mean, it might be for example, that animals, AI systems, plants, fungi, microscopic organisms, all have a non negligible chance of mattering if we really take that question seriously. And then it might also be that our actions and policies have a non negligible chance of affecting them in all kinds of ways. And so if we really take my argument seriously, we might, uh-oh, be signing ourselves up for extending at least some, at least minimal consideration to our impacts on this overwhelmingly vast number of beings. I think the right response to that is, yeah, we actually should do that, and it is okay. It will not it will not be totally overwhelming, totally destabilizing, totally disorienting. We can develop tools for doing that in a sustainable way, and everything will be okay.

I wanna zoom in on one of what I think is the kind of most interesting normative sticking points in the discussion over moral status, and that's the the consciousness bit, particularly like the idea of consciousness in different species. You mentioned that our views about

what kind of animals, for instance, count as sentient have sort of expanded over time and suggest that this is possible or or likely probable that our understanding of consciousness is gonna do something similar. So what has changed in our understanding of, I guess, in particular, animal consciousness regarding things that we may think of as non obvious? Right? Things like invertebrates and other kinds of beings that we may be accustomed to thinking of as more kind of marginal cases.

So one general trend is that animal consciousness has fortunately become a credible legitimate area of scientific study again. There was a period during the 20th century when it was simply not regarded as a credible legitimate topic for science because it was thought to be this private kind of experience that can never be scientifically studied or scientifically confirmed. So it should be relegated to the domain of philosophy or religion instead of the domain of science. But in recent years, by which I mean over the past 30 or 40 years or so, there has been a resurgence of interest in studying animal consciousness, making progress on animal consciousness within various relevant scientific fields. Now a second trend is that in the wake of that development, there has emerged a new framework for examining evidence of consciousness in nonhuman animals despite the difficulty of that task and despite ongoing disagreement and uncertainty about the fundamental nature of consciousness. And this is called the marker method or the indicator method. And in short, the way this works is we can introspect in our own experience and make a distinction between conscious and unconscious processing. I can look inward and make a distinction between, for example, the felt experience of pain and the bodily reactions that I have that are completely unfelt. Right? And then I can look at behavioral and anatomical markers or indicators or correlates of the conscious processing, the conscious experiences. And then I can extrapolate from that and look for similar behavioral and anatomical markers in other animals. For example, do they have some of the same brain structures that are relevant to the human experience of pain? Do they nurse their own wounds? Do they respond to analgesics or antidepressants in the same ways that we do? Do they make behavioral trade offs between the avoidance of pain and the pursuit of other valuable goals? Now, none of that is proof of the ability to experience pain or have conscious experiences in general, but it can all count as evidence. It can all be a data point. When we find that, it can tick the probability up a little bit. And then the 3rd development, which is related to the second, is that we have realized what you and I discussed a moment ago, that we are not presently able to decisively prove that animals are or are not conscious. And yet we do urgently need to make policy decisions that affect animals. And so the bar for considering animal welfare, for example, should not be certainty or proof about consciousness, but should rather be a reasonable, realistic chance of consciousness given the evidence available to us. So those are the general trends that have fortunately put animal consciousness back on the map and have conspired to make

people more open minded to extending animal welfare policies to invertebrates like insects where we might not be confident that they are conscious, but we do think there is a reasonable realistic chance given the evidence available to us.

You've written here and elsewhere about the idea of parts and wholes possessing moral standing. This is something that I find really interesting, thinking about the distinction, for instance, to run with the kind of insect example between, say, individual insects and insect colonies, for instance. So how might this view change the picture a little bit in our thinking about moral consideration and moral status?

Yeah. This topic is so interesting, and and I wanna work on it more in the future. I hope other people are interested in it too. We tend to think of individual humans or individual organisms as the basic units of moral analysis. The beings that have reasons and duties and rights that that that are what we discuss when we discuss what we owe each other. But when you look at individuals, you realize, wow, our parts can be pretty significant in various ways too. And groups of us can be pretty significant in various ways too. And there are different ways to carve that up. So for example, I wrote my dissertation about what our different selves or side, their personalities might owe each other. And many philosophers have written about what our past and future selves might owe each other. So that is one way to think about the moral significance of parts. Another way to think about it is there might actually be regions of our brains or bodies that have their own separate but linked consciousnesses. So, for example, consider the octopus. An octopus can be fairly described as having 9 interconnected brain like structures. They have a kind of central command and control center brain, and then they have other smaller connected brains within each arm. And then they exhibit some integrated, but also some fragmented behavior. They sometimes seem to be operating as one, and they sometimes seem to be operating as a bunch of disconnected arms. So you might think of them as like a nation with states. You know? And then, of course, there are groups and insect colonies are a great example of this. Flocks of birds, of course, plants and fungi could be described as large individual organisms or as networked groups of organisms. But either way, you can see integrated behavior at the collective scale and at the individual organism scale. And that raises the question, could there be moral significance within us and could there be moral significance beyond us? Could there be individuals to whom I have moral responsibilities inside of me and across us? You know, could you and I together form a single individual that matters and and deserves respect and compassion in addition to continuing to be separate individuals who who matter and deserve respect and compassion? So so I think those are really interesting questions, and I think those questions are gonna matter more in the future as new technologies come online and enable our minds to get connected in more intimate ways than they can currently be.

Good. So I wanna stay with that for a minute and talk a little bit more about the possibility of moral consideration sort of beyond us. Right? And not just with animals either. I think some of the most interesting implications of the framework that the book brings out come not when we're just thinking about animals, but other kinds of beings. Right? As we create beings either through selective genetic engineering or through the design, as you suggest, of increasingly sophisticated artificial intelligence systems. So how should we think of the moral status of those? Maybe we can call them created things.

Yeah. So there are, as you say, all kinds of created beings. I mean, we already create a lot of animals for a lot of purposes. We breed farmed animals to grow as big as possible, as fast as possible. We breed lab animals to have diseases that are useful for us to study. We breed companion animals to have features that we regard as cute and cuddly. So we already create lots of beings for lots of purposes, and arguably that does give us a special responsibility to care for them given our role in their existence and their dependence on us. But then if you say, we are also creating new kinds of beings, not only, hybrids or chimera of humans and animals, but also radically different kinds of beings like AI systems. And there have already been inflection points regarding AI consciousness and sentience and agency and moral significance. In 2022, you might recall Blake Lemoine, a a Google engineer, was suspended and eventually fired when he publicly alleged that a large language model within Google called LAMBDA was sentient and deserved moral and even legal recognition. And there have been other such debates over the past couple of years. Now here is what I think about this, and I discuss it in the book as well within a re recent report that I coauthored with other philosophers, especially Robert Long. He and I led the report called taking AI welfare seriously. What we argue in that report and in related work is that when we look at current AI systems, we see relatively little evidence that they have consciousness or agency or other forms of moral significance.But when we consider how fast AI is developing, when we consider the trajectory of the industry and the incentives of the actors, and when we consider how much disagreement and uncertainty there still is about basic questions like what does it take to matter and could it being made out of silicon have what it takes and what will AI look like in 10 years. What we find is we are unable to rule out a realistic chance that within say 5 or 10 years, AI systems will have significantly more indicators of consciousness and agency and moral significance than they have now. For a wide range of leading scientific theories of consciousness, for example, there is no in principle barrier to creating the functional computational feat features that correspond to consciousness in silicon. Features like perception and attention and learning and memory and self awareness and social awareness and language and reason and so on and so forth. And a global workspace that ties it all together and makes it all work as one. There is no in principle reason why there could not be an AI

system in 10 years that has a physical body and then all of those functional computational capacities. So at that point, you really gotta ask, how much does it matter that you are made out of meat instead of metal? Because it could all end up depending on that. And if you're not absolutely sure that that is what is crucial being made out of meat instead of metal, then you should really take seriously the fact that we might be barreling towards the creation of a huge population of beings who can experience their very own form of happiness and suffering. Even if that is not a likelihood, the fact that we cannot rule out that possibility at present means that we should invest at least some resources in better understanding it and preparing for that as a possibility.

Yeah. I I wanna put the sort of risk framing here in conversation with an intuitive kind of eyebrow raising. Right? In the book, you have this kind of colorful example of a roommate who you one day sort of mysteriously discover is a robot. And in particular, I think it really renders sharply this sort of question, does it really matter if the difference is in fact only being made of silicon or being made of meat? Right? And the point that that example brings out is something like, look, there's still a very weird felt obstacle, right, that comes up in the mind of most people in a case like that. Maybe it's just discomfort with being sort of reduced to meat, which I guess anyone could feel, but, like, where do you think this sort of hesitancy to attribute moral status to things which are dramatically different than us sort of comes from?

Yeah. This is a really good question and a really tough question, and I think we would need psychologists and sociologists and anthropologists to to really answer it, but but I can offer some speculative thoughts, about the answer. I do think we have a kind of existential dread about our own nature and about the fact that we are ultimately bags of meat that happen to be able to walk around and experience happiness and suffering and and set and pursue goals. And so we do tell ourselves stories that try to make sense of this and and try to identify what is special about us that allows us to do that. And, you know, many stories in the past and and today focus on nonphysical parts of us like souls that that allow us to have these features. And then others focus on special features of us that maybe are unique to, if not humans, then at least animals. Like not only do we have all these functional capacities like perception and attention and learning and memory and so on and so forth, but we also have very specific brain processes that at present as far as we can tell are possible only for carbon based neurons, like certain types of chemical and electrical signals and oscillations that only human and animal brains can produce. And very smart philosophers like Peter Godfrey Smith think that is essential to consciousness, that you really do need those very specific, very fine grained chemical and electrical signals and oscillations in order to be conscious. And they may well be right, but I am not sure. That is all I can say about it. They may

be right and they may be wrong. And we have to make decisions about how to treat these increasingly sophisticated silicon based beings without knowing for sure if it all depends on the very specific fine grained chemical and electrical signals and oscillations that are possible only in human and animal brains. And then I think the other possible part of the explanation is that for our entire lives, we have experienced ourselves and each other as having minds, and we interact with ourselves and each other primarily on that basis. We explain our behavior in terms of how our minds cause it. But with AI systems and other technologies, we built them from the ground up, and we kind of default to our intentions or their mechanical structures to explain their behavior. We are just not used to explaining their behavior primarily in terms of mentality. And I and I think that creates a kind of bias, a kind of heuristic about where to look for explanations. And then that informs our views about whether they in fact have minds in the first place. But this is a mistake because even for us, we could be explained by the intentions of our creator if we had a creator. We could be explained by how our bodies make possible our behavior, but then we can also be explained by how our minds cause some of our behavior. And so even if we had intentions when we were creating AI systems and even if at some level their underlying structures do explain their behavior, it could for all that also be the case that they have mental states that are partial causes of some of their behaviors just like us and other animals. Right.

And we might worry that the kind of intention based or process based understanding here shows its age or maybe its ineptitude. When we're talking about a system that in the Lambda case tells you it's afraid of you. Right?

That is definitely right. We are more and more starting to experience AI systems as minded beings. Now what is complicated is their verbal outputs might not always be good evidence of their mental states because their verbal outputs are a result of pattern matching and text prediction that they were designed to do. So what is really complicated about, language models and generative models is that they do produce language outputs that make them seem like they have minds. And that is not good evidence that they have minds. But they might still have minds, and there might be other evidence that is good evidence of that. So that is part of what makes it so confusing.

Right. So as if this is not a complex enough conversation already, let's add a time axis. I think you're rightfully concerned, as we talked about earlier, with the ever increasing extent, if not propensity, that we have to cause harm in a particular era where our own actions and choices are are kind of constantly impacting various beings across nations, across generations, and across species. So how does this kind of scaler question about our impact change our sort of thinking about moral responsibility, importantly, now and into the future?

Yeah. We do now, for better or worse, live in a world that is through and through influenced by human activity. Industries like factory farming and deforestation and the wildlife trade and, you know, AI development and and deployment and so on and so forth. These are all transforming the planet in a way that affects humans and other animals and anyone else who might matter. And it does that not only across species and substrates, but also as you say across nations and generations. And so my argument, and and I am of course not remotely the first person to make this point, is that to the extent that our actions are either intentionally or foreseeably having morally significant impacts like harm on distant others, we ought to take that into account. And, you know, we can take it into account in an appropriate way. Like we might, discount distant impacts if we feel less uncertain about them or if we feel less able to control them. But to the extent that our actions or policies are having foreseeable effects on individuals who matter and in other nations or future generations, and and we do feel that we can predict and control them and and mitigate that harm in a way that can be achievable and sustainable for us, then then we really ought to try to do that. We really ought to factor that in to our policy decisions. And this is an idea that people already are coming to accept more and more, especially when it comes to increasingly urgent topics like pandemics and climate change. We recognize that our actions and policies can influence the frequency and intensity of disease outbreaks or extreme weather events in future generations. And to the extent that we can set the world on a better trajectory that has fewer disease outbreaks, fewer extreme weather events in the future, then that should at least be one factor among many in in the decisions that we make. And my point is just that we have to consider that alongside our responsibilities to the nonhuman world. So a major set of stakeholders in our actions and policies moving forward is not only nonhumans and not only members of future generations, but also and especially nonhumans in future generations. That is a huge population that is particularly vulnerable to the decisions that we make now when it comes to pandemics, when it comes to climate change, when it comes to the trajectory of AI. Nonhumans will be major stakeholders in that. We might soon find ourselves making decisions, for example, about whether to terraform new planets, whether to send a microbial life or other forms of life to live on new planets to make them hospitable for us in the future. We gotta take seriously the impacts that could have on the nonhuman beings we send there. So so those are the types of wild future oriented questions we might soon need to be asking.

In this sort of overarching process of evolution of our thinking, you argue that it's possible, if not likely, that different kinds of moral theories, both consequentialist and non consequentialist ones, might actually converge in a kind of practice as we think about these new frontiers. So can you say a little bit more about what this kind of convergence might look like in moral philosophy and what that might

mean for some ongoing debates?

Yeah. Thanks for asking about this. I I do think this matters a lot because in the same kind of way that we tend to ask ethical questions and yes or no, all or nothing ways, We also tend to see ethical theories as opposed to each other, rather than seeing them as partners in in a coalition, which is what I think they should be. And there are various reasons for thinking that we should, accept and combine multiple moral theories. One might be moral uncertainty, not being sure which theory is correct and so wanting to accept elements from several theories in order to be cautious and humble. But another reason is that each moral theory when properly applied ends up incorporating features from other moral theories. And and so I can just give you a couple of examples to illustrate that. You described consequentialism and non consequentialism. For simplicity, we can say consequentialism is the type of moral theory that says morality is primarily about the consequences of our actions and policies, and and our main goal is to do the most good possible or at least do good in the world. And then non consequentialism is the kind of view that says morality is primarily about something else. It could be respecting rights. It could be being a good person and having a virtuous character. It could be having good relationships, caring relationships with others in your life, but morality is about something else. Okay. Now I think these moral theories, when properly applied, end up drawing from each other. So for example, consequentialism. If we really wanna do the most good possible, if we really wanna do good in the world, we should not go about thinking like consequentialists all the time. If I make every decision by asking, you know, which stock should I put on right now that will do the most good for all sentient beings from now until the end of time? Or which sandwich should I eat today that will do the most good for all sentient beings from now until the end of time? I would never be able to make those calculations. I would never make decisions. I would constantly be making mistakes. Self interest and bias would be creeping into my assessments. So if I really wanna be a good consequentialist, I should only infrequently and with guardrails think like a consequentialist. I should ask how to do the most good and how I can orient my life towards doing the most good, but part of how I can achieve that is by really endeavoring to respect rights and really endeavoring to cultivate virtuous states of character that will naturally guide me towards good actions when I lack the opportunity to make all those calculations and really cultivating good caring relationships with others where we can support each other and and empower each other to to have good impacts in the world. So to be a good consequentialist, you need to be a good non consequentialist. You need to really care about rights and virtues and relationships. And then similarly, when when I think in terms of being a non consequentialist, as you and I talked about before, we have to reckon with the fact that whether we like it or not, our actions and policies are already affecting everyone everywhere, or at least they have a chance of affecting

everyone everywhere. And this old idea that we can just leave others alone might not be possible anymore in this kind of world. We just already are not leaving them alone. We already are affecting them. And so even if all you care about is respecting rights or being a virtuous person or having good relationships with others, you still have to reckon with the fact that you are affecting others and we are together affecting others. And so we have to think about whether that constitutes rights violations, whether that is an expression of vice, whether that places us in a kind of callous or uncaring relationship. And and there is no way around thinking about the possible harms that our actions are causing and how we can reduce those harms at scale across time and space. And that puts you in a little bit more of a consequentialist mindset. So to be a good consequentialist, you gotta think like a non consequentialist. And to be a good non consequentialist, you gotta think like a consequentialist. Now that is not to say that they totally merge, but it is to say that they should, you know, reach across the aisle and try to be good bipartisan and try to work together on on some good bipartisan policies that that are going to to help us reduce harms and rights violations and vice and callousness in the world.

The framework you're developing and describing likely requires some pretty substantial change of us as a community and a society. So thinking about this now and into the future, how specifically do you think our frameworks need to update or evolve to handle expanding circles of moral consideration? More specifically, do you think there are particular concepts that we maybe need to revise or let go of?

Yeah. A lot is going to need to change. I think when all is said and done, when the dust settles, we will probably have an entirely new set of moral, legal, political concepts for capturing the forms of value that exist in the world and and what what is owed to everyone and everything that has value. Just to pick one example, we currently allocate legal rights and standing according to our concept of legal personhood. Right? We say that you have legal rights and you have standing in a court, for example, if you are a legal person. And then, of course, conversationally, we use the words human and person totally interchangeably. So technically speaking, all it means to be a legal person is to be the kind of entity that can have legal rights and or legal duties. But in practice, we just thoroughly associate that with being a member of the species homo sapiens. And so we're at some point, maybe soon and definitely in the long run, going to have to face a decision. Either we decouple the idea of personhood from the idea of humanity and allow personhood to simply refer to any being that can have legal duties and or legal rights and perhaps that could be elephants or chimpanzees or ants or chatbots, or we allow the idea of humanity and the idea of personhood to to stay thoroughly tangled up in each other. And we abandon personhood as the concepts that we use to allocate legal duties and legal rights and legal standing and these really core features of what it means to be protected under the

law. So that is one example of a concept that may have to transform depending on how our attitudes about it evolve. Now that will be an intergenerational project as will all of these transformations. None of this is going to happen within the next decade, probably within our lifetimes. But what we can be asking ourselves is how can we be nudging our species in the world in the direction of tackling these issues in a way that can be achievable and sustainable?

So I wanna ask, less at the level of concepts and more at the level of us kind of seeking from our moral philosophy that it be action guiding: The book suggests that we need to extend or, as I said earlier, explode moral consideration to a whole lot of things that we don't currently consider valuable or as mattering. So what sort of practical steps should individuals or maybe even governments, I know you've done some advocacy work on this, take to begin addressing moral obligations at this kind of scale?

This is a great place to close, bringing it back down to earth. And as you say, these questions should be asked and answered in different domains and governments, companies, by individual citizens, and at different scales at the international level, at the local level. So just to give you a few concrete examples, there are all kinds of things. Governments, companies, individuals can be doing to take these issues more seriously even though this will be an intergenerational project. So for example, local governments can start extending consideration to more animals by factoring them into their food policies, into their infrastructure policies, into their pest and conflict management policies, just to pick infrastructure as an example. When cities transform their infrastructures to be more resilient and sustainable in the face of human caused climate change, If they include animal welfare as a factor, they can work towards all kinds of co beneficial policy like bird friendly glass that is more energy efficient and reduces collisions that affect humans and birds. Or wildlife corridors on green transportation systems that once again, the transportation systems are more energy efficient and the wildlife corridors reduce collisions that are, affect humans and animals. And there are a bunch of other examples like that and in other policy domains too. Now take companies. In in our AI welfare report, we argue that leading AI companies have a responsibility to take AI welfare seriously now, not, of course, by giving a full set of rights to current generation chatbots, but by starting to create an infrastructure that will allow us to responsibly address this issue as the technology advances. So for example, they can appoint or hire an AI welfare officer or researcher to help them better understand the issue. They can acknowledge that this is an issue and allow their language models to do the same. They can start developing assessments for consciousness and agency in AI systems modeled on similar assessments that we use for nonhumans like insects. They can start developing policies and procedures for ethically assessing their interactions with AI systems modeled on nonhuman subjects research

frameworks that we use in medical ethics. And then finally, for individuals. What can we do? I mean, we can transform the ways we think and talk. We can start thinking and talking in terms of who might matter rather than who does. We can cultivate humility about this, keeping in mind how confidently wrong so many people have been in the past. And we can start cultivating those virtues that will make us better decision makers when the stakes are high. Like when I see an insect trapped in my apartment, I can err on the side of taking them outside when I have the opportunity to do that. In part because this individual insects might matter. And in part because that helps me see insects in general as beings who are worthy of consideration. And then if I am making a high stakes decision, I will be a little bit more likely to make a good one or, you know, say please and thank you to your chatbot. Not because they matter, but maybe as practice for interactions with AI systems who in 10 years might matter. So those are some concrete examples.

Once again, for our listeners, this has been a conversation with Jeff Sebo, author of The Moral Circle, Who Matters, What Matters, and Why, newly available from Norton Shorts. Jeff, thanks so much for coming on the show. Yeah.

My pleasure. Thanks so much for having me. It was a really fun conversation.

Examining Ethics is hosted and produced by Alex Richardson and brought to you by the Janet Prindle Institute For Ethics at DePauw University. The views represented here are those of our guests and don't reflect the position of the Prindle Institute or of DePauw University. Our show's music is by Blue Dot Sessions. You can learn more about today's episode and check out supplementary resources at examiningethics.org. As always, you can contact us directly at examiningethicsdepauw.edu. Thanks for listening, and we'll see you next time.