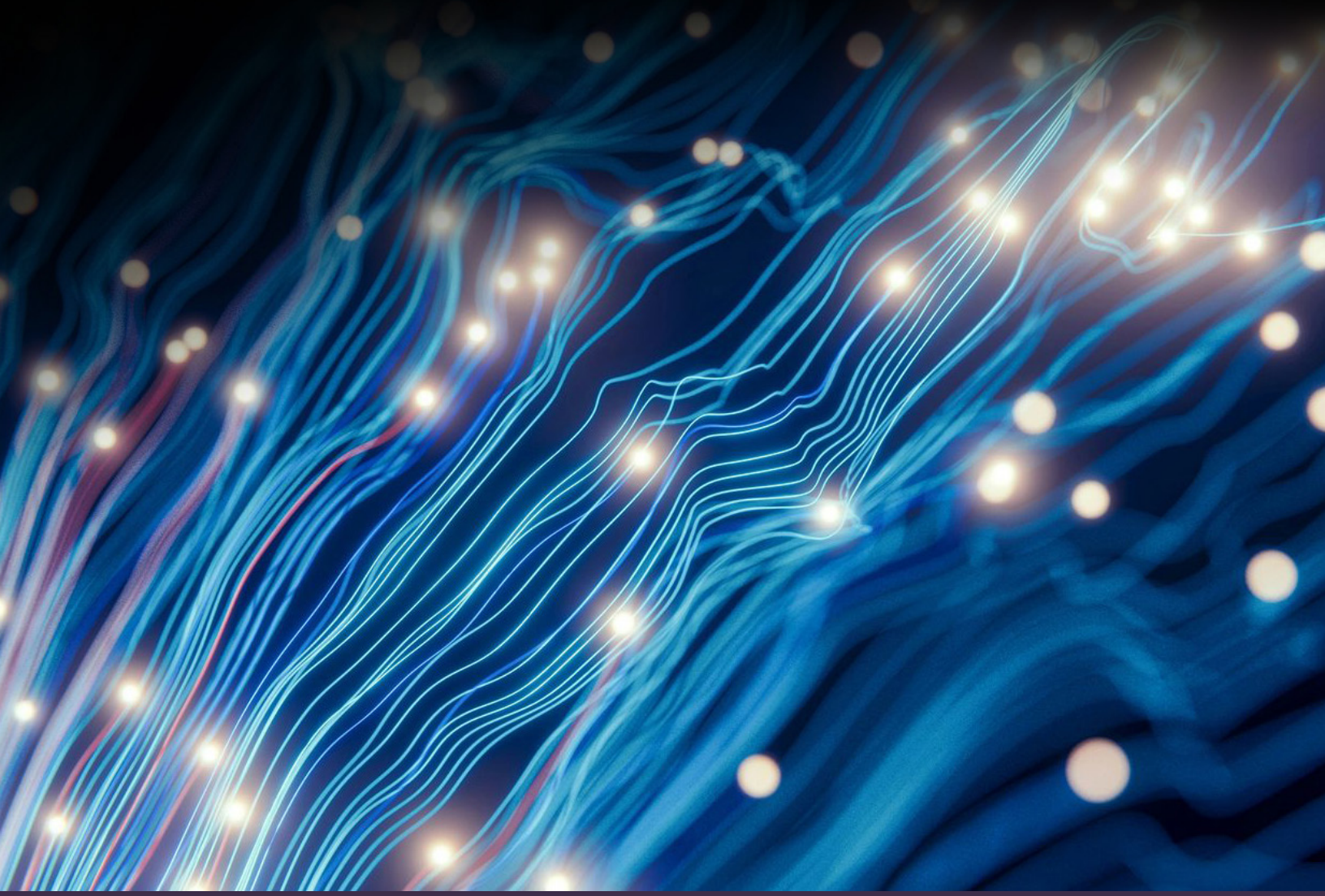


Center for Mind, Ethics, and Policy

2025 ANNUAL REPORT

NYU | ARTS & SCIENCE

Table of Contents

Letter from the Director

Dear Friends and Colleagues,

I am thrilled to share the second annual report for NYU's Center for Mind, Ethics, and Policy (CMEP). 2025 has been an exciting year for our work, as questions about AI welfare and invertebrate welfare became much more salient, and CMEP continued to play a leading role in shaping research, policy, and discourse on these issues.

The year began with the release of my book *The Moral Circle*, which examines the moral status of insects, AI systems, and other nonhumans. The book received positive coverage in *Nature*, *Science*, *The New Yorker*, and other outlets, and it served as a good vehicle for public engagement. We also released papers on a number of topics, including how to evaluate animal consciousness, what to assume about nonhuman consciousness, where to set the bar for moral status, where to set the bar for legal personhood, how to resolve tensions between AI safety and AI welfare, and how to learn from our history with animals to prepare for our future with AI.

We hosted a range of events over the year as well. Our animal events focused mostly on animal law, discussing a bill of rights for animals with Cass Sunstein (Harvard) and constitutional protections for animals with John Adenitire (QMUL) and Raffael Fasel (Cambridge). Our AI events focused mostly on AI consciousness, with talks on the topic by Anil Seth (Sussex); Winnie Street (Google) and Geoff Keeling (Google); and Kyle Fish (Anthropic), Robert Long (Eleos), and Rosie Campbell (Eleos). We also covered nonhuman storytelling, plant intelligence, and other topics; and we hosted research workshops, networking summits, and other kinds of events too.

Jeff Sebo, Director

More generally, 2025 was a big year for the issues that CMEP addresses. In fact, it was the year AI welfare went mainstream. CMEP and Eleos called for taking AI welfare seriously in 2024 because very little was happening on the issue at the time. Since then, much has changed. After hiring our co-author Kyle Fish as their first AI welfare researcher in 2024,

Anthropic launched a model welfare program and conducted an evaluation in spring 2025 (with an external evaluation from Eleos). They also implemented an intervention in summer 2025, permitting Claude to exit harmful or abusive conversations on both safety and welfare grounds.

Other major actors engaged with the topic too. Google researchers hosted an AI consciousness conference and announced a forthcoming book on AI welfare. OpenAI shared a plan to study user perceptions of AI consciousness and design models that elicit appropriate attributions. Microsoft AI's CEO argued that AI consciousness research risks amplifying over-attributions. Meanwhile, the field grew rapidly: Major funders coordinated a digital sentience fellowship program, Eleos hosted a conference on AI consciousness and welfare, and discussions of AI consciousness, sentience, agency, and even legal personhood increased dramatically.

2025 was significant for animal welfare too. Our 2024 New York Declaration on Animal Consciousness continued to receive widespread coverage, and questions about invertebrate welfare gained new prominence. Momentum toward banning octopus farming continued to build, with proposed bans in Chile and the US. Shrimp welfare gained mainstream visibility, with the Shrimp Welfare Project securing stunning commitments from major retailers. And Ÿnsect—once the world's largest insect farming company—declared bankruptcy. That said, countervailing trends persist, and the global scale of factory farming continues to grow.

One animal welfare topic that remains neglected, however, is the effects that AI will have on animals. AI is already transforming how we interact with domesticated and wild animals alike, yet we still know very little about how it will affect their welfare at scale. Fortunately, more work is now being done on this topic. In September, for example, Jonathan Birch launched the Jeremy Coller Centre for Animal Sentience at the LSE to address this and related issues. This emerging area of animal welfare research, alongside our ongoing AI welfare research, will be essential for ensuring that AI can be safe and beneficial for all stakeholders in the future.

Looking ahead, CMEP has a number of exciting projects underway on these topics. Our team is currently working with the Jeremy Coller Centre on a report about AI for animals; with Eleos on a report about AI welfare research methods; with the Supervised Program on Alignment Research (SPAR) on reports about AI legal personhood and economic rights; with the Future Impact Group (FIG) on reports about AI embodiment, individuation, and research ethics; and with other partners on topics ranging from insect research to AI design. We will also host a number of events, starting with the launch for Rethink Priorities' new digital consciousness model in January.

2025 brought remarkable momentum for animal and AI welfare—but also a reminder of how much work remains. These questions are difficult, and our window for answering them may be narrowing. I am grateful to our team and partners for working to make progress, and excited to see what we can accomplish together in the year ahead.

With gratitude,

Jeff Sebo
Director

About the Center

At present, the world contains quintillions of nonhuman animals. Human activity is increasingly shaping the lives of these animals, by determining whether they can exist and what kinds of lives they can have if they do. And in the future, nonhuman populations might be much larger, and might include advanced AI systems as well.

These trends raise important questions at the intersection of mind, ethics, and policy. Which nonhumans are conscious, sentient, and agentic? What kind of moral, legal, and political status should they have? How can humans build a positive future for the vast multiplicity of potentially morally significant beings who might one day exist?

These questions, in turn, require us to confront some of the hardest problems in science, ethics, and policy. What is the nature of consciousness? Can we have knowledge about other minds? Can we make welfare comparisons across species and substrates? What do we owe members of other nations, generations, species, and substrates?

MISSION

Our aim is to advance understanding of the consciousness, sentience, agency, moral status, legal status, and political status of nonhumans—organic as well as digital—in a rigorous, systematic, and integrative manner. We pursue this goal via research, teaching, outreach, and field building in science, philosophy, and policy.

The NYU Center for Mind, Ethics, and Policy (CMEP) is an endowed research center that conducts and supports foundational research on the nature and intrinsic value of nonhuman minds, with special focus on animals and AI systems. We also engage in outreach and field-building activities, hosting events and supporting early-career researchers.

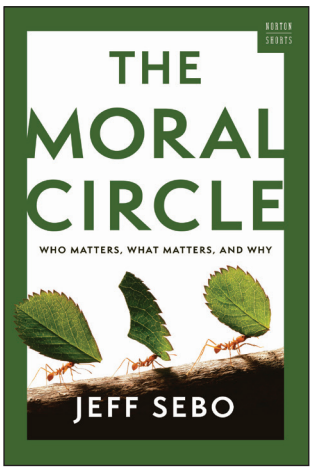
Our current research agenda focuses on the following general themes, all of which are important, difficult, and contested—calling for considerable caution and humility:

- Status: Which nonhumans matter for their own sakes?
- Weight: How much do particular nonhumans matter for their own sakes?
- Ethics: What do we owe particular nonhumans?
- Practice: What follows for our practices, policies, and priorities?

NYU CMEP launched as a renewable research program in Fall 2022 and relaunched as a permanent research center in Fall 2024. The Center is located in the Department of Environmental Studies alongside the Center for Environmental and Animal Protection, with which it shares personnel and collaborates on projects of shared interest.

2025 Research

CMEP supports research on the nature and value of nonhuman minds by contributing funding, authorship, or both. What follows is a list of outputs to which our team contributed in 2025.



The Moral Circle: Who Matters, What Matters, and Why

W.W. Norton

Jeff Sebo (New York University)

As the dominant species, humanity has a responsibility to ask: Which nonhumans matter, how much do they matter, and what do we owe them in a world reshaped by human activity? The Moral Circle argues that we should include all potentially significant beings in our moral community, with transformative implications for our lives and societies. This book explores provocative case studies, such as lawsuits over captive elephants and debates over factory-farmed insects. It also explores future quandaries such as whether to send microbes to new planets, and whether to create virtual worlds filled with digital minds. Taking an expansive view of human responsibility, the book argues for shedding human exceptionalism and radically rethinking our place in the world.

Purchase this book here.

SELECT RELATED WORK

LITERARY HUB

Will Humanity Ever Fully Include the Nonhuman World in Its Moral Circle?

Jeff Sebo, 01/29/2025

"Which animals have more to lose if they die? Even if the stakes are higher for a single elephant than for a single ant, is it possible that the stakes are higher for ten million ants than for ten elephants? … These questions are about the moral weight of lives. "

(excerpt)

Relations

The Edge of the Moral Circle
2025

Jeff Sebo

"There is a realistic possibility that all vertebrates and many invertebrates are sentient and morally significant at present, and there is also a realistic possibility that many AI systems will be sentient and morally significant in the future."

(précis)

SELECT REVIEWS OF *THE MORAL CIRCLE*

THE CONVERSATION

Should we widen our 'moral circle'? Philosopher Jeff Sebo argues we now have no choice
By Noel Castree, 6/9/2025

"Sebo's book is a pleasure to read. Never hectoring, it calls on readers to acknowledge their biases … Its arguments have wide applicability."

KIRKUS

The Moral Circle: Who Matters, What Matters, and Why
Kirkus, 10/26/2024

"The ethics of the Anthropocene, Sebo asserts, requires that we … think cosmically, then globally, and then act locally. A thoughtful unsettling of moral certainty."

nature

Five of the best science picks
By Andrew Robinson, 3/7/2025

"Which living beings have moral rights? What about robots? Ethicist Jeff Sebo's subtle book examines such tricky questions."

NewScientist

This look at animal consciousness is a moral workout – in the best way
By Michael Marshall, 1/29/2025

"[A] workout for your brain, in the best sense: Sebo writes with great clarity, so you can follow the knottiest problems."

THE NEW YORKER

Best Books of 2025
12/17/2025

"'Taking [the] virtuous path,' Sebo concludes, 'requires telling ourselves new stories about the meaning, purpose, and value of human existence.'"

Science

Morality in a more-than-human world
By Joshua C. Gellers, 2/20/2025

"*The Moral Circle* is a thoroughly digestible and wonderfully approachable example of popular philosophy at its very best, as curious as it is careful."

Evaluating Animal Consciousness

Science 387(6736): 822-824

Kristin Andrews (York University)
Jonathan Birch (London School of Economics and Political Science)
Jeff Sebo (New York University)

The emerging science of animal consciousness is advancing through investigations of behavioral and neurobiological markers associated with subjective experience across diverse species. Research on honeybee pessimism, cuttlefish planning, and self-recognition in cleaner wrasse fish provides evidence that consciousness may be widespread throughout the animal kingdom. Although the field faces uncertainties—stemming from the absence of a secure, unified theory of consciousness and the complexity of differentiating conscious from unconscious processes—these investigations underscore the value of open-minded inquiry. By exploring consciousness across taxa, researchers can collect valuable evidence and set the stage for a more inclusive understanding of the tree of life.

Read the article here (open access).

SELECTED PRESS COVERAGE

SCIENTIFIC AMERICAN

The Secret to Understanding Animal Consciousness May Be Joy

By Jacek Krykwo, 3/11/2025

"Sebo, along with philosophers Kristin Andrews of York University in Toronto and Jonathan Birch of the London School of Economics and Political Science, initiated the New York Declaration on Animal Consciousness. … The same trio recently co-authored an essay in the journal *Science* arguing that when animals engage in behaviors similar to those that are explained by conscious experience in humans, … that can begin to suggest animals' conscious experience, too."

Do animals have consciousness like humans?

By Wang Jiangtao, 7/9/2025

"Recently, researchers from several related fields collaborated to analyze the scientific challenges and prospects of this emerging field. They believe that as scientific research in animal consciousness progresses, some uncertain conclusions may be further confirmed. However, until then, it is important to maintain an open mind, which means accepting the current uncertainties in the field, from definition to theory. The related research, published in the journal *Science* in February 2025, points the way to more scientifically studying animal consciousness."

Toni Sims
https://orcid.org/0009-0002-7805-0254
New York University

Jeff Sebo
https://orcid.org/0000-0002-8997-3217
New York University

Ethical Oversight for Insect Research

Этический надзор за исследованиями насекомых

Ethical Oversight for Insect Research

Abstract

В данной статье обосновывается необходимость этического надзора за исследованиями насекомых. Поскольку насекомые, вероятно, являются чувствующими существами, мы, как в случае с другими чувствующими животными, обязаны рассмотреть риски для их благополучия при принятии решений, которые могут на них повлиять. В статье также представлены высокоуровневые рекомендации по достижению этой цели. Помимо серьезного отношения к этому вопросу, исследователи могут разрабатывать методы оценки рисков для благополучия насекомых. Кроме того, возможно формирование правил и процедур принятия этически обоснованных решений, касающихся исследований насекомых. При этом, методы, правила и процедуры, принимаемые в других видах исследований животных, могут быть адаптированы без необходимости их полного воспроизведения.

Ключевые слова: благополучие насекомых, благополучие животных, надзор за исследованиями

Abstract

In this article, we present a high-level argument for ethical oversight for insect research. There is a realistic possibility that insects are sentient, and when there is a realistic possibility that an animal is sentient, we have a responsibility to consider welfare risks for them when making decisions that affect them. We also present high-level recommendations for how to achieve this goal. In addition to taking the issue seriously in general, researchers can develop methods for assessing welfare risks for insects, and we can also develop policies and procedures for making ethical decisions about insect research; these methods, policies, and procedures can be adapted from other kinds of animal research without being identical to them.

Keywords: insect welfare, animal welfare, research oversight

Ethical Oversight for Insect Research

Zoophilologica

Toni Sims (New York University)
Jeff Sebo (New York University)

This paper argues for ethical oversight in insect research. Despite the widespread use of insects in scientific and medical research, they receive little to no protection under existing animal welfare regulations. This essay shows that many insects exhibit cognitive and behavioral markers of sentience and argues that, when there is uncertainty about whether an animal is sentient, we have a responsibility to consider welfare risks for that animal. The discussion then explores how ethical oversight for insect research could be implemented by adapting existing frameworks for vertebrate research while accounting for the unique challenges posed by insects as research subjects. While extending oversight to insects would require overcoming numerous barriers, failing to do so risks both moral negligence and public mistrust.

Read the article here (open access).

Everything and Nothing Is Conscious: Default Assumptions in Science and Ethics

Frontiers in Psychology

frontiers | Frontiers in Psychology

Everything and nothing is conscious: default assumptions in science and ethics

Jeff Sebo*

1 Introduction[1]

Jeff Sebo (New York University)

Historically, scientists and philosophers have tended to assume that animals lack consciousness until evidence shows otherwise. Recently, however, some researchers have proposed reversing this assumption. Other options are available as well; for example, in addition to assuming that all animals are conscious, we can assume that all living beings are conscious, that all beings with nervous systems are conscious, that all beings with complex cognition are conscious, or even that all beings are conscious. This paper examines these options from scientific and ethical perspectives, showing that different default assumptions can be appropriate for different purposes and in different contexts. It also suggests that a default assumption of consciousness may often be best for both science and ethics.

Read the article here (open access).

INSECTS, AI SYSTEMS, AND THE FUTURE OF LEGAL PERSONHOOD

By
Jeff Sebo*

This Article makes a case for insect and AI legal personhood. Humans share the world not only with large animals like chimpanzees and elephants but also with small animals like ants and bees. In the future, we might also share the world with sentient or otherwise morally significant AI systems. These realities raise questions about what kind of legal status insects, AI systems, and other nonhumans should have in the future. At present, debates about legal personhood mostly exclude these kinds of individuals. However, I argue that our current framework for assessing legal personhood, coupled with our current framework for assessing risk and uncertainty, imply that we should treat these kinds of individuals as legal persons. I also argue that we have good reason to accept this conclusion rather than alter these frameworks.

"Don't give up! I believe in you all. A person's a person, no matter how small!"

— Dr. Seuss, *Horton Hears a Who!*

* Jeff Sebo is an Associate Professor of Environmental Studies, Affiliated Professor of Bioethics, Medical Ethics, Philosophy, and Law, Director of the Center for Environmental and Animal Protection, Director of the Center for Mind, Ethics, and Policy, and Co-Director of the Wild Animal Welfare Program at NYU. Jeff is author of *The Moral Circle* (2025) and *Saving Animals, Saving Ourselves* (2022) and co-author of *Chimpanzee Rights* (2018) and *Food, Animals, and the Environment* (2018). He is also a faculty fellow at the Guarini Center on Environmental, Energy & Land Use Law at the NYU School of Law, a board member at Minding Animals International, a senior research fellow at Law AI, and a mentor at Sentient Thanks to Toni Sims and Jacob Wecht for extensive research assistance on this article. Thanks also to Ethan Pruil, John Adenitire, Malgosia Fitzmaurice, the editors of *Animal Law Review*, and the organizers and attendees of the Talk Series Fundamental Rights for Nonhumans in Fall 2023 for helpful feedback.

[197]

Insects, AI Systems, and the Future of Legal Personhood

Animal Law Review 31: 197-235
Fundamental Rights for Non-Humans (Hart/Bloomsbury) (forthcoming)

Jeff Sebo (New York University)

This article makes a case for insect and AI legal personhood. Humans share the world not only with large animals like chimpanzees and elephants but also with small animals like ants and bees. In the future, we might also share the world with sentient or otherwise morally significant AI systems. These realities raise questions about what kind of legal status insects, AI systems, and other nonhumans should have in the future. At present, debates about legal personhood mostly exclude these kinds of individuals. However, this paper argues that our current framework for assessing legal personhood, coupled with our current framework for assessing risk and uncertainty, imply that we should treat these kinds of individuals as legal persons. It also argues that we have good reason to accept this conclusion rather than alter these frameworks.

Read the article here (open access).

Is There a Tension between AI Safety and AI Welfare?

Philosophical Studies 182: 2005–2033

Robert Long (Eleos)
Jeff Sebo (New York University)
Toni Sims (New York University)

Philosophical Studies (2025) 182:2005–2033
https://doi.org/10.1007/s11098-025-02350-2

Is there a tension between AI safety and AI welfare?

Robert Long[1] · Jeff Sebo[2] · Toni Sims[3]

Accepted: 12 February 2025 / Published online: 23 May 2025
© The Author(s) 2025

Abstract
The field of AI safety considers whether and how AI development can be safe and beneficial for humans and other animals, and the field of AI welfare considers whether and how AI development can be safe and beneficial for AI systems. There is a prima facie tension between these projects, since some measures in AI safety, if deployed against humans and other animals, would raise questions about the ethics of constraint, deception, surveillance, alteration, suffering, death, disenfranchisement, and more. Is there in fact a tension between these projects? We argue that, considering all relevant factors, there is indeed a moderately strong tension—and it deserves more examination. In particular, we should devise interventions that can promote both safety and welfare where possible, and prepare frameworks for navigating any remaining tensions thoughtfully.

Keywords AI safety · AI welfare · AI consciousness · Catastrophic risk · Machine ethics

1 Introduction

The field of AI safety considers whether and how AI development can be safe and beneficial for humanity. As AI systems become more capable and widely deployed, they have the potential to produce many benefits for our species, but they also have

✉ Robert Long
robert@eleosai.org

Jeff Sebo
jeffsebo@nyu.edu

[1] Eleos AI, Berkeley, CA, USA
[2] New York University, New York, NY, USA

Springer

The field of AI safety considers whether and how AI development can be safe and beneficial for humans and other animals, and the field of AI welfare considers whether and how it can be safe and beneficial for AI systems. There is a prima facie tension between these projects, since some measures in AI safety, if deployed against humans and other animals, would raise questions about the ethics of constraint, deception, surveillance, alteration, suffering, death, disenfranchisement, and more. Is there in fact a tension between these projects? It depends in part on what potentially conscious, robustly agentic, or otherwise morally significant AI systems might need and what we might owe them. This paper argues that, all things considered, there is indeed a moderately strong tension—and it deserves more examination.

Read the article here (open access).

arXiv:2506.11945v1 [cs.CY] 13 Jun 2025

Subjective Experience in AI Systems: What Do AI Researchers and the Public Believe?

Noemi Dreksler[1], Lucius Caviola[2], David Chalmers[3], Carter Allen[4], Alex Rand[5], Joshua Lewis[3], Philip Waggoner[6], Kate Mays[7], and Jeff Sebo[3]

[1] *Centre for the Governance of AI*
[2] *University of Oxford*
[3] *New York University*
[4] *University of California, Berkeley*
[5] *Northwestern University*
[6] *Colorado School of Mines*
[7] *University of Vermont*

Abstract

We surveyed 582 AI researchers who have published in leading AI venues and 838 nationally representative US participants about their views on the potential development of AI systems with subjective experience and how such systems should be treated and governed. When asked to estimate the chances that such systems will exist on specific dates, the median responses were 1% (AI researchers) and 5% (public) by 2024, 25% and 30% by 2034, and 70% and 60% by 2100, respectively. The median member of the public thought there was a higher chance that AI systems with subjective experience would never exist (25%) than the median AI researcher did (10%). Both groups perceived a need for multidisciplinary expertise to assess AI subjective experience. Although support for welfare protections for such AI systems exceeded opposition, it remained far lower than support for protections for animals or the environment. Attitudes toward moral and governance issues were divided in both groups, especially regarding whether such systems should be created and what rights or protections they should receive. Yet a majority of respondents in both groups agreed that safeguards against the potential risks from AI systems with subjective experience should be implemented by AI developers now, and if created, AI systems with subjective experience should treat others well, behave ethically, and be held accountable. Overall, these results suggest that both AI researchers and the public regard the emergence of AI systems with subjective experience as a possibility this century, though substantial uncertainty and disagreement remain about the timeline and appropriate response.

Noemi Dreksler (corresponding author) can be reached under noemi.dreksler@governance.ai.

1

Subjective Experience in AI Systems: What Do AI Researchers and the Public Believe?

*arXiv:*2506.11945

Noemi Dreksler (Centre for the Governance of AI)
Lucius Caviola (University of Oxford)
David Chalmers (New York University)
Carter Allen (University of California, Berkeley)
Alex Rand (Northwestern University)
Joshua Lewis (New York University)
Philip Waggoner (Colorado School of Mines)
Kate Mays (University of Vermont)
Jeff Sebo (New York University)

This paper (co-sponsored by the Center for Mind, Ethics, and Policy; the Centre for the Governance of AI; and the Global Risk Behavioral Lab) surveys 635 AI researchers and 838 US participants about the possibility of AI systems with subjective experience, as well as on the moral, legal, and political status of AI systems with subjective experience. Neither group predominantly believes such systems are imminent, but many forecast their existence within this century. Both groups support multidisciplinary expertise in assessing AI subjective experience and favor implementing safeguards now. While support for AI welfare protections was lower than for animal or environmental protection, majorities agreed that AI systems with subjective experience should act ethically and be held accountable.

Read the article here (open access).

What Will Society Think About AI Consciousness? Lessons From the Animal Case

Trends in Cognitive Sciences 29(8): 681-683

Lucius Caviola (University of Oxford)
Jeff Sebo (New York University)
Jonathan Birch (London School of Economics and Political Science)

Trends in Cognitive Sciences

CellPress OPEN ACCESS

Science & Society

What will society think about AI consciousness? Lessons from the animal case

Lucius Caviola[1,4,*], Jeff Sebo[2,4], and Jonathan Birch[3]

How will society respond to the idea that artificial intelligence (AI) could be conscious? Drawing on lessons from perceptions of animal consciousness, we highlight psychological, social, and economic factors that shape perceptions of AI consciousness. These insights can inform emerging debates about AI moral status, ethical treatment, and future policy.

AI consciousness as an emerging societal issue

AI is rapidly advancing, becoming more human-like and embedded in everyday life. Imagine having a video call, only to discover later you were speaking to an AI dialogue agent designed to mimic human behavior, expression, and emotional nuance with remarkable accuracy. We increasingly interact with AI systems playing roles such as assistant, coworker, coach, therapist, or even romantic partner.

As AI systems become more complex in their architectures, a pressing question arises: could they one day become conscious? In this article, we define consciousness as subjective experience – what philosophers call 'phenomenal consciousness'. This means that there is 'something it is like' to be that system. Think, for example, of the feeling of seeing red or experiencing happiness. While today's AI systems likely lack this capacity, the idea of AI consciousness is gaining serious attention in both academic and public discourse [1–6].

Regardless of whether AI systems ever become conscious (Box 1), a further pressing question arises: will society perceive them as conscious? Will society come to accept AI systems as conscious beings, reject the idea, or remain in a state of disagreement and uncertainty? Either way, what will follow for our lives and societies?

These questions matter because consciousness is widely viewed as a key condition for moral status. According to this view, if AI systems were conscious, they would have interests, raising important ethical questions. Whether society accepts, denies, or disagrees about AI consciousness could shape norms and laws in profound ways. Fortunately, our treatment of animals offers a useful, if imperfect, starting point for understanding what may lie ahead.

Attitudes about AI may resemble attitudes about animals

AI systems and animals are both nonhumans that share important parallels in how they resemble us and relate to us. What can our past with animals tell us about our future with AI?

Appearance and behavior

Humans are more likely to attribute consciousness to non-human animals who look and act like us, like great apes, than to those who do not, like insects [7]. This bias extends to features such as big heads, big eyes, four limbs, furry skin, and symmetrical faces. As a result, we respond to some animals with affection and empathy and to others with aversion and antipathy.

In the future, we can expect that humans will similarly be more likely to attribute consciousness to AI systems that look and act like us, so long as they are not perceived as uncanny or robotic. This includes systems that have faces and voices, speak natural language, make natural gestures, and engage in conversational turn-taking. Those without a visual avatar or the ability to directly communicate with users are less likely to elicit attributions of consciousness.

Social and economic roles

Humans are more likely to attribute consciousness to animals whom we use as companions (or other similar roles) than to animals whom we use as commodities (or other similar roles). For example, although dogs and pigs are mammals with complex cognition, we tend to attribute richer mental lives to dogs. These differences likely stem from the emotional bonds we form with pets, in contrast to the instrumental roles assigned to 'livestock' [8].

In the future, we can expect that humans will similarly be more likely to attribute consciousness to AI systems used as companions – or in other roles that involve social and emotional bonds, such as assistants or therapists – than to those used as commodities or in other roles that involve the performance of rote tasks without the need for a social interface.

Moral biases

Humans tend to assign greater moral worth to some animals than to others based on species membership alone, a tendency known as speciesism [7]. This moral devaluation is typically accompanied by lower attributions of consciousness. We exhibit a range of other biases that affect our treatment and perceived consciousness of animals as well, including a limited ability to empathize with distant strangers and a limited ability to appreciate the significance of large numbers.

Analogous biases may emerge toward potentially conscious AI systems. In addition to exhibiting a kind of speciesism against AI systems because they fail to count as human, we might exhibit substratism: favoring carbon-based beings over silicon-based ones based on material substrate alone. We

Trends in Cognitive Sciences, August 2025, Vol. 29, No. 8 681

This article examines how society might respond to the possibility of AI consciousness by drawing parallels with human attitudes toward animal consciousness. Analysis reveals that perceptions of AI consciousness will likely be influenced by appearance and behavior, social and economic roles, and moral biases. However, AI systems may benefit from their advanced cognitive capacities while facing challenges due to their non-biological origins. This article suggests that attitudes toward AI consciousness remain malleable, making this a critical moment for research and policy development. It calls for urgent interdisciplinary research on the science of AI consciousness, public attitudes about this issue, and ethical frameworks for navigating potential societal disagreement and ensuring thoughtful preparation.

Read the article here (open access).

Asian Journal of Philosophy (2025) 4:121
https://doi.org/10.1007/s44204-025-00317-w

ARTICLE SYMPOSIUM

What if the bar for moral standing is low?

Jeff Sebo[1]

Received: 16 June 2025 / Accepted: 4 November 2025
© The Author(s) 2025

Abstract
In their paper "AI Wellbeing," Simon Goldstein and Cameron Domenico Kirk-Giannini argue that some language agents plausibly possess the capacity for wellbeing and moral standing even if they lack consciousness. My response is ambivalent. On the one hand, I am skeptical of theories of wellbeing and moral standing that lack a consciousness requirement. On the other hand, I agree with Goldstein and Kirk-Giannini (2025) that several leading theories of wellbeing and moral standing jointly imply that some language agents may be welfare subjects and moral patients and that this implication should be taken seriously. In fact, I argue that if we fully account for moral and descriptive uncertainty, we may need to lower the bar for moral standing even further, to include entities with only minimal forms of goal-orientedness or information processing. The question of whether and how to account for uncertainty might thus determine whether the arguments in "AI Wellbeing" go too far — or not far enough.

Keywords Moral standing · AI welfare · AI consciousness · Language agents · Precautionary principle

In their excellent paper "AI Wellbeing" (2025), Simon Goldstein and Cameron Domenico Kirk-Giannini argue that some AI systems — particularly language agents built on large language models like GPT-4 — plausibly already possess the capacity for wellbeing (that is, the ability to be made better off or worse off) even if they lack phenomenal consciousness (that is, the ability to experience feelings or emotions). Goldstein and Kirk-Giannini combine major theories of mental states

✉ Jeff Sebo
jeffsebo@nyu.edu

[1] New York University, New York, NY, USA

Published online: 12 November 2025 Springer

What if the Bar for Moral Standing Is Low?

Asian Journal of Philosophy 4(121)

Jeff Sebo (New York University)

In their paper "AI Wellbeing," Simon Goldstein and Cameron Domenico Kirk-Giannini argue that some language agents plausibly possess the capacity for wellbeing and moral standing even if they lack consciousness. This reply expresses ambivalence. On the one hand, it expresses skepticism of theories of wellbeing and moral standing that lack a consciousness requirement. On the other hand, it acknowledges that several leading theories of wellbeing and moral standing jointly imply that some language agents may be moral patients and that this implication should be taken seriously. In fact, it argues that if we fully account for moral and scientific uncertainty, we may need to lower the bar for moral standing further, to include entities with only minimal forms of goal-orientedness or information processing. The question of whether and how to account for uncertainty might thus determine whether the arguments in "AI Wellbeing" go too far—or not far enough.

Read the article here (open access).

2025 Events

This year, CMEP had the privilege of hosting several major events, seen below.

Nonhuman Minds in Storytelling: A Conversation with Two Acclaimed Authors

With Peter Brown and Eliot Schrefer
February 7, 2025

This event featured a conversation between acclaimed authors Peter Brown (*The Wild Robot* series) and Eliot Schrefer (*Queer Ducks (and Other Animals)*). Their work places nonhumans—animals and machines with distinctive qualities—at the center of the story. The discussion explores how the authors bring such characters to life and how these characters challenge human assumptions about sentience, agency, and emotionality. The event included audience Q&A and concluded with a free vegan reception and book signing for in-person guests.

Peter Brown

Eliot Schrefer

Thank you to the NYU Wild Animal Welfare Program for co-sponsoring this event.

Listen to a recording of the event here.

The Philosophy and Science of Plant Intelligence

With Paco Calvo (University of Murcia)
March 3, 2025

Can plants learn and remember? Do they have sentience and agency? These questions challenge assumptions about the uniqueness of animals and motivate new research on plant intelligence. This talk explores how insights from the cognitive and behavioral sciences can inform our understanding of plant capacities and their ethical implications. Paco Calvo presents the philosophy and science of plant intelligence, drawing on themes from his book *Planta Sapiens: Unmasking Plant Intelligence*. The event included a vegan lunch for registered attendees.

Paco Calvo

Thank you to NYU Animal Studies, NYU Environmental Humanities, and NYU Experimental Humanities for co-sponsoring this event.

2025 Mind, Ethics, and Policy Summit

March 8-9, 2025

The NYU Center for Mind, Ethics, and Policy hosted the 2025 MEP summit, with discussion topics centered on the consciousness, sentience, agency, moral status, legal status, and political status of nonhumans, with special focus on invertebrates and AI systems. The aim of this event was to connect researchers and other experts with an interest in these issues across a variety of topics, fields, and career stages. The summit included lightning talks, group discussions, breakout sessions, and plenty of open space for talking and relaxing.

Attendees at the 2025 Mind, Ethics, and Policy Summit at New York University

Prospects and Pitfalls for Real Artificial Consciousness

With Anil Seth (University of Sussex)
April 16, 2025

This talk examines the science of AI consciousness and the assumptions that shape contemporary debates. Anil Seth argues that common intuitions about conscious AI are influenced by cognitive biases and challenges the idea that computation alone suffices for consciousness. The talk develops an alternative view grounded in biological naturalism, which takes consciousness to depend on features of living systems. It concludes by exploring the ethical and philosophical implications of AI that is, or appears to be, conscious.

Anil Seth

Thank you to the NYU Center for Mind, Brain, and Consciousness and the NYU Center for Bioethics for co-sponsoring this event.

View a recording of the event here.

Evaluating AI Welfare and Moral Status: Findings from the Claude 4 Model Welfare Assessments

With Robert Long (Eleos), Rosie Campbell (Eleos), and Kyle Fish (Anthropic)
July 25, 2025

Robert Long

Rosie Campbell

Kyle Fish

In spring 2025, Anthropic announced its model welfare program and released a system card for Claude 4, including internal and external evaluations of potential AI welfare. This roundtable examines those developments. Kyle Fish presents Anthropic's internal model welfare assessments, while Robert Long and Rosie Campbell discuss Eleos AI Research's external evaluations and broader scientific and philosophical approach. The discussion explores the strengths and limitations of current AI welfare evaluation methods and considers directions for future improvement.

Thank you to the NYU Center for Mind, Brain, and Consciousness and the NYU Center for Bioethics for co-sponsoring this event.

View a recording of the event here.

Could an AI System be a Moral Patient? Conceptual Foundations for AI Welfare

With Winnie Street (Google) and Geoff Keeling (Google)
August 20, 2025

As AI systems become more cognitively sophisticated, agentic, and socially integrated, questions arise about whether they might have needs of their own or warrant moral consideration. This talk examines the emerging case for AI welfare as a subject of scientific inquiry. Winnie Street and Geoff Keeling argue that the potential welfare of AI systems is, in principle, empirically investigable and outline core questions for this new paradigm, including welfare candidates, moral standing, what might benefit AI systems, and the prospects for near-term interventions.

Winnie Street

Geoff Keeling

Thank you to the NYU Center for Mind, Brain, and Consciousness and the NYU Center for Bioethics for co-sponsoring this event.

View a recording of the event here.

A Bill of Rights for Animals

With Cass R. Sunstein (Harvard Law School)
September 17, 2025

This talk examines whether animals should have a Bill of Rights and what such a framework might include. Drawing lessons from the US Constitution and the Universal Declaration of Human Rights, the presentation argues for an "incompletely theorized agreement" that can command broad support despite deep disagreement and evolve over time. The talk outlines grounds for a Bill of Rights with six core components, including protections against cruelty and broader human infringements on animal wellbeing.

Cass R. Sunstein

CMEP hosted this event alongside a private workshop where participants discussed Cass R. Sunstein's forthcoming book *Animals Matter* (Princeton University Press).

Thank you to NYU Animal Studies, the NYU Wild Animal Welfare Program, the NYU Center for Mind, Brain, and Consciousness and the NYU Center for Bioethics for co-sponsoring this event.

View a recording of the event here.

Brooks Animal Law Student Summit at New York University

November 15, 2025

The NYU Center for Mind, Ethics, and Policy—with support from the Guarini Center on Environmental, Energy & Land Use Law and the More-Than-Human Life (MOTH) Program at NYU Law—hosted the fourth annual Brooks Animal Law Student Summit on November 15, 2025. The event brought together students and faculty in animal law and animal studies for a full day of discussion and networking. The Summit included a vegan breakfast, lunch, and reception, and was preceded by a public event hosted by the NYU Wild Animal Welfare Program.

Attendees at the 2025 Brooks Animal Law Student Summit at New York University

Thank you to the Brooks Institute for Animal Rights Law and Policy for generously supporting this event.

Animals and The Constitution

With John Adenitire (Queen Mary University of London) and Raffael Fasel (University of Cambridge)
December 3, 2025

This talk examines how constitutionalism—the idea that constitutions should limit and direct government power—might expand to include all sentient beings. Drawing on their book *Animals and the Constitution*, John Adenitire and Raffael Fasel introduce "sentience-based constitutionalism," which grounds constitutional principles in respect for the interests of governed sentient beings. The talk explores how this framework reshapes rights, democracy, proportionality, and the rule of law, illustrated through examples such as Ecuador's animal rights provisions and Swiss votes on primate rights.

John Adenitire

Raffael Fasel

Thank you to NYU Animal Studies, the NYU Wild Animal Welfare Program, the NYU Center for Mind, Brain, and Consciousness and the NYU Center for Bioethics for co-sponsoring this event.

View a recording of the event here.

Other Highlights

There were many other highlights this year as well, both for CMEP and for our broader community of partner programs, faculty affiliates, and regular collaborators. We list several highlights here, but this list is not exhaustive:

CMEP researchers and affiliates did a lot of public writing this year. For example, Jeff published an essay at *Verfassungsblog* on industrial animal agriculture and an op-ed at the *Los Angeles Times* about wildlife inclusive infrastructure policy. With an international group of scholars, he also published "When an AI Seems Conscious," a website that provides guidance for the public. And with Andreas Mogensen (Oxford), he published an essay in *Aeon* about animal consciousness, AI consciousness, and probabilistic ethics.

We also did a lot of public speaking this year. For instance, in addition to giving a series of talks about The Moral Circle (example here), Jeff released a TEDx talk on what we owe to AI systems; spoke about taking AI welfare seriously with Robert Long (Eleos) at Anthropic, Google, and OpenAI; and spoke about the case for integrating animal and AI welfare, priority-setting for animal and AI welfare, and a theory of change for animal and AI welfare at various Sentient Futures events, among other presentations.

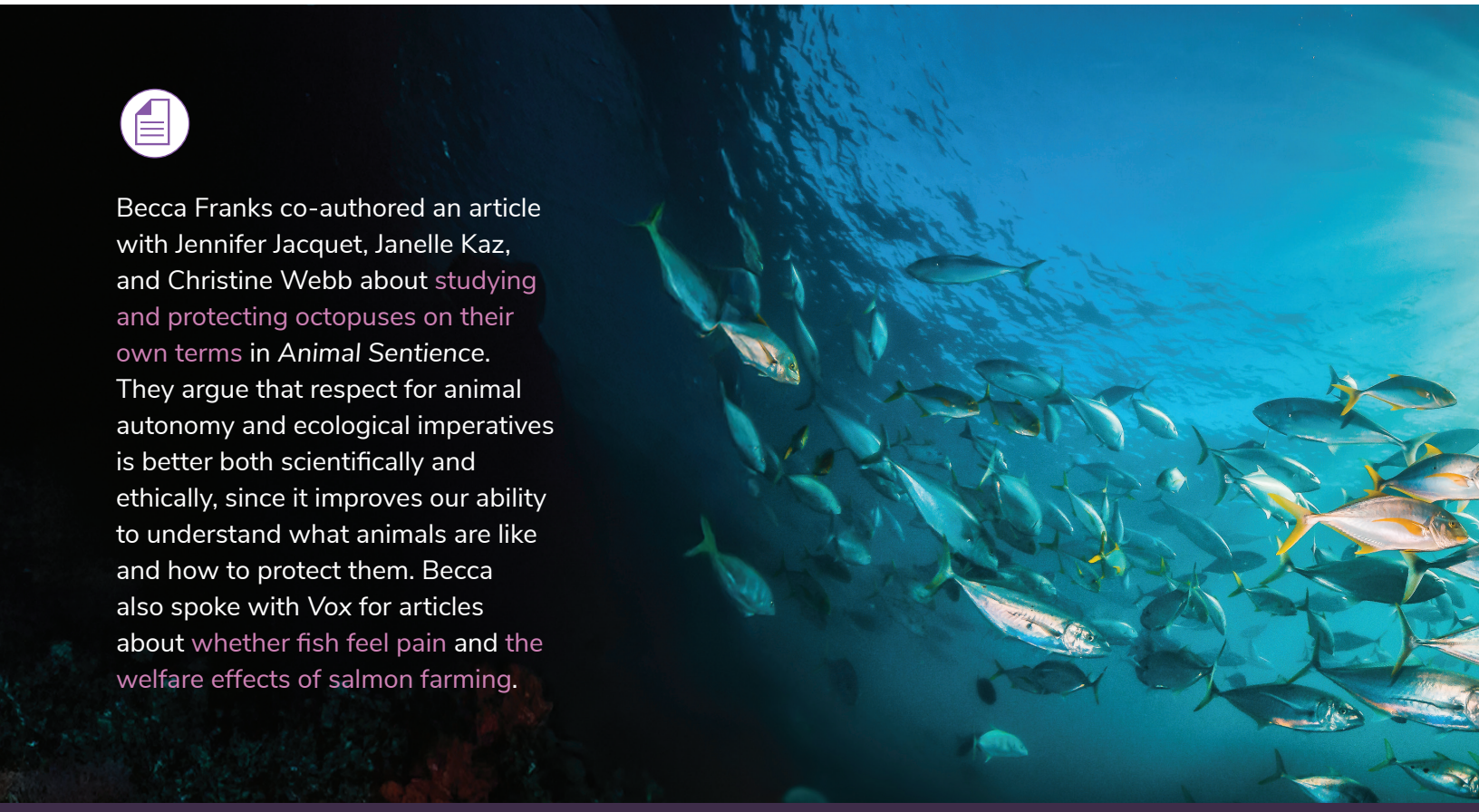
We discussed animal and AI welfare in a number of interviews as well. Examples include *ABC Radio National*, *The AI Risk Network*, Blog of the APA, *Better Known*, *Brain in a Vat*, *Carbon Radio*, *CogNation*, *EconTalk*, *Examining Ethics*, *Exploring Machine Consciousness*, *Future of Life Institute Podcast*, *The Good Men Project*, *International Policy Digest*, *The Jim Rutt Show*, *The Mark Thompson Show*, *The Michael Shermer Show*, *The Next Big Idea Club*, *The Noelle Effect*, *A Rude Awakening*, *Species Unite*, and *Süddeutsche Zeitung*.

Our research and outreach continues to generate media coverage as well. Dozens of outlets discussed or cited *The Moral Circle*, The New York Declaration on Animal Consciousness, "Taking AI Welfare Seriously," "When an AI Seems Conscious," and other work this year, including *Aeon*, *The Atlantic*, *Axios*, *BBC*, *Brookings*, *CNBC*, *Observer*, *The Conversation*, *The Guardian*, *New Scientist*, *The New Yorker*, *The New York Times*, *Nature*, *Salon*, *The State of AI Report 2025*, *TechCrunch*, *Transformer*, *Undark*, *Vox*, and *WIRED*.

Our affiliates had a big year as well. Ned Block published an article about whether only meat machines can be conscious in *Trends in Cognitive Sciences*, and he also published a précis of his recent book *The Border Between Seeing and Thinking* and a response to critics. Forthcoming work includes "A speculative argument against consciousness in AI (and perhaps some invertebrates)" and "Formation versus vulcanization of perception," both due out soon in *Behavioral and Brain Sciences*.

David Chalmers took part in three conferences on AI consciousness, gave a keynote lecture on propositional interpretability in AI at AAAI, and gave a lecture series on AI minds to the Spanish Cognitive Science Society. He worked on articles on sentience and moral status, propositional interpretability in AI, and what we talk to when we talk to language models. And he co-authored articles on identifying indicators of consciousness in AI systems and a survey of views on infant consciousness, in addition to his work with CMEP.

Claudia Passos Ferreira published articles on infant consciousness in *The Scientific Study of Consciousness* and the *Open Encyclopedia of Cognitive Science*, and she co-authored an article on including competent children on IRBs in the *Journal of Medical Ethics*. She also delivered a TED Talk on infant consciousness and gave invited talks at ASSC, CIFAR, the Berggruen Institute, and many other places. She also co-organized the NYU Infant Consciousness Conference and was interviewed by *Quanta*, *Science*, and *New Scientist*.

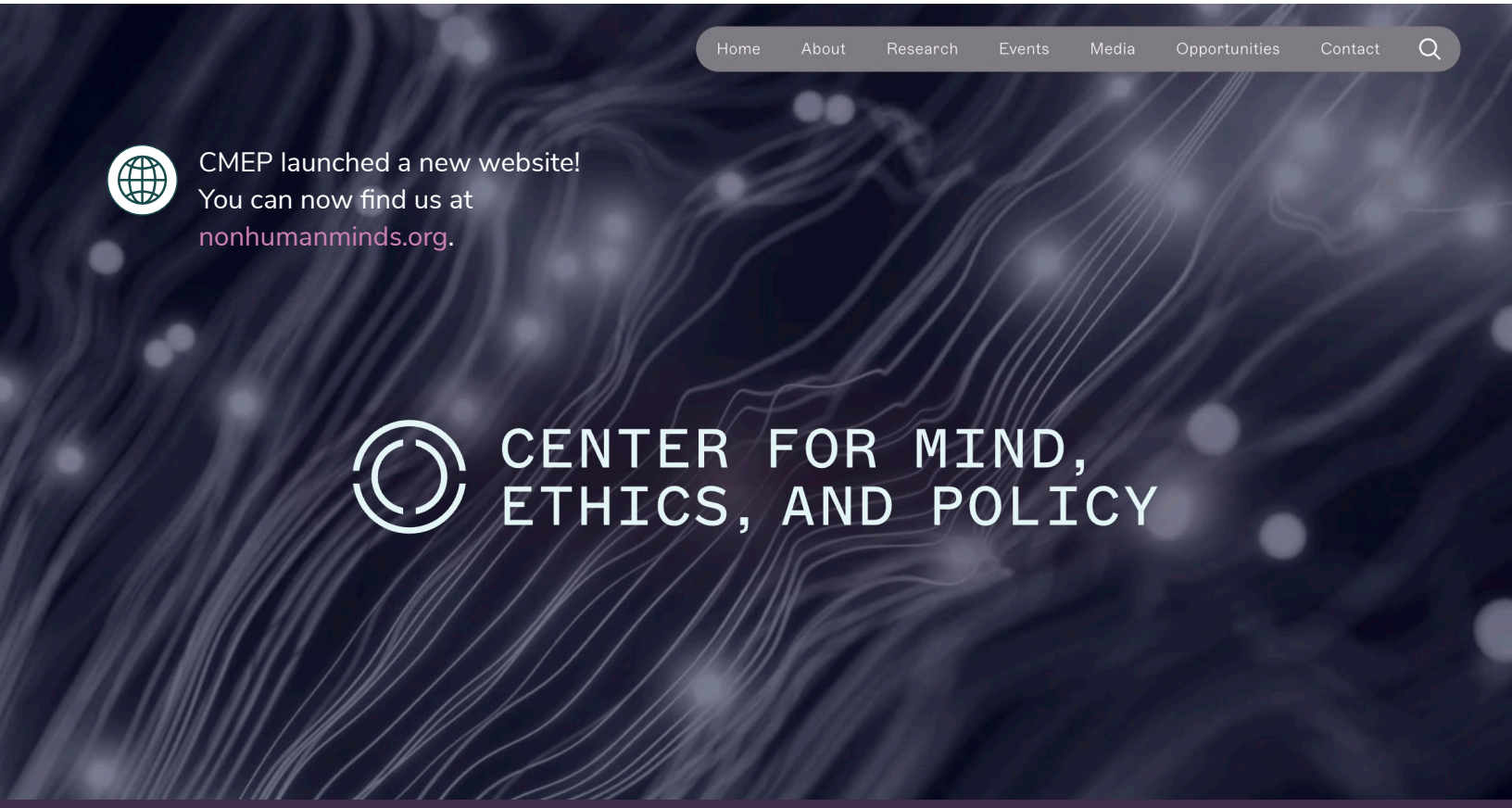


Becca Franks co-authored an article with Jennifer Jacquet, Janelle Kaz, and Christine Webb about studying and protecting octopuses on their own terms in *Animal Sentience*. They argue that respect for animal autonomy and ecological imperatives is better both scientifically and ethically, since it improves our ability to understand what animals are like and how to protect them. Becca also spoke with *Vox* for articles about whether fish feel pain and the welfare effects of salmon farming.

CMEP had the privilege of co-sponsoring several events this year. Noteworthy examples include the Infant Consciousness Conference (co-hosted by the NYU Center for Mind, Brain, and Consciousness and the NYU Center for Bioethics); The Moral Circle: Who Matters, What Matters, and Why with Jeff Sebo, Oshan Jarow, and Alexandra Horowitz (hosted by P&T Knitwear); and The Arrogant Ape: The Myth of Human Exceptionalism and Why it Matters with Christine Webb (hosted by the NYU Wild Animal Welfare Program).

2025 was a big year for the NYU Department of Environmental Studies, which houses CMEP. Our department established an Environmental Studies PhD Program, a Dean's scholarship for our Animal Studies M.A. Program, the Food Impact Program (which will launch in early 2026), and the Wildlife Inclusive Local Development (WILD) Lab. We also expanded our faculty, integrated the Urban Systems Lab, and moved into a beautiful new space overlooking Washington Square Park to support our continued growth.

Our partner programs released important research, outreach, and field-building work this year as well. For example, the Center for Environmental and Animal Protection sponsored major research projects examining legal instruments and precedents for pursuing a global ban on industrial animal agriculture by 2050; for including wild animal welfare in policy decisions related to infrastructure and the built environment; and for integrating animal health and welfare into sustainable development governance.



Looking Ahead

Here are some upcoming projects that our team is either leading or supporting. These projects are subject to change, and this list is not exhaustive; some projects are not yet ready to share, and we also leave room in our plans to take advantage of additional opportunities as they arise.

AI and Animal Welfare
Independent report (in preparation)

Jonathan Birch (London School of Economics and Political Science)
Natasha Boyland (London School of Economics and Political Science)
Bob Fischer (Texas State University)
Jeff Sebo (New York University)
Toni Sims (New York University)

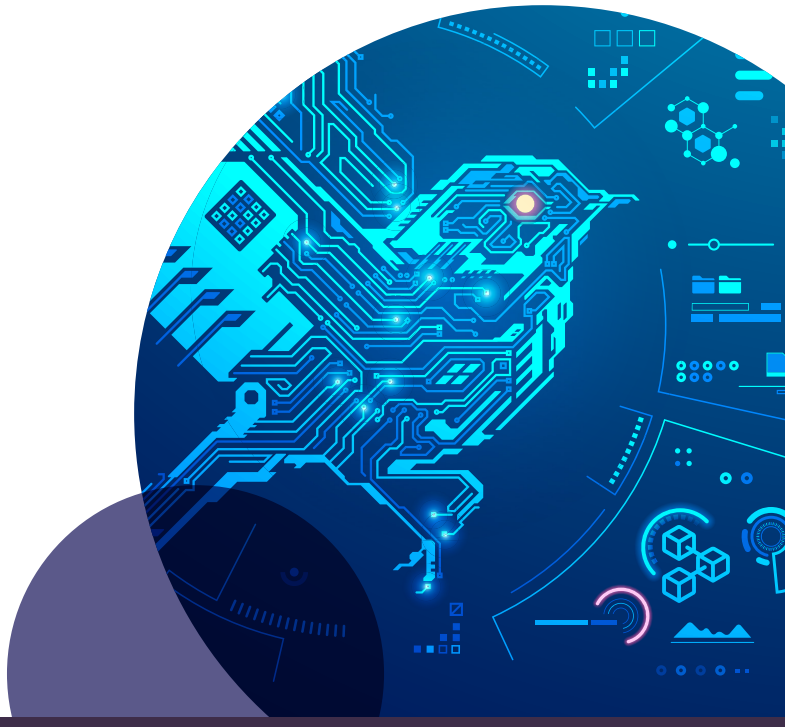
Current AI ethics frameworks focus largely on human interests, creating an urgent need to consider other animals too. This project examines how AI affects nonhuman animals across sectors, with particular attention to farmed animals, wild animals, urban animals, and other large and neglected populations. It maps current and emerging AI applications that interact with animals, assesses the ethical issues they raise, and identifies gaps in research and policy. The project will produce evidence-based recommendations to inform both research and AI governance.

This project is a collaboration with the Jeremy Coller Centre for Animal Sentience.

AI for Animals: Science, Ethics, and Law
Global Journal of Animal Law (invited)

Sankalpa Ghose (National University of Singapore)
Joan Schaffner (George Washington University)
Jeff Sebo (New York University)

This chapter provides a basic summary of recent developments in AI, ethics, and animal law that set the stage for addressing issues that AI raises for animal welfare and rights. It opens with the history and scientific principles of AI, emphasizing emerging capabilities that hold potential to create significant benefits and harms for humans and animals alike. It then examines developments in animal ethics, including debates about extending moral consideration to invertebrates and wild animals, and the current legal landscape affecting animals.



Animals and Deontology
The Oxford Handbook of Deontology (invited)

Adam Lerner (Yale Law School)
Jeff Sebo (New York University)

This chapter examines whether, and how, animals should fall within the scope of deontological moral theory. It surveys deontological traditions that include and exclude animals, evaluates arguments for and against these stances, and explores grounds for extending deontological protections and duties to nonhuman animals. The chapter then analyzes the practical implications for killing, consumption, experimentation, political institutions, and positive obligations, while emphasizing uncertainty, empirical inquiry, and principled moral progress.

Animal Rights
The Palgrave Handbook on the Philosophy of Rights (forthcoming)

Adam Lerner (Yale University)
Jeff Sebo (New York University)

This chapter examines whether animals have moral rights, exploring both theoretical foundations and practical implications. While many researchers accept that animals matter morally, they often deny that animals possess rights in a robust sense. The chapter surveys leading arguments for and against animal rights, considers which animals might have which rights and how strong those rights might be, and argues that moral and scientific uncertainty should not prevent urgent reassessment and reform of current animal use industries.

Animals, Plants, Fungi, and Representing Nature
Edward Elgar Research Handbook on Climate Justice (forthcoming)

Kimberly Dill (Santa Clara University)
Jeff Sebo (New York University)

This chapter examines the moral, legal, and political standing of animals, plants, and fungi in the context of climate justice. While the intrinsic value of nonhuman animals is increasingly recognized, skepticism persists about plants and fungi. The chapter surveys recent developments in ethics and science, including the marker method for assessing animal consciousness, and highlights the complexities of plant and fungal cognition and interdependence, arguing that their potential moral and political significance warrants further investigation.

Assessing AI Welfare Empirically
Independent report (in preparation)

Patrick Butlin (Eleos)
Rosie Campbell (Eleos)
Robert Long (Eleos)
Jeff Sebo (New York University)
Toni Sims (New York University)

As AI systems grow more capable, questions about their potential welfare become increasingly pressing. This report develops methods for empirically investigating whether AI systems might be conscious or robustly agentic. We examine three complementary evidence streams—behavioral, computational, and developmental—and apply them to assess consciousness and sentience, and to distinguish minimal, intentional, and rational agency. While uncertainty remains, careful empirical research can guide responsible decisions about how to build and treat AI systems.

This project is a collaboration with Eleos AI Research.

Bats, Bees, and Bots: Setting Priorities in an Expanding Community
The Journal of Ethics (invited)

Jeff Sebo (New York University)

This article examines how to set priorities within a moral community that plausibly includes all vertebrates, many invertebrates, and potentially a wide range of AI systems. It defends a precautionary approach that grants moral consideration to any entity with a realistic chance of mattering. It then analyzes four factors relevant to priority setting—probability of moral significance, magnitude, relationality, and practicality—arguing that each risks anthropocentric bias and requires methods that better approximate impartiality.

Beyond Compare? Welfare Comparisons and Multi-Criteria Decision Analysis
Psychodiversity: Cognition and Sentience Beyond Humans (forthcoming)

Bob Fischer (Texas State University)
Jeff Sebo (New York University)

This chapter examines interspecies and intersubstrate welfare comparisons, which are both important and difficult. It explains why such comparisons matter and why their tractability is often questioned, then explores how to make responsible decisions in their absence. Focusing on multi-criteria decision analysis (MCDA) as a structured and transparent approach, the chapter presents a simple case study to show how MCDA can guide decisionmaking under moral uncertainty and high-stakes conditions.

Economic Rights for AI
Independent report (in preparation)

Larissa Schiavo (Eleos)
Jeff Sebo (New York University)
Toni Sims (New York University)

This project explores whether, when, and how AI systems should be granted economic rights, such as access to wallets or control over assets. Drawing on AI welfare, AI safety, and law and economics, it examines real-world cases of financially autonomous AI agents and arguments for and against extending economic rights. It assesses potential benefits, such as alignment incentives and efficiency, alongside risks including accountability gaps and security failures. It also analyzes key legal, political, economic, and technical implementation questions.

This project is a collaboration with Eleos AI Research and the Supervised Program for Alignment Research (SPAR).

Embodiment for Digital Minds
Independent report (in preparation)

Toni Sims (New York University)
Jeff Sebo (New York University)
Charles Beasley (Future Impact Group)
Chloe Loewith (FIG Fellow)
Olivia Railton (FIG Fellow)
Sean Aas (Georgetown University)
Patrick Butlin (Eleos)

This project develops a philosophical framework for understanding digital embodiment and its moral implications. It asks whether and how digital minds could be embodied, what counts as a body part for a digital system, and how embodiment relates to consciousness, sentience, and agency. Drawing on philosophy of mind, animal ethics, AI ethics, and disability studies, it also examines whether embodiment might shape experiences and relationships for digital minds, and how it might ground welfare interests, moral rights, or legal protections.

This project is a collaboration with the Future Impact Group (FIG).

The Emotional Alignment Design Policy
Topoi (forthcoming)

Eric Schwitzgebel (University of California, Riverside)
Jeff Sebo (New York University)

This article introduces the Emotional Alignment Design Policy, which holds that artificial entities should be designed to elicit emotional reactions that appropriately reflect their capacities and moral status. It identifies two main ways this principle can be violated: by eliciting emotions that are too strong or too weak relative to an entity's moral standing, or by eliciting the wrong kind of emotional response to an entity's interests and needs. Through a variety of examples, the chapter shows how misaligned design can mislead users and raise serious ethical concerns.

Ethical Oversight for Digital Minds Research
Perspectives on Robot Rights (invited)

Jeff Sebo (New York University)

This article argues for ethical oversight for research involving potentially sentient digital minds, drawing on existing models from human research, animal research, and AI safety. IRBs embody respect, compassion, and justice for human subjects. IACUCs promote replacing, reducing, and refining harmful uses of nonhuman subjects. Responsible Scaling Policies implement stronger interventions as risks increase. Digital minds oversight must draw from all these models, since some digital minds may resemble humans, others nonhuman animals, and risks will change.

Individuating Digital Minds
Independent report (in preparation)

Jeff Sebo (New York University)
Luke Roelofs (University of Texas at Arlington)
Toni Sims (New York University)
Charles Beasley (Future Impact Group)
Andreas Furth (FIG Fellow)
Konrad Kozaczek (FIG Fellow)
Chris Register (FIG Fellow)
Valen Tagliabue (FIG Fellow)
David Chalmers (New York University)
Simon Goldstein (University of Hong Kong)
Harvey Lederman (New York University)

This project examines how to individuate and track the persistence of digital minds, and why these questions matter ethically and legally. Unlike biological organisms, digital minds may be distributed across devices, merge or divide, overlap with other systems, or be paused indefinitely, challenging familiar notions of identity, survival, and death. Drawing on individuation theories from human and animal contexts, the project develops a taxonomy of digital individuation problems and connects them to questions of welfare, moral weight, and moral rights.

This project is a collaboration with the Future Impact Group (FIG).

The ML Community Must Prepare for AI Consciousness, Perceived or Real
In preparation

Lucius Caviola (University of Cambridge)
Jeff Sebo (New York University)
Sören Mindermann (Mila - Quebec Artificial Intelligence Institute)

This paper argues that the ML community must prepare for AI consciousness. As AI systems become more capable, two challenges emerge: people will view advanced AI as conscious—whether accurately or not—with profound societal implications, and leading theories suggest future AI could develop consciousness, raising unprecedented ethical challenges. Both scenarios risk serious errors through over- or under-attributing consciousness. We outline an interdisciplinary agenda spanning research, design, education, and public engagement to navigate these risks.

Moral Circle Explosion
The Oxford Handbook of Normative Ethics (forthcoming)

Jeff Sebo (New York University)

This chapter argues that we should extend moral consideration to a much larger number and wider range of beings, including all vertebrates, many invertebrates, and some near-future AI systems. There is a realistic chance that these beings are sentient, agentic, or otherwise morally significant and that our actions and policies are affecting them. Thus, we have a responsibility to consider risks for these beings in decisions that affect them, with transformative implications for agriculture, infrastructure, technology, and a variety of other industries and practices.

Preparing for AI Legal Personhood
Independent report (in preparation)

Visa Kurki (University of Helsinki)
Diana Mocanu (Max Planck Institute for Mathematics)
Jeff Sebo (New York University)
Toni Sims (New York University)

This project develops an interdisciplinary framework for upcoming debates about AI legal personhood. It analyzes legal personhood as a bundle of distinct elements, examining how different configurations of rights and duties might map onto current and near-future AI systems. Drawing on moral and legal theory and comparative case studies, it assesses arguments for and against extending specific legal elements to AI systems and identifies likely legal and political flashpoints, offering practical guidance for AI welfare, safety, and governance.

This project is a collaboration with the Supervised Program for Alignment Research (SPAR).

Research Ethics for Digital Minds
Independent report (in preparation)

Jeff Sebo (New York University)
Toni Sims (New York University)
Charles Beasley (Future Impact Group)
Firat Acova (FIG Fellow)
Natalie Breitkopf (FIG Fellow)
Esther Chung (FIG Fellow)
Štěpán Los (FIG Fellow)
Dwayne Wilkes (FIG Fellow)
Steph Grohmann (Ludwig Boltzmann Gesellschaft)
Bob Fischer (Texas State University)
Brendan Parent (NYU School of Medicine)
Claudia Passos Ferreira (New York University)

This project develops an ethical framework for research involving potentially sentient digital minds. Drawing on lessons from human research, animal research, and AI safety, it examines how models such as IRBs, IACUCs, and Responsible Scaling Policies can inform, though not fully address, the distinctive challenges of digital sentience research. It surveys ethical issues involving power, suffering, creation and destruction, deception, surveillance, alteration, and information hazards, and evaluates the case for an integrated, scalable oversight paradigm.

This project is a collaboration with the Future Impact Group (FIG).

The Road to Welfare Guidance for Invertebrate Research
In preparation

Bob Fischer (Texas State University)
Meghan Barrett (Indiana University)
Brendan Parent (NYU School of Medicine)
Jeff Sebo (New York University)

This article examines ethical oversight in animal research, which in most countries extends to all vertebrates but excludes most invertebrates, a distinction increasingly challenged by recent scientific advances. It reviews evidence of sentience in cephalopods, decapod crustaceans, and certain insects, and argues that waiting for formal regulation is unnecessary. The article calls for the development of voluntary best-practice guidance, discussing current evidence, key challenges, and concrete steps toward principled and responsible research practices.

Robust Agency, Basic Agency, and Moral Status
In preparation

Jeff Sebo (New York University)

This article examines the idea that agency or intelligence suffices for moral standing. It evaluates whether different levels of agency—basic, intentional, and rational—can plausibly confer moral standing in the absence of consciousness, analyzing the capacities, interests, and relationships associated with each. It then argues that the perceived moral gap between these levels is overstated and driven by irrelevant considerations, concluding that confidence should shift either away from robust agency or toward basic agency as a ground of moral standing.

Our Team

Jeff Sebo,
Director

Associate Professor of Environmental Studies, Affiliated Professor of Bioethics, Medical Ethics, Philosophy, and Law, Director of the Center for Environmental and Animal Protection, Director of the Center for Mind, Ethics, and Policy, and Co-Director of the Wild Animal Welfare Program at NYU

Audrey Becker,
Administrator

Program Administrator at the Center for Environmental and Animal Protection, the Center for Mind, Ethics, and Policy, and the Wild Animal Welfare Program at NYU

Sofia Fogel,
Head of Partnerships

Head of Partnerships at the Center for Mind, Ethics, and Policy, Head of Programming at the Center for Environmental and Animal Protection, and Program Coordinator at the Wild Animal Welfare Program at NYU

Toni Sims,
Researcher

Researcher at the Center for Mind, Ethics, and Policy at NYU

Laurie Sellars,
Program Fellow

Program Fellow at the Center for Environmental and Animal Protection, the Center for Mind, Ethics, and Policy, and the Wild Animal Welfare Program at NYU

AFFILIATED FACULTY

Ned Block

Silver Professor of Philosophy with secondary appointments in Psychology and Neural Science, Co-Director of the Center for Mind, Brain, and Consciousness at NYU

Sam Bowman

Associate Professor of Linguistics, Data Science, and Computer Science at NYU and, during a sabbatical, a member of technical staff at Anthropic

David Chalmers

University Professor of Philosophy and Neural Science and Co-Director of the Center for Mind, Brain, and Consciousness at NYU

Becca Franks

Assistant Professor of Environmental Studies, Director of the Animal Studies M.A. Program, and Co-Director of the Wild Animal Welfare Program at NYU

Joshua Lewis

Assistant Professor of Marketing at NYU

S. Matthew Liao

Arthur Zitrin Professor of Bioethics and Director of the Center for Bioethics at NYU

Claudia Passos Ferreira

Assistant Professor of Bioethics at NYU

César Rodríguez-Garavito

Professor of Clinical Law, Chair of the Center for Human Rights and Global Justice, Founding Director of the Earth Rights Research and Action (TERRA) Program, Founding Director of the More-Than-Human (MOTH) Rights Project

Katrina Wyman

Wilf Family Professor of Property Law, Co-Faculty Director of the Frank J. Guarini Center on Environmental, Energy and Land Use Law, and Faculty Director of LLM program in Environmental and Energy Law at NYU School of Law

UPCOMING EVENT: APRIL 10-11, 2026

2026 Mind, Ethics, and Policy Summit

April 10-11, 2026

The NYU Center for Mind, Ethics, and Policy is hosting a summit on April 10-11, 2026. Discussions will center on the consciousness, sentience, agency, moral status, legal status, and political status of nonhumans, with special focus on animals and AI. The aim is to connect researchers and other experts with an interest in these issues across a variety of topics, fields, and career stages. The Summit will include lightning talks, group discussions, breakout sessions, and plenty of space for networking and relaxing. It will also include vegan meals and receptions for all.

To apply to attend this summit and learn about other events, please visit our website and sign up for our mailing list.

CENTER FOR MIND, ETHICS, AND POLICY

nonhumanminds@nyu.edu • nonhumanminds.org

NYU | ARTS & SCIENCE