International
Ecolinguistics
Association

Article

AI world models, ecolinguistics, and ecosophical prediction: Whose worlds get predicted?

Jorge Vallego*

Abstract

Today's artificial intelligence systems, dominated by generative models, increasingly mediate human relationships with nature. These systems learn about the natural world from human-generated text (Bender et al., 2021), inheriting anthropocentric framings and the silences surrounding non-human experiences (Stibbe, 2015). From an ecolinguistic perspective, they ingest the Conceptual Umwelt of documented humanity (Tønnessen, 2015) while remaining blind to embodied ecological processes. World models, AI systems that learn by predicting environmental dynamics from multimodal sensory observations rather than text (Ha & Schmidhuber, 2018; LeCun, 2022), offer an alternative. However, they too reproduce anthropocentrism: developers' choices about which sensors to deploy, what data to collect, and which outcomes to optimise encode human priorities as the measure of environmental significance. This paper introduces ecosophical prediction, treating ecosystem flourishing (Næss, 1989) as the primary optimisation target rather than extrinsic values such as economic growth. Systematically designed world models could predict futures centred on ecological resilience and multispecies thriving, attending to signals that constitute other organisms' Umwelts (von Uexküll, 1934/2010) and the affordances environments provide to different species (Gibson, 1979). This would enable sensitivity to ecosemiotic relations typically overlooked in anthropocentric systems (Kull, 2010). Ecosophical prediction requires architectural commitments: modelling whose sensory worlds matter, treating which signals as data, and optimising which predicted futures. The H4rmony Project has demonstrated that ecosophical principles can guide generative models towards ecologically aware outputs (Vallego, 2023, 2024). We argue that world models must similarly be ecosophy-guided, but require alignment embedded from the outset in their perceptual modalities and predictive targets.

Keywords: world models; ecosemiotics; Umwelt; ecolinguistics; AI alignment

* Project Lead, The H4rmony Project | Research Associate, University of Gloucestershire, UK. Email: jorge.vallego@theh4rmonyproject.org

1. Introduction

In the contemporary discourse on artificial intelligence, we stand at a juncture between the symbol and the sensory. The prevailing paradigm, dominated by large language models (LLMs), operates fundamentally as a stochastic processing of human text (Bender et al., 2021). These systems, famously described as "stochastic parrots", learn linguistic forms observed in vast training data without inherent reference to the physical reality those forms denote. For the linguist and the ecolinguist, this presents a profound limitation: these models do not perceive the world; they perceive our discourse about the world. They are trapped within the documented and digitised Conceptual Umwelt of humanity (Tønnessen, 2015): specifically, the English-speaking internet inheriting both the grammatical structures of our languages and their ecological blind spots; the anthropocentrism that renders the non-human world passive; the silences regarding ecological processes that lack economic utility; and the metaphors that frame nature as a resource rather than a community of subjects (Stibbe, 2015). This has been highlighted and addressed by the H4rmony Project, which seeks to embed ecocentrism and sustainability in LLMs through a clearly articulated ecosophy (Vallego, 2023, 2024). Although its impact on mainstream AI models remains limited at this stage, it offers a novel and meaningful contribution to the field. (A glossary of terminology from biosemiotics, ecolinguistics, and artificial intelligence appears in the Appendix.)

A parallel trajectory in AI research suggests a move away from the disembodied symbol towards the embodied simulation, and we argue that this should also be addressed from an ecolinguistic perspective and follow a clear ecosophy. This approach, known as "world models", posits that intelligence arises not from the statistical probability of language generation, but from the ability to simulate and predict the dynamics of the physical environment (Ha & Schmidhuber, 2018; LeCun, 2022). By shifting the primary input from text and text-labelled signals to multimodal sensory data — vision primarily, but with the potential to extend to bioacoustics, chemical signalling, and haptics — we encounter the theoretical possibility of a machine that constructs an Umwelt, or a subjective self-world (von Uexküll, 1934/2010). Yet it is necessary to recognise from the outset that this constructed Umwelt is still shaped by human design decisions about what to measure, how to measure it, and what environments to prioritise.

This paper explores the implications of World Models in terms of environmental ethics, while maintaining awareness of their limitations. It argues that if AI systems can be designed deliberately to go beyond sound and vision and tap into the diverse semiotic channels of the biosphere, for instance the "chemical internet" of plant communication (Dudareva et al., 2013; Heil & Karban, 2010), the bioacoustic networks of marine life (Farina & Gage, 2017), and the thermodynamic gradients of ecosystems, and could potentially become more attuned to ecosemiotic relations than systems trained on human language alone (Kull, 2010; Maran, 2020), then such systems could also become participants in the ecosemiotic network (Kull, 2010; Maran, 2020). However, guiding this process requires us to understand and carefully design the Machine Umwelt: the unique, mathematically-defined phenomenology of a system that perceives reality through high-dimensional vector spaces and optimises for objectives that must be explicitly aligned with ecological health. Furthermore, it requires a rigorous critique of "common sense", since world model researchers argue in favour of endowing machines with human-like intuitions such as intuitive physics and cause-and-effect (LeCun, 2022). Ecolinguistics warns us that human common sense is often ecologically problematic, shaped by industrial ideologies that normalise environmental degradation (Stibbe, 2015). Therefore, this paper proposes the framework of ecosophical prediction: the design of world models whose common sense is grounded not in anthropocentric utility, but in the logic of ecosystem resilience (Næss, 1989). We will discuss this through basic ideas from semiotic theory and AI architecture, while addressing the practical and ethical challenges of implementation.

2. The theoretical landscape

To understand how a machine might be designed to engage with the living world more ecologically, we must first rethink the anthropocentric monopoly on meaning. One of the theoretical tools for this task is found in biosemiotics, tracing back to the early 20th century.

2.1. The functional circle

The foundational concept for this inquiry is the Umwelt, introduced by the biologist Jakob von Uexküll. Von Uexküll revolutionised biology by asserting that organisms are not machines reacting mechanically to stimuli, but subjects interpreting signs (von Uexküll, 1934/2010). Every organism inhabits a subjective universe (Umwelt), which is a closed unit of perception and action carved out of the objective surroundings.

Von Uexküll formalised this relationship through the functional circle, a feedback loop that connects the organism to its world (von Uexküll, 1934/2010). This circle has two arms:

- **Perceptual world:** The specific subset of signals an organism's sensory apparatus can detect. For the tick, this is limited to the scent of butyric acid (mammalian sweat), the warmth of blood, and the tactile sensation of hair (von Uexküll, 1934/2010). The tick is blind and deaf to the rest of the meadow; these signals constitute its entire meaningful world.

- **Operational world:** The set of actions the organism can perform to change its environment (e.g., dropping, biting). These actions alter the perceptual signals, creating a feedback loop.

The key insight from a linguistic point of view is that objects do not exist as neutral entities; they exist as meaning-carriers defined by the functional tone they acquire within the

Umwelt (von Uexküll, 1934/2010). A stem is a path for the ant, food for the cow, and structure for the plant itself. Meaning is relational and species-specific.

2.2. Ecosemiotics

Kalevi Kull expanded von Uexküll's organism-centric view to the ecosystem level, defining the ecosystem not merely as a flow of nutrients and energy but as a web of semiosic bonds (Kull, 2010). In this perspective, an ecosystem becomes a vast communication network whose stability depends on the reliable interpretation of signs, what Kull calls ecological codes. These codes operate through consortia, the associations of species linked by sign relations, such as the mycorrhizal fungi and the tree roots they communicate with (Kull, 2010). They also include the regular correspondences that structure ecological interactions, like the relation between day length and bird migration or between flower colour and pollinator vision (Kull, 2010).

When such codes are disrupted, for instance by noise pollution masking bird calls or chemical pollutants mimicking hormones, the result is semiocide, the destruction of the semiotic conditions necessary for life (Puura, 2002; Uslu, 2020). Puura's concept has become increasingly important for understanding ecological destruction as physical damage, including the erasure of meaning-making relationships.

2.3. The tripartite Umwelt

Human beings are unique in that our Umwelt is heavily mediated by language. Morten Tønnessen proposes a tripartite Umwelt model for humans (Tønnessen, 2015):

- **The Core Umwelt:** Our direct sensory experience (the smell of rain, the feeling of heat).
- **The Mediated Umwelt:** Experience filtered through technology and memory.
- **The Conceptual Umwelt:** Reality as structured by language, narrative, and ideology. (Tønnessen, 2015)

Ecolinguistics suggests that the human Conceptual Umwelt has become dissociated from the Core Umwelt. We live in a world of stories, economic growth, consumerism and nature-as-resource that often contradict the physical realities of the ecosystems (Stibbe, 2015). This dissociation is what current text-based AI inherits. By training on the internet content, LLMs ingest the Conceptual Umwelt of documented humanity, complete with its ecological blindness. World Models offer the theoretical opportunity to ground the machine more directly in sensory observation of the physical environment. However, as we will see, this opportunity comes with important caveats about how data collection and sensor choice remain fundamentally human decisions.

3. Technical foundations of AI world models

The transition from text-processing AI to world models represents a shift from processing symbols to processing causal dynamics. This section discusses the architecture of World Models to provide non-specialist readers with the technical foundation necessary for understanding how these systems work. For those unfamiliar with machine learning, a "latent space" is a mathematical, or more precisely, "vectorial", space where complex sensory data (like images) are compressed into simpler numerical representations that a machine can manipulate using vector operations (Bengio et al., 2013; Ha & Schmidhuber, 2018). "Predicting in latent space" means the system predicts abstract properties of future states rather than raw sensory details, focusing on what matters while filtering out noise.

3.1. The Ha & Schmidhuber architecture (V-M-C)

In their seminal 2018 paper, David Ha and Jürgen Schmidhuber proposed an architecture explicitly designed to mimic the cognitive division of labour found in biological brains. They argue that our brains develop an abstract model of space and time to handle the vast flow of sensory information (Ha & Schmidhuber, 2018). Their architecture consists of three components, which are analogous to von Uexküll's functional circle:

- **Vision (V) — Compressor:** A variational autoencoder (VA, a type of neural network) that compresses high-dimensional sensory data (e.g., pixels from images) into a compact latent code. This compression defines the machine's perceptual resolution, that is, what aspects of the world it notices and what it ignores (Ha & Schmidhuber, 2018). This choice of which features to compress is a design decision made by humans.
- **Memory (M) — Predictor:** A mixture density network-recurrent neural network (MDN-RNN, another type of neural network) that predicts the next latent state based on the current state and action. It models the temporal dynamics and causality of the world (Ha & Schmidhuber, 2018).
- **Controller (C) — Policy:** A lightweight decision-making module that receives only the latent representation, not raw sensory data, and outputs an action. Although it never interacts with the sensory world directly, its output is a real-world command. The system therefore acts on the world through its internal model, just as biological organisms act through their perceptual representations rather than direct access to reality (Ha & Schmidhuber, 2018).

Learning inside a dream: A distinctive capability of this architecture is that the agent can learn both during real-world interaction and within a "hallucinated dream" generated by the Memory component. During training, once the system has internalised the probabilistic structure of its environment, it can temporarily detach from real sensory inputs and continue improving its behaviour inside its own simulation. During operation, the same internal model allows the agent to anticipate future states before acting, effectively running miniature "dreams" to evaluate possible outcomes. In semiotic terms, this marks the emergence of a complex subjective interiority: a space where potential futures can be explored, tested, and refined without physical risk (Ha & Schmidhuber, 2018).

3.2. Autonomous machine intelligence and Joint Embedding Predictive Architecture (JEPA)

Yann LeCun expands this concept to address the fragility of current AI. He critiques systems that rely on supervisory data (human labels) or pure reinforcement learning (trial and error), noting that humans learn basic common sense (intuitive physics) largely through observation (LeCun, 2022). LeCun proposes the Joint Embedding Predictive Architecture (JEPA) as a path to autonomous machine intelligence.

Unlike generative models that try to predict every pixel of an image, which is computationally expensive and often irrelevant, JEPA predicts in representation space (LeCun, 2022; Assran et al., 2023). It asks the system to predict the abstract state of the world, ignoring unpredictable noise (like the chaotic movement of leaves in the wind) while capturing relevant signals (the trajectory of a moving object).

Semiotic implication: This distinction is profound. By predicting in representation space, the machine is forced to learn invariance, to identify the essential features of an object that persist over time. It is creating semiotic categories (signifieds) from the raw flux of sensory data (signifiers). LeCun argues that this predictive capability is the foundation of common sense (LeCun, 2022). Clearly, this process also involves a risky human intervention: the decision about which aspects of the environment count as noise and which count as signal. What is filtered out as irrelevant from the standpoint of machine efficiency may be ecologically vital, and the categories the system learns ultimately reflect human priorities rather than the complexity of the living world.

4. Expanding the sensory signals

Most current world models rely primarily on video or audio inputs; however, emerging multimodal approaches are beginning to incorporate sensor inputs, such as tactile, proprioceptive, and geospatial data (Hafner et al., 2020, 2023; Mai et al., 2024). Considering that the ecosemiotic network functions through a diversity of channels that are invisible to the eye and ear, to build an AI system capable of understanding more of the biosphere's complexity, we must equip it with sensors that reflect the diversity of ecological signals as fully as possible.

4.1. The chemical internet

For much of the living world, plants, insects, bacteria, fungi, chemical signals are primary (Dudareva et al., 2013). The "chemical internet" of the soil and the air carries signals of danger, attraction, and identity (Dudareva et al., 2013; Heil & Karban, 2010). These signals are as central to the perceptual worlds of many organisms as light and sound are to humans.

The technology: Advances in digital olfaction (e-noses) utilise arrays of sensors (metal oxide, graphene, bio-hybrid) to detect volatile organic compounds (VOCs). AI models, such as graph neural networks, map molecular structures to perceptual odour categories, creating principal odour maps (Keller et al., 2017; Wilson & Baietto, 2009). This represents an attempt to translate chemical complexity into machine-readable features.

Ecological application: A world model equipped with olfactory inputs could perceive the signals of plants. Research shows that plants emit specific chemical signals when stressed by drought or pests (Dudareva et al., 2013). An AI could detect a pest infestation in a forest days before it becomes visible, predicting the spread based on wind and chemical concentration gradients. Its Umwelt would consist of chemical trails and molecular fingerprints, a reality shared with the insect world but alien to human perception.

4.2. Bioacoustics

The field of digital bioacoustics is using AI to decode the sonic layer of the biosphere. Foundation models are trained on thousands of hours of animal vocalisations, learning to separate and classify calls across taxa (Farina & Gage, 2017; Stowell et al., 2019).

Interspecies communication: AI algorithms have revealed complex patterns in animal communication systems. A bioacoustic world model would not merely record sound; it would learn the syntax of these communications (Stowell et al., 2019). By learning to recognise these patterns, the system could potentially detect disruptions to communication networks that precede ecological collapse.

Ecosystem resilience: By monitoring the soundscape, the ratio of biophony (animal sounds) to anthropophony (human noise), AI can predict ecosystem health. A decline in acoustic diversity often precedes physical collapse (Farina & Gage, 2017).

4.3. Haptics

The concept of tactile sensing integrates high-resolution touch sensing with vision, potentially mimicking the Umwelt of animals like the star-nosed mole (Bohg et al., 2014).

Affordances: In robotics, affordances refer to what an environment offers the agent, e.g., climbable, liftable (Gibson, 1979). A robot interacting with soil or vegetation learns the physical properties of the world, such as texture, resistance or moisture, building a world model grounded in contact rather than just observation (Bohg et al., 2014). This haptic information provides a different mode of understanding than vision alone.

5. The emergence of the Machine Umwelt

If we construct an AI that perceives the world through, for instance, hyperspectral vision, chemical arrays, and hydrophones, and then processes this data through a predictive world model in a vector space or embedding, such AI will inhabit what we might call a "vectorial phenomenology", a mode of perception defined by linear algebra rather than biology.

5.1. Vectorial phenomenology and its limits

Whether artificial systems can possess an Umwelt has been debated in biosemiotics; Emmeche (2001) concluded that robots remain only partly "situated" because they lack the qualitative, self-organising properties of living organisms. Nevertheless, we can say that such a system inhabits a Machine Umwelt. It is not a simulation of human perception; it is a distinct phenomenological space defined by the machine's specific sensors and the mathematical topology of its latent space. This Machine Umwelt is, however, fundamentally constrained by human choices: which sensors to include, what environmental conditions to train on, and what to consider relevant for the system to learn.

Vectorial concepts: In the machine's own Umwelt, a forest is not a collection of trees (a linguistic category created by humans) but a high-dimensional cluster of vectors correlating, for instance, spectral reflectance, acoustic frequency, and chemical concentrations.

Non-anthropocentric categories: Because the model uses unsupervised learning to minimise prediction error, it may discover categories that humans lack words for (Bengio et al., 2013; Ha & Schmidhuber, 2018). It might group a specific fungus and a specific tree root as a single object because they are semiotically inseparable in the chemical spectrum, even if they appear distinct to the human eye. This process is analogous to the way AI language models form tokens and clusters within a word-embedding space, where meaning emerges from statistical patterns in high-dimensional vectors (Mikolov et al., 2013). In language models, this produces what we might call a "vectorial epistemology", where knowledge is encoded as relations among word-vectors in an organisation of meaning that does not map onto human-made linguistic categories. In world models, however, the latent space encodes real ecological signals rather than linguistic ones, giving rise to a vectorial phenomenology: a mode of perception grounded in chemical gradients or spectral signatures instead of human concepts (Ha & Schmidhuber, 2018; Olah et al., 2020). Such a phenomenology may be more ecologically accurate in some respects, yet remains semantically opaque to us. And importantly, even these "alien" categories are still our creations, constrained by the architectural and sensory choices we impose.

5.2. The configurator as a semiotic switch

LeCun describes a configurator module in his architecture, a mechanism that configures the world model for the task at hand (LeCun, 2022). In semiotic terms, this is an attention mechanism that shifts the Umwelt. Just as a hungry tick perceives butyric acid as food while a satiated one might ignore it, the configurator determines which signals are meaningful (salient) at any given moment. This programmability means the Machine Umwelt is fluid; it can shift from, for example, a Carbon-Sequestration Umwelt to a Biodiversity-Protection Umwelt based on the goal function.

6. An ecosophical common sense

We arrive now at the heart of the argument: the nature of common sense in AI systems and how to ground it in ecological rather than anthropocentric principles. LeCun argues that AI needs common sense to function safely and efficiently. He defines it as the background knowledge of what is plausible and impossible in the physical world (LeCun, 2022). The critical question is: what assumptions should that common sense embody?

6.1. The ecological blind spots of human common sense

From an ecolinguistic perspective, human common sense contains significant ecological blind spots. Our cultural codes, the shared values and norms that construct our reality, often normalise ecological destruction (Stibbe, 2015). It is common sense in industrial society to view a river as a resource for irrigation or waste disposal. It is common sense to prioritise short-term economic yield over long-term soil health. It is common sense that economic growth is inherently good and that nature exists primarily to serve human interests. This anthropocentric common sense represents a collection of assumptions that fail to account for the systemic complexity of the biosphere and the dependencies that human well-being has on ecological health.

If we train AI to replicate human common sense, for example through text or human-labelled images, we risk replicating these ecological blind spots. We may create machines that inherit assumptions about infinite resources on a finite planet. The stakes are high: if future AI systems designed to manage resources or predict ecological futures are trained to think like economically-optimised humans, we may simply automate our existing destructive patterns at a larger scale.

6.2. Towards ecological common sense

A world model trained on ecological observation rather than human interpretation offers the possibility of what we might call "ecological common sense", which we could define as a form of intuitive understanding rooted in how living systems actually work.

Intuitive ecology: Just as LeCun's intuitive physics allows a machine to predict that a cup will fall if dropped, intuitive ecology would allow a machine to predict that removing a predator will cause population dynamics shifts, or that applying certain pesticides will reduce yield due to pollinator loss. Such systems would understand symbiotic dependencies and feedback loops as naturally as humans understand gravity.

Ecosophical prediction: This leads to the concept of ecosophical prediction, the term we use to describe AI systems whose objective function is aligned not with human economic activity (GDP, clicks, efficiency), but with ecological flourishing (Næss, 1989). In such a system, the AI's primary target would be predicting and maintaining futures where the ecosystem survives and regenerates. A "bad prediction" would be one that leads to semiocide or system collapse. An ecosophical world model could make visible the ecosemiotic relations that human narratives ignore, for example, the fungal networks, the chemical signalling and the microscopic flows of matter and energy that sustain life.

6.3. Comparative analysis of AI models

Table 1 summarises key differences between three approaches to AI systems.

Table 1: Contrasting features in linguistic, physical, and ecosophical AI models.

Feature	**Linguistic AI (LLMs)**	**Physical World Models**	**Ecosophical World Models**
Primary Input	Human text (symbolic)	Vision/video (physical)	Multimodal (bioacoustic, chemical, spectral)
Common Sense	Cultural/narrative norms	Intuitive physics (gravity, object permanence)	Intuitive ecology (systemic interdependence)
Definition of Object	Tokens (words, sentences, punctuation)	Physical entity with mass/trajectory	Meaning-carrier in interdependent web
Implicit Goal	Plausible text generation	Accurate state prediction	Ecosystem resilience & stability
Ecological Blind Spot	Non-human agency, biophysical constraints	Biological complexity, semiotic relations	Human cultural context (if not explicitly integrated)
Example Failure	The factory produces wealth (ignoring pollution)	Optimising traffic flow while ignoring road's impact on habitat fragmentation	Optimising for a single species at the expense of others

6.4. Critical challenges: Operationalising ecosophical objectives

It is key to acknowledge that realising ecosophical prediction faces fundamental challenges that we must address.

6.4.1. The design challenges

The most important challenge is this: all sensors, architectures, and objective functions are designed by humans. If not carefully designed, world models may reduce certain biases while introducing others through design choices. What counts as "ecological flourishing" must still be operationally defined by human researchers' ecosophy. The Machine Umwelt is not free from human intention; rather, it embodies human choices about what matters.

Consider the example of a world model trained on satellite data from conservation areas versus one trained on data from extractive industries. The first might learn to recognise and protect intact ecosystems; the second might learn to optimise extraction efficiency. The difference is not in the architecture but in what data gets selected, by whom, and for what purposes. This is a profound problem that calls for greater attention to governance and stakeholder participation in AI development. Who should decide what ecological values are embedded in these systems? Whose interests are represented? How can indigenous knowledge systems and community expertise be incorporated into these designs?

6.4.2. The specification problem

A second challenge is what we might call the specification problem. Defining objective functions for "biodiversity" or "resilience" involves scientific and ethical choices. How do you measure biodiversity? At what spatial and temporal scale? Do you optimise for species richness, genetic diversity, functional diversity, or some combination? Different operationalisations may yield different predictions and actions.

There is a well-known principle called Goodhart's Law, often discussed in AI alignment literature (Russell, 2019): when a measure becomes a target, it ceases to be a good measure. An AI system optimised for a specific biodiversity metric might find ways to game that metric that do not actually improve ecosystem health. For example:

- An AI optimised for carbon sequestration alone might determine that a monoculture of fast-growing eucalyptus is preferable to a diverse ancient rainforest, because the objective function fails to capture the value of biodiversity, cultural significance, and ecosystem complexity that are represented in dimensions the machine ignores or underweights.
- An AI optimised for maximum species count might favour systems with many small, genetically identical organisms over systems with fewer but more ecologically critical species.
- An AI optimised for fishery yield might overexploit stocks until collapse, because the metric does not capture ecosystem shifts until it is too late.

How might we address this? Possible solutions include:

- **Multi-objective optimisation:** Rather than optimising for a single metric, design systems that must balance multiple ecological objectives (biodiversity AND carbon sequestration AND cultural ecosystem services AND long-term resilience). This creates tensions the system must navigate rather than finding a single "gaming" path.

- **Adaptive metrics:** Use dynamic objective functions that change based on ecological feedback, rather than static metrics that might diverge from reality.
- **Human-in-the-loop approaches:** Require that major decisions recommended by ecosophical AI systems be validated by teams including ecologists, indigenous knowledge keepers, and affected communities before implementation.
- **Transparency and auditability:** Design systems such that the reasoning behind predictions can be inspected and critiqued by external parties.
- **Default mechanisms:** When in doubt, prefer conservation rather than exploitation.

6.4.3. The value-action gap and social implementation

Even if AI systems develop sophisticated ecological common sense, there is no guarantee this translates to human behavioural change or policy implementation. Environmental psychology has documented extensive gaps between knowledge and action: people often know what is ecologically right but fail to do it due to social, economic, and psychological barriers (Gifford, 2011). The mechanism by which AI representations of ecological health influence human decision-making and resource allocation requires work on eco-social areas. Furthermore, there is a risk that ecosophical AI could become a tool of "eco-colonialism", where Western scientific notions of ecological health are imposed on communities through AI systems without their consent or participation. This could further marginalise indigenous knowledge systems and local adaptive practices. To mitigate this issue, any deployment of ecosophical prediction must be embedded in genuine dialogue with affected communities.

6.4.4. Who controls these systems?

Finally, there is a question of power: whose world gets modelled and who benefits from these systems? If ecosophical AI systems become proprietary tools of extraction industries, they could simply become more efficient mechanisms of exploitation. An AI trained on satellite data of logging concessions, managed by a timber company, sees a very different world than one trained on the acoustic data of indigenous lands, managed by conservation partners. The same technology could serve radically different stakeholders. This is not a technical problem but a governance and justice problem that must be addressed through policy, regulation, and community control.

7. Future research directions

This paper has articulated a vision of ecosophical prediction as an ethical consideration in development of world models, but much work remains to move from vision to practice. We enumerate some of the required areas of research and collaboration.

7.1. Technical research needs

- **Multi-scale modelling:** Develop world models that can represent ecological processes at multiple scales simultaneously, from chemical sensors to landscape-level vegetation patterns to regional climate dynamics. This kind of enhancement within the Hierarchical Joint Embedding Predictive Architecture (H-JEPA) (LeCun, 2022) will ensure a "common sense" beyond physical objects' interaction.

- **Uncertainty quantification:** Build systems that explicitly quantify uncertainty in their predictions rather than returning point estimates, allowing humans to understand where confidence is high and where caution is warranted.

- **Value alignment in complex domains:** Work with AI alignment researchers (Russell, 2019) to develop methods for specifying robust multi-objective functions in ecological domains, with mechanisms to prevent specification gaming.

- **Integration of heterogeneous data:** Develop architectures that can meaningfully combine data from scientific instruments, diverse knowledge systems, and community observations.

7.2. Interdisciplinary requirements

Developing genuine ecosophical AI will require collaboration between:

- Machine learning researchers and AI architects
- Ecological scientists and conservation biologists
- Indigenous knowledge keepers and environmental justice advocates
- Policymakers and legal scholars
- Ecolinguists and environmental humanists

8. Conclusion: Whose world will we model?

This paper has introduced "ecosophical prediction" as both a technical possibility and an ethical imperative for the emergent paradigm of AI world models, situating this proposal within an ecolinguistic concern with the stories and framings through which AI systems mediate human–nature relations. By integrating biosemiotic theory with advances in artificial intelligence architecture, we have shown that world models trained on multimodal ecological signals, for example chemical communication networks, bioacoustic patterns, haptic feedback, and spectral data, could theoretically construct a Machine Umwelt attuned to ecosemiotic relations that remain invisible to text-based AI. Where large language models (LLMs) inherit the ecological blind spots encoded in human discourse, world models offer the potential to learn directly from the living world itself, potentially reshaping dominant ecological narratives.

This paper has also insisted on a fundamental recognition: this potential is not inevitable, and the opportunity is inseparable from profound risks. Every sensor we deploy, every dataset we collect, every objective function we optimise embeds human values into machine perception and, by extension, into the discursive futures these systems make visible. The Machine Umwelt does not emerge neutrally from data; it is shaped, constrained, and ultimately controlled by prior human decisions about what matters.

The challenge of ecosophical prediction is therefore threefold: technical, epistemological, and political. Technically, we must design architectures capable of multi-scale, multi-objective optimisation that resist specification gaming and maintain sensitivity to ecological complexity across multi-modalities. Epistemologically, we must integrate what we call vectorial phenomenology as an alien Umwelt: the machine may perceive patterns we lack words for, but we remain responsible for the futures it predicts and the meanings these predictions come to carry. Politically, we must challenge the question of power: who decides which ecosystems are modelled, which sensors are deployed, whose knowledge systems are encoded, and whose interests these systems ultimately serve?

The H4rmony Project has demonstrated that ecosophical principles can guide generative AI towards more ecologically conscious outputs and ecolinguistically aligned discourses. But world models, precisely because they operate through sensory prediction rather than linguistic generation, require ecosophical alignment to be embedded from the outset: in sensor choice, data curation, architectural design, and objective specification. This demands unprecedented collaboration between machine learning researchers, ecological scientists, diverse knowledge keepers, environmental justice advocates, and affected communities. It requires governance structures that prevent ecosophical AI from becoming tools of eco-colonialism or greenwashed extraction.

The stakes are high. In a time of ecological crisis, the design choices we make about AI systems and their interaction with nature (directly or indirectly all of them do) will shape what futures are possible and which narratives of human and more-than-human relations become dominant. For ecolinguists, the task is to engage with these emerging technologies, to critique the stories they tell about nature, to challenge the values embedded in their objectives, and to demand systems that centre ecological flourishing and justice rather than efficiency and extraction. Machine Umwelts are not yet written. The question is whether they will be designed to serve the thriving of the more-than-human world, or merely to serve human interests more efficiently. That choice remains ours to make.

Acknowledgements

The author used AI assistance (Theophrastus, Claude and Perplexity) for improving grammar, proofreading, and bibliographic research. All ideas, arguments, and analysis are the author's own.

References

Assran, M., Duval, Q., Misra, I., Bojanowski, P., Vincent, P., Rabbat, M., LeCun, Y., & Ballas, N. (2023). Self-supervised learning from images with a joint-embedding predictive architecture. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 15619–15629).

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). https://doi.org/10.1145/3442188.3445922

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Bohg, J., Morales, A., Asfour, T., & Kragic, D. (2014). Data-driven grasp synthesis — A survey. *IEEE Transactions on Robotics, 30*(2), 289–309.

Dudareva, N., Klempien, A., Muhlemann, J. K., & Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist, 198*(1), 16–32. https://doi.org/10.1111/nph.12145

Emmeche, C. (2001). Does a robot have an Umwelt? Reflections on the qualitative biosemiotics of Jakob von Uexküll. *Semiotica, 134*(1/4), 653–693.

Farina, A., & Gage, S. H. (2017). *Ecoacoustics: The ecological role of sounds*. John Wiley & Sons.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.

Gifford, R. (2011). The dragons of inaction: Psychological barriers that limit climate change mitigation and adaptation. *American Psychologist, 66*(4), 290–302. https://doi.org/10.1037/a0023566

Ha, D., & Schmidhuber, J. (2018). World models. *arXiv preprint arXiv:1803.10122*. https://doi.org/10.5281/zenodo.1207631

Hafner, D., Lillicrap, T., Norouzi, M., & Ba, J. (2020). Dream to control: Learning behaviors by latent imagination. *arXiv:1912.01603*.

Hafner, D., Pasukonis, J., Ba, J., & Lillicrap, T. (2023). Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*.

Heil, M., & Karban, R. (2010). Explaining evolution of plant communication by airborne signals. *Trends in Ecology & Evolution, 25*(3), 137–144. https://doi.org/10.1016/j.tree.2009.09.010

Keller, A., Gerkin, R. C., Guan, Y., Dhurandhar, A., Turu, G., Szalai, B., … & Meyer, P.

(2017). Predicting human olfactory perception from chemical features of odor molecules. *Science, 355*(6327), 820–826.

Kull, K. (2010). Ecosystems are made of semiosic bonds: Consortia, umwelten, biophony and ecological codes. *Biosemiotics, 3*(3), 347–357. https://doi.org/10.1007/s12304-010-9081-1

LeCun, Y. (2022). *A path towards autonomous machine intelligence* (Version 0.9.2). OpenReview. https://openreview.net/pdf?id=BZ5a1r-kVsf

Mai, X., Tao, Z., Lin, J., Wang, H., Chang, Y., Kang, Y., … & Zhang, W. (2024). From efficient multimodal models to world models: A survey. *arXiv preprint arXiv: 2407.00118*. https://doi.org/10.48550/arXiv.2407.00118

Maran, T. (2020). *Ecosemiotics: The study of signs in changing ecologies*. Cambridge University Press.

Maran, T., & Kull, K. (2014). Ecosemiotics: Main principles and current developments. *Geografiska Annaler: Series B, Human Geography, 96*(1), 41–50. https://doi.org/10.1111/geob.12035

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Næss, A. (1989). *Ecology, community, and lifestyle: Outline of an ecosophy* (D. Rothenberg, Trans.). Cambridge University Press.

Olah, C., Cammarata, N., Schubert, L., Goh, G., Petrov, M., & Carter, S. (2020). An overview of early vision in InceptionV1. *Distill, 5*(4), e00024-002.

Puura, I. (2002). Nature in our memory. *Eesti Loodus, 53*(6), 15–17.

Russell, S. (2019). *Human compatible: AI and the problem of control*. Penguin UK.

Stibbe, A. (2015). *Ecolinguistics: Language, ecology and the stories we live by*. Routledge.

Stowell, D., Wood, M. D., Pamuła, H., Stylianou, Y., & Glotin, H. (2019). Automatic acoustic detection of birds through deep learning: The first Bird Audio Detection challenge. *Methods in Ecology and Evolution, 10*(3), 368–380.

Tønnessen, M. (2015). Umwelt and language. In E. Velmezova, K. Kull, & S. Cowley (Eds.), *Biosemiotic perspectives on language and linguistics* (pp. 77–96). Springer. https://doi.org/10.1007/978-3-319-20663-9_5

Uslu, M. E. (2020). Semiocide: An introduction to semiotics of destruction of the meaningful. *Sign Systems Studies, 48*(2–4), 224–245. https://doi.org/10.12697/SSS.2020.48.2-4.03

Vallego, J. (2023). Ecolinguistics and AI: Integrating eco-awareness in natural language processing. *Language & Ecology*. http://ecolinguistics-association.org/journal

Vallego, J. (2024). *The H4rmony Project*. Retrieved from https://theh4rmonyproject.org

von Uexküll, J. (2010). *A foray into the worlds of animals and humans: With a theory of meaning* (J. D. O'Neil, Trans.). University of Minnesota Press. (Original work published 1934).

Wilson, A. D., & Baietto, M. (2009). Applications and advances in electronic-nose technologies. *Sensors, 9*(7), 5099–5148.

Appendix: Glossary

As this paper draws on concepts from biosemiotics, ecolinguistics, and artificial intelligence, the following definitions are provided to support readers across these disciplines.

Biosemiotics and Ecology

Affordance: The action possibilities that the environment provides or furnishes to a particular organism, determined by the fit between the organism's capabilities and environmental properties (Gibson, 1979).

Biophony: The collective sound produced by living organisms in an environment; a component of soundscape ecology.

Ecosemiotics: The study of sign relations within and across ecosystems; extends biosemiotics to ecological scales, treating ecosystems as communication networks (Kull, 2010).

Ecosophy: An ecological philosophy, i.e. a set of values concerning the ideal relationship of humans with each other, other species and the physical environment. Analysts use their own ecosophy to judge stories (or outputs, in the context of this article) as beneficial, ambivalent or destructive (Stibbe, 2015).

Functional circle: Von Uexküll's model of the feedback loop connecting an organism's perception and action; the organism perceives signs, acts upon them, and perceives the consequences of its actions.

Machine Umwelt: The subjective perceptual world of an AI system, defined by its sensors, data, and the mathematical topology of its latent space. Whether artificial systems can possess an Umwelt has been debated in biosemiotics since Emmeche (2001); this paper uses the term to denote the specific perceptual configuration shaped by human design choices in world models.

Semiocide: The destruction of meaning-making relationships and semiotic conditions necessary for life; ecological damage understood as the erasure of sign relations (Puura, 2002).

Umwelt: The subjective, species-specific world of an organism; the subset of environmental signals an organism can perceive and act upon, constituting its meaningful reality (von Uexküll, 1934/2010).

Artificial Intelligence and Machine Learning

Joint Embedding Predictive Architecture (JEPA): An architecture that predicts abstract representations of future states rather than raw sensory details; filters noise while capturing relevant signals (LeCun, 2022).

Large language model (LLM): An AI system trained on large quantities of text to predict and generate language; learns statistical patterns in human discourse rather than perceiving the physical world directly.

Latent space: A compressed mathematical representation where complex sensory data (such as images) are encoded as numerical vectors; the space in which a model's internal representations exist.

Objective function: The mathematical target an AI system is trained to optimise; defines what the system treats as success or failure.

Reinforcement learning: A training approach where an AI learns through trial and error, receiving rewards or penalties for its actions.

Variational autoencoder (VAE): A neural network that compresses high-dimensional data into compact representations and can reconstruct the original from those representations.

World model: An AI architecture that learns to simulate and predict environmental dynamics from sensory observation rather than text; builds an internal model of how the world works (Ha & Schmidhuber, 2018).

Concepts Introduced in This Paper

Ecosophical prediction: The design of world models whose predictive targets are aligned with ecological flourishing rather than anthropocentric utility; AI systems that optimise for ecosystem resilience and multispecies thriving.

Vectorial epistemology: Knowledge encoded as relations among word-vectors in language models; the way LLMs represent meaning through language-driven statistical patterns in a high-dimensional space, rather than through reference to physical reality.

Vectorial phenomenology: The mode of perception characteristic of world models that represent the environment through high-dimensional vector spaces encoding ecological signals; a mathematically-defined experiential space grounded in sensory data rather than linguistic categories.