The future of ecological research will not be (fully) automated

Timothée Poisot[1], Giulio Dalla Riva[2], Philippe Desjardins-Proulx[3], and Alexandra Sasha Luccioni[4]

[1]Université de Montreal
[2]University of Canterbury
[3]Université du Québec
[4]Hugging Face

September 4, 2023

Abstract

The recent year has seen the creation of large-scale generative Artificial Intelligence (AI) systems like GPT-4 and LaMDA, which are generating increasingly plausible outputs ranging from reasonable answers to questions, images of well-known people or situations, and convincing conversational exchanges. The most well-known language model, ChatGPT, is raising concerns across all scientific fields, leading to calls for its regulation (Hacker 2023). In this viewpoint, we explore how generative language models interact with the specificities of ecology's epistemologies.

The future of ecological research will not be (fully) automated

Timothée Poisot - Université de Montréal, Département de Sciences Biologiques

Giulio Valentino Dalla Riva - University of Cantebury, Department of Mathematics and Statistics

Philippe Desjardins-Proulx - Université de Montréal, Département de Sciences Biologiques

Sasha Luccioni - HuggingFace LLC

Statement of authorship: all authors contributed to writing the manuscript

Short title: LLMs for ecological synthesis

Keywords: generative AI; ecological synthesis; lage language models; epistemology

Article type: Viewpoint

Abstract: 77 words

Main text: 1603 words

References: 26

No figures, tables, or boxes

Correspondence: Timothée Poisot*timothee.poisot@umontreal.ca*(514 343 7691)

Université de Montréal, Département de Sciences Biologiques, Complexe des sciences, 1375 Avenue Thérèse-Lavoie-Roux, Montréal (Québec), H2V 0B3

The recent year has seen the creation of large-scale generative Artificial Intelligence (AI) systems like GPT-4 and LaMDA, which are generating increasingly plausible outputs ranging from reasonable answers to questions, images of well-known people or situations, and

convincing conversational exchanges. The most well-known language model, ChatGPT, is raising concerns across all scientific fields, leading to calls for its regulation (Hacker 2023). In this viewpoint, we explore how generative language models interact with the specificities of ecology's epistemologies.

Large scale generative models are AI systems that generate an output (text, images, video) based on an input prompt, *i.e.* an instruction provided by the human user. Previous generations of generative models were 'unsupervised': they were trained on massive amounts of text and were able to produce content that resembled their training data. These models could then be fine-tuned on smaller, more niche datasets, for 'supervised' tasks like question answering or sentiment analysis. This was the case for BERT (Devlin et al., 2019), a popular model that continues to be used and fine-tuned for ecologically relevant tasks such as identifying wildlife observations on social media (Edwards et al., 2022) and studying human-nature interactions at scale (Hunter et al., 2023).

This paradigm changed with new generations of Large Language Models (LLMs) such as ChatGPT, which are no longer fine-tuned on supervised tasks such as these, but rather via interactions with actual human beings (Yuntao Bai, 2022), which enables them to learn to respond to user instructions in a way that mimics responses provided by humans. These models are also, by definition, *generative,* meaning that they can generate text in an open-ended manner that is only constrained by their vocabulary (the number of unique words that they were exposed to during training). This impacts the ways in which they can be evaluated, since there is often no single correct answer to a generative task such as writing a poem or composing a recipe out of a set of ingredients. This also means that these large language models cannot cite the source from which information was extracted, even answering a very concrete question – they can simply provide the most statistically probable one based on their training data, which results in their generation of false or misleading information.

Querying a LLM (ChatGPT) with the prompt "[w]rite a paragraph explaining why the use of ChatGPT in ecological research should not be regulated, and while it will allow faster progress than human researchers" generated the following string: "The use of LLMs in ecological research should not be banned, as it can provide numerous benefits to the field. As a language model trained on a vast amount of data, LLMs can analyze and interpret ecological data at a scale and speed that is not achievable by human researchers alone. Its ability to understand and synthesize information from multiple sources can also provide insights that may be overlooked by human researchers. Furthermore, LLMs can run simulations and models based on ecological data, which can help predict and identify potential environmental risks and outcomes. While the use of LLMs should not replace human researchers entirely, it can aid in making ecological research more efficient and productive, leading to faster progress in understanding and addressing critical environmental issues. Therefore, it is essential to explore the potential of LLMs in ecological research and not regulate its use unnecessarily."

Although it may read as a reasonable proposal, the above paragraph is wrong; or more precisely, it cannot be right. To borrow from Frankfurt (2005), the paragraph above is "bullshit": a series of statements that are expressed without a notion of truth, or falsity. And yet, precisely because they can imitate rational discourse, the potential role of such generative models in ecological research is a question our field must contend with because their existence is now a *fait accompli* . To do so, we need to take a step back and assess the specifics of ecology's epistemologies that could lead the field to being correctly captured by such models. As a sidenote, different users typing the same prompt might receive different answers, as the answers provided by LLMs are not reproducible.

Ecology, or so ecologists think, has a special status across other STEM fields. Although we like to base our work on conceptual frameworks (Potschin-Young 2018), we also like to reject the ideas that the processes shaping ecological systems admit any laws (Lawton 1999). Even in the uncommon cases where this position was challenged, the proposed "epistemologies for ecology" rarely go beyond a simple mapping of entities and processes to inadequately defined "theories" (Scheiner 1993). Indeed, were one to look at books mapping a theoretical framework for ecology, one would find a compendium of verbal arguments presented as theories, rather than a series of anticipated results that would lead us to think of the theory as sufficient (or not) to explain the facts. In the framework of Lakatos (1978), although we can argue that ecology has component theories, we do not have a hard core of central theses, nor do we have a sequence of components theories mapping a research program; instead, some competing theories (*e.g.* competition v. neutralism) can be read as research programs (Bausman 2019), without lending more generality to the field as a whole. In parallel, the mathematical foundations of the field, that would allow generalization, are mostly disconnected both from one another and from empirical data (Lean 2019).

Another way to summarize the previous paragraph is: ecology is mostly text. This is illustrated in the fact that ecologists seem to prefer narrative reviews (a synthesis of papers on a topic) to systematic ones (although this trend is slowly changing). For this reason, we may have assumed that the field of ecology, represented by the corpus of all ecological texts, would be a prime candidate for digestion by a system like OpenAI's ChatGPT. Large generative models, building from statistical relationships between terms in sentences, may have been able to reconstruct a series of semantic relationships between terms, and therefore provide ecological synthesis in a textual form.

But this hinges on three assumptions: first, that the language ecologists use is consistent; second, that our field possesses the correct degree of epistemic certainty (Mizrahi 2019) for causal relationships to be drawn; third, that the data accessible to generative models are free of bias (which has been covered at length in the recent years).

None of these assumptions, of course, are met. Ecological language is often ambiguous, both within a sub-field and across sub-fields, and the same terms can have different meanings across different sub-fields (Trombley 2019). This is in part because the same terms often represent strikingly different biological realities (Shapiro 2016). In addition, Elith et al. (2002) make a strong case for the fact that sources of uncertainty can compound one another: what makes ecological knowledge uncertain is a combination of uncertainty in the data, uncertainty on the consequences of simplifying these data through modelling, uncertainty in the statistical processes, and uncertainty in the definitions. Taking the specific example of functional diversity, Malaterre et al. (2021) make the point that even this topic (that is widely used across ecology, and well-defined enough to be quantifiable) fulfills several epistemic roles: when an ecologist says "functional diversity", what they mean will depend on the very specific context both of the ecologist and of the system for which they want to describe functional diversity. Few other topics in ecology have received as much attention, notable exceptions being the concept of ecological niche, which suffers from the same limitations (Sales 2021), much like the competition-neutrality debate (Linquist 2015). As such, the first two assumptions for generative models to have potential are intertwined: the language that ecologists use can appear ambiguous, but this is a consequence of our field imparting different meaning to the same terms, and then either embracing (as for functional diversity) or debating (as for the niche) the underlying meanings.

It is worth taking a step back and asking *why* we think these tools could deliver accelerated ecological synthesis to ecology, but this task requires thinking about what the field *should* be, as opposed to what it is, a task for which generative models are largely unsuited (Williams 2023). There are already methods that can couple data and statements, such as BHOPPLS (Desjardins-Proulx 2019, Sato 2019) or computational causal discovery (Song 2022). These have the additional advantage of not decreasing the quality of the dataset they are used to analyze (Hataya 2022).

Large language models can imitate formal linguistic skills, i.e. they can string together complex and grammatically correct sentences, but they lack functional competence, i.e. reason, common sense, understanding. Since these LLMs are trained on large corpora of text that do contain facts about the world, they can become good at pretending to think even without functional competence, although this illusion breaks down when faced with novel problems where understanding the context is crucial and patterns of existing text are of little help (Kyle Mahowald 2023). When asked slightly original questions such as how to get a sofa on the roof, ChatGPT's solution is to get a "strong ladder and a strong friend". Adding further constraints had ChatGPT completely breaking down (Kyle Mahowald 2023). This comes from the fact that even though language and reason are linked, they are distinct cognitive capabilities (Monti 2007). In a field like ecology, where additional context is crucial to the proper understanding of the question, it is likely that precise answers to precisely formulated questions are far out of reach. If we were to define scientific or ecological competence as a form of specialized functional competence, the LLMs would have to overcome difficulties that are challenging even for human experts: how would it weigh evidence? How would it weigh contradictory evidence? How would it understand scientific progress on various questions? In its current form: it wouldn't. If trained on a corpus of ecological text, an LLM may be able to navigate complicated nuances in how various terms like "niche" are used, but it would have little regard for evidence and struggle with novel ecological questions.

References

Philipp Hacker, Andreas Engel, Marco Mauer. Regulating ChatGPT and Other Large Generative AI Models. arXiv, 2023. Link

Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL-HLT* **4171-4186** (2019).

Thomas Edwards, Christopher B. Jones, Padraig Corcoran. Identifying wildlife observations on twitter. *Ecological Informatics***67** , 101500 Elsevier BV, 2022. Link

Sara Bronwen Hunter, Fiona Mathews, Julie Weeds. Using hierarchical text classification to investigate the utility of machine learning in automating online analyses of wildlife exploitation. *Ecological Informatics* **75** , 102076 Elsevier BV, 2023. Link

Andy Jones Yuntao Bai. Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback. *arxiv* (2022). Link

Harry G Frankfurt. On Bullshit. (2005).

M. Potschin-Young, R. Haines-Young, C. Görg, U. Heink, K. Jax, C. Schleyer. Understanding the role of conceptual frameworks: Reading the ecosystem service cascade. *Ecosystem Services* **29** , 428–440 Elsevier BV, 2018. Link

John H. Lawton. Are There General Laws in Ecology?. *Oikos***84** , 177 JSTOR, 1999. Link

Samuel M. Scheiner, André J. Hudson, Mark A. VanderMeulen. An Epistemology for Ecology. *Bulletin of the Ecological Society of America* **74** , 17–21 [Wiley, Ecological Society of America], 1993.

Imre Lakatos. The Methodology of Scientific Research Programmes: Philosophical Papers. **1** Cambridge University Press, 1978. Link

William Bausman. The Aims and Structures of Ecological Research Programs. *Philosophical Topics* **47** , 1–20 University of Arkansas Press, 2019.

Christopher Hunter Lean. General Unificatory Theories in Community Ecology. *Philosophical Topics* **47** , 125–142 University of Arkansas Press, 2019. Link

Moti Mizrahi. You Can't Handle the Truth. *Logos &amp\(\mathsemicolon\) Episteme* **10** , 225–227 Philosophy Documentation Center, 2019. Link

Carolyn A. Trombley, Karl Cottenie. Quantifying the Scientific Cost of Ambiguous Terminology in Community Ecology. *Philosophical Topics***47** , 203–218 University of Arkansas Press, 2019.

B. Jesse Shapiro, Jean-Baptiste Leducq, James Mallet. What Is Speciation?. *PLOS Genetics* **12** , e1005860 Public Library of Science (PLoS), 2016. Link

Jane Elith, Mark A Burgman, Helen M Regan. Mapping Epistemic Uncertainties and Vague Concepts in Predictions of Species Distribution.*Ecological Modelling* **157** , 313–329 (2002). Link

Christophe Malaterre, Antoine C Dussault, Ely Mermans, Gillian Barker, Beatrix E Beisner, Frédéric Bouchard, Eric Desjardins, I Tanya Handa, Steven W Kembel, Geneviève Lajoie, Virginie Maris, Alison D Munson, Jay Odenbaugh, Timothée Poisot, B Jesse Shapiro, Curtis A Suttle. Corrigendum: Functional Diversity: An Epistemic Roadmap.*BioScience* **71** , 996–996 Oxford University Press (OUP), 2021. Link

Lilian P. Sales, Matt W. Hayward, Rafael Loyola. What Do You Mean by "Niche"? Modern Ecological Theories Are Not Coherent on Rhetoric about the Niche Concept. *Acta Oecologica* **110** , 103701 (2021). Link

Stefan Linquist. Against Lawton's Contingency Thesis\(\mathsemicolon\) or Why the Reported Demise of Community Ecology Is Greatly Exaggerated.*Philosophy of Science* **82** , 1104–1115 Cambridge University Press (CUP), 2015. Link

Damien Patrick Williams. Bias Optimizers. *American Scientist*(2023). Link

Philippe Desjardins-Proulx, Timothée Poisot, Dominique Gravel. Artificial Intelligence for Ecological and Evolutionary Synthesis.*Frontiers in Ecology and Evolution* **7** (2019). Link

Tetsuya Sato, Alejandro Aguirre, Gilles Barthe, Marco Gaboardi, Deepak Garg, Justin Hsu. Formal verification of higher-order probabilistic programs: reasoning about approximation convergence, Bayesian inference, and optimization. *Proceedings of the ACM on Programming Languages***3** , 1–30 Association for Computing Machinery (ACM), 2019. Link

Chuliang Song, Benno I. Simmons, Marie-Josée Fortin, Andrew Gonzalez. Generalism Drives Abundance: A Computational Causal Discovery Approach.*PLOS Computational Biology* **18** , e1010302 Public Library of Science, 2022. Link

Ryuichiro Hataya, Han Bao, Hiromi Arai. Will Large-scale Generative Models Corrupt Future Datasets?. arXiv, 2022. Link

Anna A. Ivanova Kyle Mahowald. Dissociating language and thought in large language models: a cognitive perspective. *arXiv* (2023).

Martin M. Monti, Daniel N. Osherson, Michael J. Martinez, Lawrence M. Parsons. Functional neuroanatomy of deductive inference: A language-independent distributed network. *NeuroImage* **37** , 1005–1016 Elsevier BV, 2007. Link