# Questions for the Open-Ended Evolution Community: Reflections from the 2021 Cross Labs Innovation Science Workshop

Kevin Frans[1,2], L. B. Soros[2] and Olaf Witkowski[2,3,4]

[1]Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Cross Labs, Cross Compass Ltd., Tokyo, Japan
[3]Earth-Life Science Institute, Tokyo Institute of Technology, Japan
[4]College of Arts and Sciences, University of Tokyo, Japan
info@crosslabs.org

## Introduction

Open-ended evolution (OEE) research is exciting because it encompasses a unifying concept that points to fundamental questions in a wide range of fields. Concretely, though, what do researchers and practitioners want to get out of OEE? What can be tried that hasn't been tried already? What open questions do we want to see answered first?

In April 2021, we held the inaugural Cross Labs Workshop on Innovation Science. This workshop focused on algorithmic approaches to open-ended creativity and discovery, from procedural content generation to evolution in virtual worlds. Our primary goal was to foster cross-disciplinary conversation among a small group of academic researchers via a mix of moderated discussions and open dialogue. A main focus of the workshop was identifying similarities and differences between approaches to open-endedness in different domains, particularly from an AI and cognitive science perspective, in the hope of identifying questions of common interest.

The workshop days centered on five primary themes: 1) Finding common threads across interdisciplinary bounds, 2) Generative systems in virtual worlds, 3) Open-ended technology and society, 4) Evolutionary innovations and artificial life, and 5) Roadmapping the future.

This abstract summarizes the primary topics of conversation at the workshop, in addition to highlighting open questions. Our hope in sharing these reflections is to identify opportunities for future work and to open the door for new collaborations on mutual topics of interest. While some of the topics listed in this abstract are stalwarts of the OEE community, and of the OEE workshop series in particular, others suggest new perspectives that may be useful.

We first share our current view of open-endedness research, and discuss how this workshop relates to the field as a whole. We present a set of practical goals related to open-endedness, as suggested by participants, and we elaborate on open questions that were discussed during the workshop. Finally, we highlight topics of interest for future pursuit.

| Goal | Success Looks Like |
| --- | --- |
| Studying open-ended systems in real life. | We understand OEE in biology, human culture, technological innovation, etc. |
| Building deeply interesting OEE simulations. | We build a simulation that you would never want to stop watching. |
| Using OEE as a training bed for AI. | We use OEE to generate vastly diverse data for training strong AI models. |

Figure 1: **What results do we want to see in the near future?** The practical goals of workshop participants largely fell into the three categories shown above. All three goals share a common requirement that we understand the principles of open-endedness, so as to put them to applied use.

## Status of Open-Endedness Research

Since the 1950's, OEE has been a central topic of research for artificial life approaches to the fundamental principles of life. The pioneer John Von Neumann has contributed to the issue as well, with his early model of self-reproducing automata (Von Neumann, 1951). Historically, an evolutionary system would be considered open-ended if it is able to endlessly generate diverse novel entities of growing complexity. Engineering open-ended systems in the lab is not an easy task. The main obstacle may be that the designed evolutionary systems are subject to a thermodynamic drift making them collapse into equilibrium states. Once local optima are reached, they do not produce novelty anymore, which bounds their complexity and diversity.

More recently, since 2015, a series of workshops have been taking place at Artificial Life conferences (Taylor et al., 2016), the last of which – at ALIFE 2018, in Tokyo – has resulted in two special issues on the topic (Packard et al., 2019a,b). A primary takeaway includes taxonomies synthesizing topics of interest for the OEE community. The York categories, named after the inaugural workshop in York, consist of two primary categories: 1) the ongoing generation of adaptive novelty (ongoing generation of new adaptations, ongoing generations of new kinds of entities, emergence of major transitions, and evolution of evolvability) and 2) the ongoing growth of complexity (ongoing growth of entity complexity and ongoing growth of interaction complexity). The Tokyo categories, which came out of the 2018 workshop, removed the evolution of complexity as an explicit desiderata for OEE, calling for ongoing generation of 1) interesting new entities and interactions, 2) evolution of evolvability, 3) major transitions, and 4) semantic evolution.

The Cross Labs Innovation Science workshop focused primarily on ongoing generation of new entities and interactions viewed through the lens of innovation science, artificial intelligence, learning, and games. Attendees' backgrounds reflected this emphasis. Discussions also frequently centered on cognitive aspects of OEE such as the utility (or lack thereof) of subjectivity and its implementation in OEE systems. Unlike OEE1–3 workshops, we did not focus on many topics related to e.g. biology, chemistry, or linguistics.

## What do we want to see in five years?

At the beginning of the workshop, we asked participants what results they'd like to see in the near future. Emphasis was placed on applied goals, with the aim of grounding the theoretical discussion within a set of practical objectives.

### Studying open-ended systems in real life.

Open-ended systems exist all around us. Most famous is the development of life on Earth, which has undergone transition after transition, each which radically changed the evolutionary process thenceforward. In more abstract spaces, the evolution of human culture, art, and technological innovation all seemingly display the capability to increase complexity indefinitely (Ruiz-Mirazo et al., 2008), and continually producing more novelty over time (Standish, 2003; Soros and Stanley, 2014).

What makes these real systems open-ended? Can we predict the open-ended behavior of such systems, or categorize how they operate? What principles do we need to develop, so that we can measure things about these systems, understand them, and analyze emergent behavior? On a practical level, we want to learn from and understand existing open-ended systems, and to do this we need to come up with a common framework for discussing their properties and dynamics.

### Building deeply interesting OEE simulations.

Is it possible to build a simulation that someone would never want to stop watching, or that they'd be thrilled to check in on and the end of the day? Could the interactions between agents have enough depth that histories could be written about them? What are the principles behind designing OEE systems that are constantly interesting to a human viewer?

While judging open-endedness from a human perspective is inherently biased, OEE systems capable of consistently surprising humans would have many applications in building rich multi-agent simulations or in procedural content generation. In addition, human interest may be a suitable proxy for ongoing novelty in a greater sense, as humans are remarkably capable pattern-finders and will quickly get bored of predictable designs.

### Using OEE as a training bed for AI.

One application of OEE with tremendous practical value lies in their use as generative systems for AI training data. Deep learning methods have shown a remarkable capability to scale in performance, and results such as GPT-3 (Brown et al., 2020) show that the path towards strong AI may be bottlenecked not by algorithm design, but by the accessibility of data. In many cases, diverse data is valuable data, presenting a problem which aligns well with open-endedness. OEE systems that produce unbounded amounts of interesting designs, all of which follow common core principles, may prove crucial as a vast source of structured data on which large models can be trained.

As a prime example, the desire to train an *artificial game intelligence*, a general AI which can play any game, is largely bottlenecked by the lack of a rich dataset of games. The recent trend of meta-learning methods, or AI that learn to learn, also requires a wide distribution of tasks in order to reliably function. Thus, OEE systems that can produce vast amounts of interesting games would be of great value.

## Questions From the Workshop

This section reviews the main questions we discussed.

### What is the role of exaptation in OEE?

Exaptations (Gould and Vrba, 1982), also known as *pre-adaptations* or *co-options*, were an exciting topic of discussion at the workshop. In essenence, an exaptation is an evolutionary innovation that develops in some evolutionary context but then becomes useful for a completely different purpose. An example from evolutionary biology is the feather, which initially primarily served the function of insulation but then turned out to be useful for flight. Another example from the domain of technological evolution is the screw press for making wine, which was eventually co-opted into Gutenberg's printing press. It is not immediately clear what role exaptation serves for OEE, if all adaptations are indeed exaptations, or if we might be able to predict or discover novel exaptations.

### What simulated worlds do we care about?

The artificial life community has given rise to a variety of worlds intentionally designed to be open-ended, including (but not limited to) Tierra (Ray, 1992), Avida (Ofria and Wilke, 2004), Evosphere (Miconi and Channon, 2005), PolyWorld (Yaeger, 1994), Geb (Channon, 2001), and Chromaria (Soros and Stanley, 2014). However, there may be utility in widening the purview of open-endedness research. One day of the workshop focused on generative systems in virtual worlds, with many discussions centering on what features of the video game Minecraft might be open-ended even though there is no evolution. An important feature of Minecraft that makes it *not* open-ended is that the world only changes substantially when the player acts in the world and the effects of local interactions have very limited effects.

**Is an open-ended world the same as one displaying open-ended evolution?** An important outcome of the workshop was clarifying the difference between open-endedness in general and open-ended evolution in particular. Open-endedness seems to be a property of search processes that divergently explore a space. Alternatively, if we consider a space itself to be open-ended, we might say that it would be impossible for a search algorithm to fully exhaust the space, such as humans exploring the space of all possible works of art. Open-ended evolution, then, requires specific operations such as mutation, selection, reproduction, etc. Making the distinction between open-endedness and open-ended evolution more precise requires ongoing effort.

**Do constraints make environments more interesting?** We frequently discussed the effect of introducing *limitations* to a world. For example, compare Minecraft's Creative Mode (where players have unlimited resources) versus Survival Mode (where resources must be collected). The space of possible creations in Creative Mode is much higher. On the other hand, one can argue that certain creations are more interesting in Survival Mode, e.g. a house of diamonds implies the player had to go and collect those diamonds. On a high level, this discussion questioned whether open-ended worlds should be designed such that creations are easily made, therefore more creations are possible, or if limitations and constraints make creations more interesting.

## How can we measure open-ended qualities?

As discussion advanced towards future goals, and on referencing past literature, a common question arose: *How can we measure progress?* Talks often converged to this point, with three main directions standing out.

**Do task-agnostic metrics exist?** A recurring barrier when discussing progress towards open-endedness was our lack of task-agnostic metrics. Is there some way to measure complexity, novelty, or open-endedness that is applicable across many kinds of environments? As researchers, many of us have qualitative measurements of success, or specific features we want to see, but without a concrete measurement it is hard to compare results across various bodies of work. An additional desire brought up was the desire for modularity in experiments. In many works, the agent, environment, and method are tightly coupled, thus it is hard to tell which aspect is the primary contributor to any new results. Frameworks to reduce this confusion, such as a common set of environments to test on, expressed high desire.

**What marks end of an OEE system?** Much work focuses on the question of what makes a system open-ended, but as suggested by Dolson et al. (2015), it may be helpful to reframe discussions in terms of what open-endedness is *not*. One potentially new way of thinking about the end of OEE might be that the system becomes predictable to some observer of the system. The observer might be inside the system (an endophysics point of view) or outside of it. Relatedly, many discussions addressed the question of to what extent open-endedness might be subjective.

**How can we judge OEE without being biased by the human mind?** An open-ended system can be seen as a procedure that continuously generates interesting designs. But what makes something interesting? Discussion around this question often arrived at the conclusion that interestingness is highly dependent on the viewpoint taken. Interesting to humans may be different from interesting to an AI or interesting in a universal sense.

As humans, we have an inherent metric for interestingness, developed through biological and cultural evolution. A hefty discussion took place on what *humans* find interesting, and if they were relevant only subjectively or on a more fundamental level. At some level, humans value novelty, but we don't stand around looking at random number generators all day. A key reference was the "10,000 bowls of oatmeal" problem (Compton, 2016), which states that 10,000 bowls of oatmeal provide large amounts of variation on a granular level, but are basically identical on a conceptual level. One thread proposed that humans inherently construct a model of our world, and novel information is viewed as interesting if it allows us to update our predictive models.

As researchers, we naturally judge results from our own viewpoints, thus the questions rise: Are we being fair when judging open-ended systems? Are we only looking at the open-ended behavior that is recognizable to humans? If so, is human judgement a good proxy for fundamentally interesting open-endedness, or are our biases limiting the scope of our work?

# Moving Forward

Many interesting questions were raised over the course of the week-long workshop, and not all were easily or readily answered. Moving forward, here are some topics that we at Cross Labs are interested in pursuing in the near future:

**Evolution of Evolution.** How can evolutionary systems *emerge*, without a human designer? In our world, we can identify many systems which we would describe as displaying open-ended evolution: biological evolution, the progression of culture, the development of technology. What do these systems have in common, and how did they emerge from one another? What principles are required for an evolutionary system to support the development of a higher-level evolutionary system on top of it? Would it be more feasible to build off a rich evolutionary system (i.e, developing cultural evolution on top of biological evolution) than from a less structured system (i.e. developing biological evolution on top of chemical autopoesis)?

**Empowerment** *Empowerment* (Klyubin et al., 2008) is a concept from information theory for quantifying the degree to which an embodied agent knows that it can take actions in its environment. Because of the strong ties to embodied, embedded, and enactive theories of cognition, we believe there is strong potential in exploring this concept more in the context of artificial life and open-ended evolution.

**Artificial life & video games** Artificial life research and games research share a common theme of computational creativity, albeit with different approaches and goals. This overlap was particularly apparent when discussing evolution and generative systems in virtual worlds on Day 2 of the workshop. On a practical level, many ALife-inspired techniques are now used for e.g. procedural content generation in games, but other hybrid research endeavors such as the Open-Endedness in Minecraft challenge and the Generative Design in Minecraft Competition could potentially inspire new perspectives on the theoretical science as well.

## Conclusion

This abstract presented takeaways from the 2021 Cross Labs Innovation Science Workshop. In particular, we focus on a set of open questions centered on how best to encourage the emergence of new kinds of entities and individuals in an evolving system. These questions were addressed from a multi-disciplinary perspective, with emphasis on contributions from AI and cognitive science.

## Acknowledgments

These reflections are based on discussions with the workshop participants, which included Kai Aruklumaran, Martin Biehl, Bert Chan, Javier Fernandez, Nicholas Guttenberg, Giovanni Iacca, Takashi Ikegami, Sina Khajehabdollahi, Miklos Kepes, Mahdi Khosravy, Joel Lehman, Jan Prosi, Penny Faulkner Rainford, Oneris Rico, Sebastian Risi, Christoph Salge, Lana Sinapayen, Elena Tapia, Tim Taylor, Nathaniel Virgo, Jaroslav Vitku, and Dennis Wilson.

## References

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.

Channon, A. (2001). Passing the alife test: Activity statistics classify evolution in geb as unbounded. In *Advances in Artificial Life*, pages 417–426. Springer.

Compton, K. (2016). So You Want To Build A Generator. https://galaxykate0.tumblr.com/post/139774965871/so-you-want-to-build-a-generator.

Dolson, E., Vostinar, A., and Ofria, C. (2015). What's holding artificial life back from open-ended evolution? *The Winnower*.

Gould, S. J. and Vrba, E. S. (1982). Exaptation-a missing term in the science of form. *Paleobiology*, pages 4–15.

Klyubin, A. S., Polani, D., and Nehaniv, C. L. (2008). Keep your options open: An information-based driving principle for sensorimotor systems. *PloS one*, 3(12):e4018.

Miconi, T. and Channon, A. (2005). A virtual creatures model for studies in artificial evolution. In *The 2005 IEEE Congress on Evolutionary Computation*, volume 1, pages 565–572. IEEE.

Ofria, C. and Wilke, C. O. (2004). Avida: A software platform for research in computational evolutionary biology. *Artificial life*, 10(2):191–229.

Packard, N., Bedau, M. A., Channon, A., Ikegami, T., Rasmussen, S., Stanley, K., and Taylor, T. (2019a). Open-Ended Evolution and Open-Endedness: Editorial Introduction to the Open-Ended Evolution I Special Issue. *Artificial Life*, 25(1):1–3.

Packard, N., Bedau, M. A., Channon, A., Ikegami, T., Rasmussen, S., Stanley, K. O., and Taylor, T. (2019b). An Overview of Open-Ended Evolution: Editorial Introduction to the Open-Ended Evolution II Special Issue. *Artificial Life*, 25(2):93–103.

Ray, T. S. (1992). An approach to the synthesis of life. In *Proc. of Artificial Life II*, pages 371–408.

Ruiz-Mirazo, K., Umerez, J., and Moreno, A. (2008). Enabling conditions for 'open-ended evolution'. *Biology & Philosophy*, 23(1):67–85.

Soros, L. B. and Stanley, K. O. (2014). Identifying necessary conditions for open-ended evolution through the artificial life world of chromaria. In Sayama, H., Rieffel, J., Risi, S., Doursat, R., and Lipson, H., editors, *Proceedings of the Fourteenth International Conference on the Simulation and Synthesis of Living Systems (Artificial Life 14)*, pages 793–800.

Standish, R. K. (2003). Open-ended artificial evolution. *International Journal of Computational Intelligence and Applications*, 3(02):167–175.

Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., McMullin, B., Packard, N., Rasmussen, S., Virgo, N., Agmon, E., Clark, E., McGregor, S., Ofria, C., Ropella, G., Spector, L., Stanley, K. O., Stanton, A., Timperley, C., Vostinar, A., and Wiser, M. (2016). Open-ended evolution: perspectives from the oee workshop in york. *Artificial life*, 22(3):408–423.

Von Neumann, J. (1951). The general and logical theory of automata. In Jeffress, L. A., editor, *"Cerebral Mechanisms in Behavior. The Hixon Symposium*, pages 1–31. John Wiley Sons, New York, NY.

Yaeger, L. (1994). PolyWorld: Life in a new context. *Proc. Artificial Life*, 3:263–263.