



Autonomous Knowledge Formation in Artificial Intelligence Systems Under Open-Ended Environments

Aman Gupta

Research Associate, India.

Sai Prakash Narasingu

Software Engineer –Cloud Observability, USA.

Abstract

The emergence of open-ended environments in artificial intelligence (AI) presents a paradigm shift in how intelligent agents acquire, represent, and utilize knowledge. Unlike conventional task-specific systems, open-ended systems require agents to autonomously form knowledge through continuous interaction with dynamic, uncertain, and evolving environments. This paper explores recent advances in autonomous knowledge formation, highlighting the architectural, cognitive, and computational mechanisms that enable self-driven learning and adaptation. We integrate insights from developmental robotics, lifelong learning, and intrinsic motivation frameworks, and propose an updated conceptual and practical framework for knowledge emergence. A comprehensive literature review is included to establish the historical trajectory of research, alongside visual illustrations of knowledge dynamics and architectures.

Keywords:

Autonomous knowledge formation, open-ended learning, cognitive autonomy, adaptive AI, artificial general intelligence, knowledge emergence.

How to cite this paper: Aman Gupta, Sai Prakash Narasingu. (2026). Autonomous Knowledge Formation in Artificial Intelligence Systems Under Open-Ended Environments. *ISCSITR-International Journal of Artificial Intelligence (ISCSITR-IJAI)*, 7(1), 1–7.

URL: https://icsitr.com/index.php/ISCSITR-IJAI/article/view/ISCSITR-IJAI_2026_07_01_001/ISCSITR-IJAI_2026_07_01_001

Published: 08th January 2026

Copyright © 2026 by author(s) and International Society for Computer Science and Information Technology Research (ISCSITR). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. Introduction

The capacity for autonomous knowledge formation is essential for building artificial agents capable of long-term adaptation in open-ended environments. These environments are inherently unstructured, unpredictable, and continually evolving, necessitating systems that can learn not only from data but from the dynamic affordances of interaction itself. Traditional AI systems—designed for predefined tasks—fall short when exposed to contexts that evolve beyond their training distributions.

Autonomous knowledge formation refers to an agent's ability to generate, update, and refine internal representations of the world without direct supervision. This includes forming abstract concepts, recognizing environmental patterns, and adapting strategies in real-time. It is not merely about learning from data but about continuously constructing meaning from experience. This paradigm aligns closely with biological cognition, where knowledge is developed through embodied experience and situated interaction.

2. Literature Review

The foundations of autonomous knowledge formation were laid by early work on cognitive architectures and symbolic AI. *Vernon et al.* discussed the importance of developmental robotics in supporting emergent cognitive capabilities. *Fujita* proposed open-ended learning agents that simulate neural dynamics. *Angelov* emphasized the role of real-time data streams in evolving autonomous learning systems.

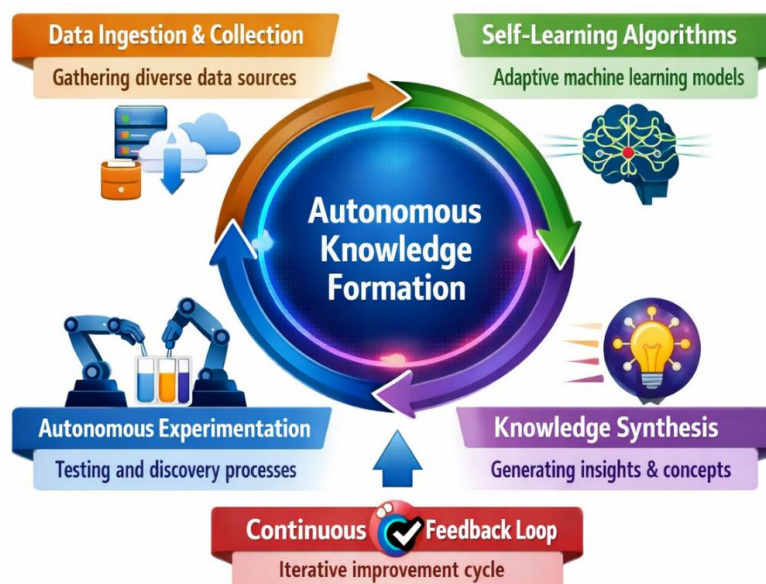
Weinbaum and Veitas developed the concept of open-ended intelligence, drawing from individuation theory, while *Santucci et al.* introduced intrinsic motivation as a driver of lifelong exploration. *Kunze et al.* surveyed robotic autonomy, highlighting persistent learning in dynamic environments. *Cartoni et al.* introduced Real-x, a fully autonomous architecture

for sensorimotor exploration.

In collaborative learning systems, *Ogata and Yano* examined how information filtering and awareness mechanisms affect knowledge building in open-ended environments. *Oudeyer* emphasized self-organization and skill emergence in developmental agents. Collectively, these works provided a theoretical and experimental basis for open-ended AI systems.

3. Conceptual Framework for Knowledge Formation

Autonomous knowledge formation in AI requires a synthesis of four core mechanisms: intrinsic motivation, sensorimotor integration, self-organization, and symbolic abstraction. These components interact within an agent's control architecture to continuously update and refine internal representations based on external feedback.



Intrinsic motivation, modeled as curiosity or uncertainty reduction, guides exploration in the absence of extrinsic rewards. Agents engage in exploratory behaviors that maximize learning progress, leading to novel concept formation. Sensorimotor integration enables agents to associate bodily actions with environmental changes, forming grounded knowledge representations.

Table 1: Comparison Between Supervised Learning and Autonomous Knowledge Formation Systems

Feature	Supervised Learning	Autonomous Knowledge Formation
Environment	Closed	Open-ended
Learning Objective	Task-specific	Self-defined
Knowledge Representation	Fixed models	Emergent, evolving
Supervision	Human-labeled	Self-organized
Adaptability	Limited	High

4. Methodological Approaches

AI landscape includes several notable approaches to support autonomous knowledge formation. Chief among them is Autotelic Reinforcement Learning, where agents define and revise their own goals based on interaction history. This method supports skill hierarchies that are contextually adapted.



Figure 2: Open-Ended Knowledge Formation Process

Continual Learning (CL) is also foundational, particularly when enhanced with memory consolidation mechanisms to reduce catastrophic forgetting. Meta-learning strategies enable agents to learn how to learn, optimizing knowledge formation strategies over time.

Another technique is Open-Ended Evolutionary Developmental Architectures (Eco-Evo-Devo) that simulate natural evolutionary processes to yield adaptive, flexible cognition. These systems use simulated ecosystems to encourage adaptation through selective pressures, mutation, and self-organization.

5. Challenges in Open-Ended Environments

Open-ended environments present challenges in stability, control, and evaluation. One key difficulty lies in the **non-stationarity** of these environments—agents must constantly recalibrate their internal models to reflect novel affordances. Another issue is **knowledge validation** in the absence of ground truth.

Moreover, **catastrophic interference** remains a critical problem, where new knowledge overwrites old representations. Techniques such as elastic weight consolidation and experience replay mitigate these effects but remain limited in fully autonomous systems.

Table 2: Challenges in Open-Ended Environments and Mitigation Techniques

Challenge	Description	Mitigation Technique
Catastrophic Forgetting	Loss of previous knowledge	Elastic Weight Consolidation, Replay
Non-stationarity	Constantly changing environment	Online adaptation, meta-learning
Sparse Feedback	Limited external reinforcement	Intrinsic motivation, curiosity models
Evaluation	Lack of performance benchmarks	Self-assessment metrics, behavioral logs

6. Applications and Future Trends

Autonomous AI systems are increasingly applied in robotic systems, autonomous vehicles, smart tutoring systems, and scientific discovery engines. In robotics, self-exploration enables adaptive locomotion and manipulation across terrains. In education, AI tutors adjust pedagogy based on student behavior patterns.

Looking forward, the convergence of neuromorphic computing and biological modeling will foster energy-efficient, real-time autonomous knowledge formation. Likewise, symbolic-neural hybrids may bridge the gap between statistical learning and structured reasoning.

Autonomous agents will play vital roles in scientific discovery, generating hypotheses, conducting simulations, and formulating theories without human input.

7. Conclusion

The shift toward autonomous knowledge formation under open-ended environments redefines the essence of intelligence in machines. By moving beyond static datasets and fixed goals, AI systems begin to emulate cognitive processes of learning, adaptation, and abstraction akin to human and animal cognition. Integrating dynamic interaction with self-driven exploration, these systems evolve with their environments, unlocking unprecedented levels of generalization and autonomy. Future research must address the scalability, ethics, and explainability of such agents to ensure their responsible deployment.

References

- [1] Vernon, D., Metta, G., & Sandini, G. (2007). A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents. *IEEE Transactions on Evolutionary Computation*, 11(2), 151–180.
- [2] Fujita, M. (2009). Intelligence dynamics: A concept and preliminary experiments for open-ended learning agents. *Autonomous Agents and Multi-Agent Systems*, 19(1), 9–30.
- [3] Angelov, P. (2012). *Autonomous Learning Systems: From Data Streams to Knowledge in Real-Time*. Springer, Berlin.
- [4] Weinbaum, D., & Veitas, V. (2017). Open-ended intelligence: The individuation of intelligent agents. *Journal of Experimental & Theoretical Artificial Intelligence*, 29(2), 371–396.
- [5] Santucci, V. G., Oudeyer, P.-Y., & Barto, A. (2020). Intrinsically motivated open-ended learning in autonomous robots. *Frontiers in Neurorobotics*, 13, Article 115.
- [6] Cartoni, E., Montella, D., & Triesch, J. (2023). Real-X—Robot open-ended autonomous learning architecture: Building truly end-to-end sensorimotor autonomous learning systems. *IEEE Systems Journal*, 17(4), 6632–6644.
- [7] Ogata, H., & Yano, Y. (2000). Combining knowledge awareness and information filtering in an open-ended collaborative learning environment. *International Journal of Artificial Intelligence in Education*, 11(1), 31–52.

- [8] Kunze, L., Hawes, N., & Duckett, T. (2018). Artificial intelligence for long-term robot autonomy: A survey. *IEEE Robotics and Automation Magazine*, 25(3), 82–90.
- [9] Oudeyer, P.-Y., & Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 1, Article 6.
- [10] Bahle, G., Calma, A., & Leimeister, J. M. (2016). Lifelong learning and collaboration of smart technical systems in open-ended environments—Opportunistic collaborative interactive learning. In *2016 IEEE International Conference on Industrial Technology (ICIT)* (pp. 1150–1155). IEEE.
- [11] Srivastava, P., & Singh, J. (2025). Autotelic Reinforcement Learning: Exploring Intrinsic Motivations for Skill Acquisition in Open-Ended Environments. *arXiv preprint*, arXiv:2502.04418.
- [12] Liu, B., Mazumder, S., Robertson, E., & Grigsby, S. (2023). AI autonomy: Self-initiated open-world continual learning and adaptation. *AI Magazine*, 44(2), 36–48.
- [13] Hamon, G. (2025). *Towards open-ended dynamics in Artificial Life and Artificial Intelligence: An Eco-Evo-Devo Perspective* (Doctoral dissertation). HAL Open Science.
- [14] Moulin-Frier, C. (2022). *The ecology of open-ended skill acquisition* (Doctoral dissertation). HAL INRIA.
- [15] Golilarz, N. A., Penchala, S., & Rahimi, S. (2025). Bridging the gap: Toward cognitive autonomy in artificial intelligence. *arXiv preprint*, arXiv:2512.02280.