ARTICLE IN PRESS

Artificial Intelligence xxx (xxxx) xxx

Contents lists available at ScienceDirect

Artificial Intelligence

journal homepage: www.elsevier.com/locate/artint

Introduction to open-world AI

Lawrence Holder [a,*], Pat Langley [b], Bryan Loyall [c], Ted Senator [d]

[a] *School of Electrical Engineering and Computer Science, Washington State University, Box 642752, Pullman, WA, 99164, USA*
[b] *Institute for the Study of Learning and Expertise, 2164 Staunton Court, Palo Alto, CA, 94306, USA*
[c] *Charles River Analytics, 625 Mt. Auburn St., Cambridge, MA, 02138, USA*
[d] *Defense Advanced Research Projects Agency, 675 N. Randolph St., Arlington, VA, 22203, USA*

A B S T R A C T

Open-world AI is characterized by sudden novel changes in a domain that are outside the scope of the training data, or the deployment of an agent in conditions that violate the implicit or explicit assumptions of the designer. In such situations, the AI system must detect the novelty and adapt in a short time frame. In this introduction to the special issue on open-world AI, we discuss the background and motivation for this new research area and define the field in the context of similar AI challenges. We then discuss recent research in the area that has made significant contributions to the field. Many of those contributions are reflected in the papers of this special issue, which we summarize alongside more traditional approaches to open-world AI. Finally, we discuss future directions for the field.

1. Background and motivation

> There are more things in heaven and earth, Horatio,
> Than are dreamt of in your philosophy.
> (*Hamlet 1.5.167–168*)

AI systems have become increasingly successful on tasks for which there is a well-understood domain model and reasoning approach, or significant amount of training data and time available, and the underlying assumptions of the task do not vary significantly over time. Recent examples include the application of deep convolutional networks on image classification [1], deep reinforcement learning on games [2], and large language models on knowledge retrieval and reasoning [3]. Longer standing approaches often taken for granted include routing and scheduling, e.g., where AI software allows UPS to save $300 to $400 million annually in fuel, wages and vehicle running costs alone [4], tax software that millions of people rely on annually, and NASA mission software [5]. Yet, when the tasks change rapidly or assumptions are violated, AI systems can fail in scenarios where humans can robustly adapt [6–9]. Specifically, the AI system might be deployed in conditions that violate the implicit or explicit assumptions of the designer or that are outside the scope of the training data. An example in the domain of image classification is the open-set scenario, in which previously-unseen classes appear (e.g., a new disease in medical images [10], or a new species in wildlife images [11]). A non-open-world AI system incorrectly classifies the new image as one of its known classes. Preferably, the system would first detect the presence of a novel class and then adapt its model to distinguish better the new class from others going forward. Another example is the appearance of novel adversarial behaviors in games (e.g., cyclic attacks against deep Go models [12]). The commensurate decrease in game-play performance assists with the detection of novelty, but adaptation to combat the strategy is challenging given the lack of experience with the strategy combined with the need for quick adaptation to regain performance or even to exploit the novel behavior. Such

* Corresponding author.
E-mail address: holder@wsu.edu (L. Holder).

https://doi.org/10.1016/j.artint.2025.104393

Available online 4 July 2025
0004-3702/© 2025 Published by Elsevier B.V.

Please cite this article as: Lawrence Holder et al., *Artificial Intelligence*, https://doi.org/10.1016/j.artint.2025.104393

adversarial attacks are becoming increasingly difficult to anticipate given the adversary's ability to use AI systems to find weaknesses in existing AI systems [13].

The challenges demonstrated above are the hallmarks of Open-World AI: sudden changes in the domain that impact performance due to use of the agent in conditions that violate design assumptions or that are outside the scope of training data, and the need for detection and adaptation in a short time frame. In addition, the agent would ideally characterize the novelty, which can assist in the selection of an appropriate adaptation technique. Open-world AI systems need mechanisms to deal quickly with novel situations, where collecting and retraining on large amounts of data is costly, if not impossible. The development of such mechanisms will have a significant impact on the capability and robustness of AI systems as we enter an era of AI systems combining general intelligence and adaptive decision making that "could enable fully autonomous decision-making in novel or unpredictable situations, such as coordinating multiple unmanned systems in complex, dynamic environments." [14]

This special issue presents several approaches for designing AI systems that adapt effectively to open-world novelty along with techniques for evaluating such systems. First, we provide a more precise definition of open-world AI and the challenges presented by open worlds. Then, we discuss the Science of Artificial Intelligence and Learning for Open-world Novelty (SAIL-ON) program that facilitated significant advances in open-world AI. Next, we discuss approaches to open-world AI with a focus on the papers in this special issue. Finally, we discuss future directions for the field.

2. Open-world AI

Open-world AI is characterized by sudden and unannounced changes in the environment that degrade an agent's performance due to deployment or use in conditions or environments unanticipated by the designers or not covered by training data. In order to regain acceptable performance, the agent must detect, characterize and accommodate the novelty with limited time and experience. This differs from environments in which the agent has sufficient time and experience to retrain from scratch or adapt over the long term, where techniques such as reinforcement learning have been successful. This also differs from non-real time adaptation mediated by developers. For example, autonomous driving collects data from large numbers of initially equivalent agents and which then results in retraining and version updating mediated by developers. By contrast, in open-world AI, the agents must recognize and adapt to novelty by themselves and in real time. New approaches are needed to rapidly repair expertise with limited experience, drawing from multiple paradigms, such as model diagnosis and repair, plan monitoring, meta-cognition, problem reformulation, change detection, theory revision, scientific discovery, and transfer of learned expertise. Of particular interest are approaches that span both perceptual and interactive settings and that support both reactive and deliberative behavior. In addition to the development of new approaches, research is needed into the sources and classes of novelty, methods for generating and evaluating approaches, and theoretical frameworks for understanding open-world AI.

The term "open world" has been used in the literature to describe various specializations of open-world AI. For example, "open world" has been used to describe out of distribution learning tasks [15] and learning in the presence of previously-unseen classes [16]. Other terms have been proposed to generalize over these tasks, such as "open-environment machine learning" [17]. But "open world" has been used in a more general context in the last few years [18,19] to also encompass the challenges mentioned above, leading to integrated approaches [20] and novel theoretical frameworks [21,22].

Open-world AI is a critical real-world capability needed by AI systems to adapt to novel environments. The timeliness of the topic prompted the 2022 AAAI Spring Symposium on "Designing Artificial Intelligence for Open Worlds" that included 22 papers and other talks and panels on the topic [23]. Other special issues have appeared on this topic. The *Pattern Recognition* journal's special issue on "Open World Robust Pattern Recognition" [24] focuses on pattern recognition, which is a subset of the types of open-world tasks confronting AI systems, in particular, decision-making and planning tasks. The *Neural Networks* journal's special issue on "Lifelong Learning" [25] is related in terms of the need to learn new tasks and adapt behavior to changing tasks over time, but many open-world scenarios require immediate adaptation, and approaches encompass a much larger set of paradigms beyond just neural networks. The need for advances in open-world AI is highlighted in the area of defense by the Defense Advanced Research Project Agency program on the Science of Artificial Intelligence and Learning for Open-world Novelty (SAIL-ON), which resulted in significant advances in the field.

3. SAIL-ON

For AI to be used widely for defense applications, the ability to recognize and act in novel contexts is essential, as is a rigorous engineering methodology for scaling to many diverse applications. For this reason, the Defense Advanced Research Project Agency (DARPA) formulated and executed the Science of Artificial Intelligence and Learning for Open-world Novelty (SAIL-ON) program, starting in 2019 and finishing in 2023 [26]. The SAIL-ON program had three goals: (1) develop scientific principles to quantify and characterize novelty in open-world domains; (2) create AI systems that act appropriately and effectively in open-world domains; and (3) demonstrate and evaluate these systems in multiple domains. The SAIL-ON program wanted to ensure that novelties were truly unknown and unknowable to the agent and its designer. Frequently, the agent evaluator and agent designer had been one and the same [27], which means that novelties are not truly unanticipated. These experiments misrepresent the open world capabilities of the agent and often lead to over-engineered AI/ML applications failing when deployed to the real-world [28,29]. The program addressed these issues by (i) instituting a firewall between the evaluators and the agent designers and (ii) requiring the agents to be evaluated in multiple domains. The program selected ten agent design teams and six evaluation teams. Papers from three of the agent design teams [30–32] and one of the evaluation teams [33] appear in this special issue. Many of the results from the SAIL-ON program

Table 1
The open-world novelty hierarchy developed under the SAIL-ON program describing different categories and types of novelties with examples from chess.

Category	Type	Description	Chess Example
Single Entities	Objects	New classes, attributes, or representations of non-volitional entities.	Number of board columns increased.
	Agents	New classes, attributes, or representations of volitional entities.	Opponent uses new opening strategy.
	Actions	New classes, attributes, or representations of external agent behavior.	Opponent takes more time per move.
Multiple Entities	Relations	New classes, attributes, or representations of static properties of the relationships between multiple entities.	Rooks and bishops swap starting positions.
	Interactions	New classes, attributes, or representations of dynamic properties of behaviors impacting multiple entities.	Capturing opponent's piece removes agent's piece.
Complex Phenomena	Environment	New classes, attributes, or representations of elements independent of specific entities.	All pieces pushed forward one square every three turns.
	Goals	New classes, attributes, or representations of external agent objectives.	One of each type of piece must be captured to win.
	Events	New classes, attributes, or representations of series of state changes that are not each the direct result of volitional action by an agent.	Captured pieces reenter the game after delay.

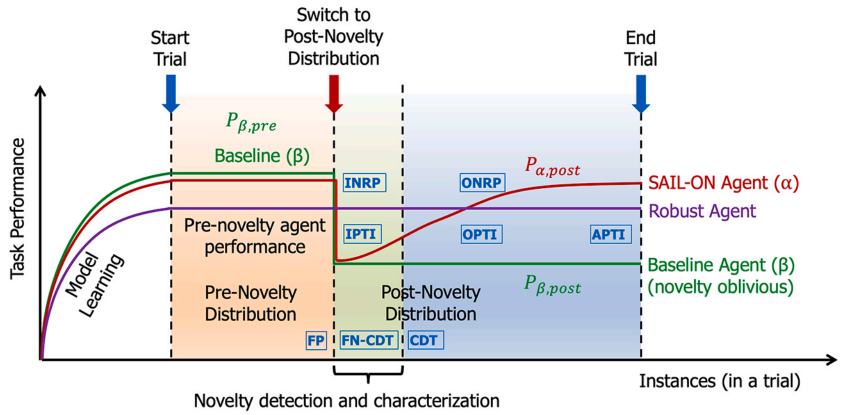

Fig. 1. The SAIL-ON experiment methodology, where the SAIL-ON agent and a non-novelty-aware Baseline agent are given a series of instances. At some point, novelty is introduced into the instances, and the SAIL-ON agent is assessed according to its ability to recover performance and detect novelty. While a Baseline agent is unlikely to recover, a Robust agent can be resilient to the novelty. Metrics (in bold) measure various aspects of the agent's novelty detection and adaptation based on pre and post performance (see Table 2).

were published at the AAAI Spring Symposium on Designing Artificial Intelligence for Open-Worlds [23] and the SAIL-ON GitHub repository [34].

The SAIL-ON program achieved several major accomplishments in the field of open-world AI. The evaluation teams created novelty generators in one or more domains, including interactive tasks (Angry Birds, CartPole, MineCraft, Monopoly, ViZDoom) and perceptual tasks (image recognition, activity recognition). Interactive tasks provide goals to the agent, allow the agent to affect the environment through actions, and facilitate exploration in the environment. Perceptual tasks provide feedback to the agent regarding perception performance, but do not support agent interaction. The program also developed a separate military-relevant domain for evaluating AI agents [35,36]. The design teams developed several different types of agents to detect and accommodate novelty, including capabilities in perception, learning, planning, reasoning, and combinations. The program sought to formalize different categories and types of novelty, called the *novelty hierarchy*. The hierarchy underwent several revisions, which resulted in a formal framework for characterizing novel environment transformations [37]. The final version of the novelty hierarchy is shown in Table 1 with examples from the domain of chess. By the end of the program each AI agent had to handle all eight different types of novelty for three different domains and the military-relevant domain. In the final phase of the program, the agents were also expected to characterize the novelties according to the novelty hierarchy in a language agreed upon between the agent and domain teams.

The SAIL-ON program developed an experiment methodology and set of metrics to assess different aspects of an agent's ability to detect and adapt to novelty [38]. Fig. 1 shows the progress of an experiment trial in the SAIL-ON methodology. Agents are assumed to have done some pretrial model learning on non-novel instances from the domain. A state-of-the-art non-novelty-aware Baseline agent is also trained on the pretrial data. A trial begins with non-novel data presented to both the SAIL-ON and Baseline agents. The relevant task performance (e.g., accuracy, percent games won) is collected throughout the trial. At some point during, unknown to the agent, novelty is introduced. Typically, both the Baseline and SAIL-ON agents experience a decline in performance. The Baseline agent, possessing no novelty adaptation apparatus, does not recover, but a novelty-robust agent will regain performance, ideally reaching,

Table 2
SAIL-ON performance metrics used for evaluating novelty detection (FN-CDT, CDT, FP) and reaction (ONRP, INRP, OPTI, IPTI, APTI). Metrics marked with ↓ indicate that lower values are better.

Metric	Description
FN-CDT ↓	Number of instances before novelty detection (for correctly detected trials).
CDT	Fraction of trials that are correctly detected trials (CDTs).
FP ↓	Fraction of trials with at least one false positive (FP).
ONRP	Overall Novelty Reaction Performance (NRP) relative to baseline pre-novelty performance.
INRP	NRP for the first 10% of instances after novelty introduction.
OPTI	Overall Performance Task Improvement (PTI) across all post-novelty instances.
IPTI	Initial PTI based on the first 10% of post-novelty instances.
APTI	Asymptotic PTI based on the last 10% of post-novelty instances.

or even better, exceeding pre-novelty performance levels. Fig. 1 also depicts a Robust agent, representing non-adaptive agents that are resilient to the novelty, which the program found to occur in some cases due to the agent's approach being independent of the details of the novelty. The SAIL-ON agent is also called upon to identify the presence of novelty as soon as possible after introduction, as well as characterize the novelty using the agent's internal representation of the domain. Trials are repeated many times with different novelty introduction points during the trial to ensure statistical validity.

The SAIL-ON program developed several metrics to assess different aspects of agent performance, as outlined in Table 2. These metrics are domain-independent functions of the domain and task dependent metrics that are typically used to evaluate performance of AI agents. The first three metrics evaluate the agent's ability to detect novelty. One key aspect of detection is identifying a Correctly Detected Trial (CDT) - a scenario where the agent detects novelty only after it has been introduced, not before. The CDT metric quantifies the fraction of trials that meet this criterion. The FN-CDT detection metric measures the number of instances that pass after novelty is introduced before the agent successfully detects it. This is calculated only for trials where novelty is eventually recognized (CDTs). The third detection metric, FP (False Positives), determines the proportion of trials in which the agent falsely detects novelty before it actually occurs. Collectively, these three metrics capture both the accuracy and speed of novelty detection. Ideally, a well-performing agent should achieve a high CDT score while keeping FN-CDT and FP scores low.

The remaining five metrics assess how well the agent adapts to novelty in terms of its task performance. A distinctive feature of these metrics is that they evaluate the novelty-aware SAIL-ON agent's performance relative to a state-of-the-art, non-novelty-aware Baseline agent. Before novelty is introduced, the novelty-aware agent should perform at least as well as the Baseline agent. After novelty is introduced, both agents typically experience a performance drop. However, the expectation is that the novelty-aware agent will adapt and eventually surpass the non-novelty-aware Baseline agent.

Two key metrics, ONRP (Overall Novelty Reaction Performance) and INRP (Initial Novelty Reaction Performance), quantify this adaptation. The NRP metric is defined as the ratio of the SAIL-ON agent's performance after novelty is introduced (post-novelty) to the Baseline agent's performance before novelty is introduced (pre-novelty):

$$\mathrm{NRP} = \frac{P_{\alpha,\mathrm{post}}}{P_{\beta,\mathrm{pre}}} \tag{1}$$

where P_α represents the SAIL-ON agent's performance and P_β represents the Baseline agent's performance. ONRP measures this ratio over all post-novelty instances, whereas INRP focuses on the first 10% of post-novelty instances.

Additionally, three Performance Task Improvement (PTI) metrics assess the SAIL-ON agent's relative advantage over the Baseline agent post-novelty. While NRP compares pre- and post-novelty performance, PTI only evaluates post-novelty performance. PTI is defined as:

$$\mathrm{PTI} = \frac{P_{\alpha,\mathrm{post}}}{P_{\alpha,\mathrm{post}} + P_{\beta,\mathrm{post}}} \tag{2}$$

which represents the fraction of post-novelty performance attributed to the SAIL-ON agent. PTI is further categorized into: OPTI (Overall PTI) computed over all post-novelty instances, IPTI (Initial PTI) based on the first 10% of post-novelty instances, and APTI (Asymptotic PTI) based on the last 10% of post-novelty instances.

Three major evaluations were conducted during the SAIL-ON program, each one adding additional novelty types and domains to the agent requirements. While a complete review of the results is beyond the scope of this article, several lessons were learned from the program. *First,* robust accommodation can be achieved by a mix of adaptability and resilience, as depicted by the robust agent in Fig. 1. Reaction performance does not directly depend on accurate novelty detection. Therefore, agents must trade-off the resources necessary to do detection and accommodation and find a balance between exploring the domain for novelty and exploiting the novelty. *Second,* although agent teams developed methods for novelty characterization, this capability did not significantly enhance an agent's ability to adapt to the novelty. *Third,* interactive domains are fundamentally different from perceptual domains in terms of detecting and adapting to novelty. Agents in interactive domains are able to actively seek information and decide whether to explore or exploit. Agents in perceptual domains require feedback from the domain, e.g., requesting shared labels of novel images. This contrasts with interactive domain agents for which feedback is received as they act. *Fourth,* the SAIL-ON evaluation paradigm of separating domains from agents (e.g., hiding novelty and schemes for generating novelties) makes it difficult for peer reviewers to verify results but is essential to build open-world agents. *Fifth,* there is no clear way to measure the difficulty of task a priori, i.e., there is no clear set of independent variables that determine a task's difficulty [39]. Also, the difficulty of a task depends heavily on the abilities of the agent performing the task. Still, identifying characteristics of domains and novelties that fundamentally increase the difficulty of adaptation, or increase the required complexity of the agent, is an interesting future direction [40].

4. Approaches

In this section we discuss existing methods that address some aspects of open-world AI challenges, but fail to address others. We then discuss the specific approaches proposed in the special issue articles. One existing approach to tackle open-world settings is through reinforcement learning (RL), which has demonstrated remarkable success in game-playing tasks (e.g., [41]). However, RL methods typically require millions of iterations, making them impractical for real-world scenarios where large-scale simulations are unavailable. Open-world AI emphasizes limited exposure to the novel environments and the need for rapid adaptation. Several approaches have been proposed based on a system's ability to modify itself to overcome some gap in its expertise for handling open-world settings. Model-based diagnosis and repair [42] can modify an agent's expertise to accommodate novelty, though it is designed for handling specific cases rather than developing general models. Integrated planning, execution, and monitoring [43] enables agents to dynamically generate and adjust plans to achieve objectives. Metacognition [44] analyzes cognitive processes to identify gaps or errors in knowledge. Problem reformulation [45] modifies state and operator representations to improve task feasibility. Change detection in streaming data [46] is primarily used for classification tasks in dynamic environments. Theory revision [47] refines predictive models based on new training data. Scientific discovery [48] uncovers underlying laws and models to explain observed phenomena. Transfer learning [49,50] facilitates the adaptation of previously acquired expertise to new contexts once a change has been identified. While these approaches are valuable, none alone are sufficient to fully address open-world AI. However, a combination of these techniques, structured within an integrated agent architecture, holds promise for a more effective solution.

The papers in the special issue mostly focus on methods for detection and adaptation in open-world environments, although one paper focuses on the challenge of evaluating open-world AI systems. The first two papers focus on open-world classification. Zhao et al. [51] consider the scenario in which unknown classes exist in the domain, and instances of these classes are incorrectly classified using known class labels, i.e., unknown unknowns. Their approach first detects the presence of unknown unknowns using a rejection model [52] to identify unknown instances, then identifies new features that can better distinguish unknown instances from known instances, and then learns an augmented model that better classifies these instances. Both empirical and theoretical analyses confirm the superiority of their approach for the unknown unknowns scenario. Continual or lifelong learning [53] is one approach for adapting to change, although the change is typically more gradual and still within the scope of abilities of the AI agent. Kim et al. [54] extend the continual learning approach into the open-world scenario to accommodate new items in the image classification task. They further show both theoretically and empirically that the closed-world detection and adaptation techniques are necessary for good performance in open-world settings.

The next three papers focus on robust AI approaches to handling novelty in open-world settings. Goel et al. [30] take a more robust approach to a novelty-aware AI agent using a framework that combines symbolic planning, counterfactual reasoning, reinforcement learning, and deep computer vision, where violations of planner expectations initiate detection and adaptation. The framework synergizes lower-level perception and learning with higher-level planning and reasoning, and thus can be applied to a broad spectrum of domains. Extensive empirical evaluation demonstrates the strengths of the approach in a Minecraft-like domain called PolyCraft [55]. Mohan et al. [31] propose the HYDRA framework that augments a planning module with visual reasoning, task selection and action execution modules that allow the agent to monitor its own behavior and detect a divergence from expectations. HYDRA then identifies model changes that can realign the agent with observed behavior and adapt to novel situations. The framework is evaluated on numerous discrete-continuous domains, including PolyCraft, a three-dimensional cartpole world [56], and an Angry Birds novelty generator called Science Birds [57]. Loyall et al. [32] propose the Coltrane planning based system that is designed to rapidly detect, characterize and adapt to novelty, and continues to improve over time as more observations of the environment become available. Major components of their approach include using probabilistic program synthesis to learn minimal changes to the internal domain model to characterize the novelty, and a Monte Carlo tree search (MCTS) [58] that finds new heuristics and feature weighting to improve planning performance post-novelty in the expanded domain model. They demonstrate their approach on several novel changes to the Monopoly game [59] and the VizDoom domain [60,61].

The next two papers focus on more task-specific approaches. For the task of knowledge-based visual reasoning, in which the AI system has to both understand and answer questions about visual scenes, the appearance of new objects and concepts in the scenes is a difficulty challenge. Zheng et al. [62] propose a method for open-world knowledge representation learning (OWKRL) that transfers previously-learned knowledge to new scenes using a novel graph-based self-cross transformer network that learns how to transfer attention across knowledge graph networks to identify new concepts and relationships in a novel visual scene. Experimentation validates the approach and suggests its applicability to related tasks, such as science and medical question answering. Transportation services (e.g., public transport, ride sharing) and their inter-operation represent a complex domain that can be significantly impacted by novel events, such as natural disasters or epidemics. Wang et al. [63] propose an open-world spatio-temporal network (OWST-Net) that can adapt to unexpected changes to the multi-modal network of mobility services. They validate their approach on several real-world mobility datasets and show superior performance over state-of-the-art methods.

Evaluating open-world AI systems is challenging, because true open-world tasks are those we have not yet considered, even in a simulated setting. Still some progress has been made in designing systems for generating novelties and in designing metrics for assessing agent performance. In the final paper of the special issue, Pinto et al. [33] developed the NovPhy system that generates physics-based novelties in the Science Birds simulator [57] and evaluated several AI agents using these novelties. They also assessed human performance on these novel tasks. Results found that despite improvements in AI agents' abilities to detect and adapt to novelties, humans are still superior.

5. Future directions

Open-world AI remains one of the most significant challenges in artificial intelligence, requiring systems that can dynamically adapt to unexpected changes, novel scenarios, and evolving environments. As AI applications extend beyond controlled, well-defined domains into complex, real-world settings - such as autonomous robotics, cybersecurity, and adaptive healthcare - the need for more robust, generalizable, and self-improving AI systems becomes increasingly critical. Future research in this area is likely to focus on several key advancements.

One promising direction is the development of AI architectures that integrate multiple learning paradigms to handle novelty detection, adaptation, and knowledge transfer more effectively. Current approaches often rely on either deep learning or symbolic reasoning, but combining neuro-symbolic AI with meta-learning and continual learning may provide a more holistic approach to handling novelty. Hybrid models could allow AI systems to detect novelties, assess their significance, and modify their internal representations accordingly, leading to better generalization across a wide range of tasks. Some of the papers in this special issue, and further work from the SAIL-ON program [64–69], are taking this hybrid approach, but more work is needed to explore the space and identify best practices.

Another important area of focus is few-shot and zero-shot adaptation [70]. Real-world environments do not provide the vast amounts of labeled training data that AI models typically require. Future AI systems must become proficient at recognizing novel situations with minimal prior experience. Self-supervised learning and unsupervised reinforcement learning are expected to play a larger role in enabling agents to infer patterns, learn new rules on the fly, and generalize across different domains without relying on explicit supervision.

Catastrophic forgetting [71] is a significant challenge in open-world AI, which refers to the tendency of AI systems to abruptly and drastically forget previously learned information upon learning new information. This phenomenon poses a substantial obstacle for AI systems that must adapt continuously to new data and tasks in dynamic environments. To mitigate catastrophic forgetting, researchers have proposed various strategies, e.g., regularization techniques to prevent significant updates to previously acquired knowledge [72], and replay techniques to integrate data from past tasks alongside new data to reinforce earlier learning [73]. Catastrophic forgetting has also been addressed in the context of lifelong learning as an emphasis in the Lifelong Learning Machines (L2M) program [74], highlighting both algorithmic approaches [75] and biological underpinnings [76]. Given open-world AI's emphasis on sudden change and quick response, more robust approaches to catastrophic forgetting are essential for advancing the field.

Additionally, the field is moving towards AI systems that engage in self-explanation and causal reasoning to enhance transparency and adaptability. Many current deep learning models struggle with interpretability, making it difficult to understand how they react to novelty. Future AI should not only be able to identify when it encounters something unexpected but also characterize why a situation is novel and what changes need to be made to adapt successfully. This capability is particularly relevant in safety-critical applications, such as autonomous vehicles and medical diagnosis, where understanding AI decisions is just as important as achieving high accuracy.

Large Language Models (LLMs) can significantly enhance open-world AI by enabling real-time knowledge integration, supporting adaptive reasoning, facilitating agent communication, and improving explainability. Future research should focus on hybrid AI architectures that combine LLM-driven reasoning with real-time sensor-based adaptation, allowing AI systems to not only react to novel situations but also understand and explain them. As AI continues to advance, LLMs will play an increasingly critical role in making open-world AI systems more intelligent, flexible, and human-aligned.

Finally, open-world AI must become more adept at collaborative learning and knowledge sharing. This is particularly useful in dynamic, interconnected environments such as disaster response, industrial automation, military engagements, and global cybersecurity, where multiple AI systems must coordinate in real-time to handle unforeseen challenges. The Shared-Experience Lifelong Learning (ShELL) program [77] advanced this area in the context of lifelong learning [78]. Future research should explore distributed AI networks where multiple AI agents can share knowledge, exchange learned experiences, and collectively adapt to new challenges.

Overall, the next generation of open-world AI will need to move beyond static models trained on predefined datasets and towards truly autonomous learning systems capable of understanding and reasoning about novel situations in real time. By integrating advances in meta-learning, causal inference, neuro-symbolic reasoning, and collaborative AI, researchers can develop more adaptive, reliable, and generalizable AI systems that thrive in complex and unpredictable environments.

Acknowledgements

The authors would like to thank the numerous reviewers who contributed to the special issue. All information in this article comes from publicly-available sources. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency (DARPA) or Army Research Office (ARO), or the U.S. Government.

References

[1] J. Plested, T. Gedeon, Deep transfer learning for image classification: a survey, arXiv:2205.09904, 2022.

[2] X. Wang, S. Wang, X. Liang, D. Zhao, J. Huang, X. Xu, B. Dai, Q. Miao, Deep reinforcement learning: a survey, IEEE Trans. Neural Netw. Learn. Syst. 35 (4) (2024) 5064–5078, https://doi.org/10.1109/TNNLS.2022.3207346.

[3] Y. Zhu, H. Yuan, S. Wang, J. Liu, W. Liu, C. Deng, H. Chen, Z. Liu, Z. Dou, J.-R. Wen, Large language models for information retrieval: a survey, arXiv:2308.07107, 2024.

[4] C.E. Staff, Why ups trucks don't turn left, https://www.cnn.com/2017/02/16/world/ups-trucks-no-left-turns, 2017. (Accessed 16 April 2025).

[5] NASA, Artificial intelligence at nasa, https://www.nasa.gov/artificial-intelligence, 2024. (Accessed 16 April 2025).

[6] A. Nguyen, J. Yosinski, J. Clune, Deep neural networks are easily fooled: high confidence predictions for unrecognizable images, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 427–436.

[7] A. Olmo, S. Sengupta, S. Kambhampati, Not all failure modes are created equal: training deep neural networks for explicable (mis)classification, arXiv:2006.14841, 2021.

[8] L.L.D. Langosco, J. Koch, L.D. Sharkey, J. Pfau, D. Krueger, Goal misgeneralization in deep reinforcement learning, in: K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, S. Sabato (Eds.), Proceedings of the 39th International Conference on Machine Learning, in: Proceedings of Machine Learning Research, vol. 162, PMLR, 2022, pp. 12004–12019, https://proceedings.mlr.press/v162/langosco22a.html.

[9] X. Chen, L. Li, L. Chang, Y. Huang, Y. Zhao, Y. Zhang, D. Li, Challenges and contributing factors in the utilization of large language models (llms), arXiv:2310.13343, 2023.

[10] A. He, T. Li, Y. Zhao, J. Zhao, H. Fu, Open-set semi-supervised medical image classification with learnable prototypes and outlier filter, in: Proceedings of Medical Image Computing and Computer Assisted Intervention – MICCAI 2024, in: LNCS, vol. 15011, Springer Nature Switzerland, 2024, pp. 492–501.

[11] N. Lang, V. Snæbjarnarson, E. Cole, O.M. Aodha, C. Igel, S. Belongie, From coarse to fine-grained open-set recognition, in: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024, pp. 17804–17814.

[12] T. Tseng, E. McLean, K. Pelrine, T.T. Wang, A. Gleave, Can go ais be adversarially robust?, arXiv:2406.12843, 2025.

[13] I.Y. Tyukin, D.J. Higham, A.N. Gorban, On adversarial examples and stealth attacks in artificial intelligence systems, in: 2020 International Joint Conference on Neural Networks (IJCNN), IEEE, 2020, pp. 1–6.

[14] G. Reim, The state of ai in the aerospace industry, Aviat. Week Space Technol. (January 2025), https://aviationweek.com/aerospace/emerging-technologies/state-ai-aerospace-industry.

[15] V. Sehwag, A.N. Bhagoji, L. Song, C. Sitawarin, D. Cullina, M. Chiang, P. Mittal, Analyzing the robustness of open-world machine learning, in: Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, AISec'19, Association for Computing Machinery, New York, NY, USA, 2019, pp. 105–116.

[16] J. Parmar, S.S. Chouhan, V. Raychoudhury, S.S. Rathore, Open-world machine learning: applications, challenges, and opportunities, arXiv:2105.13448, 2022.

[17] Z.-H. Zhou, Open-environment machine learning, Natl. Sci. Rev. 9 (8) (2022) nwac123, https://doi.org/10.1093/nsr/nwac123, https://academic.oup.com/nsr/article-pdf/9/8/nwac123/45957092/nwac123.pdf.

[18] Z. Chen, B. Liu, Open-world learning, in: Lifelong Machine Learning, Springer International Publishing, Cham, 2018, pp. 77–89.

[19] P. Langley, Open-world learning for radically autonomous agents, Proc. AAAI Conf. Artif. Intell. 34 (09) (2020) 13539–13543, https://doi.org/10.1609/aaai.v34i09.7078.

[20] J. Niehaus, B. Loyall, A. Pfeffer, An integrated architecture for online adaptation to novelty in open worlds using probabilistic programming and novelty-aware planning, in: Proceedings of the AAAI Spring Symposium on Designing Artificial Intelligence for Open Worlds, 2022, pp. 1–9, https://usc-isi-i2.github.io/AAAI2022SS/papers/SSS-22_paper_51.pdf.

[21] T. Boult, P. Grabowicz, D. Prijatelj, R. Stern, L. Holder, J. Alspector, M.M. Jafarzadeh, T. Ahmad, A. Dhamija, C. Li, S. Cruz, A. Shrivastava, C. Vondrick, S. Walter, Towards a unifying framework for formal theories of novelty, Proc. AAAI Conf. Artif. Intell. 35 (17) (2021) 15047–15052, https://doi.org/10.1609/aaai.v35i17.17766.

[22] T. Boult, W. Scheirer (Eds.), A Unifying Framework for Formal Theories of Novelty, 1st edition, Synthesis Lectures on Computer Vision, vol. 1, Springer, Cham, 2023, The Editor(s) and Author(s), under exclusive license to Springer Nature Switzerland AG, 2024.

[23] AAAI Spring Symposium, Designing artificial intelligence for open worlds, https://usc-isi-i2.github.io/AAAI2022SS/, 2022. (Accessed 22 February 2025).

[24] C.-L. Liu, C. Suen, Special issue on open world robust pattern recognition, Pattern Recognit. 130 (2022).

[25] S. Ramasamy, H. Fayek, V. Lomonaco, L. Xiaoli, S. Sundaram, Special issue on lifelong learning, Neural Netw. 167 (2023).

[26] Defense Advanced Research Projects Agency, Science of Artificial Intelligence and Learning for Open-world Novelty (SAIL-ON), https://sam.gov/opp/88fdca99de93ddbb74cd8fb51916ceaa/view, 2019. (Accessed 21 February 2025).

[27] B. Recht, R. Roelofs, L. Schmidt, V. Shankar, Do cifar-10 classifiers generalize to cifar-10?, arXiv:1806.00451, 2018.

[28] B. Chakravorti, Why ai failed to live up to its potential during the pandemic, Harv. Bus. Rev. (March 2022), https://hbr.org/2022/03/why-ai-failed-to-live-up-to-its-potential-during-the-pandemic.

[29] R. Ramakrishnan, E. Kamar, D. Dey, J. Shah, E. Horvitz, Discovering blind spots in reinforcement learning, in: Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '18, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2018, pp. 1017–1025.

[30] S. Goel, P. Lymperopoulos, R. Thielstrom, E. Krause, P. Feeney, P. Lorang, S. Schneider, Y. Wei, E. Kildebeck, S. Goss, M.C. Hughes, L. Liu, J. Sinapov, M. Scheutz, A neurosymbolic cognitive architecture framework for handling novelties in open worlds, Artif. Intell. 331 (2024) 104111, https://doi.org/10.1016/j.artint.2024.104111.

[31] S. Mohan, W. Piotrowski, R. Stern, S. Grover, S. Kim, J. Le, Y. Sher, J. de Kleer, A domain-independent agent architecture for adaptive operation in evolving open worlds, Artif. Intell. 334 (2024) 104161, https://doi.org/10.1016/j.artint.2024.104161.

[32] B. Loyall, A. Pfeffer, J. Niehaus, M. Harradon, P. Rizzo, A. Gee, J. Campolongo, T. Mayer, J. Steigerwald, Coltrane: a domain-independent system for characterizing and planning in novel situations, Artif. Intell. 345 (2025) 104336, https://doi.org/10.1016/j.artint.2025.104336.

[33] V. Pinto, C. Gamage, C. Xue, P. Zhang, E. Nikonova, M. Stephenson, J. Renz, Novphy: a physical reasoning benchmark for open-world ai systems, Artif. Intell. 336 (2024) 104198, https://doi.org/10.1016/j.artint.2024.104198.

[34] DARPA, SAIL-ON GitHub repository, https://github.com/darpa-sail-on, 2023. (Accessed 22 February 2025).

[35] T. Chadwick, J. Chao, C. Izumigawa, G. Galdorisi, H. Ortiz-Pena, E. Loup, N. Soultanian, M. Manzanares, A. Mai, R. Yen, D.S. Lange, Characterizing novelty in the military domain, arXiv:2302.12314, 2023.

[36] J. Chao, W. Piotrowski, M. Manzanares, D.S. Lange, Novelty accommodating multi-agent planning in high fidelity simulated open world, in: F. Amigoni, A. Sinha (Eds.), Autonomous Agents and Multiagent Systems. Best and Visionary Papers, Springer Nature Switzerland, Cham, 2024, pp. 201–216.

[37] M. Molineaux, D. Dannenhauer, E. Kildebeck, A framework for characterizing novel environment transformations in general environments, arXiv:2305.04315, 2023.

[38] V. Pinto, J. Renz, C. Xue, P. Zhang, K. Doctor, D.W. Aha, Measuring the performance of open-world ai systems, in: Proceedings of the AAAI Spring Symposium on Designing Artificial Intelligence for Open Worlds, 2022, pp. 1–10, https://usc-isi-i2.github.io/AAAI2022SS/papers/SSS-22_paper_73.pdf.

[39] K. Doctor, C. Task, E. Kildebeck, M. Kejriwal, L. Holder, R. Leong, Toward defining a domain complexity measure across domains, in: Proceedings of the AAAI Spring Symposium on Designing Artificial Intelligence for Open Worlds, 2022, pp. 1–10, https://usc-isi-i2.github.io/AAAI2022SS/papers/SSS-22_paper_79.pdf.

ARTICLE IN PRESS

[40] M.M. Baker, A. New, M. Aguilar-Simon, Z. Al-Halah, S.M. Arnold, E. Ben-Iwhiwhu, A.P. Brna, E. Brooks, R.C. Brown, Z. Daniels, A. Daram, F. Delattre, R. Dellana, E. Eaton, H. Fu, K. Grauman, J. Hostetler, S. Iqbal, C. Kent, N. Ketz, S. Kolouri, G. Konidaris, D. Kudithipudi, E. Learned-Miller, S. Lee, M.L. Littman, S. Madireddy, J.A. Mendez, E.Q. Nguyen, C. Piatko, P.K. Pilly, A. Raghavan, A. Rahman, S.K. Ramakrishnan, N. Ratzlaff, A. Soltoggio, P. Stone, I. Sur, Z. Tang, S. Tiwari, K. Vedder, F. Wang, Z. Xu, A. Yanguas-Gil, H. Yedidsion, S. Yu, G.K. Vallabha, A domain-agnostic approach for characterization of lifelong learning systems, Neural Netw. 160 (2023) 274–296, https://doi.org/10.1016/j.neunet.2023.01.007.

[41] C. Clark, A. Storkey, Training deep convolutional neural networks to play go, in: F. Bach, D. Blei (Eds.), Proceedings of the 32nd International Conference on Machine Learning, in: Proceedings of Machine Learning Research, vol. 37, PMLR, Lille, France, 2015, pp. 1766–1774, https://proceedings.mlr.press/v37/clark15.html.

[42] W. Hamscher, L. Console, J. de Kleer (Eds.), Readings in Model-Based Diagnosis, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1992.

[43] P. Langley, D. Choi, M. Barley, B. Meadows, E. Katz, Generating, executing, and monitoring plans with goal-based utilities in continuous domains, Adv. Cogn. Syst. (May 2017), http://www.cogsys.org/proceedings/2017/paper-2017-1.pdf.

[44] M.T. Cox, Perpetual self-aware cognitive agents, AI Mag. 28 (1) (2007) 32–45, https://doi.org/10.1609/aimag.v28i1.2027.

[45] P. Riddle, M. Barley, S. Franco, Problem reformulation as meta-level search, in: Second Annual Conference on Advances in Cognitive Systems, 2013, pp. 199–216, http://www.cogsys.org/posters/2013/poster-2013-14.pdf.

[46] R. Pears, S. Sakthithasan, Y.S. Koh, Detecting concept change in dynamic data streams, Mach. Learn. 97 (3) (2014) 259–293, https://doi.org/10.1007/s10994-013-5433-9.

[47] S.D. Bay, D.G. Shapiro, P. Langley, Revising engineering models: combining computational discovery with knowledge, in: Proceedings of the 13th European Conference on Machine Learning, ECML'02, Springer-Verlag, Berlin, Heidelberg, 2002, pp. 10–22.

[48] S. Džeroski, P. Langley, L. Todorovski, Computational discovery of scientific knowledge, in: S. Džeroski, L. Todorovski (Eds.), Computational Discovery of Scientific Knowledge: Introduction, Techniques, and Applications in Environmental and Life Sciences, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 1–14.

[49] T.E. Senator, Transfer learning progress and potential, AI Mag. 32 (1) (2011) 84–86, https://doi.org/10.1609/aimag.v32i1.2333.

[50] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proc. IEEE 109 (1) (2021) 43–76, https://doi.org/10.1109/JPROC.2020.3004555.

[51] P. Zhao, J.-W. Shan, Y.-J. Zhang, Z.-H. Zhou, Exploratory machine learning with unknown unknowns, Artif. Intell. 327 (2024) 104059, https://doi.org/10.1016/j.artint.2023.104059.

[52] C. Cortes, G. DeSalvo, M. Mohri, Learning with rejection, in: R. Ortner, H.U. Simon, S. Zilles (Eds.), Algorithmic Learning Theory, Springer International Publishing, Cham, 2016, pp. 67–82.

[53] L. Wang, X. Zhang, H. Su, J. Zhu, A comprehensive survey of continual learning: theory, method and application, IEEE Trans. Pattern Anal. Mach. Intell. 46 (8) (2024) 5362–5383, https://doi.org/10.1109/TPAMI.2024.3367329.

[54] G. Kim, C. Xiao, T. Konishi, Z. Ke, B. Liu, Open-world continual learning: unifying novelty detection and continual learning, Artif. Intell. 338 (2025) 104237, https://doi.org/10.1016/j.artint.2024.104237.

[55] S.A. Goss, R.J. Steininger, D. Narayanan, D.V. Olivença, Y. Sun, P. Qiu, J. Amato, E.O. Voit, W.E. Voit, E.J. Kildebeck, Polycraft world ai lab (pal): an extensible platform for evaluating artificial intelligence agents, arXiv:2301.11891, 2023.

[56] T.E. Boult, N.M. Windesheim, S. Zhou, C. Pereyda, L.B. Holder, Weibull-open-world (wow) multi-type novelty detection in cartpole3d, Algorithms 15 (10) (2022), https://doi.org/10.3390/a15100381.

[57] C. Gamage, V. Pinto, C. Xue, M. Stephenson, P. Zhang, J. Renz, Novelty generation framework for ai agents in angry birds style physics games, in: 2021 IEEE Conference on Games (CoG), 2021, pp. 1–8.

[58] C.B. Browne, E. Powley, D. Whitehouse, S.M. Lucas, P.I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, S. Colton, A survey of Monte Carlo tree search methods, IEEE Trans. Comput. Intell. AI Games 4 (1) (2012) 1–43, https://doi.org/10.1109/TCIAIG.2012.2186810.

[59] M. Kejriwal, S. Thomas, A multi-agent simulator for generating novelty in monopoly, Simul. Model. Pract. Theory 112 (2021) 102364, https://doi.org/10.1016/j.simpat.2021.102364.

[60] M. Kempka, M. Wydmuch, G. Runc, J. Toczek, W. Jaśkowski, ViZDoom: a Doom-based AI research platform for visual reinforcement learning, in: IEEE Conference on Computational Intelligence and Games, IEEE, Santorini, Greece, 2016, pp. 341–348, The Best Paper Award.

[61] L. Holder, VizDoom novelty generator, https://github.com/holderlb/WSU-SAILON-NG, 2022. (Accessed 16 April 2025).

[62] W. Zheng, L. Yan, F.-Y. Wang, Knowledge is power: open-world knowledge representation learning for knowledge-based visual reasoning, Artif. Intell. 333 (2024) 104147, https://doi.org/10.1016/j.artint.2024.104147.

[63] Z. Wang, R. Jiang, H. Xue, F.D. Salim, X. Song, R. Shibasaki, W. Hu, S. Wang, Learning spatio-temporal dynamics on mobility networks for adaptation to open-world events, Artif. Intell. 335 (2024) 104120, https://doi.org/10.1016/j.artint.2024.104120.

[64] D.J. Musliner, M.J.S. Pelican, M. McLure, S. Johnston, R.G. Freedman, C. Knutson, Openmind: planning and adapting in domains with novelty, in: Ninth Annual Conference on Advances in Cognitive Systems, 2021, pp. 1–20, https://www.sift.net/sites/default/files/publications/main_5.pdf.

[65] G. Burachas, S. Grigsby, B. Ferguson, J. Krichmar, R. Rao, Metacognitive mechanisms for novelty processing: lessons for ai, in: Proceedings of the AAAI Spring Symposium on Designing Artificial Intelligence for Open Worlds, 2022, pp. 1–9, https://usc-isi-i2.github.io/AAAI2022SS/papers/SSS-22_paper_47.pdf.

[66] T.G. Dietterich, A. Guyer, The familiarity hypothesis: explaining the behavior of deep open set methods, Pattern Recognit. 132 (2022) 108931, https://doi.org/10.1016/j.patcog.2022.108931.

[67] J.C. Balloch, Z. Lin, X. Peng, M. Hussain, A. Srinivas, R. Wright, J.M. Kim, M.O. Riedl, Neuro-symbolic world models for adapting to open world novelty, in: Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems, AAMAS '23, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2023, pp. 2848–2850, https://dl.acm.org/doi/10.5555/3545946.3599099.

[68] C. Zhao, D. Du, A. Hoogs, C. Funk, Open set action recognition via multi-label evidential learning, in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 22982–22991.

[69] P.A. Grabowicz, C. Pereyda, K. Clary, R. Stern, T. Boult, D. Jensen, L.B. Holder, Novelty in 2d cartpole domain, in: T. Boult, W. Scheirer (Eds.), A Unifying Framework for Formal Theories of Novelty: Discussions, Guidelines, and Examples for Artificial Intelligence, Springer Nature Switzerland, Cham, 2024, pp. 5–19.

[70] Z. Lin, W. Yang, H. Wang, H. Chi, L. Lan, J. Wang, Scaling few-shot learning for the open world, Proc. AAAI Conf. Artif. Intell. 38 (12) (2024) 13846–13854, https://doi.org/10.1609/aaai.v38i12.29291.

[71] E.L. Aleixo, J.G. Colonna, M. Cristo, E. Fernandes, Catastrophic forgetting in deep learning: a comprehensive taxonomy, arXiv:2312.10549, 2023.

[72] R.O. Malashin, M.A. Mikhalkova, Avoiding catastrophic forgetting via neuronal decay, in: 2024 Wave Electronics and Its Application in Information and Telecommunication Systems (WECONF), 2024, pp. 1–6.

[73] A. Kumar, S. Agarwal, Incremental learning in generative adversarial networks (gans) using pseudo replay to overcome catastrophic forgetting, in: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2024, pp. 1–6.

[74] Defense Advanced Research Projects Agency, Lifelong Learning Machines (L2M), https://sam.gov/opp/1237b294850e7b12c2baec95fb8498c6/view, 2017. (Accessed 21 February 2025).

[75] J. Dick, P. Ladosz, E. Ben-Iwhiwhu, H. Shimadzu, P. Kinnell, P.K. Pilly, S. Kolouri, A. Soltoggio, Detecting changes and avoiding catastrophic forgetting in dynamic partially observable environments, Front. Neurorobot. 14 (2020), https://doi.org/10.3389/fnbot.2020.578675.

[76] D. Kudithipudi, M. Aguilar-Simon, J. Babb, M. Bazhenov, D. Blackiston, J. Bongard, A.P. Brna, S. Chakravarthi Raja, N. Cheney, J. Clune, A. Daram, S. Fusi, P. Helfer, L. Kay, N. Ketz, Z. Kira, S. Kolouri, J.L. Krichmar, S. Kriegman, M. Levin, S. Madireddy, S. Manicka, A. Marjaninejad, B. McNaughton, R. Miikkulainen, Z. Navratilova, T. Pandit, A. Parker, P.K. Pilly, S. Risi, T.J. Sejnowski, A. Soltoggio, N. Soures, A.S. Tolias, D. Urbina-Meléndez, F.J. Valero-Cuevas, G.M. van de Ven, J.T. Vogelstein, F. Wang, R. Weiss, A. Yanguas-Gil, X. Zou, H. Siegelmann, Biological underpinnings for lifelong learning machines, Nat. Mach. Intell. 4 (3) (2022) 196–210, https://doi.org/10.1038/s42256-022-00452-0.

[77] Defense Advanced Research Projects Agency, Shared-Experience Lifelong Learning (ShELL), https://sam.gov/opp/1afbf600f2e04b26941fad352c08d1f1/view, 2021. (Accessed 21 February 2025).

[78] A. Soltoggio, E. Ben-Iwhiwhu, V. Braverman, E. Eaton, B. Epstein, Y. Ge, L. Halperin, J. How, L. Itti, M.A. Jacobs, P. Kantharaju, L. Le, S. Lee, X. Liu, S.T. Monteiro, D. Musliner, S. Nath, P. Panda, C. Peridis, H. Pirsiavash, V. Parekh, K. Roy, S. Shperberg, H.T. Siegelmann, P. Stone, K. Vedder, J. Wu, L. Yang, G. Zheng, S. Kolouri, A collective AI via lifelong learning and sharing at the edge, Nat. Mach. Intell. 6 (3) (2024) 251–264, https://doi.org/10.1038/s42256-024-00800-2.