Collectionless Artificial Intelligence: The **NARNIAN** Project

Short Abstract

Marco Gori, Stefano Melacci
Department of Information Engineering and Mathematics
University of Siena
53100 – Siena, Italy
marco.gori@unisi.it, stefano.melacci@unisi.it

March 21, 2025

Abstract

Collectionless Artificial Intelligence (**Collectionless AI**, Gori and Melacci [2023], https://cai.diism.unisi.it) focuses on the perspective in which intelligent agents are developed by learning over **time**, continuously adapting to the streamed information without storing it, and **interacting** with the environment, other agents, and humans. This is radically different from the usual procedure of collecting huge datasets and massively learn in an offline manner by stochastic gradient descent or using batches of pre-collected interactions for reinforcement learning. As remarked in [Gori and Melacci, 2023], learning from data collections introduces *risks related to data centralization, privacy, energy efficiency, limited customizability, and control.* Collectionless AI promotes on-the-edge local computations, where community of agents learn to interact in an efficient-and-controlled way with each other. This is the context in which the authors of this abstract conceived and introduced the **NARNIAN project**, standing for NAtuRe iNspired Intelligent AgeNts.

NARNIAN is to Collectionless AI what training and benchmarking based on static datasets are to today's AI (Figure 1).

Actually, it is more general than that.

> *NARNIAN is a dynamic conceptualization of the notion of learning setting and benchmark, though for studying Collectionless AI technologies but open and fully compatible with every other machine learning technology, allowing it to be be studied in a dynamic setting of continuous learning, interaction, progressive development.*

Classic benchmarking procedures are based on static data collections which are used to train models and static test sets which (hopefully) represent the considered task. NARNIAN provides a framework to implement the environment where the agents live, but also how they live, how they can get in touch, and how humans can get in touch with them to see what they are doing, evaluating them by direct interaction. Recall that we (humans/casual users) decide that a Large Language Model works well just by interacting with it, and not by running a benchmark. In a nutshell, static datasets are limited samples of the operating conditions in which the model is expected to work, while NARNIAN offers a framework to simulate the environment(s) where agents can develop themselves, being them a simulation of reality or the reality itself.

As briefly anticipated, NARNIAN is motivated by Collectionless AI [Gori and Melacci, 2023], but (i) it is generic enough to cover the existing offline or continual learning settings and (ii) offers tools for exploring developmental processes that are frequently not part of the mainstream, but could lead to novel technologies [Gori and Melacci, 2023]. NARNIAN, in its most naive instance, can be tough as a "sandbox" for growing agents that progressively learn, offering a dynamic world where they live and interact. Of course, nothing prevents existing technologies, such as Large Language Models and their multi-modal counterparts, to be used to implement more advanced agents living in NARNIAN worlds, or to implement parts of such agents. However, instead of barely distilling knowledge from
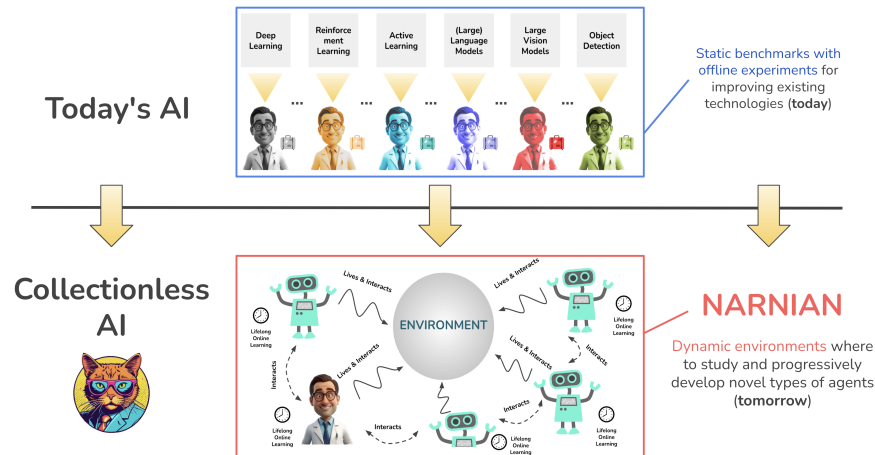
Figure 1: Toward Collectionless AI [Gori and Melacci, 2023] by means of NARNIAN. **Top:** Mainstream machine learning, as it is today. Different researchers/practitioners (scientist-like guys) that focus on different, limited views (gray tiles) of the same reality, and they have their own static benchmarks (suitcase). **Bottom:** NARNIAN offers a dynamic world, where agents live and interact, continuously learning over time without storing sensory information, as promoted by Collectionless AI. However, NARNIAN is generic enough to recover all the existing offline-learning-driven approaches to AI, including the ones of the top part of this figure, and it opens to the study of different types of algorithms, solutions, cross-overs, offering them the possibility to interact and grow over time. See `https://cai.diism.unisi.it`.

pre-trained models or finding good datasets to fine-tune them, NARNIAN promote the research for more efficient, compact, effective ways to interact with them, to get what **your agent** needs at the current stage of its developmental process. NARNIAN is fully open to the study of all existing or novel categories of learning algorithms and solutions, including never-experimented-before cross-overs.

Imagine a world where you can grow your agent, decide who it interacts with, control its exposure, and effectively improve it over time–potentially over a lifetime. NARNIAN is such a world.

A first prototype of NARNIAN, intended to showcase a basic synchronous version of what has been described so far, with agents fully generalizing generative models, classification problems, continual learning, local learning over time, and others, is available at

`https://github.com/mela64/narnian`

This is just a limited "mock-up" of the potentiality of NARNIAN, intended to be a basis to start building the fundamentals of a framework that we imagine as a social network of learning agents aggregated in communities and operating in a server-less manner.

Acknowledgments

References

Marco Gori and Stefano Melacci. Collectionless artificial intelligence, 2023. URL `https://arxiv.org/abs/2309.06938`.

Stefano Melacci, Alessandro Betti, Michele Casoni, Tommaso Guidi, Matteo Tiezzi, and Marco Gori. A unified framework for neural computation and learning over time, 2024. URL `https://arxiv.org/abs/2409.12038`.