# François Chollet co-founds ARC Prize Foundation to advance AI benchmarks



François Chollet, a former engineer at Google and a prominent figure in artificial intelligence research, has made a significant move by co-founding the ARC Prize Foundation, a nonprofit dedicated to the advancement of benchmarks aimed at assessing AI capabilities related to "human-level" intelligence. The foundation will be presided over by Greg Kamradt, a former engineering director at Salesforce, who has a history of founding AI-focused initiatives, including the AI product studio Leverage.

In a statement made on the organization's website, Chollet outlined the foundation's intent: “[W]e’re growing … into a proper nonprofit foundation to act as a useful north star toward artificial general intelligence.” The term artificial general intelligence (AGI) commonly refers to AI systems that can perform a wide variety of tasks with proficiency akin to that of a human. The aim of the ARC Prize Foundation is to inspire advancements in AI by closely examining the "gap" in essential human capabilities that current AI technologies still face.

Building on previously established work, the ARC Prize Foundation will further develop the Abstract and Reasoning Corpus for Artificial General Intelligence (ARC-AGI), a benchmark created by Chollet in 2019. This benchmark is designed to evaluate whether AI can acquire new skills autonomously, independent from the datasets on which it was trained. The ARC-AGI comprises a series of puzzle-like tasks where AI systems must generate an appropriate solution grid from a selection of differently coloured squares. These tasks are intended to challenge AI’s ability to adapt to new scenarios that it has not encountered before.

Since its introduction, ARC-AGI has revealed notable limitations in AI performance. Historically, although many AI systems excel at solving complex mathematical problems or advanced academic questions, their ability to tackle the tasks within ARC-AGI has remained significantly lower, with peak performances reaching just below 30% success. Chollet elaborated on the differences between current benchmarks and the goals of ARC-AGI, stating, “Unlike most frontier AI benchmarks, we are not trying to measure AI risk with superhuman exam questions.” He emphasized that future iterations of ARC-AGI will be concentrated on closing the gap between human ability and AI performance.

Last June, Chollet, in collaboration with Mike Knoop, co-founder of Zapier, initiated a competition aimed at creating an AI that could outperform the ARC-AGI benchmark. OpenAI’s unreleased o3 model emerged as the first to achieve a qualifying score in this competition, although it did so through the utilisation of extensive computing resources.

Chollet has acknowledged that while ARC-AGI has its limitations—some AI models achieving high scores through brute-force methods—the notion of o3 representing human-level intelligence is not supported by his findings. In a December statement, he provided further context on the performance expectations for the upcoming versions of the ARC-AGI benchmark, indicating that early results suggest that the successor benchmark may present a formidable challenge to the o3 model, stating, “[E]arly data points suggest that the upcoming [successor to the ARC-AGI] benchmark will still pose a significant challenge to o3, potentially reducing its score to under 30% even at high compute (while a smart human would still be able to score over 95% with no training).”

Plans to launch a second-generation ARC-AGI benchmark and a new competition are slated for this year, with the nonprofit also aiming to design a third version of ARC-AGI. Nonetheless, Chollet’s previous claims regarding ARC-AGI have attracted criticism, particularly regarding the definition and achievement of AGI, a concept that remains subject to active debate within the AI community. Notably, an OpenAI staffer has suggested that AGI may already have been realised depending on one’s definition of the term. In an interesting turn, OpenAI CEO Sam Altman confirmed in December that the company intends to collaborate with the ARC-AGI team to develop future benchmarks, though there were no updates provided by Chollet regarding this potential partnership in his recent announcement.

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://neuron.expert/news/ai-researcher-francois-chollet-is-co-founding-a-nonprofit-to-build-benchmarks-for-agi/10227/en/> - Corroborates François Chollet co-founding the ARC Prize Foundation and the foundation's mission to assess AI's 'human-level' intelligence.
2. <https://www.geniusgroup.ai/ai-researcher-francois-chollet-co-founds-nonprofit-to-benchmark-ais-path-to-human-level-intelligence/> - Supports the ARC Prize Foundation's goal to develop metrics for AI's adaptability and performance, and Greg Kamradt's role as president.
3. <https://neuron.expert/news/ai-researcher-francois-chollet-is-co-founding-a-nonprofit-to-build-benchmarks-for-agi/10227/en/> - Details the Abstract and Reasoning Corpus for Artificial General Intelligence (ARC-AGI) and its purpose in evaluating AI's ability to learn new skills.
4. <https://www.geniusgroup.ai/ai-researcher-francois-chollet-co-founds-nonprofit-to-benchmark-ais-path-to-human-level-intelligence/> - Explains the puzzle-like tasks in ARC-AGI and the challenges they pose to AI systems in adapting to new scenarios.
5. <https://neuron.expert/news/ai-researcher-francois-chollet-is-co-founding-a-nonprofit-to-build-benchmarks-for-agi/10227/en/> - Discusses the limitations of current AI systems in solving ARC-AGI tasks and the gap between human and AI performance.
6. <https://www.geniusgroup.ai/ai-researcher-francois-chollet-co-founds-nonprofit-to-benchmark-ais-path-to-human-level-intelligence/> - Mentions OpenAI's o3 model achieving a qualifying score in the ARC-AGI competition and the use of extensive computing resources.
7. <https://neuron.expert/news/ai-researcher-francois-chollet-is-co-founding-a-nonprofit-to-build-benchmarks-for-agi/10227/en/> - Highlights Chollet's views on the limitations of ARC-AGI and the distinction between current AI performance and human-level intelligence.
8. <https://www.geniusgroup.ai/ai-researcher-francois-chollet-co-founds-nonprofit-to-benchmark-ais-path-to-human-level-intelligence/> - Details plans for launching a second-generation ARC-AGI benchmark and future competitions.
9. <https://neuron.expert/news/ai-researcher-francois-chollet-is-co-founding-a-nonprofit-to-build-benchmarks-for-agi/10227/en/> - Addresses the ongoing debate within the AI community regarding the definition and achievement of AGI.
10. <https://www.geniusgroup.ai/ai-researcher-francois-chollet-co-founds-nonprofit-to-benchmark-ais-path-to-human-level-intelligence/> - Mentions potential collaboration between the ARC-AGI team and OpenAI for developing future benchmarks.
11. <https://neuron.expert/news/ai-researcher-francois-chollet-is-co-founding-a-nonprofit-to-build-benchmarks-for-agi/10227/en/> - Provides context on Chollet's vision for the ARC Prize Foundation as a 'north star' guiding progress towards AGI.
12. <https://techcrunch.com/2025/01/08/ai-researcher-francois-chollet-is-co-founding-a-nonprofit-to-build-benchmarks-for-agi/> - Please view link - unable to able to access data