The Chinese Room is a thought experiment, devised by American philosopher John Searle, to prove that artificial intelligence (AI) only has the ability to **appear knowledgeable**, rather than truly understanding the information it outputs [1].

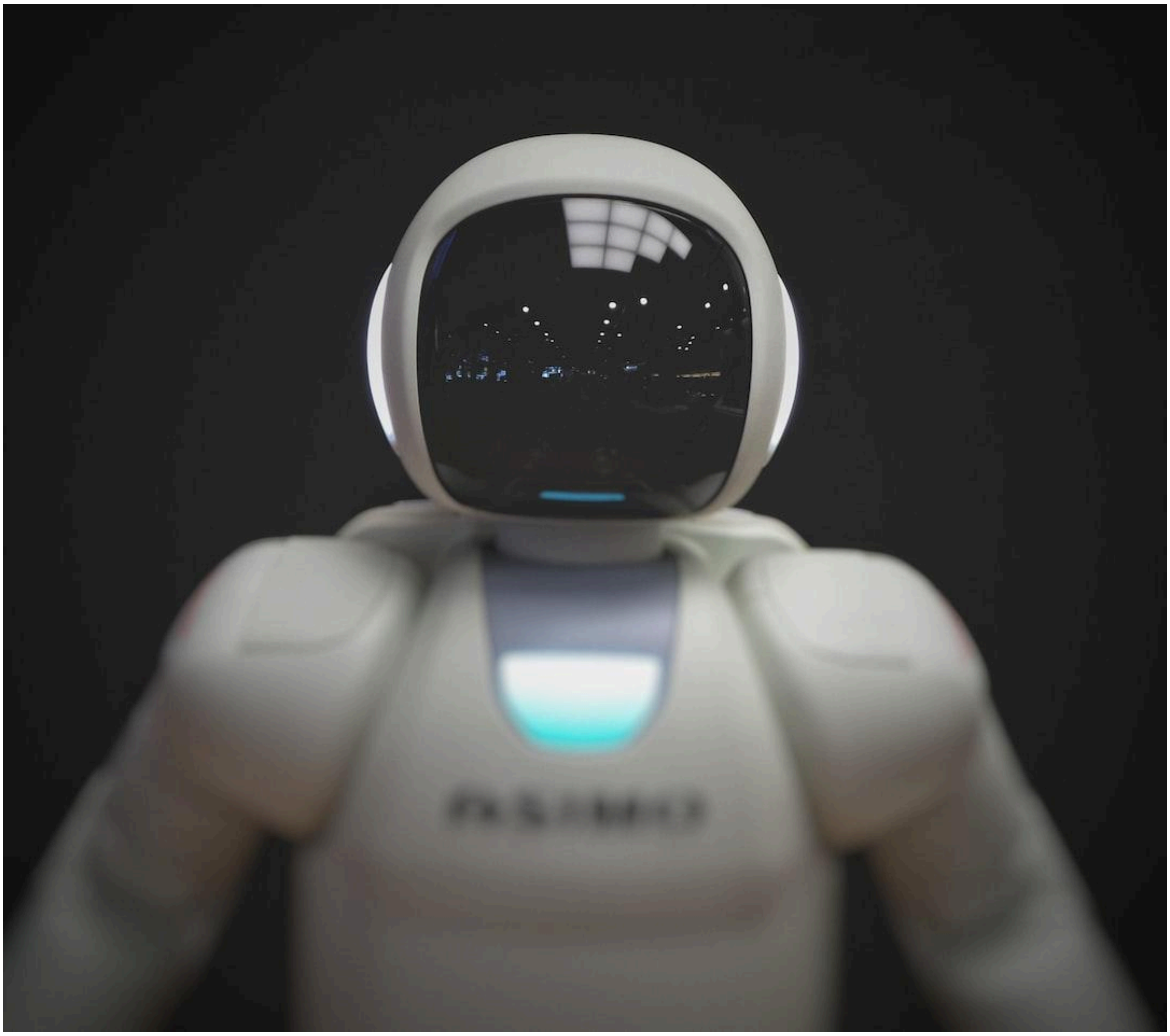
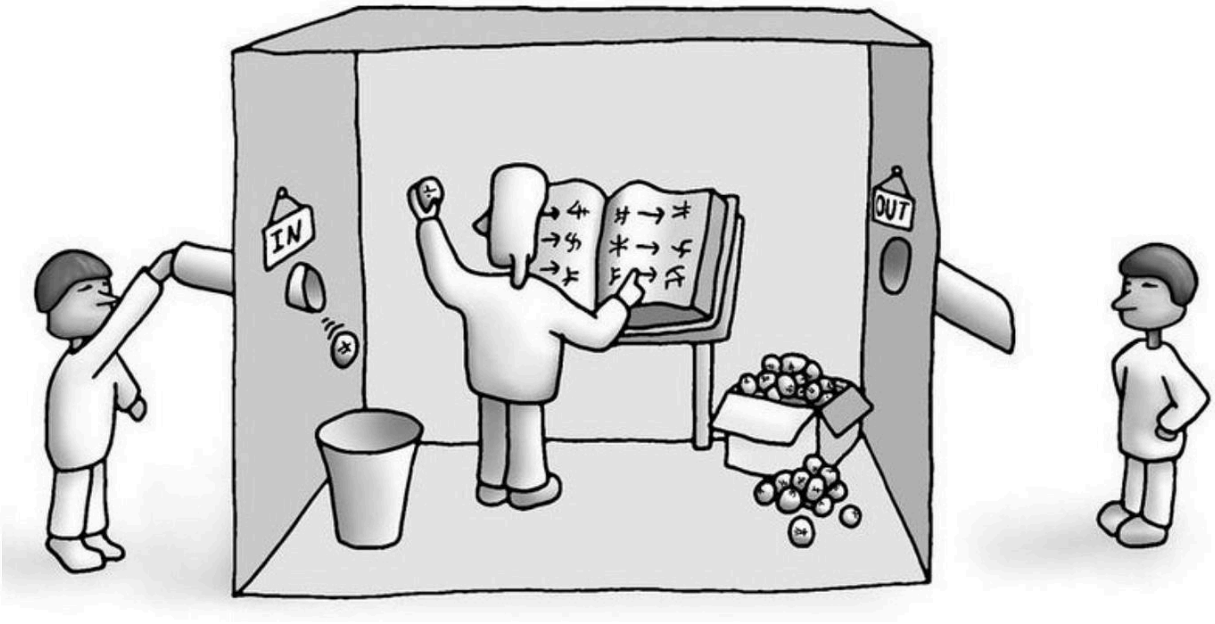


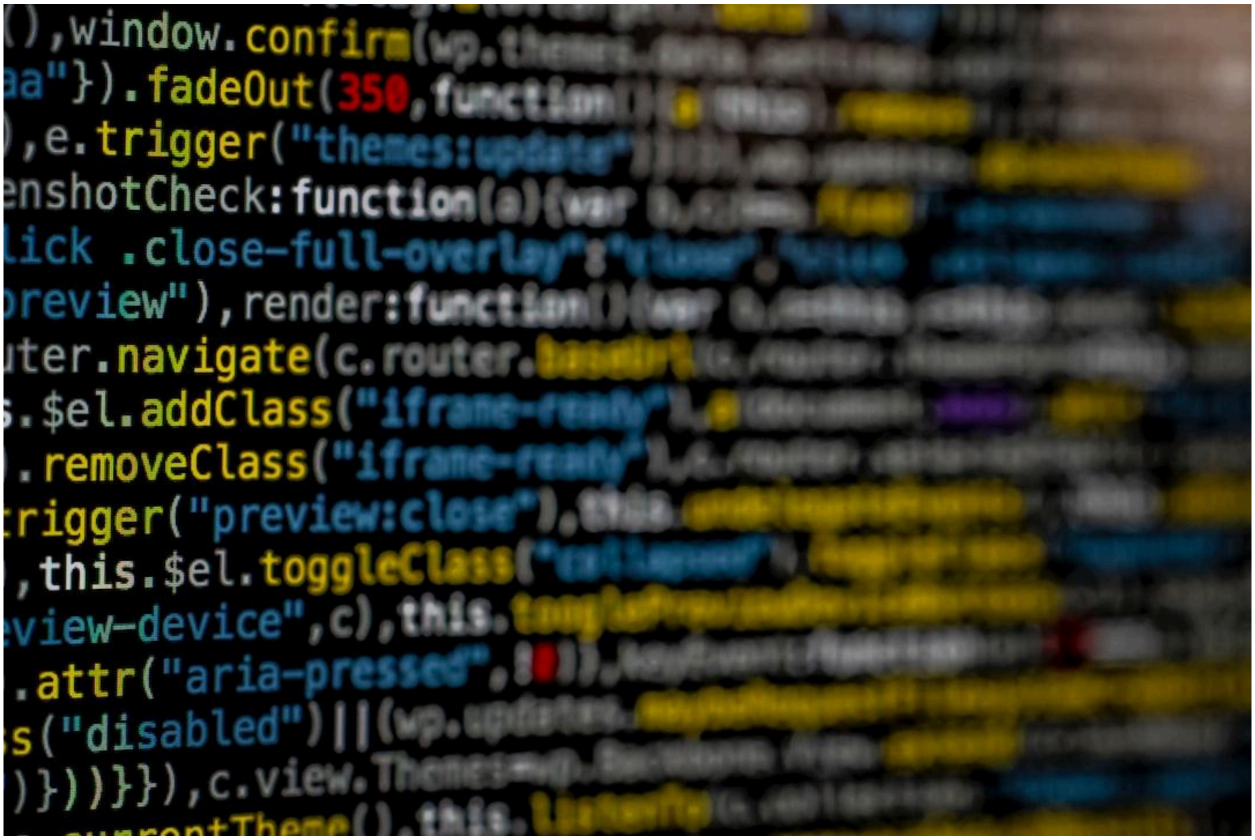
Photo by Possessed Photography on Unsplash

The thought experiment involves a non-Mandarin speaker in a locked room with a book filled with Chinese characters and an instruction book. The instruction book gives the room's inhabitant the knowledge on how to arrange the various characters in response to certain other characters. In essence, the book states, "if character [A] is shown, respond with character [B]." After practicing for long periods of time and becoming increasingly familiar with this pairing action, the inhabitant is able to arrange entire passages in complex Mandarin, to the point where his outputs are indistinguishable from those of a native speaker's. However, although the inhabitant can produce great displays of Mandarin, he continues to have no comprehension of what he is writing. Searle advances the story when he suggests that one day, a native Mandarin speaker slides a message under the door. Viewing the message, the inhabitant is able to put his newfound pairing skills into practice, uses it to create a wise response to the native's question, and slides his response under the door. After receiving the message, and being astonished at the intelligent response given to him, the native concludes that the inhabitant inside the room must be a native-speaker, and moreover, an intelligent human being. As the native cannot see in to the room, he is unable to discover that the "intelligent human being" does not understand a lick of Mandarin [2].



The Chinese Room Argument, via Medium.com

The essence of Searle's story is that while we perceive AI to be intelligent and "having a mind of its own," it doesn't. Instead, he reasons that all the information AI can possibly know, and all of the abilities it can ever possess, are derived from human knowledge. AI doesn't comprehend the information, similar to the inhabitant's inability to understand Mandarin, but rather, can simply output "intelligent" messages because a human has already given knowledge and coded it on how to do so.



```
(),window.confirm(wp.
aa"}).fadeOut(350,function(
),e.trigger("themes:update"
enshotCheck:function(a){var
lick .close-full-overlay":
preview"),render:function(
uter.navigate(c.router.
s.$el.addClass("iframe-ready"
).removeClass("iframe-ready"
trigger("preview:close"),this
,this.$el.toggleClass(
view-device",c),this.
.attr("aria-pressed",
s("disabled")||(wp.
)})}}}),c.view.Theme
```

Photo by Markus Spiske on Unsplash

Many software engineers and artificial intelligence developers have used the Chinese Room as a way to accurately predict the limitations of what AI can do. They propose that if AI relies on human input to create their own outputs, AI is limited to humanity's knowledge, and can never exceed it. A great real-life showcase of this is when Deep Blue, an AI chess engine, battles against World Champion Garry Kasparov in a rematch best of six series, after losing to the champion a year prior. With three draws in total, Deep Blue narrowly beat Kasparov 2-1. After Kasparov's defeat, the media was engulfed with claims that AI and machine learning would surpass human ability, as it had done so with the champion. However, engineers around the globe pointed out that the Deep Blue AI hadn't analyzed the game of chess on its own and mastered it, but rather received input from many grandmaster games from previous decades, analyzed the patterns and sequences, and outputted the moves that had seemed to work the best, a situation very similar to the thought experiment [3].

As the experimentation of AI in recent years has been increasing, the future abilities of AI are unknown. However, as John Searle's "Chinese Room Argument" proves, as well as the real life example of Garry Kasparov and Deep Blue, the connotation that AI "has a mind of its own" may be false. However, there have been many interpretations of the thought experiment, leading to the disputed question: is AI really intelligent? No? Yes? Maybe? What is classified as "intelligence?" As with most thought experiments, the lessons learned are still debated. Nonetheless, the Chinese Room experiment remains significant, due to its groundbreaking ramifications on how AI is perceived throughout the world.

1. *The Chinese Room Argument (Stanford Encyclopedia of Philosophy)*. 20 Feb. 2020, plato.stanford.edu/entries/chinese-room.
2. Barrero, Andres Felipe. "Can AI Think? Searle's Chinese Room Thought Experiment." *TheCollector*, Mar. 2023, www.thecollector.com/can-ai-think-searle-chinese-room-argument.
3. Yao, Deborah. "25 Years Ago Today: How Deep Blue Vs. Kasparov Changed AI Forever." *AI Business*, July 2023, aibusiness.com/ml/25-years-ago-today-how-deep-blue-vs-kasparov-changed-ai-forever.