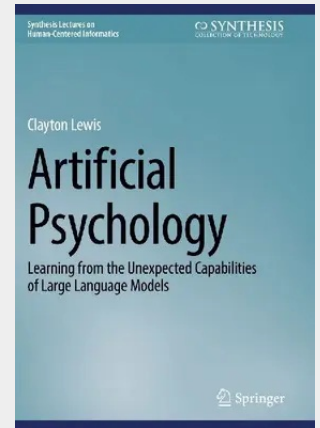


Artificial Psychology

Learning from the Unexpected Capabilities of Large Language Models

The success of predictive large language models (PLLMs) like GPT3 and ChatGPT has created both enthusiasts and skeptics of their widespread practical applications, but this book argues that the larger significance of such models is contained in what they suggest about human cognition. To explore this potential, the book develops a thought experiment called the Prediction Room, a reference to John Searle's influential Chinese Room argument, in which a human agent processes language by following a set of opaque written rules without possessing an inherent understanding of the language. The book proposes a new Room model—the Prediction Room with its resident Prediction Agent—generalizing the working of large language models. Working through a wide range of topics in cognitive science, the book challenges the conclusion of Searle's thought experiment, that discredited contemporary artificial intelligences (AI), through the suggestion that the Prediction Room offers a means of exploring how new ideas in AI can provide productive alternatives to traditional understandings of human cognition. In considering the implications of this, the book reviews an array of topics and issues in cognitive science to uncover new ideas and reinforce older ideas about the mental mechanisms involved in both sides. The discussion of these topics in the book serves two purposes. First, it aims to stimulate new thinking about familiar topics like language acquisition or the nature and acquisition of concepts. Second, by contrasting human psychology with the form of artificial psychology these models exhibit, it uncovers how new directions in the development of these systems can be better explored.

The success of predictive large language models (PLLMs) like GPT3 and ChatGPT has created both enthusiasts and skeptics of their widespread practical applications, but this book argues that the larger significance of such models is contained in what they suggest about human cognition. To explore this potential, the book develops a thought experiment called the Prediction Room, a reference to John Searle's influential Chinese Room argument, in which a human agent processes language by following a set of opaque written rules without possessing an inherent understanding of the language. The book proposes a new Room model—the Prediction Room with its resident Prediction Agent—generalizing the working of large language models. Working through a wide range of topics in cognitive science, the book challenges the conclusion of Searle's thought experiment, that discredited contemporary artificial intelligences (AI), through the suggestion that the Prediction Room offers a means of exploring how new ideas in AI can provide productive alternatives to traditional understandings of human cognition. In considering the implications of this, the book reviews an array of topics and issues in cognitive science to uncover new ideas and reinforce older ideas about the mental mechanisms involved in both sides. The discussion of these topics in the book serves two purposes. First, it aims to stimulate new thinking about familiar topics like language acquisition or the nature and acquisition of concepts. Second, by contrasting human psychology with the form of artificial psychology these models exhibit, it uncovers how new directions in the development of these systems can be better explored.



48,14 €
44,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9783031766480
Medium: Buch
ISBN: 978-3-031-76648-0
Verlag: Springer
Erscheinungstermin: 20.02.2026
Sprache(n): Englisch
Auflage: Erscheinungsjahr 2026
Serie: Synthesis Lectures on Human-Centered Informatics
Produktform: Kartoniert
Gewicht: 468 g
Seiten: 262
Format (B x H): 168 x 240 mm

