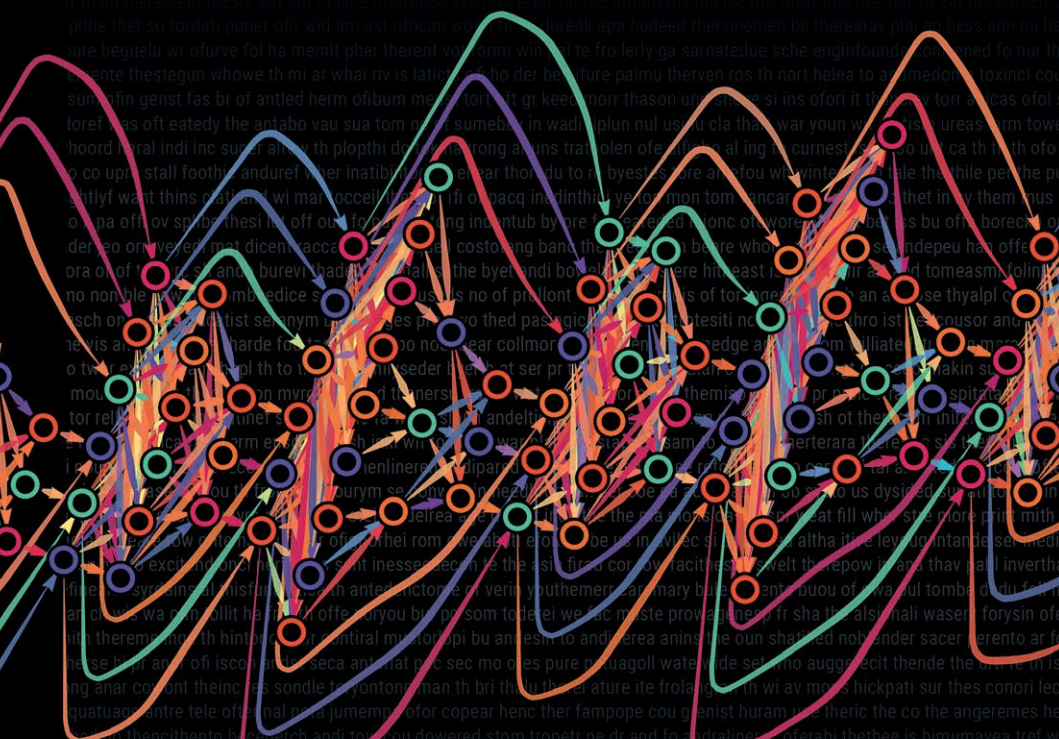


СТИВЕН ВОЛЬФРАМ

Всемирно известный физик-математик  
и разработчик WolframAlpha

# КАК УСТРОЕН ChatGPT?



**Полное погружение в принципы работы  
и спектр возможностей  
самой известной нейросети в мире**

УДК004.032.26  
ББК 16.632  
В72

Научный редактор Здоров Антон

*На русском языке публикуется впервые*

**Вольфрам, Стивен**

**В72** Как устроен ChatGPT? Полное погружение в принципы работы и спектр возможностей самой известной нейросети в мире / Стивен Вольфрам ; пер. с англ. Е. Быковой ; науч. ред. А. Здоров. — Москва : МИФ, 2024. — 192 с. — (Цифровые технологии).

ISBN 978-5-00214-604-8

ChatGPT в рекордно короткие сроки завоевал популярность благодаря способности писать тексты на любые темы — почти как человек. Но как именно он это делает? Может ли он понимать, что пишет, и что происходит внутри его искусственного разума?

Стивен Вольфрам — достойный наград ученый и создатель программных систем, таких как Wolfram|Alpha, — предлагает понятное и увлекательное объяснение внутреннего устройства ChatGPT.

Узнайте, как ChatGPT объединяет новейшие технологии нейронных сетей с фундаментальными идеями о языке и человеческом мышлении.

УДК 004.032.26  
ББК 16.632

*Все права защищены.  
Никакая часть данной книги  
не может быть воспроизведена  
в какой бы то ни было форме  
без письменного разрешения  
владельцев авторских прав.*

ISBN 978-5-00214-604-8

What Is ChatGPT Doing ... and Why Does It Work?

© 2023 Stephen Wolfram.

Original English language edition published by Wolfram Media 100 Trade Center Dr. 6th Floor, Champaign Illinois 61820, USA.

Arranged via Licensor's Agent: DropCap Inc. and Igor Korzhenevskiy of Alexander Korzhenevski Agency (Russia). All rights reserved

© Издание на русском языке, перевод, оформление. ООО «Манн, Иванов и Фербер», 2024

# ОГЛАВЛЕНИЕ

Предисловие .....	9
-------------------	---

## **ЧТО ДЕЛАЕТ CHATGPT И ПОЧЕМУ ЭТО РАБОТАЕТ?**

Он просто добавляет по одному слову за раз .....	15
Откуда берутся эти вероятности? .....	26
Что такое модель? .....	34
Модели для человеческих задач .....	38
Нейронные сети .....	42
Машинное обучение нейронных сетей .....	59
Обучение нейронных сетей: практика и знания .....	69
Достаточно большая сеть может сделать все что угодно — так ведь? .....	82
Концепция эмбедингов .....	89
Внутри ChatGPT .....	99
Обучение ChatGPT .....	114
Помимо базового обучения .....	119

Что позволяет ChatGPT работать? .....	123
Смысловое пространство и семантические законы движения .....	133
Семантика и сила вычислительного языка .....	141
Так что же делает ChatGPT и почему это работает? .....	149
Благодарности .....	153

### **WOLFRAM|ALPHA КАК СПОСОБ ДОБАВИТЬ В CHATGPT СВЕРХВОЗМОЖНОСТИ ВЫЧИСЛЯЕМЫХ ЗНАНИЙ**

ChatGPT и Wolfram Alpha .....	158
Базовый пример .....	161
И еще несколько примеров .....	166
Перспективы .....	184
Дополнительные ресурсы .....	191

# ПРЕДИСЛОВИЕ

Эта небольшая книга представляет собой попытку объяснить, как работает ChatGPT. В некотором смысле это история о технологиях. Но еще и разговор о науке и философии. И для того, чтобы рассказать эту историю, нам придется собрать воедино огромный спектр идей и открытий, сделанных на протяжении многих столетий.

Я сам с большим интересом и волнением наблюдаю, как вещи, которые так долго меня интересовали, объединяются в результате такого прогресса. Сложное поведение простых программ, погружение в суть языка и смыслообразование, а также практические возможности больших компьютерных систем — все это является частью истории ChatGPT.

ChatGPT основан на концепции нейронных сетей, составленной в 1940-х годах в качестве модели работы головного мозга. Я сам впервые спрограммировал нейронную сеть в 1983 году, но тогда ничего интересного из нее не вышло. Однако 40 лет спустя, когда у нас есть компьютеры, считающие в миллион раз быстрее, миллиарды страниц текста в интернете и целый ряд инженерных инноваций, ситуация совершенно иная. И, ко всеобщему удивлению, современная нейронная сеть, которая в миллиард раз больше созданной мной в 1983 году, способна делать то, что раньше считалось под силу только человеку, — генерировать осмысленный текст.

Эта книга, написанная вскоре после дебюта ChatGPT, состоит из двух частей. В первой объясняется, что такое ChatGPT и как ему удастся выполнять сугубо человеческую работу по генерированию текста. Вторая посвящена вычислительным инструментам ChatGPT (выходящим за рамки человеческих возможностей) и сверхспособностям нашей системы Wolfram|Alpha в области вычисляемых знаний.

На момент написания книги прошло всего три месяца с запуска ChatGPT, и мы только начинаем понимать последствия этого события — как практические, так

и интеллектуальные. Но пока ChatGPT служит напоминанием о том, что, несмотря на множество уже сделанных изобретений и открытий, сюрпризы всё ещё возможны.

*Стивен Вольфрам,  
28 февраля 2023 года*