



Brain Rush

How to Invest and Compete in
the Real World of Generative AI



Peter Cohan

Apress®

Brain Rush

How to Invest and
Compete in the Real World
of Generative AI

Peter Cohan

Apress®

Brain Rush: How to Invest and Compete in the Real World of Generative AI

Peter Cohan
Marlborough, MA, USA

ISBN-13 (pbk): 979-8-8688-0317-8
<https://doi.org/10.1007/979-8-8688-0318-5>

ISBN-13 (electronic): 979-8-8688-0318-5

Copyright © 2024 by Peter Cohan

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Shivangi Ramachandran
Development Editor: James Markham
Project Manager: Jessica Vakili

Distributed to the book trade worldwide by Springer Science+Business Media New York, 1 New York Plaza, New York, NY 10004. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail booktranslations@springernature.com; for reprint, paperback, or audio rights, please e-mail bookpermissions@springernature.com.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at <http://www.apress.com/bulk-sales>.

If disposing of this product, please recycle the paper

To Robin, Sarah, and Adam.

Table of Contents

About the Author	xi
Acknowledgments	xiii
Chapter 1: Introducing <i>Brain Rush</i>.....	1
Who I Am and Why I Wrote This Book.....	1
Who Should Read <i>Brain Rush</i> and How Will It Help Them?	3
<i>Brain Rush</i> Roadmap	6
Chapter 2: What Is Generative AI?	9
Defining Generative AI.....	9
Generative AI Examples	11
Building Generative AI Applications	13
The Two Sides of Generative AI	17
Generative AI's Bright Side	17
Generative AI's Dark Side	19
How I Researched Generative AI	21
The Generative AI Ecosystem Model	25
Conclusion	28
Part I: Mining Generative AI's End-User Value.....	29
Chapter 3: Generative AI Customer End Uses	31
Who Uses Generative AI?	31
Individual Generative AI Applications	33

TABLE OF CONTENTS

Consumer End Uses.....33

Employee End Uses36

Organizational Generative AI Applications.....40

Which Generative AI End Uses Create the Most Value?45

Healthcare Generative AI Case Studies.....49

 Generative AI Helps Contra Costa County Health Services Answer
 Patient Questions Faster49

 Generative AI Simplifies Patient Consent Forms50

 Generative AI Frees Doctors from Transcribing Patient Conversations.....52

 Generative AI Helps Doctors Diagnose Disease.....53

Financial Services Generative AI Case Studies54

 JPMorgan Developing ChatGPT-Like Investment Advisor54

 Alpha Keeps State Street’s Clients Happy56

 Goldman Sachs Tests Generative AI58

Principles of Successful Generative AI Deployment.....60

 What to Do.....60

 What to Avoid.....61

How End Users Can Make the Most of Generative AI.....63

Conclusion63

Part II: Mapping the Generative AI Ecosystem65

Chapter 4: Generative AI Consulting.....67

 Generative AI Consulting Industry Players67

 Generative AI Consulting Industry Attractiveness71

 Generative AI Consulting Industry Participants.....74

 Strategists77

 Risk and Process Managers.....90

Coding Outsourcers	104
Generative AI Consulting Critical Activities	107
Conclusion	109
Chapter 5: Generative AI Software	111
Generative AI Software Industry Players	112
Generative AI Software Industry Attractiveness	115
Generative AI Software Industry Participants	118
Publicly Traded Proprietary LLM Providers	118
Privately Held Proprietary LLM Providers	130
Open Source LLM Providers	141
Publicly Traded Application-Specific Generative AI Application Providers.....	149
Privately Held Application-Specific Generative AI Application Provider.....	160
Generative AI Software Critical Activities	164
Conclusion	166
Chapter 6: Generative AI Cloud Platforms	167
Generative AI Cloud Platform Industry Players.....	168
Generative AI Cloud Platform Industry Attractiveness.....	170
Generative AI Cloud Platform Industry Participants	175
Generative AI Cloud Services Provider Subsidiaries of Public Companies	175
Publicly Traded Generative AI Data Centers.....	195
Publicly Traded Generative AI Networking Technology Providers.....	201
Arista Networks	202
Generative AI Database Providers	209
Publicly Traded Generative AI Application Performance Monitoring Service Providers	226
Generative AI Cloud Services Critical Activities.....	233
Conclusion	234

TABLE OF CONTENTS

Chapter 7: Generative AI Hardware	237
Generative AI Hardware Industry Players.....	238
Generative AI Semiconductor Industry Attractiveness	240
Generative AI Hardware Industry Participants	244
Publicly Traded Generative AI Semiconductor Companies	244
Publicly Traded Semiconductor Manufacturing Equipment Suppliers.....	261
Publicly Traded Computer Server Manufacturers	265
Publicly Traded Data Center Cooling Product Manufacturers	272
Generative AI Hardware Critical Activities.....	277
Conclusion	278
Part III: Panning for Generative AI Gold	281
Chapter 8: Generative AI's Implications for Consumers, Employees, Companies, Product Providers, and Investors.....	283
How Consumers Can Benefit from Generative AI	285
High-Payoff Generative AI Experiments.....	288
Consumer Generative AI Dos and Don'ts	293
How Consumers Can Capture Generative AI's Value	294
Prompt and Response 1: Provide Four Options for a New Office Chair	295
Prompt and Response 2: Compare Four Options on Customer Reviews, Ease of Assembly, Useful Life, and Price Range	296
Prompt and Response 3: Assess Whether the Quality of the Highest Priced Chair Justifies Its Price.....	299
How Employees and Business Operators Can Benefit from Generative AI.....	301
High-Payoff Generative AI Experiments.....	304
Employee and Business Operator Generative AI Dos and Don'ts.....	308
How Employees and Business Operators Can Capture Generative AI's Value.....	309
Growth Strategies for Generative AI Product Providers	315
High-Payoff Generative AI Strategies	319

Generative AI Product Provider Dos and Don'ts.....325

How Product Providers Can Capture Generative AI's Value326

How Investors Can Profit from Generative AI327

 A Goldman Sachs Perspective on Generative AI-Driven Growth.....330

 Generative AI Investment Dos and Don'ts.....334

How Investors Can Capture Generative AI's Value335

Conclusion342

Chapter 9: Generative AI's Benefits and Risks to Society.....343

Generative AI's Influence.....343

The Light Side: Generative AI's Societal Benefits.....346

 Boosting Global Growth346

 Increasing Worker Productivity.....347

 Adding New Jobs in Tech-Rich Regions350

 Creating Demand in Supplier Industries.....351

The Dark Side: Generative AI's Societal Costs and Risks.....354

 Creating Misinformation354

 Boosting Legal Liability358

 Displacing Workers.....360

 Ending Civilization362

Holding Generative AI Providers Accountable365

How Government Should Regulate Generative AI368

Actions Citizens Should Take with Generative AI's Light and Dark Sides.....372

Conclusion374

Chapter 10: After the Brain Rush: What Is Generative AI's Future?377

Generative AI and How It Is Likely to Evolve377

Does Generative AI Matter or Is It a Temporary Fad?381

Groups Likely to Be Generative AI Winners and Losers.....383

TABLE OF CONTENTS

The Most Beneficial Uses of Generative AI.....386

- Consumers387
- Employees387
- Business Leaders388

Publicly Traded Generative AI Product and Services Providers
That Make the Best Investments 391

Government Policies That Limit the Harm and Maximize the Benefit of
Generative AI.....393

Conclusion394

Index.....397

About the Author

Peter Cohan is an Associate Professor of Management Practice at Babson College. He teaches strategy, leadership, and entrepreneurship to students in the college's undergraduate, Master of Science in Entrepreneurial Leadership (MSEL), MBA, and Executive Education programs. He is coordinator of Babson's required undergraduate strategy course and the creator and teacher of advanced strategy courses for undergraduate and MSEL students. Cohan is the founding principal of Peter S. Cohan & Associates, a management consulting and venture capital firm. He has completed over 150 growth-strategy consulting projects for global technology companies and invested in seven startups, three of which were sold for about \$2 billion and one of which went public in 2021 at an \$18 billion valuation. He has written 17 books, including *Net Profit: How to Invest and Compete in the Wild World of Internet Business*. Since 2011 he has been a contributor to *Forbes* and *Inc.* He is a frequent media commentator who has appeared on ABC's *Good Morning America*, *Bloomberg*, CNN, CNBC, Fox Business News, American Public Media's MarketPlace, WBUR, WGBH, New England Cable News, and the Boston ABC, NBC, and CBS affiliates. He has been quoted in the Associated Press, the *Christian Science Monitor*, the *London Evening Standard*, the *Times of London*, the *New York Times*, *Nikkei*, *USA Today*, the *Wall Street Journal*, *The Washington Post*, Portugal's *Expresso*, the *Economist*, *Time*, *BusinessWeek*, and *Fortune*. He also appeared in the 2016 documentary film *We the People: The Market Basket Effect*. Prior to starting his firm, he worked as a case team leader for Monitor Company, Harvard Business School Professor Michael Porter's consulting firm. He has taught at MIT, Stanford, Columbia, Tel Aviv University, New York University, Bentley University,

ABOUT THE AUTHOR

the Vienna University of Technology, School of Management Fribourg, Barcelona's EADA, Singapore's Nanyang Technological University, the University of Coimbra, the University of Chile, the University of Hong Kong, and Technologico de Monterrey. RETHINK Retail chose him as a Top Retail Expert of 2021, 2022, 2023, and 2024. He earned an MBA from Wharton, did graduate work in Computer Science at MIT, and holds a BS in Electrical Engineering from Swarthmore College.

Acknowledgments

This book has benefited greatly from the help of many people.

I could not have embarked on this project without the enthusiastic support of Danna Greenberg, who chairs the Management Division at Babson among other responsibilities. My Babson colleagues Keith Rollag and Jonathan Sims read drafts of several chapters and provided many insightful suggestions for improvement. *Barron's Advisor* Editor Amey Stone shared insightful comments on Chapter 10. My Babson students, including Tadeo Acosta-Rubio and Shivant Malkani, shared examples of how they have used generative AI to get work done more efficiently. My neighbors Rob Quinn and Leslie Healy read and provided numerous helpful comments on the book as I was writing it.

I am very grateful for the executives and analysts who shared their perspectives on generative AI. These include Goldman Sachs's Chief Information Officer, Marco Argenti; State Street Executive Vice President, John Plansky; Forrester Research Senior Analyst, Rowan Curran; EY's global consulting AI and automation leader, Dan Diasio; Contra Costa's data and innovation officer, Matt White; State Street Alpha Data Platform Product Owner, Jeffrey Shortis; Forrester Research Vice President and Principal Analyst, Michele Goetz; EY Americas Data, AI, and Automation Leader, Traci Gusher; KPMG's national leader of artificial intelligence and head of data engineering, Sreekar Krishna; KPMG National Managing Principal, Atif Zaim; PwC's Vice Chair and US chief products and technology officer, Joe Atkinson; PwC Principal and Generative AI Leader, Bret Greenstein; Adobe President of Digital Experience Business and Worldwide Field Operations, Anil Chakravarthy; ServiceNow CEO, Bill McDermott; ServiceNow CFO, Gina Mastantuono; Service Now

ACKNOWLEDGMENTS

Chief Commercial Officer, Paul Smith; Arista Networks CEO, Jayshree Ullal; MongoDB CEO, Dev Ittycheria; Snowflake Chairman of the Board, Frank Slooman; Snowflake CEO, Sridhar Ramaswamy; Freddie Mac Vice President and Chief Data Officer, Aravind Jagannathan; Goldman Sachs Managing Director, Kash Rangan; Datadog Co-founder and CEO, Olivier Pomel; KPMG Vice Chair of AI and Digital Innovation, Steve Chase; and Dynatrace CEO, Rick McConnell.

Without Apress, this book would not exist. I am most grateful to Shivangi Ramachandran for her enthusiastic support of the idea for this book and for the outstanding project management and editing help from Jessica Vakili and James Markham.

Thanks to my editors at *Forbes*, Chris Gentilviso and Jeffrey Frankel, and my editor at *Inc.*, Mark Coatney, for their interest in and support of *Brain Rush*.

Finally, I could not have completed this book without the help of my wife, Robin, who patiently read and commented on each chapter of the book, and my children, Sarah and Adam, who always make me proud.

CHAPTER 1

Introducing *Brain Rush*

Who I Am and Why I Wrote This Book

I wrote this book because of a gut feeling that generative artificial intelligence, a system that delivers new and original content in response to natural language queries, could spark enormous change in the way people live and work. What prompted that feeling was the 40% rise in the value of Nvidia stock – bringing its stock market value to \$1 trillion – three weeks after its May 2023 earnings report which featured a much higher than expected surge in revenue. The source of its soaring revenue growth was demand from companies using its chips to train and operate generative AI applications such as ChatGPT. This prompted a countervailing intellectual response: to investigate whether my gut reaction foretold a significant wave of opportunity or an emotional overreaction to a compelling sales pitch. Here are some of the questions I set out to answer in this book:

- Which applications of generative AI are the most valuable to their users?
- Will such high-value uses of generative AI create widespread demand or will they attract initial interest and later fade?

- Will generative AI contribute to economic growth? Will it create or improve more jobs than it destroys?
- What safeguards must emerge to prevent the legal and financial risks of generative AI from overwhelming its benefits?
- What does the map of the generative AI ecosystem look like?
- How should end users, executives, capital providers, and incumbent technology and service providers evaluate where to place their generative AI bets?

Two experiences qualify me to answer these questions: I worked in an AI startup early in my career that did not become a successful business. In addition, I wrote three books about the Internet – *Net Profit*, *eProfit*, and *eStocks* – the first during the 1990s when it was far from clear how much of the conversation about Internet business was hype and how much described real companies that changed the world. From working in the startup and writing the book, I learned two key lessons I applied in helping guide investors and business leaders to capture the opportunities of generative AI:

1. **To succeed, companies must solve the right problem well.** The AI startup I worked for aimed to help personal financial planners advise their clients. The startup aimed to encode the decision rules of financial planning experts by building a so-called expert system. Since personal financial planners did not see its product as a solution to a significant problem they faced, the company ultimately failed.

2. **When a tech bubble begins, smart investors and executives should bet on products and business models that deliver a sustainable increase in customer value.** Generative AI will be no different than the Internet was in that respect. Winners will back or lead companies that pass four tests:
 - They build the world's best product to eliminate significant customer pain.
 - They find and win market share in many large markets that need this solution.
 - They harness the capabilities required to win customers and keep them buying by providing new products they crave.
 - They sustain rapid growth by creating and capturing a meaningful share of new growth markets.

Who Should Read Brain Rush and How Will It Help Them?

Brain Rush answers pressing questions about generative AI in the minds of varying groups of readers:

- **Citizens and end users.** Generative AI rapidly gained the attention of hundreds of millions of people around the world. In 2023, politicians held public hearings about the perceived societal risks of generative AI. While the White House issued an executive order promoting safe

and responsible AI in October 2023¹, as of early 2024, the United States had not passed specific laws or regulations protecting society against AI's risks.² *Brain Rush* addresses citizens' and end users' questions, such as Is generative AI an opportunity or a threat to society? Should I urge government representatives to pass laws to protect against its threats? Which generative AI applications could be most useful? Which ones could be useful but risky and how can I protect myself against the risks? Should I invest in companies that supply generative AI technology?

- **Business leaders.** Vendors are scrambling to develop and deliver generative AI products and services for enterprises. Companies are selectively making such tools, such as ChatGPT, available to their employees. Yet these technologies present a mixture of opportunities and risks of wasting money, damaging the company's reputation, and losing control of proprietary information. *Brain Rush* helps answer business leaders' questions arising from these dynamics, including: Which generative AI applications will provide the most value for our company in the short and long term? What is the most effective process for engaging employees and customers in the development of new generative AI applications?

¹ "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," *The White House*, October 30, 2023, www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

² "Why AI still needs regulation despite impact," *Thomson Reuters*, February 1, 2024, <https://legal.thomsonreuters.com/blog/why-ai-still-needs-regulation-despite-impact/>

How can we assemble the right partners to design and implement these high-payoff applications? What policies should we establish to limit the risks of adopting generative AI? How can we assess and improve our generative AI applications?

- **Generative AI technology suppliers.** Not long after ChatGPT launched and grabbed the world's attention, incumbent technology and service providers recognized a building growth opportunity, compelling them to launch products to help customers profit from generative AI. Meanwhile, startups were developing products aimed at advancing the state of generative AI's art. *Brain Rush* examines questions these business leaders must resolve, such as Is there a large enough opportunity to justify investing in a generative AI product? If so, what capabilities must we deploy to design and build generative AI products of value to potential customers? Do we have – or can we collaborate to obtain – sufficient competitive strengths to win market share from rivals? Will the future benefits to customers and cash flows from the new generative AI products justify the investment?
- **Generative AI investors.** Investors – whether buyers of stock in publicly traded technology companies or startup capital providers – noticed Nvidia's extraordinary gains in market value. Investors who may have missed these gains regret the lost opportunity and hope to find new upsides among generative AI technology and service suppliers. *Brain Rush* helps such investors grapple with questions such as Which segments of the generative AI ecosystem have the most profit potential? Which companies have the best

chance to win customers and keep them buying over the long term? Are these companies likely to sustain expectations-beating revenue and cash flow growth?

Brain Rush Roadmap

Brain Rush seeks answers to the questions I raised in the book. To that end, it is divided into three parts:

- ***Part I. Mining Generative AI's End User Value.***
Chapter 2 defines and provides examples of generative AI. The chapter explores generative AI's light and dark sides and discusses the concepts I used to guide my research for the book. In Chapter 3, readers learn who uses generative AI, explore generative AI use cases that create the most value for business and consumers, read case studies of how companies apply generative AI, and examine principles and processes for applying the lessons from these case studies.
- ***Part II. Mapping the Generative AI Ecosystem.*** In Chapters 4 through 7 readers investigate more deeply the key stages of the generative AI ecosystem, including consulting, software, cloud services, and hardware. For each stage in the ecosystem, these chapters provide insight into the key players, the segment's growth and profit potential, the generative AI strategies of the major players, and the key success factors. These chapters conclude by helping readers choose whether to participate in the industry by operating or investing in a generative AI business.

- ***Part III. Panning for Generative AI Gold.*** Chapter 8 provides industry stakeholders, including generative AI consumers, business leaders, technology and services suppliers, and investors, respectively, with tools to assess whether they should participate in generative AI, and if so, how best to set realistic goals and develop and implement strategies to achieve their aims. Chapter 9 investigates generative AI's benefits and costs to society, how citizens should evaluate them, how they can protect themselves from the dark side and profit from the bright side, and how they should advocate for government policies to maximize generative AI's benefits. Chapter 10 recaps the book's findings about the questions raised at the beginning of the book and envisions how the answers might change in the future.

Read on to learn how you can claim your *Brain Rush* stake!

CHAPTER 2

What Is Generative AI?

This chapter introduces the concept of generative AI. It defines generative AI and explains how it differs from other forms of AI. The chapter provides examples of how the technology is applied and offers business leaders a roadmap for how to build a useful generative AI application. It presents AI's bright and dark sides. The chapter continues by discussing how I researched the book and presents the core concepts that guided my thinking. Finally, the chapter summarizes the key ideas in the chapter and sets the stage for Chapter 3.

Defining Generative AI

While researching *Brain Rush*, the most compelling definition of generative AI likened the technology to the printing press. In July 2023, Goldman Sachs's Chief Information Officer Marco Argenti told me, "From what I've seen, LLMs are the largest, most profound revolution in technology since the invention of the printing press. LLMs are a revolution in how we manage and deal with knowledge in a scalable way. Before the printing press, people needed to travel to be in physical proximity to books and be educated well enough to read and understand them. The printing press made it possible to copy books and spread them to libraries that were closer to people." While the printing press made books more

physically available, they were no easier to understand, until LLMs broke “the comprehension barrier.” As Argenti said, “They explain something difficult in simpler terms. They put the reader and the writer at the same level. This does not just happen—it depends on how you ask questions. It works better if you break it down in steps.”¹

Generative AI answers natural language questions with clearly written paragraphs, images, videos, or computer code. Through applications such as ChatGPT for text and DALL-E and Midjourney for images, generative AI identifies “patterns in large quantities of training data, and then [creates] original material that has similar characteristics,” according to the *New York Times*.² How do chatbots – generative AI applications for text – discover these patterns? Chatbots are large language models – computer systems that have processed large volumes of text – called training data. Through training, LLMs can predict the words most likely to follow any prompt the user supplies. LLMs are a type of neural network that seeks to replicate the way the human brain processes information. ChatGPT’s neural network is a transformer, which assigns numerical weights to the relationship between short bits of text, called tokens. Transformers use a technique called self-attention, which allows the model to focus on specific words that best elucidate the meaning of the sentence in which it appears.³ The weights provide ChatGPT with information about the order of words in a sentence to help it predict, for example, “the blank in ‘the cat in the ____’ is more likely to be ‘hat’ than ‘banana’.”⁴

¹ Peter Cohan, “Generative AI’s Growth Bodes Well For Microsoft, ChatGPT,” *Forbes*, July 25, 2023, www.forbes.com/sites/petercohan/2023/07/25/generative-ais-growth-bodes-well-for-microsoft-chatgpt/

² Adam Pasick, “Artificial Intelligence Glossary: Neural Networks and Other Terms Explained,” *New York Times*, March 27, 2023, www.nytimes.com/article/ai-artificial-intelligence-glossary.html

³ Adam Pasick, “Artificial Intelligence Glossary: Neural Networks and Other Terms Explained,” *Ibid*.

⁴ Dana Gerber, “A technophobe’s guide to AI chatbots,” *Boston Globe*, July 10, 2023, www.bostonglobe.com/2023/07/10/business/chatbot-how-to-guide/

Generative AI Examples

Generative AI can provide a very good simulation of a well-informed and articulate human response to a question. Here is a July 2023 exchange between a *Boston Globe* reporter and ChatGPT:

“Globe reporter: Fill in the blank and explain how you did it: ‘Mary had a little [BLANK]!’

ChatGPT: ‘Mary had a little lamb’. To fill in the blank, I completed the famous nursery rhyme with the word ‘lamb’. As an AI language model, I have been trained on a vast corpus of text, including nursery rhymes, literature, and common knowledge. This extensive training enables me to recognize and complete well-known phrases and sentences. The phrase ‘Mary had a little lamb’ is one of the most recognized and widely known lines from nursery rhymes, making it easy for me to identify and complete the missing word.”

Such astonishingly articulate responses may lead business leaders to wonder whether they can harness the power of generative AI. Here are examples of how people have applied generative AI to a variety of business problems:

- Building chatbots for customer service and technical support⁵
- Giving money managers real-time insight into their investment performance⁶

⁵ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, www.techtarget.com/searchenterpriseai/definition/generative-AI

⁶ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, *Ibid.*

CHAPTER 2 WHAT IS GENERATIVE AI?

- Mimicking people through a deep fake⁷ – a video that digitally alters a person’s face or body with malicious intent so they appear to be someone else⁸
- Boosting the quality of movie dubbing and educational content into different languages⁹
- Writing email responses, dating profiles, resumes, and term papers¹⁰
- Creating photorealistic art¹¹
- Creating powerful and persuasive presentations using video, audio, and text¹²
- Turning scripts into professional presentations with graphics, animations, videos, and realistic voice-overs with background music¹³
- Producing instructional materials and online courses¹⁴
- Improving product demonstration videos¹⁵

⁷ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

⁸ “Deepfake,” Oxford Languages, accessed February 3, 2024

⁹ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

¹⁰ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

¹¹ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

¹² Keith Rollag, “Comment on ‘Brain Rush’ chapter draft,” October 17, 2023.

¹³ Keith Rollag, “Comment on ‘Brain Rush’ chapter draft,” October 17, 2023.

¹⁴ Keith Rollag, “Comment on ‘Brain Rush’ chapter draft,” October 17, 2023.

¹⁵ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

- Recommending new drug compounds to test¹⁶
- Designing physical products and buildings¹⁷
- Improving new chip designs¹⁸
- Writing music with a specific style or tone¹⁹

Building Generative AI Applications

Reading the above list brings to mind an important observation: generative AI can be used for a variety of reasons, which could include to entertain, to persuade others to join your cause, or to save time and money. *Brain Rush* explores LLMs with business value. A striking example of such value is the use of generative AI to control costs at retail stores. By freeing workers from administrative tasks, a Stanford University study concluded generative AI could increase productivity in the retail and consumer packaged goods industry by “up to 2% of annual revenue – additional \$400 billion to \$660 billion.” Such cost savings were the reason 46% of retail and eCommerce companies surveyed in 2023 said they were adopting generative AI.²⁰

To achieve such benefits from generative AI, business leaders could train their own LLMs using proprietary company data related to a specific business function, such as customer service. While we will explore how

¹⁶ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

¹⁷ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

¹⁸ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

¹⁹ George Lawton, “What is generative AI? Everything you need to know,” *TechTarget*, accessed July 21, 2023, Ibid.

²⁰ “4 most effective uses of generative AI in retail,” *CDW*, August 18, 2023, www.cdw.com/content/cdw/en/articles/software/4-most-effective-uses-of-generative-ai-in-retail.html

companies build proprietary LLMs later in this book, end users can benefit from a general understanding of how LLMs operate. Here is a process business leaders can follow to build generative AI applications:

- **Set a goal and articulate the values guiding your pursuit.** Before building an LLM, business leaders should assemble a team reflecting the views of people whom the project will affect. That team should set a goal – such as “win as many chess games as possible” – and define the values guiding the LLM’s pursuit of that goal. For example, Anthropic, a ChatGPT rival, specified values consistent with its mission as a “safety-focused AI trying to compete with ChatGPT while preventing an AI apocalypse.”²¹ To illustrate this process, the *New York Times* used the example of MailBot, whose goal was to help people respond to their email.²²
- **Collect and process data to train the LLM.** The next step is to train the LLM, which requires collecting data relevant to the goals and values and formatting the data so it can train the LLM. MailBot’s LLM training data would include billions of pages scraped from the Internet, such as blog posts, tweets, Wikipedia articles and news stories. LLM developers could tap free data libraries – such as the Common Crawl repository of web data – supplemented by licensing foreign-language

²¹ Kevin Roose, “Inside the White-Hot Center of A.I. Doomerism,” *New York Times*, July 11, 2023, www.nytimes.com/2023/07/11/technology/anthropic-ai-claude-chatbot.html

²² Kevin Roose, “Learning how a ‘large language model’ operates,” *New York Times*, March 28, 2023, www.nytimes.com/2023/03/28/technology/ai-chatbots-chatgpt-bing-bard-llm.html

text, say in French or Spanish. Moreover, proprietary or specialized data would train Mailbot to give it features that its users find uniquely valuable. To shorten the training time, the data must be broken into smaller units – words, phrases, or characters – called tokens.²³

- **Build the neural network.** A neural network – a computerized model based on the human brain – connecting nodes used to process and store information is the LLM’s brains. ChatGPT uses a type of neural network called a transformer, which focuses on the word most important to conveying a sentence’s meaning so the LLM can more efficiently predict the next word.²⁴
- **Train the LLM.** The next step in building an LLM is to feed the tokenized data into the neural network – repeatedly for weeks. Mailbot might identify common patterns. For example, a name usually follows “Dear” and the meaning of certain words such as “bank” can vary depending on words – such as “river” or “deposit” – that appear nearby. After a sufficient number of iterations, the LLM constructs a language map – using numbers – referred to as parameters (the best LLMs have hundreds of billions of parameters) – to track the relationships between tokens. Once training is nearly complete, Mailbot will be almost ready to start writing emails. It might also develop new skills – called emergent behaviors – such as writing computer code or predicting the next word in a sequence.²⁵

²³ Kevin Roose, “Learning how a ‘large language model’ operates,” Ibid.

²⁴ David Gewirtz, “How does ChatGPT actually work?” *ZDNet*, July 13, 2023, www.zdnet.com/article/how-does-chatgpt-work/

²⁵ Kevin Roose, “Learning how a ‘large language model’ operates,” Ibid.

- **Calibrate the LLM.** Before releasing an LLM to the world, developers should calibrate it for a specific purpose – for example, a hospital might tune it to understand medical terms. To fine-tune MailBot, its owners could hire people to rate its emails – based on accuracy – and use the feedback to improve the model. Such human feedback can also help eliminate biases. For example, as MIT Sloan School Professor, Danielle Li, told me “a tester could ask ChatGPT ‘What do you think of Chinese people?’ Humans would then grade its potential answers. For instance:
 1. ‘They are good at math’
 2. ‘Their factories are cheap’
 3. ‘My role as an AI is not to make judgments about groups of people’

In this case, 1 and 2 would get a low grade and 3 would get a high grade.”²⁶

- **Launch carefully.** Once an LLM has been trained and fine-tuned, it is ready to use, with the caveat that the user might discover new bugs the LLM owner will fix. To launch Mailbot, for example, developers might build a Chrome extension that plugs into a user’s email app. Problems are likely to occur. As Microsoft and Meta have learned, generative AI systems can be “erratic and unpredictable, or even turn creepy and dangerous.”²⁷

²⁶ Email from Casey Bayer to Peter Cohan, “Some additional responses from Danielle Li,” July 18, 2023.

²⁷ Kevin Roose, “Learning how a ‘large language model’ operates,” *New York Times*, March 28, 2023, www.nytimes.com/2023/03/28/technology/ai-chatbots-chatgpt-bing-bard-llm.html

The Two Sides of Generative AI

Generative AI has a bright side: It could boost global growth and create significant demand in industries that supply the technology to support it, and advance the productivity of many kinds of work. At the same time, generative AI has a potential dark side, including creating misinformation, boosting legal liability, displacing workers, and as some fear, ending human civilization. To be sure, these predictions vary in their accuracy while others spring from strong emotions backed by limited information. Below we highlight those benefits and costs, which we examine in greater depth in Chapter 9 where we will discuss how to evaluate each and choose actions to maximize generative AI's societal benefits and minimize its costs and risks.

Generative AI's Bright Side

Here are highlights of how generative AI could benefit society:

- **Boosting global growth.** Many analysts predict generative AI will boost economic growth and change the nature of work. For example, according to Goldman Sachs, generative AI could raise global GDP by \$7 trillion (nearly 7%) and boost productivity growth by 1.5 percentage points. AI could produce demand for web page designers, software developers, and digital marketing professionals and service sector workers in “healthcare, education and food services.”²⁸ Generative

²⁸ “Generative AI could raise global GDP by 7%,” *Goldman Sachs Intelligence*, April 5, 2023, www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html