

contents

<i>foreword</i>	<i>ix</i>
<i>preface</i>	<i>xi</i>
<i>acknowledgments</i>	<i>xiii</i>
<i>about this book</i>	<i>xv</i>
<i>about the authors</i>	<i>xix</i>
<i>about the cover illustration</i>	<i>xxi</i>

1	<i>Large language models: The power of AI</i>	1
	Evolution of natural language processing	3
	The birth of LLMs: Attention is all you need	7
	Explosion of LLMs	10
	What are LLMs used for?	11
	<i>Language modeling</i>	12
	<i>Question answering</i>	13
	<i>Coding</i>	14
	<i>Content generation</i>	15
	<i>Logical reasoning</i>	17
	<i>Other natural language tasks</i>	18
	Where do LLMs fall short?	19
	<i>Training data and bias</i>	20
	<i>Limitations in controlling machine outputs</i>	23
	<i>Sustainability of LLMs</i>	24

Revolutionizing dialogue: Conversational
LLMs 25

OpenAI's ChatGPT 26 ■ *Google's Bard/*
LaMDA 27 ■ *Microsoft's Bing AI* 29
Meta's LLaMa/Stanford's Alpaca 30

2 ***Training large language models*** 34

How are LLMs trained? 35

Exploring open web data collection 36
Demystifying autoregression and bidirectional
token prediction 38 ■ *Fine-tuning*
LLMs 39

The unexpected: Emergent properties of
LLMs 40

Quick study: Learning with few examples 40
Is emergence an illusion? 44

What's in the training data? 44

Encoding bias 45 ■ *Sensitive*
information 48

3 ***Data privacy and safety with LLMs*** 52

Safety-focused improvements for LLM
generations 53

Post-processing detection algorithms 54
Content filtering or conditional
pre-training 56 ■ *Reinforcement learning*
from human feedback 57 ■ *Reinforcement*
learning from AI feedback 59

Navigating user privacy and commercial
risks 61

Inadvertent data leakage 62 ■ *Best practices*
when interacting with chatbots 64

Understanding the rules of the road: Data
policies and regulations 65

*International standards and data protection
laws 65 ■ Are chatbots compliant with
GDPR? 68 ■ Privacy regulations in
academia 70 ■ Corporate policies 70*

4 ***The evolution of created content 73***

The rise of synthetic media 74

*Popular techniques for creating synthetic
media 75 ■ The good and the bad of synthetic
media 78 ■ AI or genuine: Detecting
synthetic media 79*

Generative AI: Transforming creative
workflows 82

*Marketing applications 82 ■ Artwork
creation 84*

Intellectual property in the LLM era 88

*Copyright law and fair use 88 ■ Open source
and licenses 95*

5 ***Misuse and adversarial attacks 100***

Cybersecurity and social engineering 101

Information disorder: Adversarial
narratives 114

Political bias and electioneering 123

Why do LLMs hallucinate? 126

Misuse of LLMs in the professional world 134

6 ***Accelerating productivity: Machine-augmented work*** 142

Using LLMs in the professional space 143

LLMs assisting doctors with administrative tasks 143 ■ *LLMs for legal research, discovery, and documentation* 146

LLMs augmenting financial investing and bank customer service 148 ■ *LLMs as collaborators in creativity* 149

LLMs as a programming sidekick 151

LLMs in daily life 154

Generative AI's footprint on education 161

Detecting AI-generated text 165

How LLMs affect jobs and the economy 169

7 ***Making social connections with chatbots*** 173

Chatbots for social interaction 174

Why humans are turning to chatbots for
relationship 180

The loneliness epidemic 180 ■ *Emotional attachment theory and chatbots* 183

The good and bad of human-chatbot
relationships 186

Charting a path for beneficial chatbot
interaction 193

8 ***What's next for AI and LLMs*** 201

Where are LLM developments headed? 202

Language: The universal interface 202

LLM agents unlock new possibilities 204

The personalization wave 206

Social and technical risks of LLMs	207
<i>Data inputs and outputs</i>	207
▪ <i>Data privacy</i>	209
▪ <i>Adversarial attacks</i>	210
▪ <i>Misuse</i>	212
▪ <i>How society is affected</i>	213

Using LLMs responsibly: Best practices	215
<i>Curating datasets and standardizing documentation</i>	215
▪ <i>Protecting data privacy</i>	217
▪ <i>Explainability, transparency, and bias</i>	219
▪ <i>Model training strategies for safety</i>	223
▪ <i>Enhanced detection</i>	226
▪ <i>Boundaries for user engagement and metrics</i>	228
▪ <i>Humans in the loop</i>	230

AI regulations: An ethics perspective	231
<i>North America overview</i>	232
▪ <i>EU overview</i>	236
▪ <i>China overview</i>	240
▪ <i>Corporate self-governance</i>	243

Toward an AI governance framework	245
-----------------------------------	-----

9 ***Broadening the horizon: Exploratory topics in AI*** 251

The quest for artificial general intelligence	252
AI sentience and consciousness?	260
How LLMs affect the environment	265
The game changer: Open source community	270

<i>references</i>	277
-------------------	-----

<i>index</i>	309
--------------	-----