

Martin Musiol

GENERATIVE AI

Navigating the Course
to the Artificial General
Intelligence Future

WILEY

Praise for Generative AI

“Cutting through the clutter, Martin Musiol explains generative AI with great insight and clarity. The reader is left with a clear understanding of the technology, without the need to master complex mathematics or code. A must read for those who want to understand the future.”

— **Rens ter Weijde, Chairman & CEO of KIMO.AI**

“An illuminating guide through the evolving landscape of generative AI and AGI, this book masterfully demystifies complex concepts, making them accessible to all and ignites the imagination about the boundless possibilities of the future.”

— **David Foster, author of Generative Deep Learning, Partner at Applied Data Science Partners**

“This book is a must-read for anyone wanting to improve their understanding of where AI has come from, where it stands today, and, importantly, where it is heading. The advent of AGI and ASI is too important not to understand, and Martin meticulously explains many potential outcomes with a factual and unbiased perspective.”

— **Roy Bhasin (Zeneca), author, entrepreneur, angel investor**

“Highly recommended. Musiol deeply and expertly demonstrates how to navigate the complex, exhilarating, and essential landscape of generative AI.”

— **Katie King, published author, CEO of AI in Business**

“Generative AI by Martin Musiol offers a comprehensive overview of the GenAI technology and skillfully demystifies complex concepts of this transformative AI.”

— **Sheamus McGovern, entrepreneur, investor, Founder & CEO Open Data Science**

“Martin, my esteemed former colleague and an AI expert, has authored this crucial book designed for anyone seeking to enhance their knowledge of generative AI, autonomous AI agents, and AGI. From complex subjects to compelling and easily comprehensible, this book is invaluable for business applications and everyday life.”

— Martin Weis, Country Head Switzerland & Global Co-Lead AI, Analytics & Automation at Infosys Consulting

“Martin’s book masterfully encapsulates the transformative power of AI and provides great foundational knowledge for innovators and builders to explore the industry further.”

— Anton Volovyk, Co-CEO Reface (GenAI app, 250m downloads, backed by a16z)

“This book is akin to a comprehensive playbook, detailing strategies and rules for navigating the complex field of AI, much like a coach laying out a winning game plan. It masterfully presents the evolutionary stages, key players beyond ChatGPT, foundational technologies, and practical guidance, equipping readers to effectively ‘play’ and excel in the dynamic and competitive arena of AI.”

— Dr. Harald Gunia, Leader for Applied Artificial Intelligence Europe at Infosys Consulting

“Martin Musiol’s book on generative AI provides a compelling narrative that unveils the meticulous evolution of this groundbreaking technology. From the quiet simmering of its inception, to the carefully curated recipe of technological advancements that propelled it to unprecedented heights, Musiol carefully peels back the layers, revealing the pivotal factors that shaped the rise of generative AI.”

— Matteo Penzo, Co-Founder & CEO of zicklearn.com

“Martin’s book offers deep insights and a comprehensive overview that makes this complex subject accessible to all readers.”

— Prof. Dr. Patrick Glauner

“This book is a must-read for anyone like me captivated by artificial intelligence’s present and future implications.”

— Catherine Adenle, Senior Director, Global Employer Brand, Elsevier, top 22 AI and tech influencer

Generative AI

**Navigating the Course to the Artificial
General Intelligence Future**

Martin Musiol

WILEY

Copyright © 2024 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada and the United Kingdom.

ISBNs: 9781394205912 (Hardback), 9781394205950 (ePDF), 9781394205943 (ePub)

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at www.wiley.com/go/permission.

Trademarks: WILEY and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Generative AI tools were used by the author to research ideas for this book; however, the writing and finished text of the book are completely the work of the author and the Wiley editorial staff.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Control Number: 2023951020

Cover image: © undefined/Getty Images

Cover design: Wiley

To my parents, who have always supported me, and to my grandma Helena, whose wise words continue to echo in my ears, guiding me through life. I will be forever grateful for the deep love I have received from you, and rest assured, I feel the same for you. A truth perhaps not spoken enough, yet profoundly felt.

Contents

<i>Introduction</i>	<i>ix</i>
Chapter 1 AI in a Nutshell	1
Chapter 2 Innovative Approaches for High-Quality Data Generation	23
Chapter 3 Generative AI's Broad Spectrum of Applications	119
Chapter 4 Generative AI's Exponential Growth	219
Chapter 5 Ethical Concerns and Social Implications of Generative AI	285
Chapter 6 Artificial General Intelligence in Sight	337
<i>Acknowledgments</i>	<i>405</i>
<i>About the Author</i>	<i>407</i>
<i>Index</i>	<i>409</i>

Introduction

In the realm of technology, epochs of transformation are often ignited by the spark of human imagination, fused with the finesse of engineering artistry. We stand at the precipice of such an epoch, where the realms of generative AI unfurl into the once uncharted territories of artificial general intelligence (AGI). I am both thrilled and humbled to be your guide on this thrilling expedition into the future, a journey that begins with the pages of this book.

The technological zeitgeist of our times is one of exponential progress. A mere glimpse into the recent past reveals the embryonic stages of generative AI, yet, within a fleeting span, advancements like ChatGPT have marked a point of no return. This crescendo of innovation is not confined to textual realms alone but spans across images, videos, 3D objects, datasets, virtual realities, code, music, and sound generation, each stride accelerating our pace toward the enigmatic horizon of AGI. The rapid maturation and adoption of generative AI outshine the evolutionary arcs of many preceding technologies.

It was during the cusp of this book's creation that the concept of autonomous AI agents morphed into a tangible reality, courtesy of emerging open source frameworks. Now, a subscription away, the first AI agents are at our beck and call. This swift progression, magnifying the efficiency of AI model development,

underscores the urgency and the timeliness of delving into the discourse this book intends to foster. As you traverse through its chapters, you'll realize we are merely at the dawn of an exhilarating technological epoch with a vast expanse yet to be unveiled.

Who should venture into this exploration? Whether you're a technology aficionado, a student with a zest for the unknown, a policymaker, or someone who's merely curious, this book beckons. No prior acquaintance with AI or machine learning is required; your curiosity is the sole ticket to this expedition. As we commence, we'll demystify the essence of AI, its lexicon, and its metamorphosis over time. With each page, we'll delve deeper, yet the narrative is crafted to foster an understanding, irrespective of your prior knowledge. By the narrative's end, your imagination will be aflame with the boundless possibilities that the future holds.

The narrative arc of this book has been meticulously crafted to offer an understanding yet a profound insight into generative AI and its trajectory toward AGI. Our expedition begins with the rudiments of AI, tracing its evolution and the brilliant minds that propelled it forward. As we delve into the heart of generative AI, we'll explore its broad spectrum of applications, unraveling potential startup ideas and pathways to venture into this domain. The discussion will then transcend into the convergence of diverse technological realms, each advancing exponentially toward a shared zenith. Ethical and social considerations, indispensable to this discourse, will be deliberated upon before we venture into the realms of AGI, humanoid and semi-humanoid robotics, and beyond. Through the annals of my experience, including my tenure as the generative AI lead for EMEA at Infosys Consulting, we'll traverse through real-world scenarios, albeit veiled for confidentiality, offering a pragmatic lens to envision the theoretical discourse.

What sets this narrative apart is not merely the content, but the vantage point from which it is observed. My journey, from

advocating generative AI since 2016, founding GenerativeAI.net in 2018, to now sharing a platform with luminaries at the AI Speaker Agency, has been nothing short of exhilarating. It's through the crucible of real-world implementations and continuous discourse with global thought leaders that the insights within this book have been honed. Our conversations, a confluence of diverse perspectives, have enriched the narrative, making it a crucible of collective wisdom.

A treasure trove of knowledge awaits to equip you to navigate the complex yet exhilarating landscape of generative AI and AGI. The ethos of this narrative is to empower you to become a 10X more effective human, to harness the tools that propel you forward, and should a spark of an idea ignite within, to pursue it with vigor. Things can be figured out along the way, especially in this era equipped with generative AI tools. Remember, AI in itself won't replace us, but those wielding AI effectively certainly will have an edge.

In the words of British physicist David Deutsch, our civilization thrives on technological growth, and it's our prerogative to strive for a better future. This book is a stepping stone toward that endeavor, and I invite you to step into the future, one page at a time.

How to Contact the Publisher

If you believe you've found a mistake in this book, please bring it to our attention. At John Wiley & Sons, we understand how important it is to provide our customers with accurate content, but even with our best efforts an error may occur.

In order to submit your possible errata, please email it to our Customer Service Team at wileysupport@wiley.com with the subject line "Possible Book Errata Submission."

How to Contact the Author

I appreciate your input and questions about this book! Feel free to contact me at the following:

Martin Musiol's email: generativeai.net@gmail.com

Martin's LinkedIn profile: www.linkedin.com/in/martinmusiol1

GenerativeAI.net's web page: <https://generativeai.net>

CHAPTER

1

AI in a Nutshell

No other field of technology has such inconsistent jargon as artificial intelligence (AI). From mainstream media to tech influencers to research scientists, each layer of media has contributed to that confusion. In order of their degree of contribution and frequency, I observed mainstream media simplifying and misusing terms consistently, tech influencers misunderstanding the tech in-depth, and even some research scientists overcomplicating their model findings with fancy terms. By no means do I intend to criticize research scientists. They are the backbone of everything discussed in this book. Their work offers solutions to a plethora of problems, making AI the umbrella term for almost every intelligent problem. However, its interdisciplinary nature, the rapid advancements in this space, and AI's general complexity make it already difficult to gain a clear understanding of this field. I am convinced that consistent and clear language would help to understand this topic area.

We can see two broad classes in AI: generative AI, the subject of this book, and discriminative AI. The latter is the traditional and better-known part of AI. Before delving into both AI classes, let's take a moment to understand the broader picture of AI, machine learning (ML), deep learning (DL), and the process of training models, to avoid getting ahead of ourselves.

What Is AI?

Even though AI includes a broad spectrum of intelligent code, the term is often incorrectly used. Figure 1.1 shows how AI, ML, and DL are related. ML, a part of AI, learns from data. DL, a deeper part of ML, uses layered setups to solve tougher problems. Non-self-learning programs like expert systems don't learn from data, unlike ML and DL. We'll explore these more next.

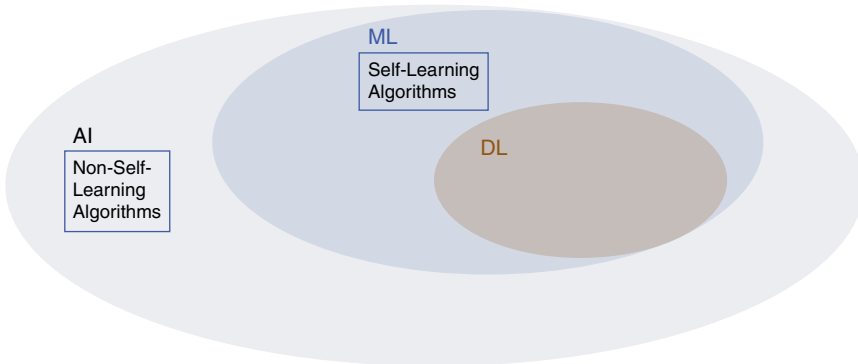


FIGURE 1.1 The relationship between AI, ML, and DL

How AI Trains Complex Tasks

AI can perform tasks ranging from predefined expert answers, also known as expert systems, to tasks that require human-level intelligence. Think about recognizing speech and images,

understanding natural language processing (NLP), making sophisticated decisions, and solving complex problems. For tasks like this, the AI has to train on a respective dataset until it is able to perform the desired activity as well as possible. This self-learning part of AI is referred to as *machine learning* (ML). Because most of the interesting applications are happening through machine learning in one way or another, and to keep it simple, we use AI and ML interchangeably.

To make it tangible, we are designing an AI system that rates the cuteness of cats from 5 (absolutely adorable) to 1 (repulsively inelegant). The ideal dataset would consist of pictures of cute kittens, normal cats, and those half-naked grumpy cats from the Internet. Further, for classifying pictures in a case like this, we would need labeled data, meaning a realistic rating of the cats. The model comes to life through three essential steps: training, validation, and evaluation.

In training, the model looks at each picture, rates it, compares it with the actually labeled cuteness of the cat, and adjusts the model's trainable parameters for a more accurate rating next time—much like a human learns by strengthening the connections between neurons in the brain. Figure 1.2 and Figure 1.3 illustrate training and prediction, respectively.

Throughout the training process, the model needs to make sure training goes in the right direction—the validation step. In validation, the model checks the progress of the training against separate validation data. As an analogy, when we acquire a skill like solving mathematical problems, it makes sense to test it in dedicated math exams.

After training has been successfully completed and respective accuracy goals have been reached, the model enters the prediction or evaluation mode. The trainable parameters are not being adjusted anymore, and the model is ready to rate all the cats in the world.

AI Model in Training (2 Steps)

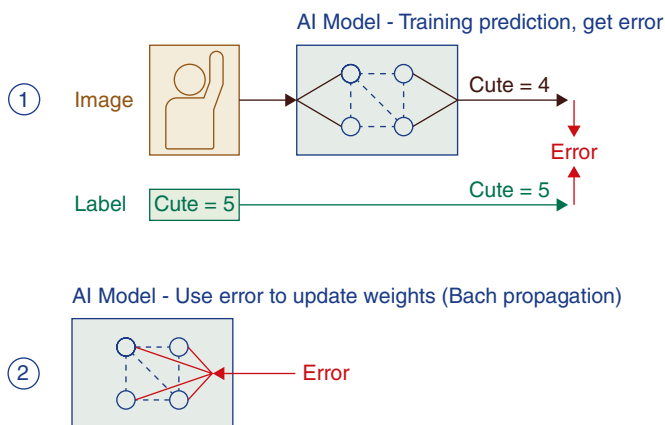


FIGURE 1.2 In supervised training of a ML model, two main steps are involved: predict the training data point, then update the trainable parameters meaningfully based on the prediction's accuracy.

AI Model in Prediction

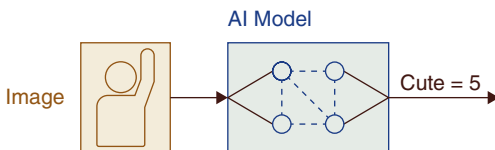


FIGURE 1.3 Prediction mode in a supervised ML model.

It is typical for a model in production mode that the accuracy gets worse over time. The reason for this could be that the real-world data changed. Maybe we are only looking at kittens and they are all cute compared to our training data. Retraining the model, whenever accuracy decreases or by scheduling retraining periodically, tackles the problem of a discrepancy between the data distribution of training data and evaluation data.

Perhaps you have a sense already that training AI models requires much more computing power than they need in prediction mode. To adjust its trainable parameters, often referred to as *weights*, we need to calculate the grade of adjustment carefully. This happens through a famous model function called

backpropagation. It entails the backward propagation of prediction errors—the learning from making mistakes in the training process. The errors are turned back to respective weights for improvement. This means that we go forward to predict a data point and backward to adjust the weights. In prediction mode, however, we don't adjust the weights anymore, but just go forward and predict. The function that has been trained through the training data is being applied, which is comparatively cheap.

Unsupervised Learning

When ML models reach a certain complexity by having many computing stages, called *layers*, we enter the realm of deep learning (DL). Most of the cutting-edge applications are at least partially drawing their algorithms from DL. Algorithms are step-by-step instructions for solving problems or performing tasks.

The preceding example of rating the cuteness of a cat was simplified drastically and didn't tell the whole story. A relevant addition to this is that as we train on labeled cat pictures, with the label being the cuteness of the cats, we call this *supervised machine learning*. With labels, we provide guidance or feedback to the learning process in a supervised fashion.

The counterpart for supervised ML is called *unsupervised machine learning*. The main difference between them is that in unsupervised ML the training data is not labeled. The algorithms ought to find patterns in the data by themselves.

For example, imagine you have a dataset of customer purchases at a grocery store, with information about the type of product, the price, and the time of day. In AI these attributes are called *features*. You could use an unsupervised clustering algorithm to group similar purchases together based on these features. This could help the store better understand customer buying habits and preferences. The algorithm might identify

that some customers tend to buy a lot of fresh produce and dairy products together, whereas others tend to purchase more processed foods and snacks. This information could be used to create targeted marketing campaigns or to optimize store layout and product placement.

Comparing the performance of unsupervised learning applications to that of supervised learning applications is akin to contrasting boats with cars—they represent distinct methodologies for addressing fundamentally diverse problems. Nevertheless, there are several reasons why we reached success years faster with supervised than with unsupervised learning methods.

In supervised learning, the model is given a training dataset that already includes correct answers through labels. Understandably, this helpful information supports model learning. It also accurately outlines the AI model's intended objective. The model knows precisely what it is trying to achieve. Evaluating the model's performance is simpler than it is in unsupervised machine learning, as accuracy and other metrics can be easily calculated. These metrics help in understanding how well the model is performing.

With this information, a variety of actions can be taken to enhance the model's learning process and ultimately improve its performance in achieving the desired outcomes.

Unsupervised models face the challenge of identifying data patterns autonomously, which is often due to the absence of apparent patterns or a multitude of ways to group available data.

Generative AI a Decade Later

Generative AI predominantly employs unsupervised learning. Crafting complex images, sounds, or texts that resemble reasonable outputs, like an adorable cat, is a challenging task compared to evaluating existing options. This is primarily due to the absence of explicit labels or instructions.