The Handbook of Data Science and AI

Generate Value from Data with Machine Learning and Data Analytics

DAS VORWORT

Preface

Time moves quickly, and we have already reached the next edition of the *Handbook of Data Science and AI*. When we published the first edition, we knew that some content would change only slowly – such as the mathematical foundations. At the same time, we were fully aware that disruptive technologies would evolve rapidly, reshape the field, and fundamentally change how systems are built.

The first edition of this book was published in 2019 (in German only). At that time, none of the authors could have imagined that by 2026 it would be possible to build complete applications within days using a paradigm now known as *vibe coding* – tasks that previously required months of work. Even more remarkably, these applications can integrate AI agents that themselves rely on services and capabilities that did not exist back then.

We are therefore particularly proud to modernize this new edition by embracing chapters on vibe coding and AI agents. Enjoy reading.

Spring 2026

Stefan Papp

This preface was NOT written by ChatGPT (or similar).

As I make this statement, I'm wondering how often it will remain true for text or even other forms of media in the future. Over the last two years, this AI-powered tool has risen to enormous popularity, and has given Data Science and AI an incredible awareness boost. As a result, the expectations for Artificial Intelligence have grown seemingly exponentially, and reached such heights that one might ask, if they can ever be achieved.

AI is following the well-known hype cycle. Some of these high expectations are well-deserved: this powerful technology will change the way we live and work in many ways. To name one example: some universities are considering not to ask their

students for seminar papers any longer, as it's not possible to check if it was written by an AI tool.

But we also must brace ourselves for some disappointment in the future, as AI inevitably fails to live up to certain people's inflated expectations.

Even when the vision is reasonable, often the timelines these people and organizations have in mind for implementing AI projects is not. This leads to further disappointment, when the hoped-for impact and value fail to materialize within the desired timeframe.

We're already seeing the beginning of this, with ChatGPT and similar tools generating plenty of eloquent and coherent – yet completely inaccurate – information. This isn't helped by the new wave of 'AI experts', who are making ever more outlandish promises about tools invented by themselves or their companies; promises which will be very hard to keep. They are, essentially, selling digital 'snake oil'.

All of this puts even more pressure on data scientists to deal with these expectations, while continuing to deliver on the same goal they've had for decades:

generating understandable answers to questions, using data.

This is what makes neutral organizations such as the Vienna Data Science Group (VDSG [*www.vdsg.at*]) – which fosters interdisciplinary and international knowledge exchange between data experts – so necessary and important. We are still highly dedicated to the development of the entire Data Science and AI ecosystem (education, certification, standardization, societal impact study, and so on), across Europe and beyond. This book represents just one of our efforts towards this goal. Because despite all the hype and hyperbole in the AI and data landscape, Data Science remains the same: an interdisciplinary science gathering a very heterogeneous crowd of specialists. It is made up of three major streams, and we are proud to have expert members in each of them:

- Computer Science and IT
- Mathematics and Statistics
- Domain expertise in the industry or field in which Data Science and AI is applied.

As a matter of fact, the VDSG [*www.vdsg.at*] has always taken a holistic approach to data science, and this book is no different: Starting at Chapter 1 we introduce a fictional company who wants to become more data driven, and we check in with them throughout the book, right up to the end of their data transformation in Chapter 28. Along the way we cover many challenges in their journey, thus providing you with practical insights which were only possible thanks to vibrant exchange among our vast Data Science and AI community.

The result is a greatly expanded edition of our Data Science & AI Handbook, with 10 new chapters covering topics like Building AI solutions (Chapter 13), Foundation Models (Chapter 15), Large Language Models and Generative AI (Chapter 16) and Climate

Change and AI (Chapter 25). This is complemented by also tackling the fundamental topics of Data Architecture, Engineering and Governance (Chapters 4, 5 and 6) and topping it off with Machine Learning Operations (MLOps, Chapter 7), which has become a very important discipline in itself.

To provide a firm foundation to help you understand all this, we've again included an introduction to the underlying Mathematics (Chapter 9) and Statistics (Chapter 10) used in Data Science, as well as chapters on the theory behind Machine Learning, Signal Processing and Computer Vision (Chapters 12, 14 and 18). We've also covered topics related to generating value from data, such as Business Intelligence (Chapter 11) and Data Driven Enterprises (Chapter 21), as well as vital information to help you use data safely, including chapters on the new EU AI Act (Chapter 23) and Trustworthy AI (Chapter 27).

This vast expansion of VDSG's Magnum Opus serves one core purpose:

to give a realistic and holistic picture of Data Science and AI.

Data Science and AI is developing at an incredibly quick pace at the moment and so is its impact on society. This means that responsibilities put on the shoulders of data scientists have grown as well, and so has the need for organizations like VDSG [*www.vdsg.at*] to get involved and tackle these challenges too.

Let's go for it!

Summer 2024

Wolfgang Weidinger

Acknowledgments

We, the authors, would like to take this opportunity to express our sincere gratitude to our families and friends, who helped us to express our thoughts and insights in this book. Without their support and patience, this work would not have been possible.

A special thanks from all the authors goes to Katherine Munro, who contributed a lot to this book and spent a tremendous amount of time and effort editing our manuscripts.

For my parents, who always said I could do anything. We never expected it would be a thing like this.

Katherine Munro

I'd like to thank my wife and the Vienna Data Science Group for their continuous support through my professional journey.

Zoltan C. Toth

Thinking about the people who supported me most, I want to thank my parents, who have always believed in me, no matter what, and my partner Verena, who was very patient again during the last months while I worked on this book.

In addition I'm very grateful for the support and motivation I got from the people I met through the Vienna Data Science Group.

Wolfgang Weidinger