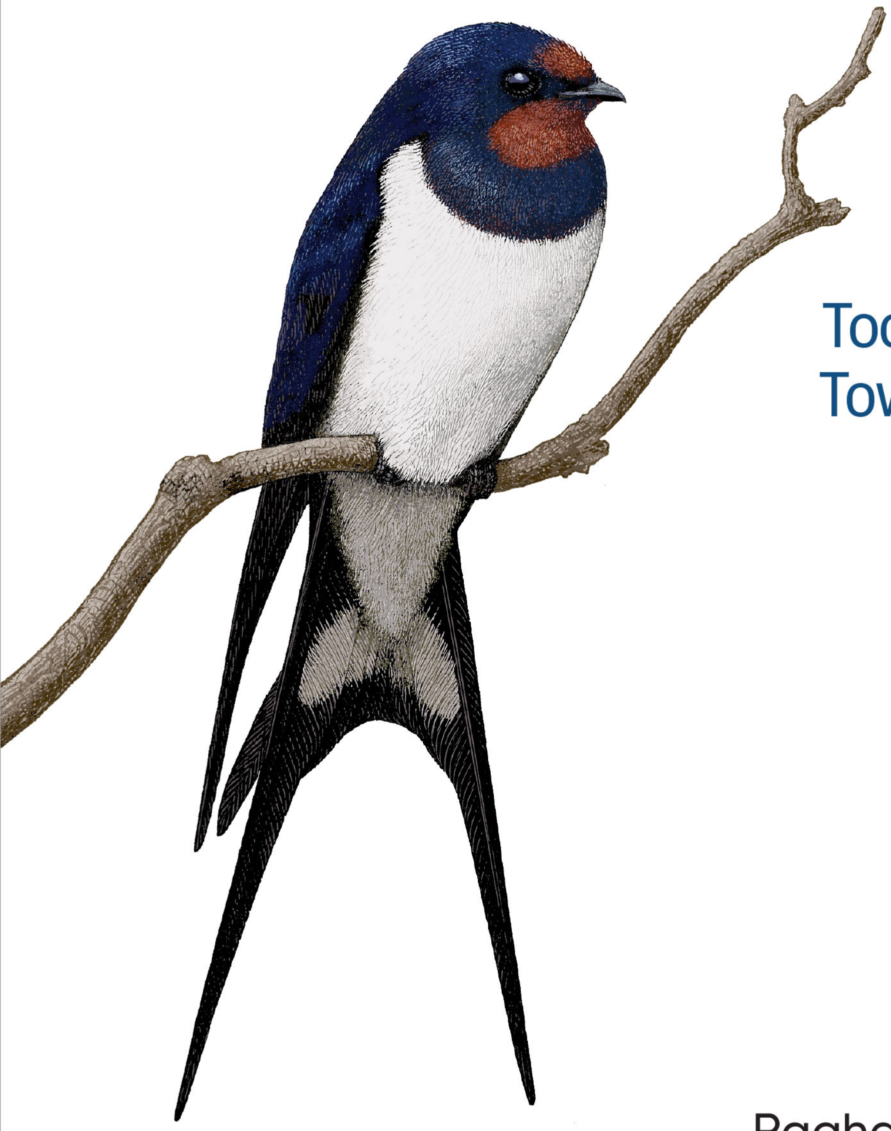
O'REILLY®

Sustainable AI



Tools for Moving Toward Green AI

Raghavendra Selvan

Sustainable AI

Tools for Moving Toward Green AI

Raghavendra Selvan

O'REILLY®

Sustainable AI
by Raghavendra Selvan

Published by O'Reilly Media, Inc., 141 Stony Circle, Suite 195, Santa Rosa, CA 95401.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*http://oreilly.com*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

Acquisitions Editor: Nicole Butterfield
Development Editor: Lauren Mine
Production Editor: Katherine Tozer
Copyeditor: Kim Wimpsett
Proofreader: Stephanie English

Indexer: Judith McConville
Interior Designer: David Futato
Cover Designer: Susan Brown
Cover Illustrator: Monica Kamsvaag
Interior Illustrator: Kate Dullea

October 2025: First Edition

Revision History for the First Edition
2025-10-08: First Release

See *http://oreilly.com/catalog/errata.csp?isbn=9781098155513* for release details.

The views expressed in this work are those of the author and do not represent the publisher's views. While the publisher and the author have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the author disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.

978-1-098-15551-3

[LSI]

Table of Contents

Preface

Climate change is at our door step. It is causing heat waves, flash floods, droughts, and other erratic weather patterns. Addressing the challenges posed by climate change will be the defining project of our times. To do this, we should employ all the tools at our disposal. And one of the most powerful tools currently is artificial intelligence (AI), which has revolutionized tasks in many application domains. As such, AI can be indispensable in our efforts to combat climate change.

The recent class of AI methods, however, is growing to be extremely resource-intensive. Developing and using them requires powerful datacenters, which consume vast amounts of energy with correspondingly high carbon emissions. In addition, the datacenters used for AI require large volumes of fresh water in their cooling systems, rely on extractive mining to manufacture the electronics, and incur additional carbon emissions along their lifecycle. These factors, and other broader environmental impacts, pose a dilemma about using AI for sustainability.

The main argument in this book is that the material basis for any technology should not be discounted even in the light of their promised benefits. This is also true for AI. Even though AI has promised—and delivered on some—solutions to the sustainability challenges, the underlying resource cost of AI should not be ignored. If we don't pay close attention to these massive costs, the supposed benefits offered might be eclipsed by the negative impacts of AI; the trade-off between the cost and benefits should always be considered.

This book is an attempt to lay out these arguments so that we can make meaningful trade-offs that advance the sustainability of AI, while using it to improve the sustainability of our planet. To do this, the book presents practical tools and conceptual frameworks that will help us assess and grapple with the complex interplay between sustainability and AI.

Who Should Read This Book?

The book is primarily aimed at machine learning (ML) practitioners, which is by now a broad definition because almost all of us are using AI in one way or another.

The majority of the book focuses on stakeholders who are responsible for developing, deploying, and assessing the impact of AI. This includes engineers who develop novel ML models, managers who commission new AI applications, and policymakers who want to obtain a better insight into the technicalities and thus assess the trade-offs when developing and using AI models. Anyone who is broadly interested in the topics of sustainability and AI but not in algorithmic development can also get a lot out of this book by skipping some of the technical sections.

Sustainability and AI are the two most important concepts that will shape our future, and this book is positioned at the intersection of these ideas. I hope that ML practitioners will get a lot out of this book, while a general audience will still find it useful to draw upon statistics and discussion points that can influence their digital culture.

What This Book Is and Is Not

This is neither a popular science book nor a graduate-level academic textbook. This book tries to balance the needs of AI stakeholders by providing key arguments, formalisms, tools, and conceptual frameworks, so that we can foster informed discussions about the sustainability of AI.

Using This Book

The main questions about the sustainability of AI are framed in Chapters 1 and 3, and revisited in Chapter 10. Anyone who is broadly interested in AI can read these chapters and get a peek into the intricacies of the questions being addressed in the book.

The book does assume some background knowledge about the inner workings of ML models; however, readers who do not have this background can bridge some of the essential concepts using Chapter 2. This chapter is not intended to be a primer on ML for AI practitioners, but I do hope they will gain new intuitions about the technology behind recent AI models that are behind generative AI.

The remaining six chapters consist of a more technical look into the algorithmic workings of modern ML models. Each chapter focuses on a step in the AI model lifecycle using the gaze of resource consumption. In doing so, these chapters identify resource bottlenecks and suggest interventions that can improve the resource efficiency, and hence advance the sustainability, of AI.

Specifically, Chapter 4 addresses the question of data used in AI, Chapter 5 explores the vast space of model selection, Chapter 6 identifies improvements during training of AI models, and Chapter 7 provides suggestions for improving the resource efficiency at deployment. These four chapters are closely tied to the algorithmic choices that ML practitioners can make to exercise high-level control. Chapter 8 addresses the question of hardware efficiency, and Chapter 9 takes a system-level view of AI.

All the technical chapters have use cases that consider a real-world application of AI for sustainability. These use cases are presented several times within a chapter after introducing technical tools to show how the concepts can be put into practice.

The questions pertaining to the limitations of resource efficiency, limitations of focusing only on environmental sustainability, and how the path to achieving broader sustainability of AI are not addressed in any single chapter but form the general themes of the book.

Conventions Used in This Book

The following typographical conventions are used in this book:

Italic
: Indicates new terms, URLs, email addresses, filenames, and file extensions.

`Constant width`
: Used for program listings, as well as within paragraphs to refer to program elements such as variable or function names, databases, data types, environment variables, statements, and keywords.

`Constant width bold`
: Shows commands or other text that should be typed literally by the user.

`Constant width italic`
: Shows text that should be replaced with user-supplied values or by values determined by context.

This element signifies a tip or suggestion.

This element signifies a general note.

This element indicates a warning or caution.

Using Code Examples

Supplemental material (code examples, exercises, etc.) is available for download at *https://github.com/raghavian/sustainable_ai*. The GitHub repository also includes a bibliography with references and further reading listed by chapter.

If you have a technical question or a problem using the code examples, please send email to *support@oreilly.com*.

This book is here to help you get your job done. In general, if example code is offered with this book, you may use it in your programs and documentation. You do not need to contact us for permission unless you're reproducing a significant portion of the code. For example, writing a program that uses several chunks of code from this book does not require permission. Selling or distributing examples from O'Reilly books does require permission. Answering a question by citing this book and quoting example code does not require permission. Incorporating a significant amount of example code from this book into your product's documentation does require permission.

We appreciate, but generally do not require, attribution. An attribution usually includes the title, author, publisher, and ISBN. For example: "*Sustainable AI* by Raghavendra Selvan (O'Reilly). Copyright 2026 Raghavendra Selvan, 978-1-098-15551-3."

If you feel your use of code examples falls outside fair use or the permission given above, feel free to contact us at *permissions@oreilly.com*.

O'Reilly Online Learning

O'REILLY®

For more than 40 years, *O'Reilly Media* has provided technology and business training, knowledge, and insight to help companies succeed.

Our unique network of experts and innovators share their knowledge and expertise through books, articles, and our online learning platform. O'Reilly's online learning platform gives you on-demand access to live training courses, in-depth learning paths, interactive coding environments, and a vast collection of text and video from O'Reilly and 200+ other publishers. For more information, visit *https://oreilly.com*.

How to Contact Us

Please address comments and questions concerning this book to the publisher:

O'Reilly Media, Inc.
141 Stony Circle, Suite 195
Santa Rosa, CA 95401
800-889-8969 (in the United States or Canada)
707-827-7019 (international or local)
707-829-0104 (fax)
support@oreilly.com
https://oreilly.com/about/contact.html

We have a web page for this book, where we list errata and any additional information. You can access this page at *https://oreil.ly/SustainableAI*.

For news and information about our books and courses, visit *https://oreilly.com*.

Find us on LinkedIn: *https://linkedin.com/company/oreilly-media*.

Watch us on YouTube: *https://youtube.com/oreillymedia*.

Acknowledgments

On a spring morning in 2019, when riding the local train to work in Copenhagen, I was reading James Bridle's thought-provoking book *New Dark Age: Technology and the End of the Future* (Verso Books). Halfway through the book was a passage about the increasing energy consumption of datacenters. That passage was the seed that got me thinking about the energy consumption, and the corresponding carbon footprint, of deep learning. I conducted a literature survey, and there were no noticeable efforts in this direction. Later that year, in collaboration with two of my motivated bachelor students—Lasse Anthony and Benjamin Kanding—we developed Carbontracker. It is now widely used by the ML community (it has been downloaded 150,000 times as of this writing).

This was my first foray into the questions of sustainability of AI, which has resulted in the book you are reading now.

A book like this is written by standing on the shoulders of passionate giants. I would like to thank all the amazing researchers, thinkers, and creators who have made it their life's objective to work on sustainability and/or AI. I have referenced all of their influential works throughout the book, and I hope you will find as much inspiration as I do when reading those works. Also, I would like to acknowledge all creators at SVG Repo for making amazing designs and permitting their usage.

Writing this book has been a tough but gratifying journey. It was tough simply due to the nature of work in academia with its ebbs and flows. There were many kind and passionate people along the way who have helped me stay afloat, navigating these tides.

I would like to thank Nicole Butterfield from O'Reilly for providing me with the opportunity to work on this book on sustainable AI. Thanks also to Katherine Tozer and the others at O'Reilly for their meticulous work in bringing the book to life. Most of all, I have the utmost gratitude to Lauren Mine—my editor—for being keen, insightful, constructive, and kind throughout the duration of the project and also for putting up with my never-ending delays.

The bulk of the content in the book is based in the research that I have been fortunate enough to be carrying out in the past years with a strong team of collaborators. I would like to thank all the former and current members of my team at SAINTS Lab at the University of Copenhagen, especially Pedram Bakhtiarifard, Rasmus Løvstad, Jonathan Wenshøj, Tong Chen, Sophia Wilson, Frederik Johansen, Julian Schön, and Dustin Wright. I would also like to thank my colleagues at the Machine Learning Section at University of Copenhagen, for the constant motivation, feedback, and enthusiasm about this book, particularly, Erik Dam, who has championed and believed in me throughout my research career. I also thank Christian Igel for his excitement about the book and his regular feedback. I also thank all my collaborators who have worked on these topics, particularly the members of the European Horizon projects, EnrichMyData, SustainML, and DataPACT. Thanks are also due to all the students I have been fortunate to know and interact with over the years, who keep me inspired. Thanks also to my reviewers Gourav Bais, Mikio Braun, Abhishek Gupta, Dan Situnayake, and Bijo Thomas.

I want to express gratitude and affection to all my friends and family, who have been supportive and proud of me always, particularly all the women in my life: my incredible mother, Chitra; my amazing sister, Rekha; my lovely niece, Nissi; my inspiring Danish godmother, Merete; and my life partner, Sneha. I would like to use this opportunity to acknowledge how amazing a researcher, person, and partner Sneha has been. I would not have been able to do this, or anything else that is remotely meaningful in my life, without her support.

CHAPTER 1 Sustainability and Artificial Intelligence

There is clear consensus among scientists about the climate trajectory of our planet—it is warming at an alarming rate. The Intergovernmental Panel on Climate Change (IPCC) in their most recent report declared, "Human activities, principally through emissions of greenhouse gases, have unequivocally caused global warming, with global surface temperature reaching 1.1°C above 1850-1900 in 2011-2020."[1] Figure 1-1 illustrates these trends of a warming planet and rising sea levels. The clear rise in global temperatures and sea levels coincides with the Industrial Revolution around the 19th century and has been accelerating in the last couple of decades.

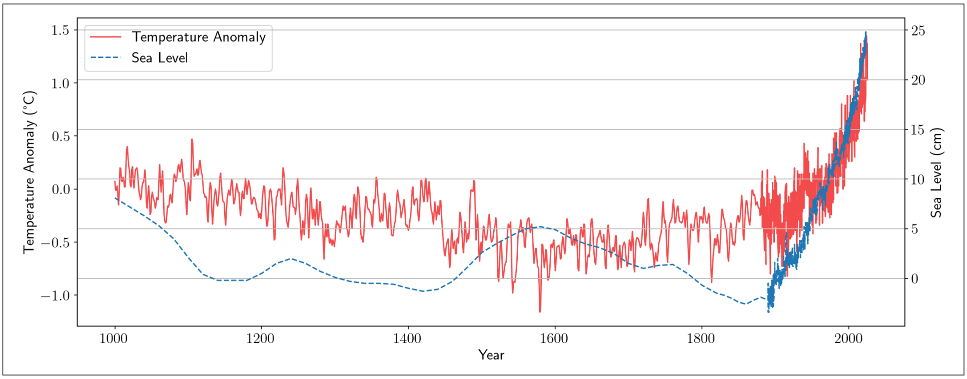


Figure 1-1. Historical trend of global temperature anomaly and sea levels. (Source: Two Degrees Institute.)

[1] IPCC, *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II, and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Geneva, Switzerland, IPCC, 2023).

The effects of planetary warming and climate change are increasingly unpredictable: erratic weather and intensifying natural disasters are becoming all too common, disproportionately harming vulnerable populations. While international agreements like the Paris Climate Accord have set emission targets, most fall short in either ambition or enforcement.[2] The result is a growing gap between what is needed and what is being done.

The Difference Between 1.5°C and 2°C Global Warming

Keeping global warming under 1.5°C, emphasized in the 2015 Paris Agreement, is considered the safer upper limit to avoid the worst impacts of climate change. At 2°C of warming, the risks grow substantially due to the nonlinear nature of climate impacts. The difference between 1.5°C and 2°C could be the difference between resilience and devastation. The planet is currently on track to exceed 1.5°C warming in the near term (by 2030).

Strategies to cope with climate change are currently categorized into two streams of efforts: climate change mitigation and climate change adaptation. *Climate change mitigation* refers to efforts of reducing or preventing greenhouse gas (GHG) emissions, aiming to limit the pace of global warming. *Climate change adaptation* involves adjusting systems, practices, and infrastructure to minimize the harm caused by the impacts of a changing climate. These are no longer future challenges; these are the defining crises of our time.

Faced with these daunting challenges, we need to draw on every tool at our disposal to advance these efforts. The scale and urgency of climate change demand a comprehensive approach, which includes policy, collective action, systems change, and, critically, technology. But for technology to play a meaningful role, it must be deeply rooted in sustainability.

We will adhere to the United Nations (UN) definition of sustainability: “meeting the needs of the present without compromising the ability of future generations to meet their own needs.”[3]

Returning to the question of using technology to combat climate change, of all the recent advances in technology, few have defined the current zeitgeist as much as artificial intelligence (AI). The rapid proliferation of AI tools across domains including science, industry, and governance has opened new possibilities for addressing

[2] The Paris Agreement was signed and adopted by 195 parties at the UN Climate Change Conference (COP2021) in December 2015.

[3] World Commission on Environment and Development, *Our Common Future* (Oxford University Press, 1987).

large-scale, complex problems, and climate change is no exception. AI can be useful to tackle the climate crisis in several ways. However, as with other technologies used to tackle climate change, AI should also be rooted in sustainability.

Making AI more sustainable is important for several reasons. The most pressing one is the large-scale resources needed to develop and use some of the recent classes of AI methods. While these large-scale AI methods are promising, their reliance on vast amounts of data, hyper-scale compute resources, massive energy consumption, and the corresponding carbon emissions are concerning as they negatively affect the environmental sustainability of AI. Resources at these scales are fiscally expensive, which also deepen the digital divide in the AI era and hence hamper the economic and social sustainability of AI.

AI/ML/DL

The terms *AI*, *machine learning* (ML), and *deep learning* (DL) are often used interchangeably, though their boundaries are neither fixed nor universally agreed upon. One useful classification frames ML as a broad class of methods designed to learn from data, encompassing everything from basic linear regression to large-scale conversational agents. DL then refers to a subset of ML methods that rely on deep neural networks as the underlying model class, ranging from simple feedforward networks to convolutional architectures and transformers. AI, in turn, subsumes both ML and DL but extends beyond them to include the broader pursuit of machine-based intelligence. This includes not only technical approaches but also the social, philosophical, and cultural dimensions of what we consider “intelligent” behavior, which is often shaped as much by industrial ambition and public imagination as by scientific consensus. Figure 1-2 visualizes the relationship between these terms as a Venn diagram.

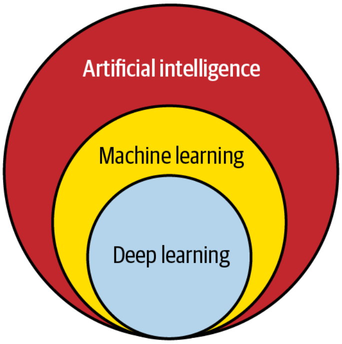


Figure 1-2. One possible classification of AI/ML/DL.

This book attempts to shine a light on the questions surrounding the sustainability of AI. It uses a lens of resource consumption, primarily energy and carbon. As AI practitioners, we can have the greatest impact by influencing the design and development of AI models to keep their resource consumption in check. While this alone will not make AI sustainable, it can be a step in the right direction.

In the remainder of this chapter, I will formalize some of the commonly used notions (including what it means to be sustainable and what we mean by AI), point out the pros and cons of pursuing resource efficiency, and outline the rest of the book. Sustainability and AI are two of the most important ideas defining our age. So, by definition, the ambitions of a book entitled *Sustainable AI* are grand. I hope this chapter will offer a glimpse into the promise of the rest of the book.

Scope of Sustainability

The most visible discussions around sustainability are focused on the environment. However, this is only part of the story as achieving true sustainability should also emphasize the economic and social aspects. We will use a real-world scenario to understand sustainability in all its complexity next.

The Samsø Sustainability Story

Step off the ferry to the Danish island, Samsø, and you will see a postcard-perfect view of Danish farms that, like any rural community, burned imported oil 25 years ago (see Figure 1-3). Then the islanders won a 1997 national contest to become Denmark's "Renewable Energy Island."[4] Within a decade they had installed cooperatively owned wind turbines and biomass district-heating plants so that today Samsø exports electricity to the mainland, and each resident averages about 3.7 tonnes (t) of carbon dioxide (CO_2) equivalent (e) of GHGs per year, which is roughly half of the Danish national average at about $7tCO_2e$.[5] For more about tCO_2e, see "GHG Emissions and Carbon Footprint" on page 68.

Denmark's energy grid is one of cleanest in the world today, primarily due to the strong investment in wind energy.

Seen through the environmental lens, the transformation is striking. The grid runs on 100% renewables, mainly wind. It has shown a carbon drop of roughly 140% from its 1997 baseline, meaning by exporting surplus renewable energy to mainland Denmark, Samsø offsets more than its total emissions.

[4] Jan Jantzen et al., "Sociotechnical Transition to Smart Energy: The Case of Samso 1997–2030," Energy 162 (August 3, 2018): 20–34.

[5] UNFCC, "Samsø: An Island Community Pointing to the Future," 2023.

improvement. Even a perfectly run server powered by renewables still sits on a global supply chain of emissions and ecological damage.[33]

The main point to bear in mind for the remainder of the book is that *efficiency is a necessary condition for sustainable AI, but it is not sufficient*. It lowers immediate emissions and proves that smarter practice is possible, but without complementary measures that look beyond resource consumption, efficiency can take us only so far in the journey toward sustainable AI.

TL;DR

So far, I have presented the context for how sustainability and AI intersect to shape our rapidly changing world. This book will explore a wide array of techniques that will help us identify resource bottlenecks in AI systems, improve upon these inefficiencies, and work toward green and ultimately sustainable AI. To do that we have to look closely at each of the complex steps involved in the algorithmic lifecycle of a DL model, as shown in Figure 1-10. I devote a chapter to each of these steps.

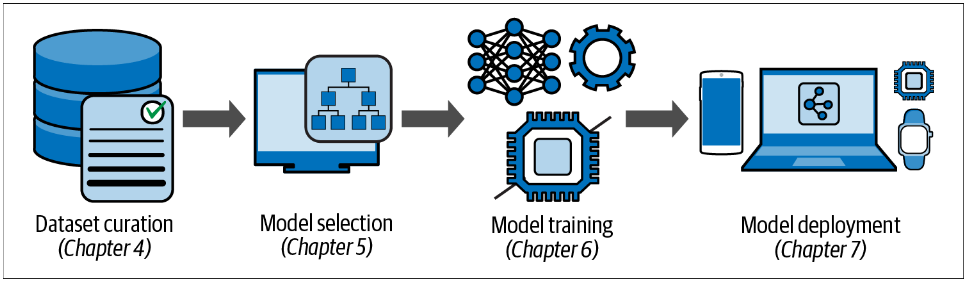


Figure 1-10. A typical AI model lifecycle.

You may already have some questions; I will try to preempt them with short answers and, in doing so, present the book's outline:

Do I have to be a machine learning expert to read the rest of the book?

No, but I am assuming you are an ML/AI practitioner with working knowledge and keen interest in these methods.

Chapter 2 aims to explore some key ML concepts to look at the foundations of AI. We will use the perspective of *representation learning*, a view of AI where algorithms learn useful features from data. This will be the basis for Chapter 2, which, while intended to be an introduction to key ML concepts, is far from comprehensive. But we will look at the relevant topics with an eye on resource

[33] Dustin Wright et al., "Efficiency Is Not Enough: A Critical Perspective of Environmentally Sustainable AI," *Communications of the ACM* 68. no. 7 (2025): 62–9.

consumption. I have not shied away from using mathematical notations, as they make the presentation of some of the concepts later easier. However, I have tried hard to not lose any readers who might not want to follow the notations with equivalent descriptions everywhere.

How do we know if the tools in this book are actually making AI more sustainable?

Sustainability is difficult to measure. But, we can use proxies that can measure resource efficiency.

Chapter 3 will introduce commonly used measures that give insight into the resource consumption of AI models. We will understand the pros and cons of measuring runtime, energy consumption, and carbon emissions. We will try some easy-to-use tools that can help us better quantify the resource consumption of AI in standardized ways. This chapter will also introduce the key concept of *AI waste*, which we will use to identify wasteful resource usage in different steps of the AI lifecycle. For instance, AI waste can manifest as training a massive model from scratch when using a smaller, pretrained model would have sufficed.

Do we always need big data to build AI models?

Some would say so. However, not all data is equally useful, and knowing this can reduce data-related costs significantly.

Availability of cheap, large-scale data has resulted in the use of more data than what might be needed to solve any given task. This abundance mindset has resulted in a lot of data-related redundancies. The process of dataset curation consists of collecting, cleaning, labeling, and preprocessing data to prepare it for training. Chapter 4 will elaborate on efficient dataset curation practices, show techniques to compress data points, and explain how to distill information in a given dataset into a few data points.

How do we decide if one model is more efficient than another?

Simpler models tend to be more efficient, but they may not perform well compared to a more complex model. Choosing models that offer the right trade-off might be the way to go.

Choosing the right class of model or its configuration is an elaborate procedure. This is dependent on the problem domain, the amount of data that is available, and the resources at our disposal. In Chapter 5 we will formalize the exploration of ML models and configurations, and then use existing techniques to efficiently explore this space so that we can identify the right model and configuration that offers the best trade-off between performance and resource consumption. The notion of Pareto optimality will be a key concept that will drive the discourse in this chapter.

Some recent models are trained for months. Is this always the case?

Yes, for the most recent class of frontier AI models such as the ones behind generative AI (GenAI). However, a broad array of very useful AI models can be trained way faster.

Model training can be one of the most resource-intensive steps. This is by design as the model parameters are updated iteratively until the model captures the desired input-output relations in the training data. Chapter 6 will cover a broad set of methods that can accelerate model training. This can be done either by "mimicking" knowledge from already trained models or by reducing the number of computations performed by modulating the number of bits being used. Recent AI models require more specialized ways to accelerate their training, which will also be discussed.

Using a model at inference does not seem expensive, compared to training it. How is model use contributing to AI being resource-intensive?

Training is done once, but models can be used millions of times. Depending on how successful a model is, its usage cost can outweigh development costs.

AI models are developed with the hope they will be used, and used widely. Once this happens, even if the energy consumption and carbon emission for a single use is small, this can become significant at scale. Chapter 7 will aim to model when the training and inference costs cross over for models. The chapter will also build upon the efficient training techniques in Chapter 6 and specialize them for deployment or inference scenarios. We'll also answer questions about how to adapt models across programming languages and hardware platforms.

As an ML/AI practitioner, I can tweak algorithms. How can this influence the resource efficiency of the AI hardware?

Underutilization of hardware is a chronic issue in AI. Many algorithmic tweaks can improve the resource efficiency of hardware.

Most recent AI models require specialized hardware for faster development. These come at huge costs (monetary and environmental) but are underutilized due to several factors. Chapter 8 will explore some easy-to-implement strategies to better utilize hardware, ranging from single computers to datacenters. That being said, there are challenges that cannot be fixed by algorithms alone. For example, e-waste is not something that can be optimized away. We will discuss these hardware-related points in depth in Chapter 8.

Building AI models involves many tedious steps. Is there a cascading effect of resource inefficiency that could percolate between steps?

Yes, certainly. AI models are developed in a long sequence of steps; poor choices in one step can blow up the resource consumption down the line.

Figures 1-6 and 1-10 show the elaborate lifecycle of recent AI models. Each of these steps requires considerable resources, and one of the main arguments of this book is that there are wasteful resource allocations everywhere. Chapter 9 points out that choices made at one step can have a huge impact down the line. Using concepts derived from systems engineering and ML operations, this chapter introduces frameworks that can be used to holistically manage and improve the resource consumption of AI models.

Let's say we do all this. Can we achieve sustainable AI?

This question lies at the heart of the matter, and the answer is necessarily complex.

CHAPTER 10
Toward Sustainable AI

For much of our history, agriculture was constrained by the natural nitrogen cycle. Usable nitrogen was scarce, and farmers relied on manure, compost, and legumes to restore soil fertility.[1] These limits kept the yields modest, and food production was tightly coupled to ecological rhythms.

With the invention of the Haber-Bosch process in the early 20th century, however, scientists unlocked the ability to produce synthetic fertilizer at industrial scale by synthesizing ammonia from atmospheric nitrogen. This had a transformative effect on food production and agriculture as crop yields soared, famine declined, and the *Green Revolution* brought this power to fields across the globe.[2]

But the efficiency gains that synthetic nitrogen unlocked came with unintended consequences. Figure 10-1 shows the historical trend of nitrous oxide (N_2O) in the atmosphere. Like other GHGs, N_2O levels rose sharply with industrialization beginning in the 19th century. What sets N_2O apart, however, is that the *vast majority* of its increase is from agriculture—primarily as a byproduct of the Haber-Bosch process, which underpins modern intensive industrial farming.[3]

[1] C. C. Delwiche, "The Nitrogen Cycle," *Scientific American*, September 1, 1970.

[2] Prabhu L. Pingali, "Green Revolution: Impacts, Limits, and the Path Ahead," *Proceedings of the National Academy of Sciences* 109, no. 31 (July 31, 2012): 12302–8.

[3] Hannah Ritchie et al., "Breakdown of Carbon Dioxide, Methane and Nitrous Oxide Emissions by Sector," *Our World in Data*, June 10, 2020.

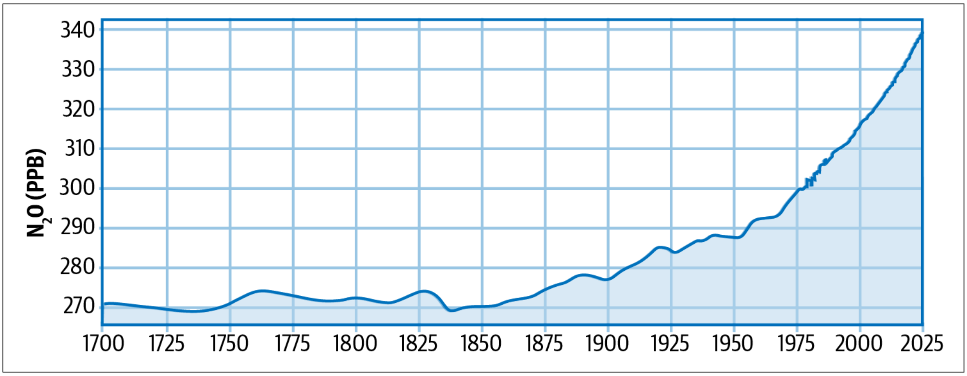


Figure 10-1. Historical trend of nitrous oxide levels measured in parts per billion (PBB). (Source: Two Degrees Institute.)

Rather than reducing environmental pressures, the rapid scaling of agriculture introduced new forms of ecological strain.[4] Now, the Green Revolution is seen as a tipping point that has made industrialized agriculture one of the largest contributors to environmental degradation.

This an example of what economists call *rebound effect*: as a technology becomes more efficient, it often also becomes cheaper or more convenient to use, which can encourage people to use it more, offsetting or even reversing the intended savings. Rebound effect is also closely related to the Jevons paradox discussed in "Energy Efficiency, Sustainable AI, and the Jevons Paradox" on page 67.

Rebound Effects and AI

The techniques presented in this book have focused on improving the resource efficiency of AI systems. We have identified various forms of AI waste ("AI Waste" on page 53), introduced the concept of environmental debt ("Environmental Debt of AI" on page 230), and outlined numerous opportunities for making AI more efficient. If resource-saving techniques make AI extremely efficient, we must still confront the rebound effects.[5] This chapter will explore how to manage this paradox toward the goal of aligning efficiency with sustainability.

As I have hinted throughout, reducing carbon emissions through resource efficiency alone has only a limited effect on the sustainability of AI. As noted in Chapter 1 ("A

[4] Harry M. Cleaver, "The Contradictions of the Green Revolution," *The American Economic Review* 62, no. 1/2 (1972): 177–86.

[5] Alexandra Sasha Luccioni et al., "From Efficiency Gains to Rebound Effects: The Problem of Jevons' Paradox in AI's Polarized Environmental Debate," in *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '25), New York, NY, 76–88.

Green Path to Sustainable AI" on page 16), we deliberately address the sustainability of AI using the narrow lens of resource efficiency. Confronting sustainability in all its dimensions—environmental, economic, and social—is a huge undertaking requiring efforts that go well beyond algorithmic improvements or efficiency hacks. I have used resource efficiency as a pragmatic entry point into the environmental dimension of AI's sustainability and to lay out the complexities that need to be addressed.

In this chapter, I will clarify why the focus on efficiency is a necessary but not a sufficient condition for achieving sustainable AI. Based on this critique, I will present suggestions based on frameworks that operate at higher abstractions (beyond algorithmic efficiency) to advance toward sustainable AI.

Efficiency Is Not Enough

The pursuit of resource efficiency is an important endeavor as it offers meaningful interventions during the development and deployment of AI systems at the level of an individual or small teams of developers.[6] However, solely obsessing over efficiency improvements by casting them as metrics to be optimized can have detrimental effects within the broader scope of sustainable AI. This is captured by the adage "When a measure becomes a target, it ceases to be a good measure," which is commonly known as Goodhart's law.[7] We need to address the broader environmental effects, economic viability, and the social impact of AI; to fully grapple with the sustainability of AI as efficiency alone is not enough.[8]

Broader Environmental Effects

In Chapter 3, we used the framework of resource pyramids (Figure 3-1) to illustrate the layered nature of resource consumption in AI models. Spanning from model complexity, we built the different levels of resource consumption leading up to the carbon footprint. As noted in the previous chapters, this carbon footprint corresponds to only the operational emissions due to the energy consumption. While we briefly touched upon the overhead due to networking costs and other IT equipment using the notion of PUE ("Estimating energy consumption" on page 64) and embodied emissions to account for the carbon emissions due to hardware manufacturing

[6] Brian R. Bartoldson et al., "Compute-Efficient Deep Learning: Algorithmic Trends and Opportunities," *Journal of Machine Learning Research* 24, no. 122 (2023): 1–7.

[7] Adrian C. Newton, "Implications of Goodhart's Law for Monitoring Global Biodiversity Loss," *Conservation Letters* 4, no. 4 (2011): 264–68.

[8] The mantra "efficiency is not enough" in the context of AI is based on a paper of the same title by Dustin Wright et al. (2025). I am the corresponding author on this paper, and several of the arguments from this work are echoed in this section.

("Embodied Emissions" on page 199), these do not fully capture the total environmental impact of AI.

To fully assess the environmental sustainability of AI, we need to look beyond operational and even embodied carbon emissions, which are mainly caused due to the energy consumed across the AI model lifecycle shown in Figure 1-6. Carbon footprint is only one factor that connects the AI model lifecycle to its environmental impact. There are multitudes of other factors that should be taken into account when discussing the true environmental impact of AI.

It is notoriously difficult to comprehensively measure the full environmental footprint of AI systems. The challenges of tracing the broader ecological consequences of AI arise not just from data scarcity but also from the sheer complexity of global supply chains and infrastructural systems.[9] For this reason, much of the existing research—including the focus of this book—has concentrated on the more readily quantifiable aspect: operational carbon emissions, or the emissions generated during the training and deployment of AI models. This approach only scratches the surface.

Beyond operational emissions, there are a number of other components that contribute to AI's environmental impact, and these are often overlooked precisely because they are harder to assess with precision. Consider, for instance, the embodied emissions associated with the manufacturing of the hardware required to run large-scale AI systems. This includes not just the energy consumed during the production of servers, GPUs, and networking equipment, but also the emissions embedded in the complex refinement processes for silicon, which are both energy-intensive and chemically hazardous.[10]

Water usage is another major factor, as large volumes are required to cool high-performance datacenters. This challenge is not easy to address directly through algorithmic improvements. While often treated as a local utility issue, the environmental costs of this water usage compound over time and are especially problematic in regions already facing water scarcity.[11]

Construction of datacenters themselves introduces yet another layer of environmental impact. The construction industry, widely recognized as one of the most carbon-intensive sectors globally, adds significantly to the AI footprint through the emissions generated in producing concrete, steel, and other building materials, as well as through land use changes.

[9] Alexandra Sasha Luccioni et al., "Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model," *Journal of Machine Learning Research* 24, no. 253 (2023): 1–15.

[10] Carole-Jean Wu et al., "Beyond Efficiency: Scaling AI Sustainably," *arXiv.org*, June 22, 2024.

[11] Pengfei Li et al., "Making AI Less 'Thirsty,'" *Communications of the ACM* 68, no. 7 (2025): 54–61.

Index

About the Author

Raghavendra (Raghav) Selvan is an assistant professor at the University of Copenhagen. His research spans sustainable machine learning, machine learning for sciences, medical image analysis, and graph neural networks. He holds a PhD from the University of Copenhagen and is affiliated with Pioneer Center for AI (Denmark) and the pan-European AI network ELLIS. Raghav was born in Bangalore, India.

Colophon

The animal on the cover of *Sustainable AI* is a barn swallow (*Hirundo rustica*), the most common swallow. Barn swallows have six subspecies and can be found on every continent around the world. They thrive in open land, such as pasture, meadow, and savanna, and build nests in human constructions.

Barn swallows are insectivores, endearing them to their human neighbors, and are known for hunting while in flight. Their long, pointed wings, forked tails, and slender bodies enable them to glide for long periods. The distinctive tail streamers, backs, and breast bands are steely blue, and the underparts are white or off-white. Above and below their short, wide beaks, barn swallows have rusty red feathers.

Many of the animals on O'Reilly covers are endangered; all of them are important to the world.

The cover illustration is by Monica Kamsvaag, based on a black-and-white engraving from *Lydekker's Royal Natural History*. The series design is by Edie Freedman, Ellie Volckhausen, and Karen Montgomery. The cover fonts are Gilroy Semibold and Guardian Sans. The text font is Adobe Minion Pro; the heading font is Adobe Myriad Condensed; and the code font is Dalton Maag's Ubuntu Mono.