

先輩データサイエンティスト からの指南書

実務で生き抜くためのエンジニアリングスキル

付録



参考文献

- 『Clean Architecture 達人に学ぶソフトウェアの構造と設計』著：Robert C. Martin、翻訳：角征典、高木正弘、2018年刊行、KADOKAWA
- 『Clean Code アジャイルソフトウェア達人の技』著：Robert C. Martin、翻訳：花井志生、2017年刊行、KADOKAWA
- 『Googleのソフトウェアエンジニアリング —持続可能なプログラミングを支える技術、文化、プロセス』著：Titus Winters、Tom Manshreck、Hyrum Wright、監訳：竹辺靖昭、翻訳：久富木隆一、2021年刊行、オライリージャパン
- 『VTuberサプーが教える！ Python 初心者のコード／プロのコード』著：サプー、2024年刊行、技術評論社
- 『アジャイルチームによる目標づくりガイドブック OKRを機能させ成果に繋げるためのアプローチ』著：小田中育生、2024年刊行、翔泳社
- 『オブジェクト指向でなぜつくるのか 第3版 知っておきたいOOP、設計、アジャイル開発の基礎知識』著：平澤章、2021年刊行、日経BP
- 『コード×AI —ソフトウェア開発者のための生成AI実践入門』著：服部佑樹、2024年刊行、技術評論社
- 『ソフトウェアアーキテクチャの基礎 —エンジニアリングに基づく体系的アプローチ』著：Mark Richards、Neal Ford、翻訳：島田浩二、2022年刊行、オライリージャパン
- 『ハイパーモダンPython —信頼性の高いワークフローを構築するモダンテクニック』著：Claudio Jolowicz、翻訳：嶋田健志、鈴木駿、2024年刊行、オライリージャパン
- 『プリンスipl オブ プログラミング 3年目までに身につけたい 一生役立つ101の原理原則』著：上田勲、2016年刊行、秀和システム新社
- 『リファクタリング 既存のコードを安全に改善する (第2版)』著：Martin Fowler、翻訳：児玉公信、友野晶夫、平澤章、梅澤真史、2019年刊行、オーム社
- 『リーダブルコード —より良いコードを書くためのシンプルで実践的なテクニック』著：Dustin Boswell、Trevor Foucher、解説：須藤功平、翻訳：角征典、2012年刊行、オライリージャパン

- 『良いコード／悪いコードで学ぶ設計入門 ―保守しやすい 成長し続けるコードの書き方』著：仙場大也、2022年刊行、技術評論社
- 『良いコードを書く技術 ―読みやすく保守しやすいプログラミング作法』著：縣俊貴、2011年刊行、技術評論社
- 『達人プログラマー(第2版)：熟達に向けたあなたの旅』著：Andrew Hunt、David Thomas、翻訳：村上雅章、2020年刊行、オーム社
- 『研鑽Rubyプログラミング 実践的なコードのための原則とトレードオフ』著：Jeremy Evans、翻訳：角谷信太郎、2023年刊行、ラムダノート
- 『データマネジメント知識体系ガイド 第二版 改定新版』著：DAMA International、監修・翻訳：DAMA 日本支部 Metafind コンサルティング株式会社、2024年刊行、日経BP
- 『実践的データ基盤への処方箋 ～ビジネス価値創出のためのデータ・システム・ヒトのノウハウ』著：ゆずたそ、渡部徹太郎、伊藤徹郎、2021年刊行、技術評論社



オススメ書籍

本章で学んだ内容をさらに深めたい方には、参考文献の書籍に加えて、以下の書籍をおすすめします（特におすすめの書籍は、参考文献と重複して掲載しています）。

Pythonの上級者になりたい方

- 『VTuberサプーが教える！ Python 初心者のコード／プロのコード』著：サプー、2024年刊行、技術評論社

動けばいいから抜け出そうという本書と類似したコンセプトで書かれた本です。Pythonのより良いコードの書き方を基礎的な部分から学びたい方におすすめです。

- 『Effective Python 第2版 ―Pythonプログラムを改良する90項目』著：Brett Slatkin、監修：石本敦夫、翻訳：黒川利明、2020年刊行、オライリージャパン

Pythonのコーディングにおける効果的なテクニックやTipsが列挙されており、上級者にもおすすめの内容です。各項目ごとに独立しているため、読みやすいのも特徴です。

- 『ロバストPython ―クリーンで保守しやすいコードを書く』著：Patrick Viafore、監修：鈴木駿、翻訳：長尾高弘、2023年刊行、オライリージャパン

型を用いて意図を明確に表現する方法について学びたい方におすすめです。Pythonの型システムの奥深さを体験することができます。

コードの書き方・設計

- 『リーダブルコード ーより良いコードを書くためのシンプルで実践的なテクニック』著：Dustin Boswell、Trevor Foucher、解説：須藤功平、翻訳：角征典、2012年刊行、オライリージャパン

言わずと知れたコードの書き方の名著です。読みやすいコードの書き方が具体的に解説されており、エンジニアやデータサイエンティストなら一度は読んでおきたい1冊です。

- 『改訂新版 良いコード／悪いコードで学ぶ設計入門 ー保守しやすい 成長し続けるコードの書き方』著：仙場大也、2024年刊行、技術評論社

オブジェクト指向を中心とした良いコードの設計について学べる良書です。初級レベルから中級レベルへステップアップするのに最適です。

- 『Clean Architecture 達人に学ぶソフトウェアの構造と設計』著：Robert C. Martin、翻訳：角征典、高木正弘、2018年刊行、KADOKAWA

SOLID原則をはじめとした設計の原則やアーキテクチャの基盤となる考え方について学びたい方におすすめです。内容は上級者向けですが、読むたびに新たな学びが得られる、何度も読み返したくなる名著です。

- 『アーキテクトの教科書 価値を生むソフトウェアのアーキテクチャ構築』著：米久保剛、2024年刊行、翔泳社

ソフトウェア開発におけるノウハウや考え方が濃縮された一冊です。ソフトウェア設計に加えて、システム全体のアーキテクチャの設計や各種開発手法について学ぶことができます。

デザインパターン

- 『Java 言語で学ぶデザインパターン入門第3版』著：結城浩、2021年刊行、SBクリエイティブ

デザインパターンについて、非常に詳しく解説されている本です。Java 言語での実装になりますが、デザインパターンの考え方や実践例を学ぶのに最適です。

- 『実践 Python 3』著：Mark Summerfield、翻訳：斎藤康毅、2015年刊行、オライリージャパン

デザインパターンについて、Pythonで学ぶことができます。『Java 言語で学ぶデザインパターン入門第3版』読了後に読むのもおすすめです。

- 『48. GoF デザインパターンと DI (前編)』fukabori.fm、2021年、URL：
<https://fukabori.fm/episode/48>

GoF デザインパターンが考え出された背景や、現代の開発におけるデザインパターンの位置づけについて解説されています。ポッドキャストなので、通勤や家事の合間に聞くこともできます。

テスト

- 『テスト駆動Python 第2版』著：Brian Okken、翻訳・監修：株式会社クワイプ、監修：安井力、2022年刊行、翔泳社

pytestを使ってPythonでテストを書く方法について書かれた本です。この本を読むことで、pytestの各種機能を使いこなすことができるようになります。

- 『**単体テストの考え方/使い方**』著：Vladimir Khorikov、翻訳：須田智之、2022年刊行、マイナビ出版

単体テストのアンチパターンとベストプラクティスについて学ぶことができます。モックを利用したテストなど高度な内容も含まれています。

コードレビュー

- 『**プリンシプル オブ プログラミング 3年目までに身につけたい 一生役立つ 101の原理原則**』著：上田勲、2016年刊行、秀和システム新社

コードを書く際だけでなく、レビューの指針にもなります。まずは取り入れたい原理・原則をピックアップして実践してみるのもよいでしょう。

- 『**アジャイルチームによる目標づくりガイドブック OKRを機能させ成果に繋げるためのアプローチ**』著：小田中育生、2024年刊行、翔泳社

本書のスコープからは外れますが、コードレビューを成立させるには土台となるチームビルディングが重要です。強いチームをつくるための理論や手法がわかりやすくまとまっている1冊です。

データマネジメント

- 『**データマネジメントが30分でわかる本**』著・編集：ゆずたそ、著：はせりよ、2020年刊行、NextPublishing Authors Press

データマネジメントに興味を持たれた方に一歩目としておすすめする書籍です。DMBOKの内容をコンパクトにまとめており、データマネジメントの全体像を把握するのに役立ちます。

- 『データマネジメント知識体系ガイド 第二版 改定新版』著：DAMA International、監修・翻訳：DAMA 日本支部 Metafind コンサルティング株式会社、2024年刊行、日経BP

データマネジメント知識体系ガイドは広範囲な内容をカバーしています。一方で、すべてを読むのには腰を据えて取り組む必要があるため、本章と特に関連の深い箇所のみをピックアップしておすすめします。

- ・第3章：データガバナンス
- ・第13章：データの品質
- ・第16章：データマネジメント組織と役割期待

Streamlit

- 『作ってわかる [入門] Streamlit ～Pythonによる実践Webサービス開発』著：豊沢聡、2025年刊行、技術評論社

サンプルのアプリケーションを作りながらStreamlitの機能を深掘りしていく書籍です。実際にアプリケーションを作ることで、Streamlitの魅力や活用イメージを具体的に理解できます。これからStreamlitを利用する方におすすめです。

- 『Streamlit入門 Pythonで学ぶデータ可視化&アプリ開発ガイド』著：山口歩夢、2024年刊行、インプレス NextPublishing

Streamlitの機能や引数などの使い方を網羅的に紹介しています。本書や『作ってわかる [入門] Streamlit』などを読んで基本的な機能やその魅力を理解した後で、より網羅的に知りたい場合や知識を整理したい場合におすすめです。



発展技術紹介

本書の内容をマスターした読者がさらにスキルを磨くために学ぶべき技術や開発手法を紹介します。

これらの内容は、機械学習エンジニアやMLOpsエンジニアが学んでいる内容であり、これらの職種の人と協働する方やデータサイエンティストとしてエンジニアリングスキルを強化したい方に向けた内容です。

Docker

Dockerに代表されるコンテナ技術は、現代の開発において欠かせない基盤技術です。Dockerを使用することで、環境の再現性を高め、開発環境の構築を効率化できます。さらに、後述するクラウドやパイプライン、CI/CDといった技術の基盤としても利用されており、これらの技術を理解する上で不可欠なスキルです。

- 『仕組みと使い方がわかる Docker & Kubernetesのきほんのきほん』著：小笠原種高、2021年刊行、マイナビ出版

Dockerの簡単な仕組みや基本コマンド、定義ファイルの書き方を初学者向けに解説した書籍です。本書のリポジトリの環境構築でもDockerおよびDocker Composeを使用していますので、この書籍を読んだ後に、本書のリポジトリ構成を見直してみるとさらに理解が深まるでしょう。なお、Kubernetesに関する内容は発展的なトピックなので読み飛ばしても問題ありません。

クラウド

クラウドとは、サーバやストレージなどのインフラを自社で保有することなく、インターネットを通じて利用することができるサービスのことです。クラウドは現在のインターネット社会を支える重要な技術であり、実際世の中のサービスの多くは、クラウドを活用して全世界のユーザーに提供されています。

- 『図解即戦力 Amazon Web Servicesのしくみと技術がこれ1冊でしっかりわかる教科書』著：小笠原種高、2025年刊行、技術評論社

クラウドの基礎知識とAWSにおける代表的なサービスについて、一通りの知識を身につけることができます。AWSは最も普及しているクラウドサービスの一つであり、クラウドを使ったことがない方が一番はじめに読む本としておすすめです。

- 『AWSの基本・仕組み・重要用語が全部わかる教科書』著：川畑光平、菊地貴彰、真中俊輝、2022年刊行、SBクリエイティブ

AWSの各種サービスについて詳しく解説されている書籍です。AWS自体は利用したことがある方が、新たなサービスを学ぶ際や、体系的な知見を得るために読むとよいでしょう。非常にボリュームな書籍なので、学びたいサービス部分についてのみ読むのも良いでしょう。

- 『図解即戦力 Google Cloudのしくみと技術がこれ1冊でしっかりわかる教科書 [改訂2版]』著：株式会社grasys、大沼翔、西岡典生、2024年刊行、技術評論社

クラウドの基礎知識とGCPにおける代表的なサービスについて、一通りの知識を身につけることができます。Google Cloudは特にAIや機械学習分野に強みを持ち、使い勝手も良いため、機械学習用途でよく利用されるクラウドサービスの一つです。

- 各クラウドの公式ドキュメント

クラウドサービスは、公式ドキュメントが充実していることが多いので、公式ドキュメントを見ながらサービスを利用することもおすすめです。また、クラウドサービスは頻繁に更新されるため、公式ドキュメントを見ることで最新の情報を得ることができます。

クラウドサービスの学習には、実際にサービスを利用することが最も効果的です。クラウドサービスは、無料枠が用意されていることが多いので、まずは無料枠を利用してサービスを利用してみingことをおすすめします。なお、クラウドサービスは料金体系が複雑で、誤って高額な料金になることがあるため、はじめに予算アラートを設定することを強く推奨します。

DevOps、CI/CD

DevOpsは、開発（Development）と運用（Operations）を統合し一連のプロセスを効率化する手法です。これにより、迅速な価値提供やシステムの品質の向上、プロジェクトの生産性の向上が期待できます。

DevOpsのプラクティスの中でも、特にデータサイエンティストが取り入れやすいのがCI/CDです。CI/CDは、継続的インテグレーション（Continuous Integration）と継続的デリバリー（Continuous Delivery）の略称です。これを実践することで、品質の向上と開発プロセスの効率化を実現することができます。

- 『The DevOps ハンドブック 理論・原則・実践のすべて』著：ジーン・キム、ジェズ・ハンブル、パトリック・ボア、ジョン・ウィリス、翻訳：長尾高弘、2017年刊行、日経BP

DevOpsのバイブルともいえる一冊。DevOpsの思想・原則から実践まで書かれており、システムの運用について考える良いきっかけになる本です。

- 『入門 継続的デリバリー ―テストからリリースまでを安全に自動化するソフトウェアデリバリーのプロセス』著：Christie Wilson、序文：Jez Humble、Eric Brewer、監訳：渡辺光一、翻訳：井上渉、庄司祐馬、遠山雄二、2024年刊行、オライリー・ジャパン

CI/CDを意義とその実践方法について解説されている書籍です。ストーリー形式で書かれており、挿し絵も多いため、非常に読みやすい内容が特徴です。

- 『GitHub CI/CD実践ガイド —持続可能なソフトウェア開発を支えるGitHub Actionsの設計と運用』著：野村友規、2024年刊行、技術評論社

GitHubを用いたCI/CDの実践方法について解説されているガイドブックです。GitHub Actionsを使ったCI/CDの実装方法やTipsが詳しく解説されているので、実務でGitHubを使っている方におすすめです。

MLOps

MLOpsは、機械学習システム特有の開発や運用における課題に対処するための手法のことです。MLOpsを実践することで、機械学習システムの開発や運用のリスクを回避し、品質向上や運用コストの削減を実現することができます。MLOpsは比較的新しい分野であり、まだその範囲や方法論が確立されているわけではないため、さまざまなソースを参考にしながら学ぶことが重要です。

- **Hidden Technical Debt in Machine Learning Systems** URL : https://proceedings.neurips.cc/paper_files/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf

機械学習システムの課題について提起した著名な論文。MLOpsの先駆けとなる論文であり、機械学習に携わる人にとっては必読の一本です。

- 『AIエンジニアのための機械学習システムデザインパターン』著：湊井雄介、2021年刊行、翔泳社

機械学習システムの構築におけるノウハウがデザインパターンとして整理されており、体系的に学ぶことができる書籍です。

- 『信頼性の高い機械学習 —SRE原則を活用したMLOps』著：Cathy Chen、Niall Richard Murphy、Kranti Parisa、D. Sculley、Todd Underwood、翻訳：井伊篤彦、張凡、樋口千洋、2024年刊行、オライリー・ジャパン

機械学習システムのライフサイクル全体にわたる信頼性を高めるための手法について解説されている書籍です。特に、機械学習の倫理的側面や安定した運用方法について重点的に学びたい方におすすめです。

- 『事例でわかるMLOps 機械学習の成果をスケールさせる処方箋』編著：杉山阿聖、太田満久、久井裕貴、2024年刊行、講談社

実践事例が豊富に紹介されているのが特徴の書籍です。実務でのMLOpsの活用方法について学ぶことができるため、これから実務でMLOpsを活用する方にとっては、非常に参考になる一冊です。

ドメイン駆動開発 (DDD)

ドメイン駆動開発 (DDD) は、ソフトウェア開発手法の一つで、ドメイン（ソフトウェアを使って問題解決しようとしている領域）を中心に設計することでソフトウェアの価値を高めることを目指す手法です。

- 『ドメイン駆動設計 モデリング/実装ガイド』著：松岡幸一郎、2020年刊行

DDDについて非常にわかりやすく解説されている書籍。コンパクトな内容でありながら、DDDの目的や考え方についての確に説明されているため、DDDについて初めて学ぶ方に読んでほしい一冊です。

- 『実践ドメイン駆動設計』著：ヴォーン・ヴァーノン、翻訳：高木正弘、2015年刊行、翔泳社

DDDについて深く掘り下げて解説した一冊。『ドメイン駆動設計 モデリング/実装ガイド』を読んだ後に、さらに深く学びたい方におすすめです。500ページ以上と分量が多いため、他の書籍で理解が進まなかったトピックだけ読むのがおすすめです。

テスト駆動開発 (TDD)

テスト駆動開発 (TDD) は、テストの作成、実装、リファクタリングというサイクルを繰り返しながらソフトウェアを開発する手法です。TDDを実践することで、実装の目的が明確化されるため開発者は目の前の課題解決に集中しやすくなります。また、テストが自動的にソフトウェアの品質を保証するため、堅牢で保守性の高いコードを効率的に開発することが可能です。

- 『テスト駆動開発』著：Kent Beck、翻訳：和田卓人、2017年刊行、オーム社

TDDの基本的な考え方や実践方法、そのメリットについて解説されている書籍です。よく誤解されがちなTDDについて、正しい理解を深めるために読んでほしい、信頼できる一冊です。

アジャイル

アジャイル開発は、変化への対応力や顧客満足度を重視するソフトウェア開発手法の一つです。柔軟性を活かし、効率的かつ迅速にプロジェクトを進めることができます。

- 『みんなでアジャイル ―変化に対応できる顧客中心組織のつくりかた』著：Matt LeMay、翻訳：吉羽龍太郎、永瀬美穂、原田騎郎、有野雅士、まえがき：及川卓也、2020年刊行、オライリージャパン

ともすれば方法論に陥りがちなアジャイルについてその本質をわかりやすく解説した書籍。アジャイルについて初めて学ぶ方や、アジャイルについてなんとなくしか理解できていない方におすすめです。