Welcome to Building LLM Apps for Data Scientists and Software Engineers: From First Principles

By Hugo Bowne-Anderson, Independent Data & AI Scientist

To help you hit the ground running, we've compiled a list of optional resources. These cover Python basics, deep learning, evaluation, and the broader field of LLM-powered systems. Feel free to check them out based on your interests and needs—nothing here is mandatory.

📚 Optional Pre-Course Resources

1. Python Catch-Up: *Python Data Science Handbook* by Jake VanderPlas
An excellent resource for brushing up on Python libraries like pandas, NumPy, and scikit-learn—fundamentals for building data-powered applications.
👉 Read it here (a primer on environments for reproducible software will also help)

2. Deep Learning Exploration: *Fastbook* by Jeremy Howard and Sylvain Gugger
A fantastic introduction to deep learning with PyTorch. While we won't cover model training in this course, this is ideal for those looking to explore deep learning further.
👉 Explore Fastbook

3. MLOps Context: *MLOps vs DevOps: Why Data Makes It Different* by Hugo Bowne-Anderson and Ville Tuulos
This article explains how managing ML systems differs from traditional software systems—a key concept for building robust LLM-powered applications.
👉 Read the article

4. Evaluation Frameworks: *LLM Evaluation and Virtuous Cycle* by Hamel Husain
A practical guide to systematically evaluating LLM outputs—central to building reliable systems.
👉 Read the blog post
👉 Watch our discussion on YouTube

5. Iterative Workflows: *AI Engineering Flywheel* by Shreya Shankar
An engineering-focused breakdown of improving AI systems iteratively through logging, evaluation, and continuous refinement.
👉 Read Shreya's post

6. Prompt Engineering: *The Prompt Report* by Sander Schulhoff and many others!
A systematic overview of prompting techniques, organizing methods and strategies to help you effectively communicate with LLMs.
👉 Read the report

7. Scaling GenAI Platforms: *Generative AI Platforms* by Chip Huyen
A strategic overview of building scalable generative AI systems, explaining key components and trade-offs.
👉 Read Chip's post

8. Applied LLM Systems: *Applied-LLMs.org*
A curated hub for tools and case studies on deploying LLM systems in production. While some content is advanced, it's a great resource for deeper exploration.

- 👉 Visit Applied-LLMs.org
- **Watch the livestream:** I hosted a 3-hour conversation with the authors, where we broke down real-world approaches to building LLM systems, explored practical workflows, and discussed key challenges.
  👉 Watch here
- **Listen to the podcast episodes:** If you prefer audio, the livestream is also available as two podcast episodes. These provide the same insights but allow you to follow along on the go:
  - Episode 1
  - Episode 2

9. Foundational Guide: *Generative AI Guidebook* by Ravin Kumar
A beginner-friendly guide covering foundational concepts and practical techniques for building with generative AI.
👉 Read the guidebook

10. AI Agent Design: *Building Effective Agents* by Anthropic
This research paper explores the principles behind designing robust, capable AI agents. It covers approaches to agentic workflows, alignment, and building systems that can perform tasks autonomously and reliably—a critical component in scaling generative AI applications.
👉 Read the paper

11. Deep Dive on Embeddings: *What Are Embeddings?* by Vicki Boykis
A comprehensive survey paper exploring embeddings, their evolution, and practical use cases in machine learning. This resource dives into the critical role embeddings play in modern AI

applications.
👉 Read the PDF

🌐 Explore Further

Many of these authors—Hamel Husain, Shreya Shankar, Eugene Yan, Chip Huyen, Ravin Kumar, Vicki Boykis, and Sander Schulhoff—share additional insights on their blogs. Exploring their work will give you broader perspectives and help you dive deeper into building with LLMs.

Note: These resources are completely optional. They're here to support you, but there's no pressure to review them before starting the course.