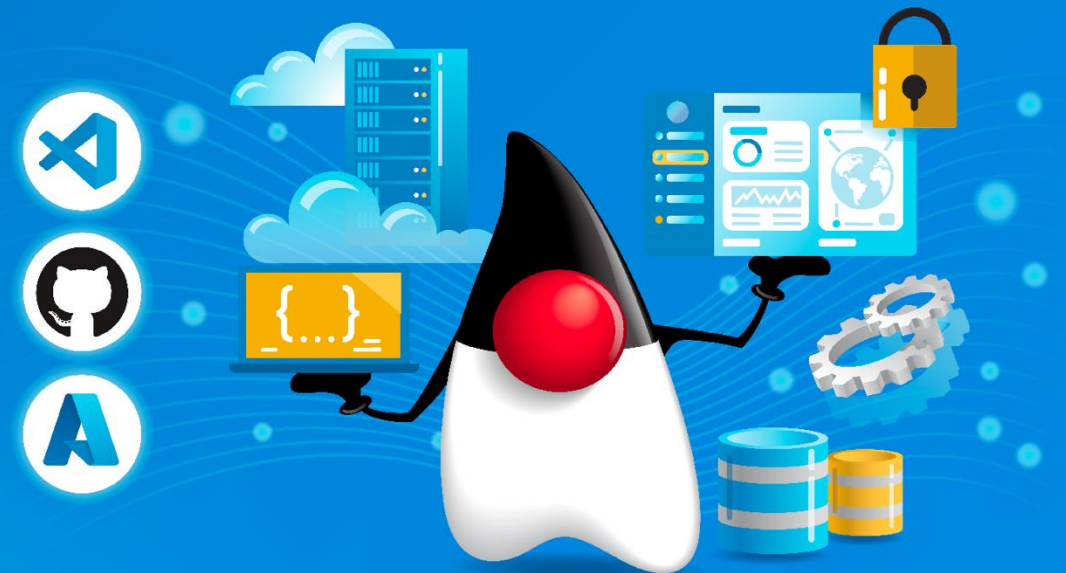




# Microsoft Java Developer Conference 2024

Code. Cloud. Community.



# Paving a Painless Path to Production AI in your Java Applications



# Paving a Painless Path to Production AI in your Java Applications

Mark Heckler

Principal Cloud Advocate, Java/JVM Languages

[markheckler@microsoft.com](mailto:markheckler@microsoft.com)

[mark@thehecklers.com](mailto:mark@thehecklers.com)

[@mkheck](#)

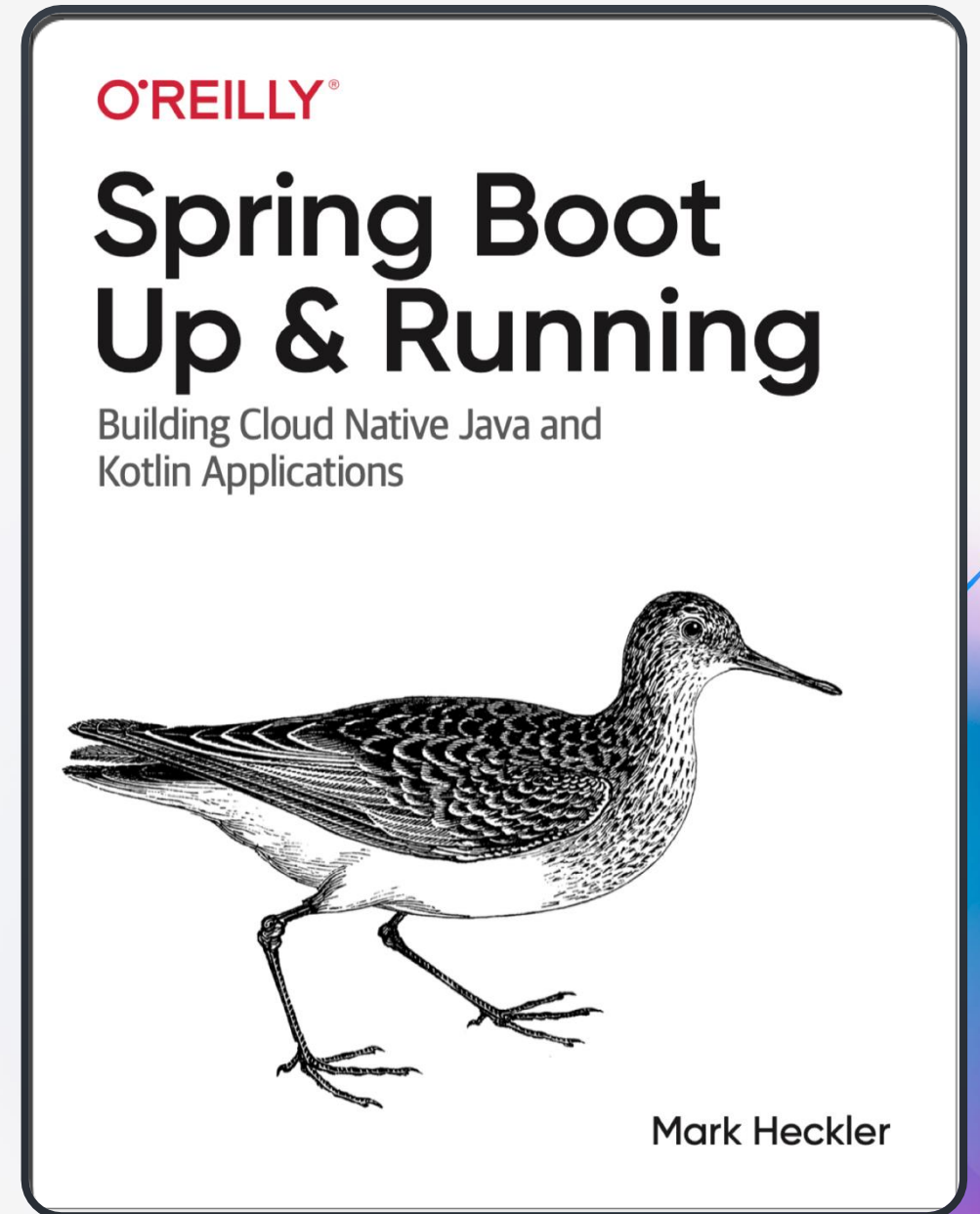
# Who am I?

- Architect & Developer
- Advocate
- Author
- Java Champion, Rockstar
- Kotlin Developer Expert
- Pilot ✈️



# The book

<https://bit.ly/springbootbook>  
@springbootbook



# What this session is NOT

It isn't an exhaustive introduction of AI, AI concepts, or AI tools. It isn't a deep dive into the details of how to use the tools. It isn't a deep dive into the details of how the tools work. It isn't a deep dive into the details of how to architect your applications to use AI. It isn't a deep dive into the details of how to deploy your applications that use AI.


We have only 30 minutes! So...

# What this session IS

- A fast-paced, selective introduction of key AI concepts
- A quick review of tools at your disposal for building AI-enabled apps
- A *live-coding adventure*<sup>™</sup> demonstrating those concepts using Spring Boot, Java, and Azure OpenAI

NOTE: Concepts are portable, tools “rhyme”

# What's the plan?

- Concepts
  - Translate concepts to reality
  - See it in action
  - Continue conversation after
- 



# Concepts

- Models\*
  - Text
  - Images
  - Audio
- Generative Pre-trained Transformer (GPT)
- Tokens, tokenization
- Embeddings
- Retrieval Augmented Generation (RAG)

\* *Currently. This is evolving.*

# Translating Concepts to Reality

- ChatGPT for text (and others, e.g. GitHub Copilot)
- DALL-E for images
- API-based models for audio
- Developers access the underlying models
- Most use cases are currently textual
- Most models expose APIs
- Useful abstractions add significant value to development

# VS Code

- Cross platform, multiple languages
- Great Java & Spring Boot plugins
- Great AI plugins (as you might expect)
- <https://code.visualstudio.com>

# GitHub Copilot

- Copilot
  - Read my comments
  - ~~Read my mind (code with me)~~
- Copilot Chat: chat with me
- <https://github.com/features/copilot>

**Also available for IntelliJ  
IDEA!**

# Azure Spring Apps, Azure Container Apps, etc.

- Natural and intelligent platform evolution for developers
- Leverages dominant framework for Java (ASA)
- Leverages containerization for portability and versatility (ACA)
- Easy to deploy, manage, monitor, scale apps
- <https://aka.ms/azurespringapps>
- <https://aka.ms/azurecontainerapps>



# Let's code!



# Thank you

Mark Heckler

markheckler@microsoft.com

mark@thehecklers.com

@springbootbook

@mkheck

<https://github.com/mkheck/production-ai>