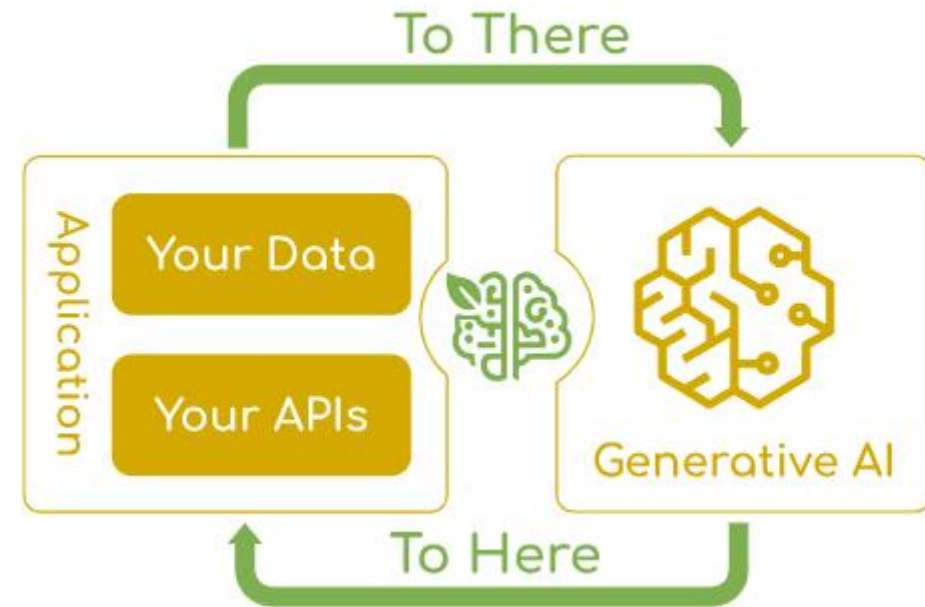The Spring AI project aims to streamline the development of applications that incorporate artificial intelligence functionality without unnecessary complexity.

The project draws inspiration from notable Python projects, such as LangChain and LlamaIndex, but Spring AI is not a direct port of those projects. The project was founded with the belief that the next wave of Generative AI applications will not be only for Python developers but will be ubiquitous across many programming languages.

Spring AI provides abstractions that serve as the foundation for developing AI applications. These abstractions have multiple implementations, enabling easy component swapping with minimal code changes.

Spring AI provides the following features:

- Portable API support across AI providers for Chat, text-to-image, and Embedding models. Both synchronous and streaming API options are supported. Access to model-specific features is also available.

- Support for all major AI Model providers such as Anthropic, OpenAI, Microsoft, Amazon, Google, and Ollama. Supported model types include:

    - Chat Completion

    - Embedding

    - Text to Image

    - Audio Transcription

    - Text to Speech

    - Moderation