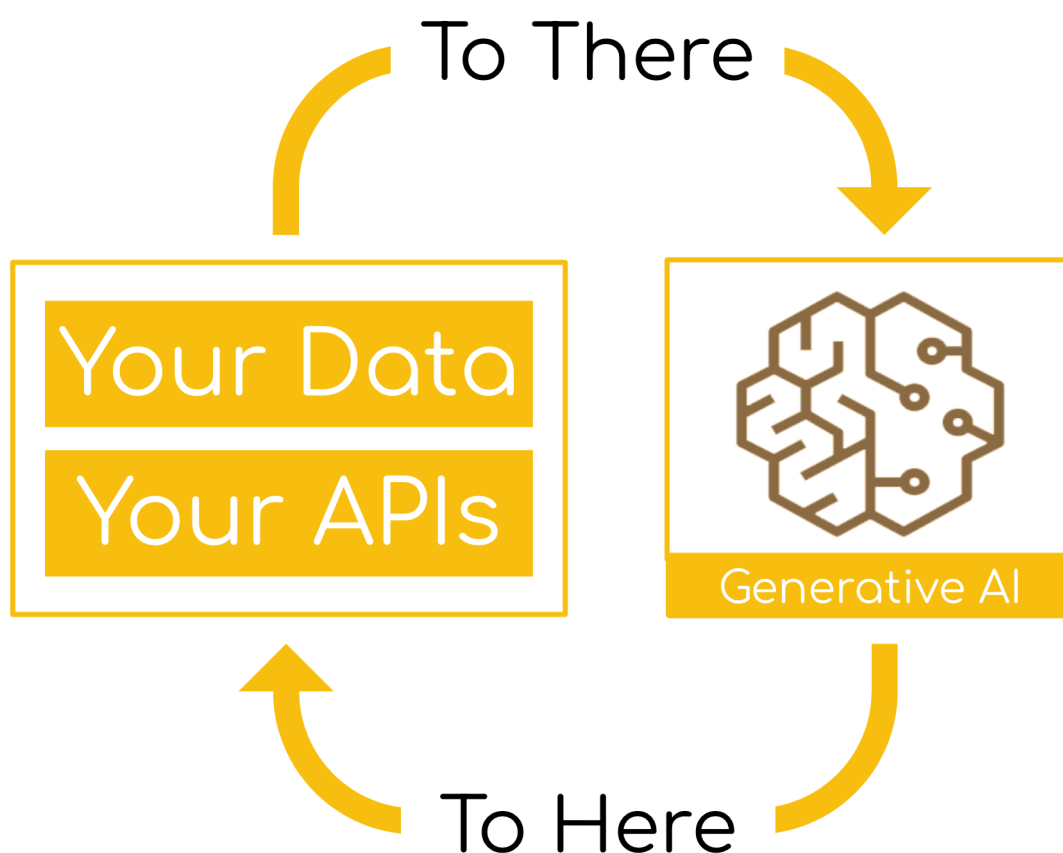


# Day1 - Spring AI 簡介

Spring AI 在2024年2月推出了0.8版本，不到半年就迅速更新到了 1.0.0-M1，1.0.0-M2 也即將推出，如同 AI 的發展迅速。實際使用後，會發現它與 **LangChain4j** 有幾分相似。讓我們來看看 Spring AI 官網是如何詮釋

Spring AI 創建的目的在簡化開發 AI 功能的應用程式，同時避免不必要的複雜性。專案汲取了 LangChain 和 LlamaIndex 等知名 Python 專案的靈感，但 Spring AI 並不是這些專案的直接移植。我們相信，下一波生成式人工智慧應用程式不僅適用於 Python 開發人員，還將廣泛適用於許多程式設計語言。

Spring AI 的核心在於解決 AI 整合的根本挑戰，即將 **企業數據 / API** 與 **AI** 模型串聯起來。



(以上由Spring AI官網翻譯而來)

原來 Spring AI 是跟 LangChain 致敬，Spring 能發展到這麼龐大，最關鍵的就是其整合以及簡化的能力，面對這麼龐大的模型以及種類，雖然跟 Python 比起來起步較慢，不過能與 Spring 其他框架整合在一起，讓憋了好久的 Java 開發人員也能大展身手了

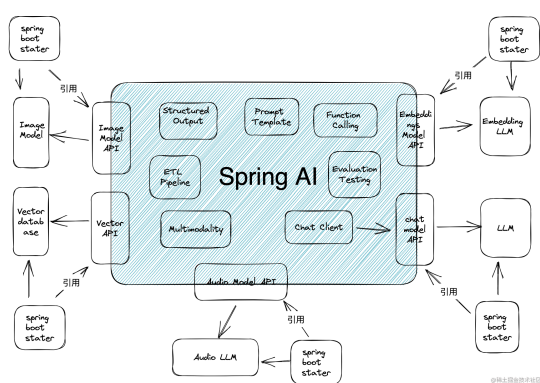
既然是參考 LangChain 的概念，Java開發人員該使用 LangChain4j 還是 Spring AI 做為開發框架？凱文大叔幫大家分析以下幾點，大家可以自行評估

	Spring AI	LangChain4j
推出時間	2024年二月推出0.8版	2023年六月推出0.1版
適用Java版本	JDK 17+	JDK 8+
更新頻率	三個月時間推出四個版本，1.0.0-M2也即將推出	截至目前已有33個版本
功能性	整合性較佳，讓不同大語言模型有相似的開發程序，也容易與 Spring 其他框架整合	個別功能較強，部分功能是針對不同大語言模型特性所開發
整合性	能輕易與 Spring 家族整合，能自動配置	雖然也能在 Spring 中使用，不過還是沒 Spring 自家的整合性高
相關資料	官網資料豐富完整	發展時間較長，社群會員眾多

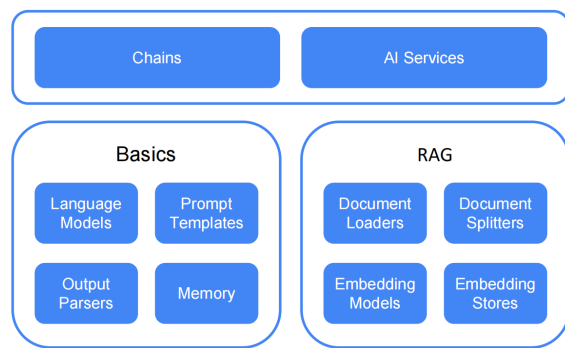
該如何選擇凱文大叔有以下幾點建議

- 如果公司還在使用 Java 8，LangChain4j 是不二選擇
- 如果公司已使用 Spring Boot 3 以上（），當然就直接使用 Spring AI
- 如果使用 Java 17+ 卻沒使用 Spring Boot 框架呢？建議你趕快學 Spring Boot :P
- 想嘗鮮或是單純開發 AI 的程式可使用 LangChain4j，若要將 AI 結合企業內部其他資訊系統則建議使用 Spring AI，因為開發企業使用的程式往往需要更高的穩定性以及精準的內容，純 AI 程式則恰恰相反

下面兩張圖大家可以感受一下兩個 AI 框架設計的理念，Spring AI 就跟 Spring 家族設計理念一樣，以整合為主，而 LangChain4j 則是各模組功能較強



取自稀土掘金



取自LangChain4j

有一點要特別說明 Spring AI 雖然版本更新較慢，不過 7/25 Ollama 才宣布支援 Tool，Spring 隔一天也宣布支援 Ollama Tool，可見得 Spring AI 的抽象可以相容不同模型

## 打造企業 RAG 知識庫 內容將分為四個章節

- ▼ 認識 Spring AI：這個章節介紹如何取得 API Key、建立 Spring starter 專案、參數配置、如何與 AI 對話以及讓對話更流暢的流式輸出
- ▼ 個性化 ChatBot：這個章節會更進一步設定對話的細節，包含系統提示詞、系統人設、提示詞範本、結構化輸出、Function Call
- ▼ 讓 ChatBot 不在金魚腦：ChatGPT 除了理解及生成，還有一個很重要的功能就是記憶，這個章節會先介紹記憶的原理，後面則會介紹 Spring AI 1.0.0 才加入的 ChatMemory
- ▼ 讓企業不再卻步的技術-RAG：前面幾個章節說穿了只是做一個聊天機器人，很多企業怕資訊外洩甚至禁用 AI，這個章節就來介紹讓 AI 融入企業的法寶 - RAG，透過 Spring 3 以後才支援的 docker compose support 可以快速地建立向量資料庫

介紹就到這邊，明天就要開始實戰演練，這次主題是打造企業 RAG 知識庫，主軸還是跟 RAG 有關的模組，沒提到的模組大家可以上 [Spring AI 官網](#) 查看