



# ICT Training Center

**Il tuo partner per la Formazione e la Trasformazione digitale della tua azienda**



## Note



# SPRING AI

## GENERATIVE ARTIFICIAL INTELLIGENCE CON JAVA

Simone Scannapieco

Corso avanzato per Venis S.p.A, Venezia, Italia

Novembre 2025

## Note



# SPRING AI E DOCKER MODEL RUNNER

## Note

## → Multi configurazione Ollama + Gemma

- 1 Pull modello LLM ed embedder in Model Runner
  - 2 Test del raggiungimento del servizio chat/embed con Postman/Insomnia
  - 3 Creazione profilo applicativo
  - 4 Modifica proprietà applicativo
  - 5 Test delle funzionalità con Postman/Insomnia

## Note

**⚠️ Cercare embedder su HuggingFace, con task NLP->Sentence similarity e Library->GGUF**

## Caricamento LLM in locale

## Note

**⚠ Utilizzare Postman/Insomnia per chiamate di test**

## Test raggiungimento servizio chat

## Test raggiungimento servizio embedding

## Note

## Configurazione applicativo

```
spring:
  ...
  profiles:
    active: rag-document-to-vector-store,docker-model-runner
  ...

```

## Note

## File application-docker-model-runner.yml

## Note

<https://github.com/simonescannapieco/spring-ai-advanced-dgroove-venis-code.git>

**Branch:** 10-spring-ai-gemma-ollama-docker-model-runner

## Note