

科研项目语义相似性检测引擎系统

用户手册

版本: 1.0.0

日期: 2025 年 9 月 15 日

目录

- [第一章：系统概况](#)
 - [1.1 功能简介](#)
 - [1.2 技术特点](#)
 - [1.3 开发语言与主要技术栈](#)
- [第二章：系统部署](#)
 - [2.1 Windows 部署方法](#)
 - [2.1.1 环境准备](#)
 - [2.1.2 安装依赖](#)
 - [2.1.3 配置环境变量](#)
 - [2.1.4 启动服务](#)
 - [2.2 Docker 部署方法 \(推荐\)](#)
 - [2.2.1 环境准备](#)
 - [2.2.2 配置文件说明](#)
 - [2.2.3 构建与启动](#)
- [第三章：系统使用 \(API 接口\)](#)
 - [3.1 认证](#)
 - [3.2 监控接口](#)
 - [3.2.1 健康检查 \(GET /health\)](#)
 - [3.3 数据入库接口](#)
 - [3.3.1 单条数据入库 \(POST /api/vectors\)](#)

- [3.3.2 批量数据入库 \(POST /api/vectors/batch\)](#)
- [3.3.3 检查数据是否存在 \(POST /api/vectors/get_status\)](#)
- [3.3.4 删除数据 \(DELETE /api/vectors/delete\)](#)
- [3.4 数据查询接口](#)
 - [3.4.1 文本相似度计算 \(POST /api/similarity\)](#)
 - [3.4.2 向量检索 \(POST /api/vectors/search\)](#)
 - [3.4.3 批量向量检索 \(相同过滤器\) \(POST /api/vectors/batch-search-same-filters\)](#)
 - [3.4.4 批量向量检索 \(独立过滤器\) \(POST /api/vectors/batch-search-individual-filters\)](#)
- [3.5 管理接口](#)
 - [3.5.1 初始化/重建数据库 \(POST /api/admin/initialize-database\)](#)
- [第四章：功能模型 HTML 页面使用方法](#)
 - [4.1 访问页面](#)
 - [4.2 页面布局](#)
 - [4.3 功能使用](#)
 - [4.3.1 文本添加](#)
 - [4.3.2 文本查询](#)
 - [4.3.3 文本相似度比较](#)

第一章：系统概况

1.1 功能简介

科研项目语义相似性检测引擎系统（CDS Semantic Engine）是一个高性能、可扩展的 Web 服务。它旨在为科研管理、文献查重、项目申报等场景提供核心的文本语义处理能力。

系统的核心功能包括：

- 文本向量化**：将任意文本（如项目摘要、论文内容、技术报告）转换为高维度的数学向量，使其可以用数学方式进行比较。
- 相似性检索**：接收一段文本，在海量数据库中快速、准确地找出语义上最相似的 N 条文本记录。
- 文本比对**：直接计算两段文本之间的语义相似度得分，用于精确判断内容相关性。
- 数据管理**：提供对向量化数据的增、删、查等管理接口，支持批量操作。

1.2 技术特点

- 高性能**：基于 FastAPI 框架构建，支持高并发异步请求，确保接口的快速响应。
- 高精度**：采用业界领先的深度学习模型进行文本嵌入，保证了语义理解的准确性。
- 海量数据支持**：集成 Milvus 向量数据库，专为大规模向量检索设计，可支持十亿级别的向量数据。
- 异步处理**：对于批量入库等耗时操作，采用后台任务处理，避免长时间阻塞 API 请求，提升用户体验。
- 易于部署**：提供 Docker 和 Docker Compose 部署方案，实现一键式环境搭建和应用启动，极大简化了部署和运维难度。

- **安全性:** 通过 API Key 机制保护接口, 确保只有授权用户才能访问服务。同时提供高危操作 (如数据库重建) 的密码保护。
- **可扩展性:** 系统采用微服务架构设计, 各个组件 (API 服务、嵌入模型、数据库) 松耦合, 便于独立升级和扩展。

1.3 开发语言与主要技术栈

- **开发语言:** Python 3.10
- **Web 框架:** FastAPI - 一个现代、高性能的 Python Web 框架。
- **向量数据库:** Milvus - 一款开源的、专为向量相似性搜索和分析而设计的云原生数据库。
- **文本嵌入模型:** 领先中文文本嵌入模型。
- **异步任务:** BackgroundTasks - FastAPI 内置的轻量级后台任务处理工具。
- **配置管理:** Pydantic-Settings - 用于从环境变量或 `.env` 文件加载和验证配置。
- **容器化:** Docker & Docker Compose - 用于构建、分发和运行应用的容器化解决方案。