

PSP - Programmers Search Platform

程序员工具搜索平台

项目概述

PSP是一个基于大语言模型(LLM)的智能工具搜索引擎，致力于让程序员找到工具不再困难。通过自然语言搜索，帮助程序员快速找到所需的开发工具、软件和技术资源。

| 特性 | 描述 |
|--------|--|
| 自然语言搜索 | 支持中文和英文自然语言查询，如 "替代Photoshop的软件"、"Mac上的视频剪辑" |
| 智能语义理解 | 使用LLM(DeepSeek + 火山方舟双API)理解用户搜索意图 |
| 结构化结果 | 返回包含安装指南、使用说明、常见问题等完整信息 |
| 相关推荐 | 智能推荐相关工具 |
| 缓存系统 | 三级缓存架构: 客户端缓存(<0.1秒) + 服务器缓存(~50-100ms) + LLM API(3-5秒) 缓存命中率优化: 通过语义归一化提高缓存命中率至96%+ 缓存版本管理: 支持双语数据格式的新旧版本兼容 |

主要功能模块

1. 智能搜索模块

- 自然语言搜索: 支持中文和英文自然语言查询，如 "替代Photoshop的软件"、"Mac上的视频剪辑"
- 智能语义理解: 使用LLM(DeepSeek + 火山方舟双API)理解用户搜索意图
- 结构化结果: 返回包含安装指南、使用说明、常见问题等完整信息
- 相关推荐: 智能推荐相关工具

2. 缓存系统模块

- 三级缓存架构: 客户端缓存(<0.1秒) + 服务器缓存(~50-100ms) + LLM API(3-5秒)
- 缓存命中率优化: 通过语义归一化提高缓存命中率至96%+
- 缓存版本管理: 支持双语数据格式的新旧版本兼容

- 智能过期策略: 服务器缓存30天过期, 客户端缓存长期有效

3. 国际化模块

- 中英文双语支持: 完整的国际化实现
- 语言切换: 实时切换, 自动保存用户偏好到localStorage
- 双语数据结构: 同时存储中英文版本, 根据用户偏好显示

4. UI组件模块

- 搜索框组件: 支持实时建议、热门搜索、GitHub Trending
- 搜索结果展示: 美观的卡片式布局, 支持展开/收起详细信息
- 深色模式: 完整的深色/浅色主题切换, 自动保存用户偏好
- 响应式设计: 完美适配桌面、平板、手机等各种屏幕尺寸

5. 速率限制模块

- 5秒冷却机制: 防止用户连续点击消耗API配额
- 每日限额: 每用户每天30次搜索限制, 控制API调用成本
- 缓存例外: 已缓存工具可无限次查询, 不受限额限制

技术栈

| 类别 | 技术 |
|--------|----------------------------|
| 前端 | React 18 + TypeScript |
| 构建 | Vite |
| 样式 | TailwindCSS |
| UI库 | shadcn/ui + Radix UI |
| 图标 | Lucide React |
| HTTP请求 | Axios |
| 云函数 | Cloudflare Pages Functions |
| 存储 | Cloudflare KV |
| LLM集成 | DeepSeek API + LLM API |
| 部署 | Cloudflare Pages |
| 域名 | douya.chat |
| 压缩 | gzip 73KB |

目标用户

| 用户群体 | 使用场景 |
|-------|------------|
| 企业级客户 | 内部搜索、知识管理 |
| 个人用户 | 日常搜索、学习助手 |
| 开发者 | API调用、工具集成 |
| 学生 | 作业助手、学习资源 |

项目特色与创新

- 智能搜索体验

无需记忆工具名称，直接描述需求；LLM理解自然语言查询意图；智能纠错和推荐

- 高性能架构

三级缓存系统减少96%+的API消耗；快速响应：缓存查询<0.1秒，API查询3-5秒；代码优化：225KB→73KB(gzip)

- 成本控制

客户端缓存优先，大部分查询不消耗API配额；速率限制机制防止恶意刷接口；双API备份提高可用性

- 开发者友好

完整的技术文档和部署指南；开源MIT许可证；一键部署到Cloudflare Pages；支持本地开发环境

- 用户体验

精美的UI设计和流畅的交互动画；深色模式支持；完整的国际化支持；实时搜索建议和热门推荐

核心技术亮点

- 语义归一化：通过智能处理提取查询核心意图，提高缓存命中率
- 双语并行生成：同时生成中英文版本，提升全球化体验
- 渐进式缓存：从本地到服务器的多级缓存策略
- 安全设计：完善的输入验证和防护机制

项目总结

PSP项目展示了一个现代化的AI驱动的搜索应用，通过巧妙的技术架构设计和用户体验优化，成功解决了开发者寻找工具的痛点。该项目结合了最新的前端技术、云原生架构和人工智能技术，为开发者提供了一个高效、智能、易用的工具搜索平台。

项目的核心优势在于其智能的自然语言处理能力、高性能的三级缓存系统、精美的用户界面设计，以及完善的成本控制机制。通过96%+的缓存命中率和73KB的优化打包大小，该项目在性能和用户体验方面都达到了业界领先水平。