

智能文档管理系统项目报告

项目名称：智能文档管理系统 v2.0
报告日期：2024年1月
项目经理：王工程师

一、项目概述

本项目旨在开发一款基于HarmonyOS的智能文档管理应用，实现文档的智能分类、内容摘要、相似度检测等功能。

二、项目进度

阶段	计划完成时间	实际完成时间	状态
需求分析	2023-11-15	2023-11-12	已完成
系统设计	2023-12-01	2023-11-30	已完成
核心开发	2024-01-15	进行中	80%
测试阶段	2024-02-01	-	待开始
上线部署	2024-02-15	-	待开始

三、已完成功能

1. 文件管理基础功能
 - 文件创建、读取、删除
 - 文件重命名和移动
 - 回收站功能
2. 智能分类功能
 - 基于关键词的本地分类（7个类别）
 - DeepSeek AI 辅助分类
 - 分类结果可视化展示
3. 内容解析功能
 - TXT文件解析
 - DOCX文件解析（使用mammoth库）
 - PDF文件解析（自研解析器）

四、技术难点与解决方案

1. PDF解析乱码问题
 - 问题：中文PDF提取文本出现乱码
 - 方案：增加Unicode文本提取、CID字体映射支持
2. 性能优化
 - 问题：大量文件分析时UI卡顿
 - 方案：使用Worker多线程处理

五、下一步计划

1. 完善PDF解析，提高中文支持
2. 增加文档相似度检测功能
3. 优化AI摘要生成质量
4. 进行全面的测试

六、风险评估

- 技术风险：PDF解析复杂度高，部分格式可能无法支持
- 进度风险：AI功能依赖外部API，需要网络稳定

项目状态：进行中
下次汇报时间：2024年1月22日