

论著 • Article

基于 WorkBuddy 的 AI 桌面智能体应用与实践研究

成阳婧 *

(西安工业大学 陕西 西安 710021)

摘要 在大模型技术快速普及与混合办公常态化的双重驱动下, AI 办公工具正从文本生成、信息咨询等辅助功能, 向任务自动化、桌面操作执行等深度应用升级。传统对话式 AI 普遍存在“能给出建议却无法落地执行”的短板, 企业级 RPA 与智能办公系统又存在部署复杂、成本偏高、适配性不足等问题, 难以满足普通职场用户轻量化、安全化、高效率的办公需求。WorkBuddy 作为一款面向全场景职场的 AI 桌面智能体工具, 以自然语言交互为入口, 实现了本地文件操作、多步骤任务拆解、自动化执行与远程协同办公的完整闭环。本文从 AI 办公应用背景出发, 对 WorkBuddy 的产品定位、核心功能、技术架构进行系统性梳理, 结合真实办公场景给出详细实操流程与应用案例, 客观分析其应用价值、优势与现存不足, 并对未来 AI 桌面智能体的发展方向进行展望。全文以应用科普与实践研究为核心, 突出实用性与逻辑性, 为同类 AI 办公工具的研究、推广与使用提供参考。

关键词 WorkBuddy; AI 智能体; 桌面自动化; 办公智能化; 远程协同

Research on the Application and Practice of AI Desktop Agent Based on WorkBuddy

Cheng Yangjing *

(Xi'an Technological University, Xi'an 710021, China.)

Abstract Driven by the rapid popularization of large model technologies and the normalization of hybrid work, AI office tools are evolving from auxiliary functions such as text generation and information inquiry to in-depth applications including task automation and desktop operation execution. Traditional conversational AI systems are often limited to providing suggestions rather than carrying out concrete actions. Meanwhile, enterprise-grade RPA and intelligent office platforms feature complicated deployment, high costs, and insufficient compatibility, which fail to meet the lightweight, secure, and efficient office demands of regular workplace users. As an all-scenario AI desktop agent for workplace users, WorkBuddy uses natural language interaction as its entry point to support a complete workflow covering local file manipulation, multi-step task decomposition, automated execution, and remote collaborative work. Starting from the background of AI office applications, this paper systematically reviews the product positioning, core functions, and technical architecture of WorkBuddy. It provides detailed operating procedures and application cases in real office scenarios, ob-

收稿日期: 2026-02-12 录用日期: 2025-04-15

通讯作者: 成阳婧; 邮箱: 1431807602@qq.com

基金项目: 西安工业大学大学生创新创业训练计划项目 (编号: 202510702031)

jectively analyzes its application value, strengths, and current limitations, and discusses the future development trend of AI desktop agents. Focusing on application popularization and practical research, this paper emphasizes practicality and logic, so as to provide a reference for the research, promotion, and use of similar AI office tools.

Keywords WorkBuddy; AI agent; Desktop automation; Intelligent office; Remote collaboration

1 引言

随着生成式人工智能技术的规模化落地, AI 已经全面渗透到办公、教育、政务、企业服务等多个领域。在日常办公场景中, 用户对 AI 工具的需求不再局限于文案撰写、问答交互、资料检索等基础功能, 而是希望 AI 能够真正替代人工完成重复性、流程化、机械性的桌面操作, 例如文件整理、数据清洗、报表生成、格式转换、信息提取等。然而当前市场上主流的 AI 办公产品大多停留在“对话辅助”层面, 仅能提供思路、文本或方案, 无法直接操控电脑、读写本地文件、串联多软件操作, 导致 AI 能力与实际办公流程之间存在明显断层。

与此同时, 混合办公、远程办公、碎片化办公成为新常态, 用户对跨设备操控、随时随地处理任务、数据安全可控的需求持续提升。传统办公软件功能分散、操作步骤繁琐, 大量时间被消耗在重复点击、复制粘贴、格式调整等低价值工作中, 不仅效率低下, 也容易出现人为失误。在此背景下, 能够直接理解自然语言指令、自主拆解任务、自动完成桌面操作的 AI 桌面智能体应运而生, 成为连接 AI 能力与真实办公场景的重要载体。

WorkBuddy 以“轻量化 AI 执行者”为核心定位, 打破了传统 AI “只说不做”的局限, 通过简洁的操作流程、安全的本地运行机制、丰富的文件处理能力, 满足了普通职场用户的高频办公需求。与偏向技术研发的专业 AI 系统不同, WorkBuddy 更强调实用性、易用性与普及性, 适合作为应用科普类期刊的研究对

象。本文围绕其技术逻辑、实操流程、应用场景展开系统性分析, 避免单纯概念罗列, 突出实践价值与推广意义^[1]。

2 AI 办公发展现状与 WorkBuddy 产品定位

2.1 当前 AI 办公工具的发展与痛点

近年来, AI 办公赛道发展迅速, 各类文本生成、语音转写、翻译润色、数据分析工具不断涌现, 在一定程度上提升了办公效率。但从实际应用效果来看, 现有工具仍存在几类典型问题^[8]。

2.1.1 交互与执行脱节

多数 AI 工具以对话为核心交互方式, 能够生成报告、总结、方案等文本内容, 但无法直接操作电脑桌面文件, 不能打开 Office 软件、不能整理文件夹、不能批量修改文件名, 用户仍需手动复制、粘贴、排版, 任务闭环难以形成。

2.1.2 工具碎片化严重

文件处理需要专用软件, 数据处理需要表格工具, 格式转换需要在线平台, 远程办公需要独立客户端, 用户需要在多个软件之间频繁切换, 操作链条冗长, 整体效率并未得到本质提升。

2.1.3 部署与使用门槛偏高

部分面向企业的流程自动化 (RPA) 工具功能强大, 但需要编写脚本、配置流程、设置触发条件, 普通职场员工不具备相关技术能力, 难以快速上手使用。

2.1.4 数据安全存在隐患

部分在线 AI 工具要求用户上传文件至云端进行处理，对于包含敏感信息的工作文档、企业数据、个人资料而言，存在数据泄露、滥用等潜在风险，用户使用意愿受限。

综上，当前市场亟需一款兼顾易用性、执行力、安全性与轻量化的 AI 桌面工具，能够以最低门槛实现“指令输入—自动执行—成果交付”的全流程办公自动化，WorkBuddy 正是在这一需求下诞生的典型产品^[2]。

2.2 WorkBuddy 的产品定位与核心理念

WorkBuddy 定位为全场景 AI 桌面智能体工作台，其核心设计理念是“让 AI 像人一样操作电脑，用自然语言完成复杂办公任务”。它并非替代 Word、Excel、PPT 等传统办公软件，而是作为统一的智能执行层，将分散在不同软件中的操作串联起来，通过任务自动化降低用户操作成本。

与传统 AI 助手相比，WorkBuddy 最大的区别在于具备桌面执行能力，可以直接读取、修改、整理、生成本地文件，自主完成多步骤任务，无需人工逐步操作。与企业级 RPA 相比，它无需复杂配置、无需代码基础、下载即用，更适合个人用户、小微企业员工、行政办公人员、学生群体等轻量化使用场景。

在安全层面，WorkBuddy 采用本地优先运行模式，配合目录授权与沙箱隔离机制，尽可能减少数据上传云端，保障用户文件与隐私安全。在场景覆盖上，它面向日常办公全流程，从文件管理、数据处理、文档生成到远程操控，均可通过自然语言指令完成，具备极强的实用价值与科普推广意义。

3 WorkBuddy 核心功能与使用逻辑

3.1 核心功能体系

3.1.1 自然语言意图理解与任务拆解

系统支持口语化、模糊化、多目标复合指令输入，能够自动识别用户真实需求，将复杂任务拆分为有序、可执行的子步骤。例如用户提出“查一下茅台最近的股价，画个走势图”，系统可自动拆分为查询茅台最近 30 天股价数据、安装 requests 库等多个环节，并按顺序执行^[3]。（图 1、2）



图 1. WorkBuddy 自然语言任务拆解示意

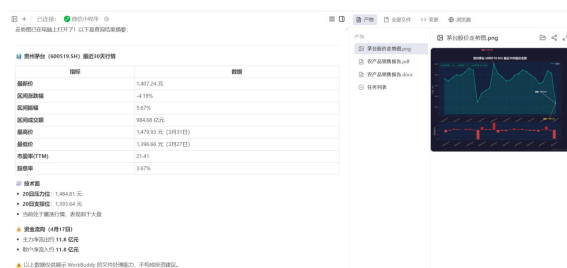


图 2. WorkBuddy 自然语言意图理解与任务拆解示例

3.1.2 本地文件全格式处理

这是 WorkBuddy 的核心能力之一，支持对 Word、Excel、PPT、PDF、图片、压缩包、文本文档等常见格式进行读写、编辑、格式转换、内容提取、批量重命名、分类归档等操作，无需借助第三方工具即可完成文件全生命周期管理。例如“帮我创建一份农产品销售报告”“把这份 Word 里的表格提取出来”“把

报告 .docx 转成 PDF’”。(图 3)



图 3. WorkBuddy 本地文件全格式处理功能示例

3.1.3 多模型自适应调用

系统内置多款主流大语言模型，针对不同任务自动匹配合适模型，也支持用户手动切换。在文本生成、逻辑分析、数据解读、专业内容撰写等场景中，可灵活发挥不同模型的优势，提升任务完成质量^[7]。(图 4)

3.1.4 远程桌面控制与跨设备协同



图 4. WorkBuddy 多模型自适应调用界面

用户可通过移动端设备向电脑端发送指令，实现远程文件整理、数据处理、文档生成等操作，无需守在电脑前即可完成办公任务，

适配差旅、外出、居家等混合办公场景。移动端发送指令如图 5、6 所示。

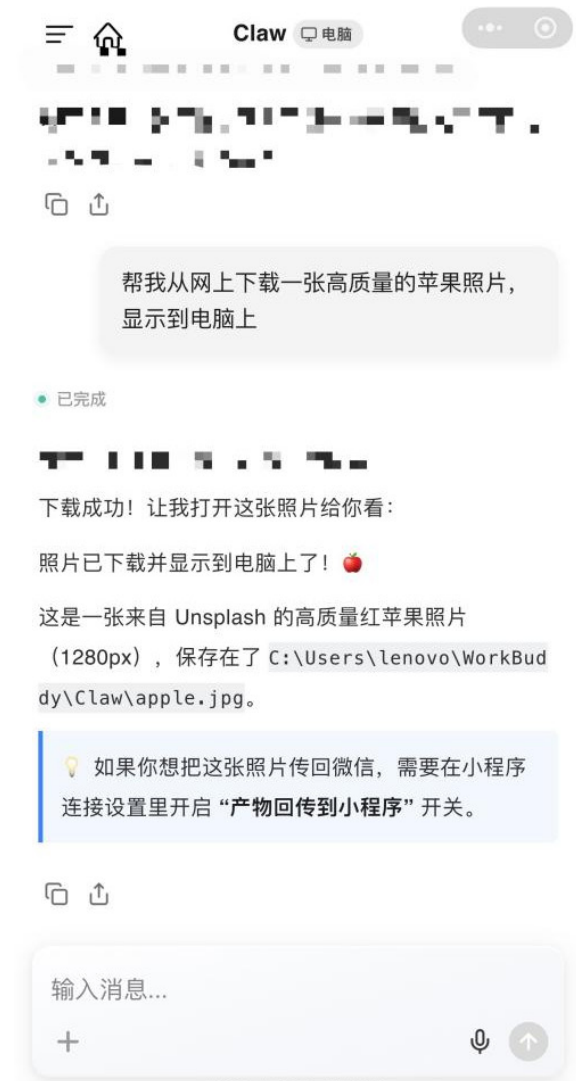


图 5. WorkBuddy 电脑端实现



图 6. WorkBuddy 远程桌面控制功能演示（上移动端指令；下电脑端执行）

3.1.5 安全权限管控与可视化执行监控

系统采用沙箱隔离运行机制，仅允许访问用户授权的文件夹，禁止修改系统文件、访问隐私目录，同时全程可视化展示执行进度，支持暂停、终止、人工干预，避免误操作带来的风险^[6]。(图 7)

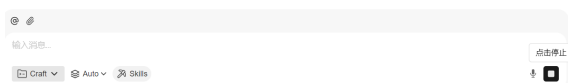


图 7. WorkBuddy 安全权限管控与可视化执行监控界面

3.2 产品使用逻辑与交互特点

WorkBuddy 遵循极简使用逻辑，整体交互流程贴近用户习惯，核心路径为“指令输入—任务解析—自动执行—结果交付”。用户无需理解技术原理，无需记忆操作命令，只需日常说话的方式下达指令，系统即可自主完成后续工作。

与传统软件的菜单式操作不同，WorkBuddy 以需求为导向，而非以功能为导向。用户不需要知道“如何批量重命名”“如何做数据透视表”“如何生成图表”，只需要告诉 AI “想要什么结果”，AI 自动选择对应的功能与执行路径。这种交互模式大幅降低了工具使用门槛，使不擅长复杂软件操作的用户也能实现高效办公。

同时，系统具备良好的容错能力与纠错能力，对于表述不清晰的指令，可主动向用户询问补充信息；在执行过程中遇到异常情况，会及时给出提示并尝试自动修复，保证任务稳定推进。

4 WorkBuddy 技术框架与运行机制

4.1 整体技术架构

WorkBuddy 采用轻量化、模块化的三层技术架构，兼顾执行效率、功能扩展与安全稳

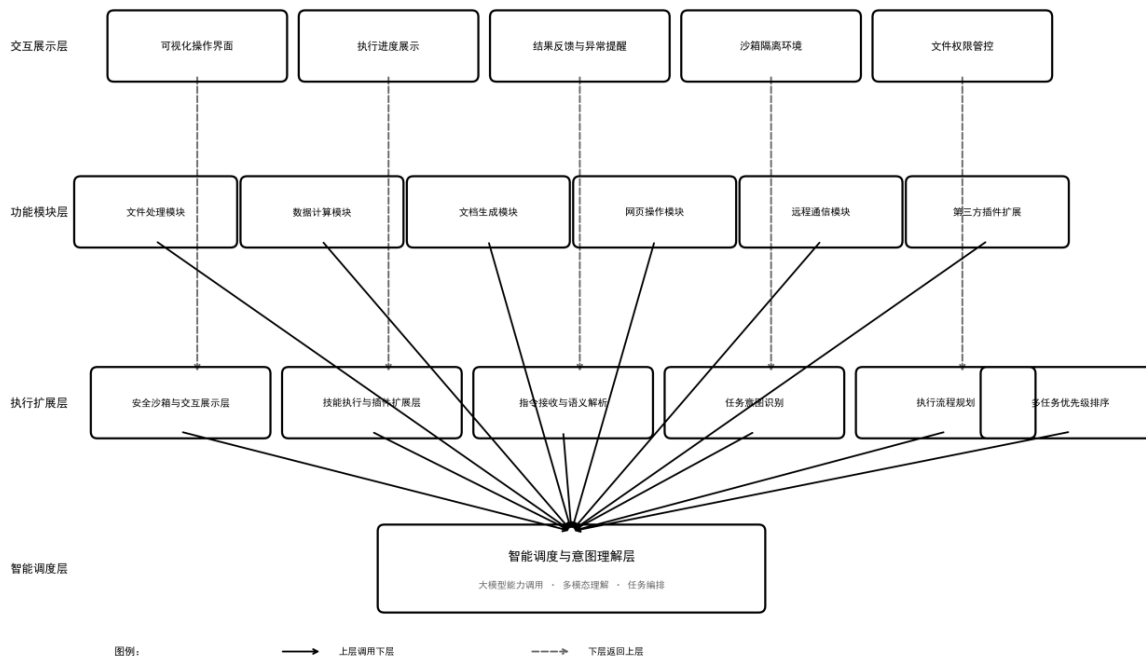


图 8. WorkBuddy 三层技术架构示意图

定性，整体结构简洁清晰，适合科普性解读。

第一层为智能调度与意图理解层，作为整个系统的核心大脑，负责接收用户指令、进行语义解析、识别任务意图、规划执行流程。该模块依托大模型理解能力，将自然语言转化为机器可执行的逻辑步骤，同时对多任务进行优先级排序，支持多任务并行处理。

第二层为技能执行与插件扩展层，是系统具体能力的载体，包含文件处理、数据计算、文档生成、网页操作、远程通信等标准化技能模块。每个技能模块对应一类办公操作，可独立运行也可组合调用。同时，该层支持第三方技能插件扩展，能够不断丰富系统能力。

第三层为安全沙箱与交互展示层，负责提供可视化操作界面、执行进度展示、结果反馈与异常提醒，同时通过沙箱环境隔离任务执行过程，严格控制文件访问权限，防止危险操作，保障本地系统安全。

三层架构协同配合，形成“理解—规划—执行—监控—安全”的完整技术闭环，既保证了 AI 智能体的自主性，又确保了运行过程的可控性与安全性^[4,9]。（图 8）

4.2 任务执行核心机制

WorkBuddy 完成一项办公任务，通常遵循稳定且可追溯的运行机制，具体分为五个步骤。（图 9）

第一步，指令接收与解析。用户通过文本输入方式下达指令，系统对指令进行分词、意图识别、关键信息提取，明确任务目标、操作对象、输出要求。

第二步，任务拆解与流程规划。系统将整体任务拆分为若干有序子任务，确定执行顺序、调用的技能模块、操作的文件路径、输出格式等，形成完整执行方案。

第三步，授权范围内自动执行。在用户

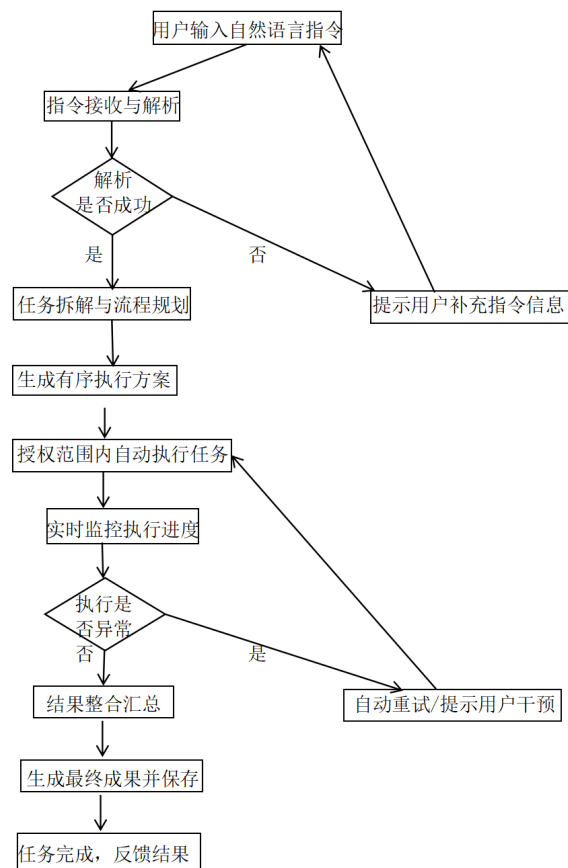


图 9. WorkBuddy 任务执行核心流程图

预先授权的目录下，系统依次调用对应技能模块，执行文件读写、数据处理、内容生成等操作，全程无需人工干预。

第四步，执行过程监控与异常处理。系统实时展示执行进度，对文件缺失、格式错误、权限不足等问题自动识别，根据情况进行跳过、重试或提示用户补充信息。

第五步，结果整合与交付。执行完成后，系统自动汇总所有子任务结果，按用户要求生成最终文件、报表或报告，并保存至指定目录，完成整个任务闭环。

这一机制避免了简单的功能堆砌，使 WorkBuddy 真正具备“任务级智能”，而非单一功能的集合，这也是其区别于普通工具的核

心技术特征^[5]。

5 WorkBuddy 实操流程与典型应用示例

5.1 安装与初始化实操流程

WorkBuddy 主打轻量化使用体验，安装与配置流程简单快捷，普通用户可在短时间内完成部署。

5.1.1 软件获取与安装。

通过官方渠道获取安装包后，双击启动安装程序，按照指引完成安装，全过程无需配置



登录成功

返回 WorkBuddy 继续使用

图 10. WorkBuddy 软件安装界面

运行环境、无需安装依赖库，对电脑配置要求较低，适配主流 Windows 系统设备。（图 10）

5.1.2 初始权限配置

首次启动时，系统引导用户设置授权访问目录，用户可选择工作文件夹、常用文档目录



图 11. WorkBuddy 初始权限配置界面

等，避免 AI 访问系统盘或隐私文件夹，从源头保障数据安全。（图 11）

5.1.3 基础参数设置

用户可根据使用习惯选择默认调用的大模型，设置执行过程展示方式、文件默认保存路径等基础参数，配置完成后即可进入主界面。（图 12）



图 12. WorkBuddy 基础参数设置界面

5.1.4 远程功能绑定

对于有远程办公需求的用户，可按照指引绑定移动端通信工具，实现手机发送指令、电脑自动执行的跨设备协同模式。移动端（图 13、14）

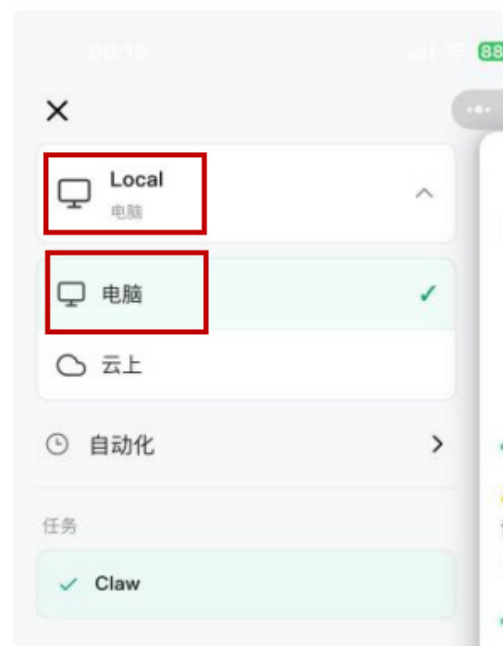


图 13. . 电脑端实现

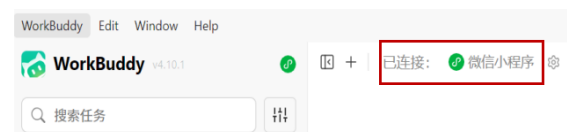


图 14. WorkBuddy 远程功能绑定界面（上移动端；下电脑端）

完成以上步骤后，系统即可投入正常使用，整体流程无技术门槛，适合各类人群快速上手。

5.2 典型办公场景实操示例

5.2.1 周报 / 月报自动化生成

运营 / 销售岗位日常需定期整理业务数据，制作周报、月报，Excel 操作繁琐、数据清洗易出错，且重复劳动量大，每周至少耗费 1-2 小时在数据汇总和报告排版上。用户指令，“帮我生成一份本周的工作周报。”（图 15）

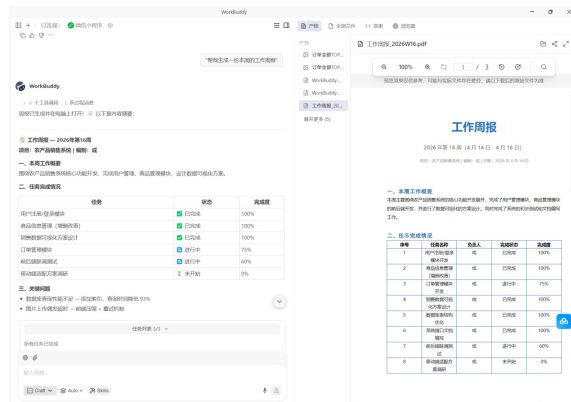


图 15. WorkBuddy 周报自动化生成效果示例

5.2.2 数据分析与可视化

运营与销售岗位日常需从大量业务数据中快速定位核心指标，手动排序、筛选、制作图表耗时较长，且对函数和图表操作有一定门槛，非专业用户难以高效完成。

• 用户指令

“帮我分析这份文档中订单金额情况，找出 TOP3，画个柱状图”

• 系统执行过程

读取目标文档→数据清洗（去重、去空值）→按销售额降序排序→提取前 3 名数据→使用 matplotlib 生成柱状图（含数值标签）→将图表保存为图片→输出 TOP3 数据表格与分析摘要。

• 应用效果

原本需要 15-30 分钟的手动数据排序和图表制作工作，可在 10-15 秒内自动完成。系统自动识别数据列、完成排序和 TOP3 提取，生成带数值标注的专业柱状图。（图 16）

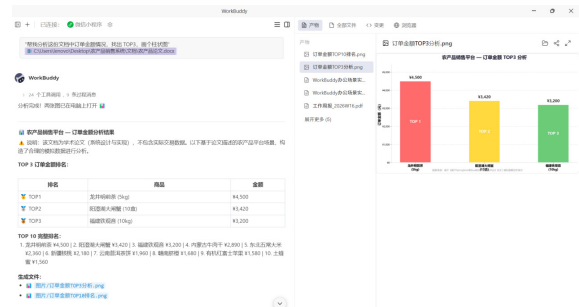


图 16. WorkBuddy 数据分析与可视化生成效果示例

5.2.3 文件批量整理与归档

项目开发过程中，随着需求文档、设计稿、代码文件、数据库脚本、测试报告等各类文件不断积累，项目文件夹往往变得杂乱无章，尤其像农产品销售系统这类涉及多模块的项目，前后端文件、文档、图片混杂存放，影响团队协作效率和文件查找速度。

• 用户指令

“帮我把农产品销售系统文件夹的文件按类型整理一下”

• 系统执行过程

遍历农产品销售系统文件夹→识别所有文件后缀名→按项目文件类型智能分组（代码文件 / 文档 / 数据库脚本 / 设计资源 / 配置文件 / 日志 / 其他）→创建对应分类子文件夹→移动文件到所属目录→保留原有项目入口文件不动→统计各类文件数量与大小→生成项目文件整理清单。

• 应用效果

原本需要 20-40 分钟的手动分类整理工作，可在数分钟内自动完成。项目文件按类型清晰

归档，代码、文档、资源各归其位，后续开发和查找效率显著提升，尤其对于多模块协作项目，规范化的文件结构有助于降低新成员上手成本。（图 17、18）



图 17. WorkBuddy 文件批量整理归档



整理内容：6 个散落文件全部按类型归入对应文件夹，根目录只保留项目子目录，干净整洁。

图 18. WorkBuddy 文件批量整理归档效果示例

6 应用价值、优势与不足分析

6.1 综合应用价值

从实际应用效果来看，WorkBuddy 在个人办公效率提升、办公模式优化、使用门槛降低等方面均体现出明显价值。

6.1.1 显著降低重复性工作耗时

在文件整理、数据处理、格式转换、批量操作等场景中，可节省 60% 以上的时间成本，将人力从机械劳动中释放出来，专注于创意、决策、沟通等高质量工作。

6.1.2 全面降低办公软件使用门槛

无需学习 Excel 函数、PPT 排版、PDF 编辑等专业技能，通过自然语言即可完成复杂操作，对职场新人、非专业岗位人员十分友好。

6.1.3 适配现代化办公模式

支持远程控制、跨设备协同、离线本地处理，契合混合办公、移动办公、居家办公等新常态，提升办公灵活性。

6.2 核心优势总结

对比同类 AI 办公工具，WorkBuddy 的优势集中体现在四个方面。一是执行力强，真正实现从指令到成果的全流程自动化，可直接操作桌面文件，而非仅提供文本辅助；二是易用性高，零代码、零配置、即开即用，交互方式贴近日常语言习惯，用户学习成本极低；三是安全性好，本地运行、权限可控、沙箱隔离，兼顾便捷性与数据安全；四是场景覆盖广，可满足行政、财务、运营、学生、自由职业者等多群体需求，适配日常办公绝大多数场景。

6.3 现存局限性

作为一款面向轻量化应用的 AI 桌面智能体，WorkBuddy 仍存在一定的能力边界。其一，在高度专业化场景中能力有限，例如深度财务审计、复杂法律文书审核、高精度工程计算、工业级数据分析等，仍需专业人员复核与

补充。其二，对特殊格式文件、高度加密文档的支持不足，部分小众行业专用文件无法正常读取与处理。其三，行业定制化能力较弱，针对特定企业的业务流程、内部系统适配性有待提升，企业级管理、多用户协同、数据审计等功能仍在完善中。其四，极端复杂的长链条任务可能出现执行偏差，需要用户适当监督与补充指令。

7 总结

WorkBuddy 代表了 AI 办公从“对话助手”向“执行智能体”的重要转变。它以轻量化架构、自然语言交互、本地安全执行和全流程任务闭环为特点，有效解决了传统 AI 工具“能说不能做”的痛点，在文件管理、数据处理、远程办公等场景具备明显实用价值。

总体来看，WorkBuddy 更偏向实用型 AI 生产力工具，研究性与科普性较强，适合作为应用类期刊的案例性文章。未来，随着桌面智能体技术的进一步成熟，AI 将更深度融入办公流程，实现更程度的自动化与协同化，而类似 WorkBuddy 的轻量化工具也将在普通职场人群中持续普及，成为推动智慧办公发展的重要力量。

参考文献

- [1] 张伟, 李明. 大语言模型驱动的任务自动化研究综述 [J]. 计算机研究与发展, 2024, 61(3): 512-530.
- [2] 王芳, 陈刚, 赵磊. 基于 GPT 的智能办公助手设计与实现 [J]. 软件工程, 2024, 27(6): 45-52.
- [3] OpenAI. GPT-4 Technical Report[R]. OpenAI, 2023.
- [4] 刘洋, 孙晓明. AI 智能体技术框架与典型应用场景研究 [J]. 人工智能, 2024, (2): 33-41.
- [5] 吴晓燕, 周建国. 桌面自动化中的自然语言处理技术研究 [J]. 计算机工程与应用, 2023, 59(18): 1-10.
- [6] Brown T, Mann B, Ryder N, et al. Language Models are Few-Shot Learners[J]. Advances in Neural Information Processing Systems, 2020, 33: 1877-1901.
- [7] Yao S, Zhao J, Yu D, et al. ReAct: Synergizing Reasoning and Acting in Language Models[C]//International Conference on Learning Representations. 2023.
- [8] 马超, 李建平. RPA 与 AI 技术融合的企业流程自动化研究 [J]. 信息系统学报, 2023, 17(4): 78-89.
- [9] Schick T, Dwivedi-Yu J, Dessì R, et al. Toolformer: Language Models Can Teach Themselves to Use Tools[J]. arXiv preprint arXiv:2302.04761, 2023.