

企业级软件 AI 功能成熟度测试方案 技术规范

一、测试方案概述

1.1 测试目的

验证企业级软件 AI 功能成熟度的各项指标是否符合要求，确保登录、用户管理、对话、任务执行、文件管理、应用市场、系统配置等模块功能正确，同时检查系统性能，保证在高负载和复杂场景下稳定运行，满足用户需求。

1.2 测试范围

涵盖所有功能模块，包括登录与权限管理、用户信息与权限配置、对话与智能助手功能、任务管理与执行、文件管理与预览、应用市场操作、系统设置与配置、一体机应用管理、语音功能、运行环境隔离及多语言支持等场景。

二、测试环境准备

准备多台不同配置的服务器，安装指定操作系统和数据库。部署软件所需中间件，搭建网络环境，确保网络稳定。准备多种浏览器和移动设备，安装测试所需工具，如性能测试工具、自动化测试框架，保证测试环境独立且数据可重置。

三、测试方案

3.1 账号与权限

3.1.1 登录认证

测试目的	用户使用账号和密码进入平台；登录状态保持有效，未登录时自动进入登录页。
测试步骤	1、打开平台登录页面，输入正确的账号和密码并点击登录按钮 2、登录成功后关闭浏览器或标签页，重新访问平台地址 3、手动清除浏览器缓存或等待会话过期后访问平台
预期结果	1、系统验证通过并成功跳转至平台首页 2、系统自动识别有效会话并直接进入首页无需再次登录 3、系统检测到未登录状态并自动重定向至登录页面

3.1.2 许可证门禁

测试目的	在启用授权校验的环境下，需先完成产品授权才能登录；未授权时会引导填写许可证。
测试步骤	1、在启用授权校验的环境中启动平台且未配置许可证 2、在引导界面输入无效的许可证密钥并提交 3、输入正确的产品授权许可证并完成验证
预期结果	1、系统拦截登录请求并弹出引导填写许可证的界面 2、系统提示授权失败并要求重新输入有效的许可证 3、授权验证通过，用户可正常进入登录流程

3.1.3 个人中心

测试目的	用户可查看和修改昵称、邮箱、登录密码，并切换界面语言与浅色/深色外观。
测试步骤	1、进入个人中心页面，修改昵称和邮箱地址并保存 2、在安全设置中尝试修改登录密码，输入旧密码和新密码 3、切换界面语言为英语并将外观主题从浅色改为深色
预期结果	1、系统提示保存成功且页面显示更新后的昵称和邮箱 2、密码修改成功，下次登录需使用新密码 3、界面文字立即变为英语且背景颜色切换为深色模式

--	--

3.1.4 用户管理

测试目的	管理员可新增、编辑、停用或删除用户，并为普通用户配置能在应用市场中看到哪些应用。
测试步骤	1、管理员进入用户管理页面，点击新增用户并填写完整信息 2、选择现有普通用户进行编辑，调整其可见的应用市场范围 3、对某个用户执行停用操作并尝试删除另一个测试用户
预期结果	1、新用户创建成功并出现在用户列表中 2、用户配置更新成功，该用户登录后仅能看到指定应用 3、停用用户无法登录，测试用户被成功移除出列表

3.1.5 引导账户保护

测试目的	系统内置的超级管理员账号不可被删除，防止误删后无人可管理系统。
测试步骤	1、尝试在用户管理列表中选中系统内置的超级管理员账号 2、右键点击超级管理员账号查看上下文菜单 3、尝试通过批量操作功能勾选超级管理员进行删除
预期结果	1、删除按钮处于禁用状态或点击后提示不可删除 2、菜单中不包含删除或停用选项 3、系统阻止操作并提示超级管理员账号受保护不可删除

3.1.6 角色与菜单权限

测试目的	普通用户仅可使用对话、应用市场、数字员工等日常功能；运维与系统配置入口仅管理员可见。
测试步骤	1、使用普通用户账号登录平台并检查侧边栏导航 2、检查普通用户界面是否存在运维管理或系统配置菜单 3、直接使用 URL 访问管理员专属的系统配置页面
预期结果	1、仅显示对话、应用市场、数字员工等功能入口 2、界面上不显示任何运维与系统配置相关的入口 3、系统返回权限不足提示或自动重定向到首页

3.1.7 产品更新说明

测试目的	点击侧栏版本号可查看本版新增、优化与修复说明，便于了解平台最新变化。
测试步骤	1、在平台侧边栏底部找到并点击版本号标识 2、检查弹窗内容是否包含新增功能、优化项和修复问题分类 3、关闭版本说明弹窗并再次点击版本号
预期结果	1、弹出窗口展示当前版本的详细更新说明 2、弹窗清晰列出各类变更详情且无乱码或缺失 3、弹窗能正常再次打开且内容保持一致

3.2 AI 对话

3.2.1 通用模式

测试目的	像日常聊天一样与所选大模型多轮对话，可上传附件、按需开启联网搜索，并管理历史会话。
测试步骤	1、选择一个已对接的大模型并开始多轮对话交互 2、在对话过程中上传一个文本附件并询问相关内容 3、开启联网搜索功能并查询实时新闻或数据
预期结果	1、模型能准确理解上下文并给出连贯的回复 2、助手成功解析附件内容并基于文件信息回答问题 3、助手检索最新网页信息并整合到回答中

3.2.2 智能体模式

测试目的	选择已对接的外部智能体进行对话，适合调用企业已有智能体应用，同样支持联网搜索。
测试步骤	1、在对话选择器中切换至已对接的外部智能体 2、向外部智能体发送需要调用企业已有应用的指令 3、在与外部智能体对话时开启联网搜索功能
预期结果	1、界面标识显示当前正在与外部智能体交互 2、智能体成功调用后端应用并返回执行结果 3、智能体能结合外部搜索信息和自身能力进行回复

3.2.3 数字员工模式

测试目的	面向复杂任务的智能助手：可绑定工作区、分步执行任务、引用文件、预览过程与成果，适合办公自动化场景。
测试步骤	1、创建一个面向复杂任务的智能助手并绑定指定工作区 2、下达分步执行任务指令并引用工作区内的文件 3、等待任务执行完成并查看最终成果展示
预期结果	1、助手创建成功且关联的工作区状态显示正常 2、助手生成任务计划并开始按步骤执行，过程可预览 3、面板展示完整的执行过程和最终生成的文档或数据

3.2.4 会话管理

测试目的	可新建、切换、重命名或删除对话；切换页面后，进行中的任务可在后台继续执行。
测试步骤	1、在对话列表中点击新建对话并输入名称 2、切换到其他对话或在后台运行任务时刷新页面 3、对某个历史对话进行重命名或删除操作
预期结果	1、新对话创建成功并成为当前活动会话 2、之前的任务在后台继续执行且进度未丢失 3、对话列表实时更新反映名称变更或条目消失

3.2.5 联网搜索

测试目的	在对话中按需开启联网，让助手能检索最新网页信息后再回答（需管理员预先配置搜索服务）。
测试步骤	1、确保管理员已预先配置好搜索服务并在对话中开启联网 2、提出一个需要最新网络信息才能回答的问题 3、在未配置搜索服务的环境下尝试开启联网功能
预期结果	1、界面显示联网搜索已激活状态 2、助手先执行搜索动作再基于搜索结果生成回答 3、系统提示无法启用或显示配置缺失的警告

3.2.6 语音输入（自动语音识别）

测试目的	在输入框用麦克风说话，语音会实时转成文字填入，便于免打字输入。
测试步骤	1、点击输入框旁的麦克风图标并开始说话 2、在嘈杂环境中测试语音转文字的准确性 3、说完话后停止录音并检查文本是否完整
预期结果	1、语音被实时录制并转换为文字填入输入框 2、系统能有效过滤噪音并较准确地识别语音内容 3、输入框内显示完整的转换文本且无截断

3.2.7 自动播报（文本转语音）

测试目的	助手回复时可按句朗读，适合听读场景或不便看屏时使用。
测试步骤	1、接收助手的长文本回复并点击朗读按钮 2、在朗读过程中暂停或停止播放 3、切换不同的音色设置后再次触发朗读
预期结果	1、系统开始按句朗读回复内容且声音清晰 2、朗读立即暂停或停止，再次点击可从断点或开头继续 3、朗读声音变为新选择的音色

3.2.8 文件引用

测试目的	输入 @ 可从本轮上传或工作区文件中选取文件，让助手明确知道要处理哪份材料。
测试步骤	1、在输入框中输入@符号 2、从菜单中选择一份特定文档并继续输入问题 3、发送包含文件引用的消息给助手
预期结果	1、弹出文件选择菜单显示本轮上传和工作区文件 2、输入框中显示被选中的文件标签且助手知晓该材料 3、助手针对引用的具体文件内容进行分析和回答

3.3 数字员工 Plan 与任务

3.3.1 Plan 执行模式

测试目的	复杂任务可先列出步骤清单再逐步执行，适合方案撰写、多文件交付等需要分阶段推进的工作。
测试步骤	1、发起一个复杂任务请求要求先列出步骤清单 2、确认计划后观察助手逐步执行每个阶段 3、任务完成后检查交付物是否符合分阶段推进的要求
预期结果	1、助手生成详细的分步执行计划并等待确认 2、系统按顺序执行每一步并实时反馈进展 3、最终成果完整覆盖所有计划步骤且逻辑连贯

3.3.2 Plan 偏好问卷

测试目的	开始重任务前，助手可先弹出选择题或填空，确认风格、范围等偏好后再动手，减少返工。
测试步骤	1、启动重任务前观察系统是否弹出偏好确认对话框 2、填写具体的偏好选项并提交 3、不填写偏好直接尝试跳过或直接执行
预期结果	1、界面弹出选择题或填空表单询问风格范围等偏好 2、助手记录偏好设置并据此开始执行任务 3、系统阻止执行并提示必须先完成偏好确认

3.3.3 Plan 计划稿同步

测试目的	任务步骤会同步保存为可读的计划文档与进度记录，必要时可从历史检查点恢复进度。
测试步骤	1、执行一个多步骤任务并观察进度记录生成情况 2、在任务执行中途模拟中断或刷新页面 3、查看保存的计划文档内容是否与实际执行一致
预期结果	1、系统同步保存可读的计划文档和实时进度记录 2、重新进入后可从历史检查点恢复任务进度 3、文档详细记录了每一步的操作和状态

3.3.4 Plan 质检与审计

测试目的	任务完成后系统可自动核对步骤是否一致，并在面板展示简要质检结果，便于发现遗漏。
测试步骤	1、等待一个复杂任务完全执行结束 2、查看任务面板展示的质检结果摘要 3、针对质检发现的遗漏项进行人工复核
预期结果	1、系统自动触发核对机制检查步骤一致性 2、面板清晰显示通过项和潜在的遗漏或错误提示 3、确认系统准确识别了未完成的步骤或偏差

3.3.5 会话任务面板

测试目的	在对话旁查看当前任务列表、每一步进展、详细说明与整体完成情况，一目了然。
测试步骤	1、在执行任务期间打开对话旁的任务列表面板 2、展开某个正在执行的任务查看详情 3、观察面板数据是否随任务执行实时更新
预期结果	1、面板展示当前所有任务列表及各自的状态 2、显示该任务每一步的详细进展说明和整体完成率 3、进度条和状态描述随执行情况动态变化

3.3.6 执行中补充说明

测试目的	任务做到一半仍可追加新要求，补充内容会在下一步执行前交给助手，并可修改或撤回排队中的说明。
测试步骤	1、在任务执行到一半时输入新的补充要求 2、在任务排队中找到刚追加的说明并进行修改 3、尝试撤回一个尚未执行的排队说明
预期结果	1、系统接收新要求并将其加入下一步执行队列 2、修改后的内容生效且将指导后续步骤执行 3、该说明被成功移除且不会影响后续流程

3.3.7 任务停止与取消

测试目的	用户可随时停止正在执行的任务；Plan 模式下停止会一并取消尚未完成的子步骤。
测试步骤	1、启动一个包含多个子步骤的 Plan 模式任务 2、在执行过程中点击停止按钮 3、检查任务状态确认是否完全停止
预期结果	1、任务开始执行并显示子步骤列表 2、当前步骤立即终止且后续未完成的子步骤被取消 3、任务状态标记为已停止且无后台进程继续运行

3.3.8 任务持久化与恢复

测试目的	任务进度会保存，刷新或稍后再进入仍可查看是否在执行；避免重复点击造成同一任务多次启动。
测试步骤	1、执行一个长时间任务并在中途刷新浏览器 2、关闭标签页后稍后重新登录平台查看任务 3、避免重复点击启动按钮防止同一任务多次运行
预期结果	1、刷新后页面仍显示任务正在执行且进度保持 2、系统保留任务状态并可查看最新进展 3、系统锁定启动按钮或提示任务已在运行中

3.4 数字员工沙箱

3.4.1 实时轨迹

测试目的	在中间工作区实时查看助手每一步在做什么、输入输出摘要，必要时可展开查看完整内容。
测试步骤	1、在中间工作区观察助手执行任务的实时动态 2、点击某一步骤的展开按钮查看详细内容 3、监控多步骤任务的整体执行流
预期结果	1、界面实时显示每一步的操作摘要和输入输出 2、面板展开显示该步骤完整的输入数据和输出结果 3、工作区清晰呈现任务流转过程无卡顿或延迟

3.4.2 链路调试

测试目的	需要排查问题时，可打开调试视图，逐条查看执行细节，便于定位卡在哪一步。
测试步骤	1、遇到任务执行异常时打开调试视图 2、逐条查看执行细节定位具体卡住的步骤 3、根据调试信息分析并解决潜在问题
预期结果	1、界面切换至调试模式并展示详细执行日志 2、日志高亮显示错误发生的位置和原因 3、能够依据详细信息找到问题根源

3.4.3 轨迹分析与总结

测试目的	将执行过程整理成通俗说明、风险点与改进建议，并支持一键复制或生成更详细解读。
测试步骤	1、任务结束后查看系统生成的执行过程总结 2、点击一键复制按钮将总结内容复制到剪贴板 3、选择生成更详细解读选项
预期结果	1、展示包含通俗说明、风险点和改进建议的报告 2、所有内容成功复制且格式保持良好 3、系统扩展报告内容提供更深入的技术分析

3.4.4 往期任务归档

测试目的	在工作区内直接预览 Word、文档、网页、PDF、表格等常见文件，少来回下载打开。
测试步骤	1、在工作区内上传一个 Word 文档并点击预览 2、上传 PDF 和 Excel 表格文件并进行预览操作 3、在预览模式下滚动浏览大文件内容
预期结果	1、工作区直接渲染显示 Word 文档内容无需下载 2、各类文件格式均能在工作区内正确显示 3、文件加载流畅且排版与原文件一致

3.4.5 文件预览

测试目的	在工作区内直接预览 Word、Markdown、网页、PDF、表格等常见文件，少来回下载打开。
测试步骤	1、在工作区上传 Markdown 文件和网页快照 2、尝试预览不同编码格式的文本文件 3、对比预览效果与本地打开的效果
预期结果	1、系统支持直接预览 Markdown 源码及渲染效果和网页内容 2、系统自动识别编码并正确显示文字无乱码 3、工作区预览内容与本地打开高度一致

3.4.6 文件共享

测试目的	把个人工作区里做好的文件发布到共享区，供团队其他成员查看或使用。
测试步骤	1、在个人工作区选择一个已完成文件点击发布 2、切换至团队成员账号查看共享区内容 3、检查发布文件的权限设置
预期结果	1、文件被成功移动到共享区并对团队成员可见 2、成员能看到并访问刚才发布的文件 3、文件权限正确配置允许团队查看或使用

3.5 工具中心

3.5.1 内置工具目录

测试目的	集中展示助手可调用的各类能力（如搜索、读写文件、计算、生成文档等），可按类别查找。
测试步骤	1、进入工具能力中心页面浏览所有可用工具 2、使用分类筛选功能查找特定类型的工具 3、点击某个工具卡片查看详细功能介绍
预期结果	1、以卡片或列表形式展示搜索、读写、计算等各类能力 2、列表仅显示符合所选类别的工具项目 3、弹出详情页展示该工具的具体用途和参数

3.5.2 工具启停控制

测试目的	用户或管理员可按需开启或关闭某些工具，控制助手能使用哪些能力。
测试步骤	1、管理员进入工具管理页面找到某个特定工具 2、关闭某个工具的启用开关并保存配置 3、普通用户尝试使用已被禁用的工具
预期结果	1、列表中显示所有可配置的工具及其开关状态 2、该工具状态变为禁用且助手无法再调用 3、系统提示该功能当前不可用或被管理员关闭

3.5.3 工具在线试调

测试目的	不进入对话也可单独试用某个工具，填写条件后看结果是否符合预期。
测试步骤	1、在不进入对话界面的情况下找到工具试用入口 2、填写工具所需的测试条件参数并执行 3、对比试用结果与预期输出
预期结果	1、独立工具测试页面打开且无聊天窗口干扰 2、系统直接运行工具并返回结果供验证 3、结果符合预期证明工具功能正常

3.5.4 工具权限与确认策略

测试目的	限制助手能调用的工具范围；对删除、发送文件等敏感操作可要求用户先确认。
测试步骤	1、配置助手可调用的工具范围限制策略 2、尝试让助手执行删除文件或发送文件等敏感操作 3、用户在确认框中点击取消
预期结果	1、策略保存成功且仅允许范围内的工具被调用 2、系统拦截请求并弹出用户确认对话框 3、敏感操作被终止且未对系统造成影响

3.5.5 工具调用审计

测试目的	查看自己在某段时间内各工具的使用次数、成功失败情况，便于了解使用习惯与排查问题。
------	--

测试步骤	1、进入个人使用统计页面选择特定时间段 2、查看统计数据中的成功与失败记录 3、分析高频使用的工具和频繁失败的工具
预期结果	1、图表或列表展示该时段内各工具的使用次数 2、清晰区分每次调用的结果状态便于排查 3、数据帮助用户了解习惯并定位问题工具

3.6 MCP 生态

3.6.1 MCP 服务管理

测试目的	接入后可查看该服务提供哪些能力与资源，并在线试用，确认可用后再用于对话。
测试步骤	1、接入一个新的外部服务后查看其能力列表 2、在页面上点击在线试用按钮测试服务连通性 3、确认服务可用后将其应用于实际对话场景
预期结果	1、界面展示该服务提供的所有资源接口和功能 2、系统调用服务并返回测试结果证明可用 3、助手能顺利调用该服务完成任务

3.6.2 MCP 发现与试调

测试目的	对接需要登录授权的外部服务时，系统自动续期授权，减少反复手工登录授权。
测试步骤	1、对接一个需要登录授权的外部服务 2、等待授权令牌过期或模拟过期场景 3、在续期后继续使用服务功能
预期结果	1、系统引导完成首次授权流程 2、系统自动检测并执行续期操作无需人工干预 3、服务调用正常无授权错误提示

3.6.3 OAuth 令牌刷新

测试目的	数字员工执行任务时，可直接调用已启用的外部工具，把其他系统能力纳入同一工作流。
测试步骤	1、配置数字员工任务流程并添加外部工具节点

	2、执行任务观察数字员工调用外部系统的行为 3、检查整个工作流的连贯性
预期结果	1、流程图显示已启用外部工具作为执行环节 2、数字员工成功调用外部工具并获取数据 3、外部系统能力无缝纳入同一工作流

3.6.4 Agent 流程集成

测试目的	数字员工执行任务时，可直接调用已启用的 MCP 工具，把外部系统能力纳入同一工作流。
测试步骤	1、在数字员工配置中启用 MCP 协议工具 2、执行涉及 MCP 工具的任务 3、验证数据在内部系统与外部 MCP 服务间的传输
预期结果	1、MCP 工具出现在可用工具列表中 2、数字员工通过 MCP 协议成功调用外部系统能力 3、数据传输准确且工作流执行顺畅

3.7 Skills 能力

3.7.1 技能中心

测试目的	浏览、搜索平台内置或导入的技能包，查看中文说明，按需启用或停用，也可试跑验证效果。
测试步骤	1、浏览技能包市场搜索感兴趣的技能 2、查看某个技能的详细信息并点击试跑 3、按需启用或停用某个已安装的 Навык
预期结果	1、展示相关技能列表含中文说明和标签 2、系统运行技能示例并展示验证效果 3、技能状态切换成功影响其在对话中的可用性

3.7.2 技能导入

测试目的	上传本地技能压缩包即可安装到平台，快速扩展助手在特定场景下的专长。
测试步骤	1、准备一个本地的技能压缩包文件 2、在技能管理页面上上传该压缩包进行安装 3、在新建对话中测试新安装的技能功能

预期结果	1、文件符合平台规定的格式要求 2、系统解析并安装成功提示扩展专长已就绪 3、助手能识别并调用新技能处理特定场景
------	--

3.7.3 SkillsHub 导入

测试目的	粘贴技能市场或代码托管平台的链接，从远端安装技能，无需手工打包上传。
测试步骤	1、复制技能市场或代码托管平台的技能链接 2、在平台粘贴链接并触发远端安装 3、检查安装日志确认远端拉取过程
预期结果	1、链接有效且指向合法的技能资源 2、系统自动下载并安装技能无需手工打包 3、日志显示下载和安装步骤顺利完成

3.7.4 技能安全与超时

测试目的	安装前检查技能包是否安全；远端安装支持中途取消与超时保护，避免长时间无响应。
测试步骤	1、上传或链接一个技能包触发安全检查 2、在远端安装过程中主动取消操作 3、模拟网络超时环境测试远端安装
预期结果	1、系统扫描并报告技能包的安全性评估结果 2、安装进程立即终止且无残留文件 3、系统在超时后自动中断并提示保护机制生效

3.7.5 自定义技能

测试目的	通过向导新建专属技能，并可自动生成中文能力说明，方便团队理解与维护。
测试步骤	1、启动新建专属技能向导并按步骤填写信息 2、利用自动生成功能创建中文能力说明 3、保存并发布新技能供团队使用

预期结果	1、向导引导顺利完成技能基础配置 2、系统生成通顺准确的中文描述文档 3、技能上架成功且团队成员可见可维护
------	---

3.7.6 技能载入审计

测试目的	统计各技能被加载使用的次数与失败情况，便于评估哪些技能常用、哪些需优化。
测试步骤	1、进入技能统计分析页面查看报表 2、筛选查看失败率较高的技能记录 3、根据统计数据评估技能优化需求
预期结果	1、展示各技能被加载使用的频次统计 2、列表突出显示失败情况及错误类型 3、数据支持决策哪些技能需改进或下线

3.7.7 导入技能删除

测试目的	对不再需要的已导入技能，可从列表中删除，保持技能库整洁。
测试步骤	1、在技能列表中找到不再需要的已导入技能 2、点击删除按钮移除该技能 3、刷新列表确认技能库整洁度
预期结果	1、目标技能清晰可见且状态正常 2、系统提示确认后成功删除技能条目 3、被删除技能不再出现在列表中

3.8 工作区与文件

3.8.1 个人工作区

测试目的	每位用户拥有独立文件空间，可新建多个工作区、设默认、重命名或归档；默认工作区删除前需先更换默认项。
测试步骤	1、进入文件空间管理创建多个新工作区 2、设置其中一个为默认工作区并重命名另一个 3、尝试删除当前默认工作区

预期结果	1、多个工作区创建成功并显示在列表中 2、默认标识更新且名称修改立即生效 3、系统阻止删除并提示需先更换默认项
------	---

3.8.2 共享工作区

测试目的	管理员维护团队共享空间；成员可查看共享内容，数字员工也可把成果发布到这里。
测试步骤	1、管理员进入团队共享空间管理界面 2、普通成员登录后查看共享内容 3、数字员工将成果发布到共享区
预期结果	1、显示共享空间列表及成员访问权限 2、成员可见授权的共享文件但不可越权修改 3、发布成功且其他成员可即时查看

3.8.3 工作区离线编辑

测试目的	在「工作区」页单独管理文件：浏览目录、在线编辑文本、上传下载，不依赖正在进行的对话。
测试步骤	1、进入工作区页面独立于对话管理文件 2、在线编辑一个文本文件并保存 3、执行文件上传和下载操作
预期结果	1、界面展示目录结构支持浏览和编辑 2、修改内容实时保存且无需依赖对话上下文 3、文件传输成功且本地与云端同步

3.8.4 预置子目录

测试目的	新建工作区时自动建好「AI 生成」「对话上传」等常用文件夹，方便助手和用户找文件。
测试步骤	1、新建一个空白工作区 2、检查工作区内是否存在 AI 生成和对话上传目录 3、在这些文件夹中进行文件操作测试
预期结果	1、系统自动初始化并创建常用文件夹 2、预置文件夹存在且命名正确 3、助手和用户能顺利存取文件到对应目录

--	--

3.8.5 路径批量检查

测试目的	助手在执行脚本或操作文件前，可先确认目标路径是否存在，减少因路径错误导致的失败。
测试步骤	1、助手准备执行脚本操作文件前检查路径 2、在路径正确的情况下继续执行脚本
预期结果	1、若路径不存在则提示错误而不盲目执行 2、脚本顺利运行并完成文件操作

3.8.6 磁盘配额

测试目的	每个工作区有容量上限；管理员可查看全站各工作区占用情况，便于容量规划。
测试步骤	1、向工作区上传大量文件直至接近容量上限 2、管理员查看全站工作区占用概况报表 3、根据报表数据进行容量规划调整
预期结果	1、系统提示剩余容量并在达到上限时阻止写入 2、报表展示各工作区大小排序和总占用量 3、管理员可依据数据决定扩容或清理

3.8.7 远程文件解析

测试目的	上传的 Word、PDF、图片等可优先交给专业解析服务转成可读内容，解析失败时再尝试本地方式。
测试步骤	1、上传一个复杂的 Word 或 PDF 文档 2、监控解析过程查看转换效果 3、模拟专业服务失败场景
预期结果	1、系统优先调用专业解析服务处理文件 2、文档被转为可读内容且格式保留良好 3、系统自动降级尝试本地方式解析

3.8.8 文档转换落盘

测试目的	文档转成 Markdown 时，正文与提取出的图片会保存到固定目录，方便后续继续提问或引用。
测试步骤	1、上传文档并触发转 Markdown 操作 2、检查提取出的图片存储位置 3、在后续提问中引用这些图片
预期结果	1、正文内容被提取并转换为 Markdown 格式 2、图片保存到固定目录且与文档关联 3、助手能准确识别并引用保存的图片

3.8.9 记忆与当日笔记

测试目的	工作区内可维护长期记忆与每日笔记；执行失败时的经验也会自动记入，供后续任务参考。
测试步骤	1、在工作区内创建长期记忆条目和每日笔记 2、执行一个失败的任务观察经验记录 3、在新任务中参考之前的经验记录
预期结果	1、笔记保存成功并可随时查阅 2、系统自动将失败原因和经验记入知识库 3、助手引用历史记录避免重复错误

3.8.10 管理员工作区治理

测试目的	管理员在确认口令后可永久删除指定工作区及其中的文件与相关对话数据，用于合规清理或回收空间。
测试步骤	1、管理员选择指定工作区发起永久删除请求 2、输入正确的确认口令并提交 3、尝试恢复已删除的工作区
预期结果	1、系统弹出确认口令输入框 2、工作区及其所有文件对话数据被彻底清除 3、系统提示数据不可恢复确认清理完成

3.9 应用市场

3.9.1 应用浏览与搜索

测试目的	以卡片形式浏览已上架应用，按名称或简介搜索，并查看版本、标签、提供方等详情。
测试步骤	1、进入应用市场浏览已上架的应用卡片 2、使用搜索框按名称或简介查找应用 3、查看应用详情页的完整元数据
预期结果	1、界面展示应用图标名称简介等信息 2、点击应用查看详情显示版本标签提供方 3、信息展示完整无误

3.9.2 应用访问

测试目的	从应用市场一键打开目标应用；访问地址支持按当前访问主机自动拼接端口，也支持填写固定网址。
测试步骤	1、在应用市场点击一键打开某个应用 2、检查访问地址是否正确拼接端口或固定网址 3、在不同主机环境下测试地址自适应
预期结果	1、新窗口或标签页加载目标应用界面 2、URL 格式正确且应用能正常访问 3、地址自动适配当前访问主机配置

3.9.3 可见范围控制

测试目的	普通用户只能看到管理员授权的应用；管理员可看到全部已上架应用。
测试步骤	1、使用普通用户账号登录并访问应用市场 2、切换至管理员账号查看应用市场 3、对比两者列表差异
预期结果	1、仅显示管理员授权可见的应用列表 2、显示全部已上架应用包括未授权的 3、权限控制生效普通用户不可见未授权应用

3.10 系统设置

3.10.1 自动语音识别配置

测试目的	配置语音识别所用服务、连接方式与识别模式，并可在保存前测试是否连通。
测试步骤	1、进入语音识别配置页面选择服务商 2、填写连接方式和密钥并点击测试连通 3、保存配置并应用到对话场景
预期结果	1、下拉列表显示支持的语音识别服务 2、系统返回测试成功或失败的具体反馈 3、语音输入功能正常调用新配置服务

3.10.2 文本转语音配置

测试目的	配置语音合成所用服务、音色、语速音量等，并可在保存前试听效果。
测试步骤	1、进入语音合成配置页面调整音色语速音量 2、保存配置
预期结果	1、音频效果符合设定的参数特征 2、设置生效且助手回复朗读时使用新配置

3.10.3 联网搜索配置

测试目的	开启或关闭联网搜索，选择搜索服务商并填写密钥，供对话与助手工具使用。
测试步骤	1、进入联网搜索配置页面开启或关闭开关 2、选择搜索服务商并填写 API 密钥 3、在对话中使用搜索功能验证配置
预期结果	1、开关状态切换并影响全局搜索功能 2、密钥保存成功且通过连通性测试 3、助手能调用配置的搜索服务获取信息

3.10.4 智能体工具 API

测试目的	集中配置搜图、生图、网页阅读、文件解析等第三方服务的开关与密钥，
------	----------------------------------

	与助手扩展能力联动。
测试步骤	1、进入第三方服务集中配置页面 2、开启某项服务并填写对应密钥 3、在助手扩展能力中联动测试该服务
预期结果	1、展示搜图生图网页阅读等服务的开关列表 2、服务状态变为激活且密钥加密存储 3、助手能成功调用新配置的第三方能力

3.10.5 数字员工模式参数

测试目的	调整数字员工的语言与时区、最多执行步数、是否自动续跑、哪些操作需用户确认等运行策略。
测试步骤	1、进入数字员工运行策略配置页 2、调整自动续跑和用户确认等操作策略 3、观察数字员工执行任务时的行为变化
预期结果	1、显示语言时区最大步数等设置项 2、策略保存成功并应用于后续任务 3、员工行为严格遵循新配置的策略限制

3.10.6 外观设置

测试目的	自定义浏览器标题、图标、侧栏与登录页的品牌名称、Logo 和欢迎文案，统一企业视觉。
测试步骤	1、进入品牌自定义配置页面 2、上传企业 Logo 并修改欢迎文案 3、保存配置并刷新登录页
预期结果	1、提供标题图标侧栏登录页等编辑入口 2、预览区域实时更新显示新品牌元素 3、全站视觉统一为企业定制样式

3.10.7 许可证配置

测试目的	查看当前授权是否有效、到期时间与授权范围；获取本机机器码并提交新的许可证密钥。
测试步骤	1、进入授权管理页面查看当前状态

	2、获取本机机器码并提交新的许可证 3、检查授权更新后的功能限制变化
预期结果	1、显示授权有效性到期时间和范围详情 2、系统验证通过并更新授权信息 3、新功能解锁或旧限制解除符合新授权范围

3.11 系统管理

3.11.1 应用管理

测试目的	维护应用市场上架的应用信息、图标、分类与访问方式，并控制上架或下架。
测试步骤	1、进入应用市场上架管理后台 2、执行上架或下架操作
预期结果	1、修改保存成功且前台展示同步更新 2、应用在前台市场的可见状态随之改变

3.11.2 模型管理

测试目的	添加和维护对话所用大模型，测试是否连通，设置默认模型，并标注是否支持图片、视频、工具调用等能力。
测试步骤	1、进入大模型管理页面添加新模型 2、设置默认模型并标注能力标签
预期结果	1、测试通过显示模型可用 2、系统默认使用该模型且能力标识准确

3.11.3 智能体管理

测试目的	维护外部智能体接入信息，测试连通性，并控制哪些智能体出现在对话页供用户选择。
测试步骤	1、进入外部智能体接入管理页面
预期结果	1、用户端列表随配置实时更新

3.11.4 工作区管理

测试目的	查看全站各工作区占用与变更概况，管理员可删除共享工作区。
测试步骤	1、管理员查看全站工作区占用概览
预期结果	1、删除成功且释放相应存储空间

3.11.5 操作日志

测试目的	查看谁在何时做了哪些关键操作（如登录、改设置、删工作区），支持按时间、用户、类型筛选。
测试步骤	1、进入审计日志页面查看操作记录
预期结果	1、列表仅显示符合条件的记录便于追溯

3.12 运维管理

3.12.1 应用商店（全部）

测试目的	浏览一体机应用商店全部应用，按标签筛选与搜索，查看详情并完成安装配置。
测试步骤	1、浏览一体机应用商店查看所有应用
预期结果	1、应用安装成功并配置完毕

3.12.2 应用商店（已安装）

测试目的	管理已安装应用：查看运行状态，启动、停止、重启、卸载，查看或修改运行参数。
测试步骤	1、进入已安装应用管理列表 2、查看或修改应用运行参数
预期结果	1、状态图标实时更新反映操作结果 2、参数保存成功且应用按新参数运行

3.12.3 应用商店（可升级）

测试目的	检查哪些应用有新版本，对比升级前后配置差异，执行升级或暂时忽略某次升级提醒。
测试步骤	1、检查应用版本更新提醒
预期结果	1、应用版本更新或提醒消失

3.12.4 设备管理

测试目的	查看一体机网络连接拓扑，配置有线网、Wi - Fi 或热点，查看硬件运行状态；部分提示可选择不再弹出。
测试步骤	1、查看一体机网络连接拓扑图 2、查看硬件运行状态并管理提示信息
预期结果	1、网络连接生效且状态正常 2、硬件指标正常且可选择不再弹出特定提示

3.12.5 大模型管理

测试目的	管理一体机预装的大模型服务：启停、连通测试、查看运行状态，配置开机自动启动顺序及语音相关参数入口。
测试步骤	1、进入大模型服务管理页面
预期结果	1、配置保存且重启后自动生效

3.13 语音能力

3.13.1 实时自动语音识别

测试目的	对话页麦克风说话时，语音实时转成文字，对接一体机语音识别服务。
测试步骤	1、在对话页点击麦克风说话
预期结果	1、延迟低体验流畅

3.13.2 自动语音识别内置工具

测试目的	数字员工可直接对工作区里的音频或视频做语音转文字，配置与系统语音识别设置一致。
测试步骤	1、数字员工处理工作区音频视频文件
预期结果	1、输出文本文件保存至工作区

3.13.3 文本转语音内置工具

测试目的	在使用通义千问实时语音识别方案时，平台已完成配套对接，保障识别效果。
测试步骤	1、使用通义千问实时语音识别方案
预期结果	1、无明显错误或延迟

3.14 沙箱运行环境

3.14.1 Docker 沙箱镜像

测试目的	助手执行脚本时使用隔离的运行环境，预装常用办公与文档处理组件，减少每次任务临时安装依赖。
测试步骤	1、助手执行脚本任务
预期结果	1、任务执行效率高

3.14.2 Python 预装库

测试目的	运行环境中已备好常用 Python 办公与数据分析库，便于直接生成 Word、Excel、图表、PDF 等成果。
测试步骤	1、运行 Python 办公数据分析脚本
预期结果	1、无缺失依赖报错

3.14.3 Node 预装依赖

测试目的	运行环境中已备好文档生成与图表渲染相关组件，便于生成复杂版式文档与插图。
测试步骤	1、执行文档生成和图表渲染任务
预期结果	1、格式美观内容完整

3.15 国际化

3.15.1 多语言国际化

测试目的	门户界面支持简体中文、繁体中文、英语、日语、韩语等，用户可在个人中心切换语言。
测试步骤	1、进入个人中心语言设置
预期结果	1、所有菜单提示均适配新语言

四、分级评价

基于企业级软件 AI 功能成熟度模型，构建了从基础级到自主级的五阶能力评价体系。该体系核心围绕架构完整性、系统协作深度、数据治理能力及自动化程度等关键维度展开评估。通过科学划分等级，界定了软件从单点辅助迈向全栈自主的演进路径，为企业的智能化转型提供参考标准，助力组织明确当前能力定位，稳步实现系统级的智能跃迁与价值重构。

级别	级别名称	成熟度分级描述
L1	基础级	单点智能，人工运维为主，数据离线或半在线。
L2	标准级	流程标准化，模型自动化部署，具备基础预测能力。
L3	融合级	全栈智能化，数据湖架构完善，具备自动化修复与韧性。
L4	协同级	跨系统协作，数据资产网格化，业务实现自迭代增强。
L5	自主级	AI 全生命周期自主管理，深度介入架构设计。