



「文枢——集团 AI 传播力中枢平台」技术实现说明

配套材料：原型演示 wenshu.html（离线可运行） | 广东省出版集团 2026 年 AI 技术创新应用大赛 · B 赛道

一、总体架构：三明治分层

事实层：结构化书目数据（书名、作者、定价、分类、在售状态），对表达层模型只读，任何环节不可写入、不可生成新事实。

表达层：大语言模型仅负责文字组织与风格变体，输入为母版事实与表达要求，输出必须回到校验层复核，无事实生成权。

校验层：确定性代码执行四道硬门禁，不依赖模型自我判断，每一步可审计、可复现、可归因。

二、原型实现

原型为单文件网页应用（wenshu.html），采用 ES5 语法与原生技术栈，零外部依赖、离线可运行，双击即可在 Chrome/Edge 等浏览器打开；数据仅存于本机浏览器（localStorage，不可用时自动降级为内存存储），不采集、不上传任何数据。

原型阶段表达层以确定性模板引擎模拟，保证演示环境合规、结果可复现；书目为演示数据，正式版接入集团真实书目库。

原型已实现五个功能模块：事实库、内容生成（五维派号）、合规拦截（四道门禁）、分发与归因（内容货架、账号登记台、归因看板）、审计留痕。其中分发模块支持机构账号排期领取与个人账号自愿取用两条通道，并可按层级汇总归因；账号登记台采集账号平台、名称与责任人，用途限于掌握底数与突发处置，不代发、不代运营。

演示省略身份认证与权限体系（属技术开发阶段实现）；实际部署时可对接集团统一身份认证，各层级、各网点按权限登录。

三、四道门禁的工程实现

①禁用语预检：维护《广告法》红线词库（如「最好」「第一」「必读」「绝对」「全网最低」等），对文稿全文做字符串匹配并定位标红，命中即拦截。

②事实一致性：抽取文稿中的书名（书名号匹配）与价格（数值+「元」），与所选书目逐项精确比对，不一致即拦截并给出证据。

③差异化度：对同书目下已放行内容计算字符二元组 Jaccard 相似度，超过阈值（0.6）判定同质化，拦截不放行。

④AIGC 标识校验：检查显式标识（「本文由 AI 辅助生成」）存在性，缺失即拦截；正式版按《人工智能生成合成内容标识办法》同时写入隐式元数据标识，导出自动注入、不提供关闭开关。

四、上线演进路径

第一步：表达层由模板引擎替换为境内合规大模型服务（如 DeepSeek、混元、讯飞星火等），调用时仅注入书目事实与表达要求，输出强制回校验层复核。

第二步：事实层升级为集团书目数据库（对接在售书目与 CNONIX 数据），并建立母版管理与五维派号调度服务。

第三步：校验层后移至服务端，防止前端绕过；接入账号体系、排期分发与归因 SDK，形成「生产—治理—度量」三环闭环。

第四步：对接集团统一身份认证（SSO），建立三级角色权限；账号登记台升级为集团账号底数台账，支持账号状态监控与突发响应联动。

五、合规与安全设计

AIGC 标识符合《人工智能生成合成内容标识办法》（2025 年 9 月 1 日施行），并按《出版业人工智能应用安全要求》行业标准方向对标落实^[1]；禁用语库随《广告法》及平台规则持续维护；书目库对模型只读，杜绝涉密内容进入公共模型；一期不涉及境外平台与数据出境，支持私有化部署；审计日志全量留痕、可追溯。

参考来源

- 靳艺昕.《出版业人工智能应用安全要求》行业标准阶段性成果发布 [N]. 中国出版传媒商报, 2026-08-28.
<https://www.cbbr.com.cn/contents/500/111647.html>