



产品白皮书

AI-OCR 智能识别系统

把文档、票据、证照、合同和报表变成可核验的结构化数据

文档定位

面向产品选型、内部立项与实施沟通。公开资料版不构成技术承诺，项目功能、接口、性能和合规边界以正式方案与合同为准。

北京开灵科技有限公司

2026 年 10 月 版本 V1.0

00

目录与阅读指引

01	一页读懂这套系统
02	现在的问题与系统边界
03	一个真实业务场景：海外商业发票自动录入
04	端到端流程与核心能力
05	数据、接口、权限与合规
06	部署方式与实施路径
07	验收口径与常见问题
08	关于开灵科技

这份材料给谁看

共享中心、运营中心、档案与数据团队、产品及信息化负责人

01

一页读懂这套系统

如果只看一页，先看这一页。它回答系统是什么、解决什么、怎么接、怎么落地。

它是什么

把文档、票据、证照、合同和报表变成可核验的结构化数据。系统以真实业务单据与过程为主线，把分散在人员、表格和不同应用中的动作组织为可执行、可追溯、可验收的闭环。

现在常见的情况	上线后要形成的状态
● 票据、证照和合同版式多，模板开发和维护成本高 ● 人工录入速度慢且难免错字，业务高峰容易积压 ● 识别结果只给文本，没有字段置信度、校验和人工复核 ● 跨境单据涉及多语种与复杂表格，难直接进入业务系统	● 业务有统一入口、状态和责任人 ● 规则前置到流程，异常单独分流 ● 关键数据一次采集、多处复用 ● 系统从结果可反查来源、审批与处理记录

与现有系统的关系	原则上不是推倒重来，而是通过接口、只读数据连接或流程协同，把现有业务系统接入统一闭环。保留哪些、替换哪些，由实施调研确定。
----------	---

02

现在的问题与系统边界

先把问题说清楚，再谈功能。系统只承诺可配置、可集成、可追溯的产品能力，不替代企业的管理判断与专业合规意见。

痛点 1 票据、证照和合同版式多，模板开发和维护成本高。	痛点 2 人工录入速度慢且难免错字，业务高峰容易积压。
痛点 3 识别结果只给文本，没有字段置信度、校验和人工复核闭环。	痛点 4 跨境单据涉及多语种与复杂表格，难直接进入业务系统。

系统解决什么

- 把高频、重复、规则明确的业务动作放进系统。
- 把关键资料、状态、例外和责任人放在同一条链上。
- 让人工集中处理判断、审批与异常，而不是反复录入和搬运。

系统不替代什么

- 不替代企业对业务真实性的负责。
- 不替代税务、法律、档案等专业判断。
- 不在缺少授权和必要材料时自动越权执行。

03

一个真实业务场景：海外商业发票自动录入

白皮书不从功能菜单开始，而从一件每天会发生的业务开始。

场景

用户上传扫描件或照片，系统先做方向、清晰度与版面检测，再识别语种、表格和关键字段。字段按业务规则校验，低置信度或勾稽失败项进入人工复核；确认后的结构化结果通过接口写入业务系统，原图和修订记录保留。

这件事为什么值得系统化

- 它跨越多个角色或系统，靠口头和表格容易断。
- 它包含可重复规则，也包含必须由人判断的例外。
- 它的结果需要追溯到来源、处理过程 and 责任人。
- 它能用清晰的输入、输出和异常清单验收。

业务闭环

1	文件接收
2	图像预处理
3	版面/语种识别
4	文字与表格抽取
5	字段结构化
6	规则校验
7	人工复核
8	接口输出

04

端到端流程

每一步都有输入、规则、输出和异常去向。下游动作只消费已经通过校验的结果。

1	文件接收
2	图像预处理
3	版面/语种识别
4	文字与表格抽取
5	字段结构化
6	规则校验
7	人工复核
8	接口输出

异常不是失败

真实业务一定可能会出现缺材料、状态冲突、接口失败或需要人工判断的情况。好的系统不是把页面都做成绿色，而是把异常原因、责任人、下一步和处理记录摆清楚。

04

核心能力矩阵

能力按业务闭环组织，不按技术名词堆砌。

通用文字 印刷体、手写体、段落、表格、印章和二维码统一解析。	证照卡证 身份证、营业执照、车辆与行业证照按字段结构化。
票据单证 发票、回单、报关单、物流单等关键字段抽取。	合同文档 条款、主体、金额、日期及印章信息抽取与比对。
多语种 多语种检测、识别与翻译，保留原文和译文对应关系。	复核工作台 低置信度、校验失败和复杂版面转人工确认并沉淀样本。

配置优先	组织、角色、规则、表单、字典、审批、提醒和接口映射优先通过配置完成。涉及外部系统限制或个性化业务逻辑时，再单独评估开发。
-----------------	---

05

和现有系统怎么接

原材料通常已经存在于业务系统。接入方式要兼顾完整性、实时性、改造成本与责任边界。

系统 / 数据源	典型接法	实施时要确认
扫描仪 / 高拍仪	开放接口 / 只读视图 / 文件交换 / 消息队列	对象、字段、频率、失败重试、责任人
移动端与文件上传	开放接口 / 只读视图 / 文件交换 / 消息队列	对象、字段、频率、失败重试、责任人
邮箱与对象存储	开放接口 / 只读视图 / 文件交换 / 消息队列	对象、字段、频率、失败重试、责任人
ERP / 业务系统	开放接口 / 只读视图 / 文件交换 / 消息队列	对象、字段、频率、失败重试、责任人
RPA / 数字员工	开放接口 / 只读视图 / 文件交换 / 消息队列	对象、字段、频率、失败重试、责任人
数据中台与档案	开放接口 / 只读视图 / 文件交换 / 消息队列	对象、字段、频率、失败重试、责任人

取数失败要可见

接口失败、账号过期、字段变化或数据不完整不能静默。系统应记录具体原因、重试结果和人工处理状态，并支持对账。

05

权限、安全与合规控制

控制点应尽量前置，并在需要人负责的地方保留确认。

控制 1 原图与结果绑定	控制 2 字段置信度展示
控制 3 业务规则勾稽	控制 4 低置信度强制复核
控制 5 修订前后留痕	控制 6 私有化与脱敏选项

通用安全原则

- 最小权限：角色只获得完成任务所需的访问范围。
- 关键操作留痕：谁在何时查看、修改、审批或导出均可追溯。
- 敏感数据保护：传输与存储按项目要求加密，可配置水印、脱敏和下载限制。
- 备份与恢复：数据库、文件和配置分别制定备份、恢复及演练机制。

06

部署方式与技术边界

部署选择由数据敏感度、既有基础设施、接口条件和运维能力共同决定。

方式	适用情况	说明
客户自有机房	有内网与运维团队	系统、数据库和文件部署在客户环境，接口按内网策略开放。
客户私有云	已有云资源和统一运维	资源弹性更好，安全、备份和监控纳入客户云平台。
混合接入	部分外部平台必须联网	核心数据留在客户环境，外部调用按最小数据原则和审批策略执行。

本产品的部署重点

识别效果与文档类型、清晰度、语言、字段定义和复核机制密切相关。正式项目应先建立样本集与验收集，按文档类型分别评测，不用单一平均准确率代替业务验收。

06

怎么落地

建议先用一条可核验的业务闭环试点，再逐步复制到更多主体、区域和场景。

1	业务调研：明确范围、角色、系统、数据、规则和例外。
2	方案确认：冻结主流程、控制点、接口清单和验收样本。
3	环境部署：完成资源、安全、账号、备份与监控。
4	配置与对接：落地组织、权限、表单、规则、字典和接口。
5	数据准备：整理主数据与必要历史资料，先校验再导入。
6	场景联调：正常、异常、回退与人工接管一起验证。
7	试运行：选定业务周期并行运行，记录问题与口径调整。
8	验收推广：按验收表关闭问题，形成运维手册后复制。

07

怎么验收

验收看业务结果、控制点和可追溯性，不只看页面是否能打开。

验收项 1 样本分类覆盖真实场景	验收项 2 关键字段单独统计
验收项 3 表格和多页关联可用	验收项 4 低置信度能转复核
验收项 5 结构化接口满足业务	验收项 6 原图与修订可追溯

建议准备三类样本

- 正常样本：完整数据能够自动跑通。
- 边界样本：金额、数量、身份、权限或状态接近规则边界。
- 异常样本：缺字段、重复、冲突、接口失败和需要人工判断。

验收记录	每个验收项应写明输入样本、预期结果、实际结果、截图或日志、责任人和关闭日期。
------	--

07

客户常问的问题

1	准确率能统一承诺吗？ 不宜脱离样本谈统一数字，应按文档类型、关键字段和实际扫描质量分别验收。
2	复杂表格能识别吗？ 支持表格结构解析，但合并单元格、遮挡和低清晰度场景需要样本验证。
3	数据必须上云吗？ 可按项目选择云端接口或私有化部署，具体依模型和算力方案确定。
4	识别错了怎么改？ 复核工作台保留原图、初始结果、修改人和最终值，并可把样本用于后续优化。

安排演示前建议准备	一条典型业务、三到五个真实但已脱敏的样本、涉及的系统清单，以及最希望系统解决的三个问题。
-----------	--

08

关于开灵科技

北京开灵科技有限公司专注企业财税数字化、档案数字化与 AI 业务应用，围绕企业已有系统提供可配置、可集成、可追溯的业务产品和实施服务。

AI 与流程自动化 数字员工、智能审核、跨系统取数与业务执行。	财税数字化 发票、乐企直连、纳税申报、费控与供应链协同。
档案与影像 电子会计档案、全门类档案、影像扫描与 OCR。	交付方式 支持私有化部署，并按项目边界提供接口、源码及实施服务。

交付上的三条原则

数据边界清楚 部署位置、外部调用、数据留存和访问权限在方案中写清楚。
异常摆在明面 不隐藏失败任务，让原因、责任人、处理结果和重试过程可见。
按业务逐项验收 用真实脱敏样本验证正常、边界和异常场景，问题关闭后再推广。

联系我们	官网 www.kailingteck.com 产品页 www.kailingteck.com/ai-ocr/ 如需系统演示或针对贵单位的实施方案，请联系开灵科技商务团队。
-------------	--

AI-OCR 智能识别系统 产品白皮书 V1.0 | 2026 年 10 月