



产品白皮书

数字化档案管理系统

把音频、视频、照片、文书、实物
和已经在跑的会计档案，管进同一套体系

北京开灵科技有限公司

2026 年 9 月 版本 V1.0

本文档中的系统界面均为产品实际截图，其中的单位名称、人员姓名、档号与业务数据为演示用示例数据，不涉及任何真实客户信息。

目录

一页读懂这套系统	03
一、现在这些资料是怎么放的	04
二、这套系统是什么	06
三、一次安全交底会，从录像到能查	08
四、档案该做的规定动作，音视频一样不少	14
五、AI 用在哪、数据去哪、花多少钱	18
六、和你现在的系统怎么接	20
七、对照国标，我们做到哪一步	22
八、怎么落地	24
九、客户常问的十个问题	26
关于开灵科技	28

这份材料是给谁看的

档案室负责人——看第三章和第四章，那里讲的是你每天要做的事在系统里怎么走。

分管领导——看"一页读懂"和第八章，一页知道这是什么，一页知道要投入什么。

信息化负责人——看第五章和第六章，数据在哪、怎么和现有系统接。

参与评审的档案专家——看第七章，那是逐条对照国标的清单。

一页读懂这套系统

如果你只肯看一页，看这一页就够。

它是什么

一套覆盖全部档案门类的数字化档案管理系统。文书、科技、会计、照片、录音录像、实物、人事、电子，八大门类（企业口径十大类）在同一套系统里管，用同一套档号规则、同一套保管期限、同一套借阅审批。

它最不一样的地方是：**音频和视频进来之后不是躺着的死文件**。系统会把里面说的话转成带时间点的文字稿，把画面里出现的字认出来，一起送进检索库。以后搜一句会上说过的话，能直接跳到那段录像的第几分几秒开始播。

它解决什么

现在的情况	用了之后
会议录像躺在经办人电脑里，人走了就没了	录完就归档，编档号、定保管期限、进库房，和纸质件一套规矩
知道录过，但翻不出来是哪一段	搜一句话，定位到第几分几秒，旁边配着转写稿
会计档案规范了，其余门类还是各管各的	全门类一套体系，会计档案是其中一类，不用推倒重来
上级来查，拿不出四性检测报告	系统按国标持续做四性检测，报告随时导出

它和你现有的系统什么关系

不替换，是**往上盖一层**。OA、财务、ERP、费控里该归档的材料，系统定时抽过来，不用业务人员再导一遍。已经接过的有泛微 OA、用友 U9、SAP S/4HANA、汇联易。

如果你已经在用我们的电子会计档案系统，这套就是它的完整版——会计档案那一类原样保留，其余七类补齐，**数据不动、流程不变**。

投入大概是什么量级

系统部署在客户自己的机房或私有云。日常运行的额外开销主要是 AI 调用费：一场 46 分钟的会议录像，转写加画面识字合计约 **1.2 元**；一个中等规模单位一年的 AI 费用在 **数百元到数千元**量级。涉密材料可以改成本地部署的模式，代价是需要一台带 GPU 的服务器。

第一章

现在这些资料是怎么放的

这一章只讲现状，不讲解决办法。如果下面四件事你觉得眼熟，这套系统才和你有关。

一、录下来了，但不知道在谁手里

安全交底会录了像，施工过程拍了视频，比价会议录了音，培训做了录播。录的时候大家都觉得“留个凭证”，录完之后文件躺在经办人的电脑、手机、或者一块移动硬盘里。

过一年再找，经办人可能已经调岗，硬盘可能已经换过，文件名叫“新建文件夹 (3)/VID_20260312_091407.mp4”。档案室不知道有这份东西，所以它从来没被归过档——它不在任何一本台账上。

二、知道录过，就是翻不出来是哪一段

出了安全事故要倒查，需要证明当时做过交底、讲过哪几条要求。大家都记得“三月份那次会上领导专门强调过整改期限”，录像也确实找到了，46 分钟。

然后就卡住了：**没人知道那句话在第几分钟**。只能派人从头听，听完一遍两个小时，还不一定听到。最后往往是放弃录像，去找会议纪要——但纪要概括的，恰恰没写那句原话。

三、会计档案规范了，其余门类还是各管各的

不少单位的电子会计档案已经做得不错：凭证、账簿、报表、发票都进了系统，有档号、有保管期限、有四性检测。

但这只是**八大门类里的一类**。文书档案还在共享盘按年份建文件夹，照片在宣传口的同事手里，录音录像在各业务部门，实物档案（印章、奖牌、证书）只有一本 Excel 台账。同一件事产生的材料散在四五个地方，**谁也说不清一件事到底留下了多少东西**。

四、上级来查，拿不出该拿的东西

电子档案要能当凭证用，国标要求做“四性检测”——真实性、完整性、可用性、安全性。检查的时候要拿出检测报告，说明这些电子件从归档到现在没被动过、该有的东西都在、格式还能打开、权限没出过问题。

现实是：文件就在共享盘上，**谁什么时候改过没有记录，副本有没有掉一份没人知道，十年前的 BMP 图片还能不能打开没人试过**。报告只能靠临时补材料，补出来的东西经不起追问。

顺带说一句保管期限

纸质件因为占地方，到期了自然会有人想起来清理。电子件不占地方，于是**存了就一直存着**，十年前的临时性材料还在盘上占着空间，真正该永久保存的反而混在里面找不到。

更麻烦的是反过来的情况：有人为了腾空间，把“看着没用的”直接删了。删的时候没有鉴定、没有审批、没有清册，**事后既说不清删了什么，也说不清是谁删的**。

这四件事的共同点

它们都不是“缺一个存文件的地方”的问题——共享盘、网盘、NAS 都能存文件。

问题在于**存进去之后，这些材料没有身份**：没有档号，所以点不着名；没有保管期限，所以不知道该留到什么时候；没有著录项，所以检索不到；内容是听的看的，所以搜索框对它无能为力；没有存证，所以证明不了它没被动过。

下一章讲的这套系统，做的就是**给每一份材料一个身份，再把听的看的变成能搜的**这两件事。

第二章

这套系统是什么

一句话：一套覆盖全部档案门类的管理系统，其中音频和视频不再是躺着的死文件。

整个系统分五段

- 1

材料怎么进来

音视频拖进来、扫描仪直连、手机拍照上传、从 OA / 财务 / ERP 定时抽取、老资料配一张 Excel 清单批量导入。五条路，进来之后一视同仁。
- 2

按门类归位

八大门类（企业口径十大类），每一类有自己的著录项模板、保管期限和归档规矩。进来的材料自动编档号、定期限、落到对应门类下。
- 3

AI 把它读懂

纸质件认字，录音录像转成带时间点的文字稿，录像画面抽关键帧再认一遍字。结果全部进检索库。这一段是音视频能被"找到"的前提。
- 4

按档案的规矩管起来

立卷归档、虚拟库房、装盒贴标、盘点、四性检测、借阅审批、到期鉴定、销毁留册。音视频和纸质件走同一套流程，一步不少。
- 5

用起来

搜内容不搜文件名。一次检索同时命中录像里说的话、画面上的字、文书正文和著录项，录像类结果直接跳到那一秒开始播。



档案总览。左边是按当前口径展开的门类树，每一类带件数。中间是门类分布，右边把"一次安全生产交底会"牵出来的八份材料摆成一棵树——这八份散在五个门类里，点任意一份能串到其余七份。

门类按国标铺，两套口径可切换

机关事业单位和企业的档案分类口径不一样。系统两套都内置了，顶栏点一下当场切换，底下的材料一份不动，只是换一种归类方式。

模板 A 机关 / 事业单位口径	模板 B 企业口径
国家档案局 8 号令 + DA/T 系列，8 个一级门类	DA/T 42-2009 企业档案工作规范，十大类
文书档案（WS）	01 党群工作类
科技档案（KJ） 下设基建、设备、产品、科研	02 行政管理类
会计档案（CW）	03 经营管理类
照片档案（ZP）	04 生产技术类
录音录像档案（LX） 下设录音、录像	05 科研类 06 基建类 07 设备类
实物档案（SW）	08 会计类
人事档案（RS）	09 人事类
电子档案（DZ）	10 声像实物类 下设录音、录像、照片、实物



顶栏切到企业口径之后。左侧菜单当场变成十大类，门类树跟着换成 01 党群工作类到 10 声像实物类。右边是这个门类的著录模板——底下的 24,806 件材料一份没动，只是换了一种归类方式。

每个门类的模板包含三部分：**建档要填什么**（哪些必填、哪些系统自动读）、**进来要做哪些 AI 加工**（转写、抽帧、识字、进检索库，逐项可开关）、**归档规矩**（一件怎么算、收什么格式、长期保存转成什么、四性怎么检、到期怎么办）。这些都在界面上配置，不用改代码。

第三章

一次安全交底会，从录像到能查

这一章不讲功能清单，讲一件事。这件事在施工企业每个季度都会发生一次。

2026 年 3 月 12 日 三季度安全生产交底会

安全环保部组织，分管领导李建国主持，37 人参加，会议 46 分 18 秒，全程录像。会上通报了上月排查出的 12 项隐患、定了整改期限、安排了新入场人员的三级教育。

这一场会开完，留下了八份材料：会议录像、会议录音、现场照片、会议纪要、隐患整改通知单、培训签到表、安全培训记录，以及一个月后那笔培训费的记账凭证。它们分别属于五个不同的门类。

第一步 录像传进来

会后经办人把 2.11 GB 的录像文件拖进系统。传的过程中系统同时在做六件事，其中只有第一件是"上传"，后面五件是人原本要手工做的事。



上传页。左边六步依次亮起：上传到档案库、读文件自带的参数、算哈希存证、送去语音转写、抽关键帧加认画面里的字、进全文检索库。右边的著录项在第二步完成时自动填好——时长 46:18、分辨率 1920×1080、码率 6.4 Mbps、拍摄设备 Sony PXW-Z150、拍摄时间 2026-03-12 09:14:07，这九项没有一项需要人输入。

大文件切片上传，断网了从断点接着传，不用从头再来。第三步的"算哈希存证"是后面四性检测的基础：这一刻文件的指纹存下来，以后每次盘点重算比对，改过一个字节都能查出来。

第二步 AI 把它读懂

上传完成之后，这份录像自动排进加工队列。**四个动作分别由不同的引擎完成**，档案员在加工中心能看到每一个任务走到哪一步、用的是谁、花了多久。

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案 1,284

科技档案 936

会计档案 12,650

照片档案 3,027

录音录像档案 148

实物档案 42

人事档案 806

电子档案 5,913

AI 加工

AI 加工中心

语音转写

AI-OCR 引擎

档案管理

立卷归档

虚拟库房

AI 加工

AI 加工中心

材料收进来之后自动排队加工。这一页是给档案员看的，那份卡住了、为什么，一眼能看见

本月转写 744 分钟录音录像 已完成 42 个任务 进行中 2 个任务 失败待处理 1 个任务 本月花费 ¥ 38.60 云端接口调用

加工动作 全部 状态 全部 搜材料题名、档号

失败的全部重试 导出加工日志

材料	加工动作	用的是	状态	耗时	操作
三季度安全生产交底会 会议录像 KL01-LX-2026-Y30-0087 46:18	语音转写	云端 阿里百炼 Paraformer	已完成	17 分钟	看结果
三季度安全生产交底会 会议录像 KL01-LX-2026-Y30-0087 46:18	抽关键帧	本机 FFmpeg	已完成	2 分钟	看结果
三季度安全生产交底会 会议录像 KL01-LX-2026-Y30-0087 46:18	帧上识字	云端 智谱 GLM-4V	已完成	3 分钟	看结果
三季度安全生产交底会 会议录像 KL01-LX-2026-Y30-0087 46:18	进检索库	本机 全文索引	已完成	12 秒	看结果
三季度安全生产交底会 会议录音（备份） KL01-LX-2026-Y30-0061 46:22	语音转写	云端 阿里百炼 Paraformer	已完成	17 分钟	看结果
三季度安全生产交底会 会议录音（备份） KL01-LX-2026-Y30-0061 46:22	进检索库	本机 全文索引	已完成	12 秒	看结果
交底会现场照片（3 张） KL01-ZP-2026-Y30-0301	帧上识字	云端 智谱 GLM-4V	已完成	3 分钟	看结果
交底会现场照片（3 张） KL01-ZP-2026-Y30-0301	进检索库	本机 全文索引	已完成	12 秒	看结果

AI 加工中心。每个任务标着走的是云端还是本机：语音转写走阿里百炼，帧上识字走智谱 GLM-4V，抽关键帧和建索引在本机完成。顶上是本月用量和花费——744 分钟录音录像，全月 AI 调用费 38.60 元。

这一页特意把失败的任务也摆在明面上。截图里那条"设备开箱验收（塔吊 QTZ80）"转写失败，原因写得很具体：音轨采样率不支持，系统已自动转码待重试。档案员不需要懂技术，只需要知道哪一份卡住了、要不要管。

这四个动作分别在做什么

动作	结果是什么	为什么需要它
语音转写	带时间点和说话人的文字稿	录像里说的话变成可以被搜索的文本，这是第三步能"点一句跳一秒"的前提
抽关键帧	每个镜头一帧，带时间点	时间轴上能预览画面，跳转时知道自己跳到了哪里
帧上识字	画面里 PPT、白板、路牌、单据上的文字	会上投影出来的整改期限、责任人、验收标准，这些字只在画面里，稿子里没有
进检索库	转写稿 + 画面文字进索引	以后搜一句话能定位到第几分几秒

这份 46 分钟的录像，转写耗时 17 分钟、抽帧 2 分钟、识字 3 分钟、建索引 12 秒，**合计不到半小时，花费约 1.2 元**。全程无人值守，档案员第二天来上班时它已经加工完了。

第三步 点一句，跳一秒

这是整套系统里最值得看的一屏。左边是录像，右边是转写稿，点稿子里任意一句，左边就跳到那一秒；画面底下同时打出那一帧里认出来的文字。

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案 1,284

科技档案 936

会计档案 12,650

照片档案 3,027

录音录像档案 148

实物档案 42

人事档案 806

电子档案 5,913

AI 加工

AI 加工中心

语音转写

AI-OCR 引擎

档案管理

立卷归档

虚拟库房

装盒与标签

盘点

四性检测

到期预警

借阅管理

鉴定与销毁

检索利用

三季度安全生产交底会 会议录像

KL01-LX-2026-Y30-0087 录音录像·录像 安全环保部

申请借阅 下载 导出著录单

录像

这一帧画面里认出来的字 05:47 整改期限（投影）
整改责任人 张×× 完成时限 2026-03-31 考核 超期扣 2 分

02:14 隐患排查通报 05:47 整改期限（投影） 08:30 培训名单 15:06 基坑监测表
24:15 整改前后对比照 43:50 会议决议

文字记录

在稿子里搜

00:00 主持人 张敏
各位同事，今天把第三季度的安全隐患排查做个交底，主要三件事：上季度问题的整改情况、这个季度的重点部位、还有新入场人员的培训安排。

01:32 主持人 张敏
先请安环部王工通报上个月的排查结果。

02:14 王海涛 安环部
上个月我们一共排查出十二项隐患，已经闭环九项，还有三项挂看。分别是三号楼东侧的临边防护、地下室的临时用电、还有塔吊的月度检测报告没归档。

03:48 王海涛 安环部
临边防护那个是材料没到位，上周已经进场，这两天就能装完。临时用电是分包单位私拉乱接，我们下了停工整改单。

05:47 李建国 分管领导
整改期限我统一一定一下，剩下这三项全部在三月三十一号之前闭环，超期的按考核办法扣分。每一项都要有整改前后的照片，归档留痕。

07:05 李建国 分管领导
点哪一句，左边就跳到哪一秒

著录项

一半系统自动填，一半档案员核对

题名	三季度安全生产交底会 会议录像	全宗	华兴建设集团有限公司	门类	录音录像·录像
形成日期	2026-03-12	形成部门	安全环保部	形成人	王海涛
保管期限	30 年	到期日	2036-12-31	密级	内部
时长	46:18	分辨率	1920×1080	文件大小	2.1 GB
格式	MP4 / H.264	拍摄设备	Sony PXW-Z150	拍摄地点	华兴建设 三号楼二层会议室

同一件事牵出来的其他材料 2026 年三季度安全生产交底会

三季度安全生产交底会 会议录音（备份）
KL01-LX-2026-Y30-0061 录音录像·录音

交底会现场照片（3 张）
KL01-ZP-2026-Y30-0301 照片档案

三季度安全生产交底会 会议纪要
KL01-WS-2026-Y30-0142 文书档案

安全隐患整改通知单（12 项）
KL01-WS-2026-Y30-0087 文书档案

安全交底培训签到表
KL01-WS-2026-Y10-0091 文书档案

2026 年三季度安全培训记录（37 人）
KL01-RS-2026-YJ-0233 人事档案·培训

记 2026-09-0117 安全培训费报销
KL01-CW-2026-Y30-0117 会计档案·记账凭证

录像详情页。当前停在 05:47——李建国说"整改期限我统一一定一下，剩下这三项全部在 三月三十一号之前闭环"。画面底下同步显示这一帧投影上的字："整改责任人 张×× 完成时限 2026-03-31 考核 超期扣 2 分"。时间轴上六个橙色竖线是关键帧，也能点。

为什么"画面里的字"单独拎出来讲

开会时投影上的内容，发言人往往不会逐字念出来。他说的是"整改期限统一一定一下"，投影上写的是"完成时限 2026-03-31，考核超期扣 2 分"。只做语音转写，这条关键信息就丢了。

所以系统把两路信息都留下来：说的话进转写稿，画面上的字进关键帧记录，两路都进检索库。第四步搜"整改期限"时，这两路会同时命中，互相印证。

再往下翻是著录项

同一页往下是这份档案的著录项：题名、全宗、门类、形成日期、形成部门、形成人、保管期限 30 年、到期日 2036-12-31、密级内部，以及时长、分辨率、大小、格式、拍摄设备、拍摄地点。灰字的那几项是系统从文件里直接读出来的，人只填了题名、密级和拍摄地点三项。

最底下是四性检测结果和存证哈希，还有"同一件事牵出来的其他材料"——那是下一节的内容。

示例数据，不涉及真实客户信息

10

第四步 半年以后，搜一句话找到它

九月份，审计要核对这笔培训费报销的真实性，需要证明这场会确实开过、确实定过整改期限。经办人只记得"领导说过三月三十一号之前"。

在检索框里输入"**整改期限**"，0.21 秒，在 24,806 件档案里命中 5 处，分布在 4 份材料上。

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案 1,284

科技档案 936

会计档案 12,650

照片档案 3,027

录音录像档案 148

实物档案 42

人事档案 806

电子档案 5,913

AI 加工

AI 加工中心

语音转写

AI-OCR 引擎

档案管理

立卷归档

检索利用

AI 全文检索

纸质件认过字、录音录像转过写，所以搜的是内容，不是文件名

🔍 整改期限

检索

换个词试试：

整改期限

基坑监测

培训签到

护栏高度

三月三十一号

✓ 录音录像的转写稿

✓ 录像画面里的字

✓ 纸质件 OCR 出来的正文

✓ 著录项（题名、关键词、摘要）

在 24,806 件档案里找到 5 处命中，分布在 4 份材料上 用时 0.21 秒

🎥 三季度安全生产交底会 会议录像

KL01-LX-2026-Y30-0087 录音录像·录像 安全环保部 2026-03-12 46:18 2 处

05:47 录像转写 ...**整改期限**我统一一下，剩下这三项全部在三月三十一号之前闭环...

05:47 画面文字 整改责任人 张×× 完成时限 2026-03-31 考核 超期扣 2 分

🎵 三季度安全生产交底会 会议录音（备份）

KL01-LX-2026-Y30-0061 录音录像·录音 安全环保部 2026-03-12 46:22 1 处

05:49 录像转写 ...**整改期限**我统一一下，剩下这三项全部在三月三十一号之前闭环...

📄 安全隐患整改通知单（12 项）

KL01-WS-2026-Y30-0087 文书档案 安全环保部 2026-03-13 1 处

第 2 页 文书正文 ...各责任单位应于 2026 年 3 月 31 日前完成整改并报安环部复查...

📄 三季度安全生产交底会 会议纪要

KL01-WS-2026-Y30-0142 文书档案 安全环保部 2026-03-13 1 处

AI 全文检索。一次搜索同时命中四种来源：录像的转写稿（05:47）、录像画面上的字（05:47 的投影）、录音备份的转写稿（05:49）、以及两份文书的正文（整改通知单第 2 页、会议纪要第 3 页）。点录像那两条，直接跳到 05:47 开始播。

这次检索为什么值钱

命中来源	位置	说明
录像转写稿	05:47	领导的原话，带语气和上下文，比纪要的概括更有说服力
录像画面文字	05:47	投影上的书面表述：责任人、时限、考核办法，是原话里没有的细节
录音转写稿	05:49	独立的第二份音源，互为印证
整改通知单正文	第 2 页	会后正式下发的文件，把口头要求落成了书面
会议纪要正文	第 3 页	决议条款，说明这不是领导临时一说，是形成了决议的

这五条合在一起，构成了一条从口头到书面、从影像到文字的完整证据链。如果没有这套系统，能拿出来的通常只有最后一份纪要。

示例数据，不涉及真实客户信息

11

第五步 一件事牵出来的全部材料

检索只解决"找到某一句话"。档案真正的价值在于把同一件事的材料连起来——这也是档案和网盘最本质的区别。

这场会牵出来的八份材料，分别落在五个门类里。在系统里它们通过同一个"事由"关联， 打开任意一份，能顺着串到其余七份。

门类	材料	档号	保管期限
录音录像	三季度安全生产交底会 会议录像 46:18	KL01-LX-2026-Y30-0087	30 年
	会议录音（备份） 46:22	KL01-LX-2026-Y30-0061	30 年
照片	交底会现场照片 3 张	KL01-ZP-2026-Y30-0301	30 年
文书	三季度安全生产交底会 会议纪要 4 页	KL01-WS-2026-Y30-0142	30 年
	安全隐患整改通知单 12 项 · 6 页	KL01-WS-2026-Y30-0087	30 年
	安全交底培训签到表 2 页	KL01-WS-2026-Y10-0091	10 年
人事	三季度安全培训记录 37 人	KL01-RS-2026-YJ-0233	永久
会计	记 2026-09-0117 安全培训费报销 8,600 元	KL01-CW-2026-Y30-0117	30 年

注意最后一行：那笔记账凭证属于会计档案。如果单位已经在用电子会计档案系统，这一类原样保留，不用重做——这套系统做的是把其余七类补齐，并且让它们和会计档案连上。

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案 1,284

科技档案 936

会计档案 12,650

照片档案 3,027

录音录像档案 148

实物档案 42

人事档案 806

电子档案 5,913

AI 加工

AI 加工中心

语音转写

AI-OCR 引擎

档案管理

立卷归档

虚拟库房

装盒与标签

档案门类·机关口径

录音录像档案

会议录音录像 访谈 工程记录

载体 全部 年度 全部 形成部门 全部 加工状态 全部 搜题名、档号，或者录像里说过的一句话

+ 上传录音 / 录像 批量导入 导出目录

题名 / 档号	时长	大小	形成日期	保管期限	密级	加工状态
三季度安全生产交底会 会议录像 主线 KL01-LX-2026-Y30-0087 录音录像 · 录像 安全环保部	46:18	2.1 GB	2026-03-12	30 年	内部	已完成
三季度安全生产交底会 会议录音（备份） KL01-LX-2026-Y30-0061 录音录像 · 录音 安全环保部	46:22	318 MB	2026-03-12	30 年	内部	已完成
3# 厂房桩基施工过程记录 KL01-LX-2026-YJ-0086 录音录像 · 录像 基建部	12:05	642 MB	2026-03-14	永久	内部	加工中 65%
钢材供应商比价会议录音 KL01-LX-2026-Y30-0085 录音录像 · 录音 采购部	1:22:30	318 MB	2026-03-15	30 年	秘密	待核对 待管理员核对
集团三十周年庆典活动全程 KL01-LX-2026-YJ-0084 录音录像 · 录像 党群工作部	2:14:08	8.4 GB	2026-02-28	永久	公开	已完成
深基坑边坡位移监测巡查 KL01-LX-2026-Y30-0083 录音录像 · 录像 基建部	08:42	396 MB	2026-03-11	30 年	内部	已完成
劳动争议调解谈话记录 KL01-LX-2026-Y30-0082 录音录像 · 录音 人力资源部	38:16	142 MB	2026-03-08	30 年	秘密	已完成
2# 地块征地拆迁现场取证 KL01-LX-2026-YJ-0081 录音录像 · 录像 经营管理部	26:31	1.3 GB	2026-03-05	永久	内部	已完成
第一季度经济活动分析会 KL01-LX-2026-Y30-0080 录音录像 · 录音 综合办公室	1:47:22	402 MB	2026-04-02	30 年	内部	加工中 31%

声像档案列表。12 份录音录像，含三种真实状态：已完成、加工中 65%、待管理员核对。加工状态摆在列表上，档案员不用点进去就知道哪些还没处理完。

这一章做完的五件事，回头看

- 1 2.11 GB 的录像有了身份**
档号 KL01-LX-2026-Y30-0087，保管期限 30 年，到期日 2036-12-31，密级内部，形成部门安全环保部。它现在在台账上，不在谁的硬盘里。
- 2 九项著录参数没有人工录入**
时长、分辨率、帧率、码率、格式、音轨、大小、拍摄设备、拍摄时间，全部从文件里直接读。人只填了题名、密级、拍摄地点三项。
- 3 46 分钟的话变成了可检索的文字**
15 段发言，标出 5 位说话人，带时间点。画面里六个关键帧上的字也认了出来。耗时不到半小时，花费约 1.2 元。
- 4 半年后一次检索 0.21 秒定位到 05:47**
同时命中转写稿、画面文字、录音备份和两份文书，构成完整证据链。
- 5 八份材料串成了一件事**
跨五个门类，打开任意一份能串到其余七份。包括那笔已经在会计档案系统里的记账凭证。

这一章有意没讲的事

整章没提“我们支持多少种格式”“并发多少”“存储多大”。那些参数在实施方案里有，但它们不是客户真正关心的问题。客户关心的是：**这件事在我这儿每季度发生一次，用了之后会怎么样。**下一章讲的是另一半——这份录像进了档案库之后，它要走完哪些档案本来就该走的规定动作。

第四章

档案该做的规定动作，音视频一样不少

上一章那份录像归档之后，它要和纸质件一样，走完档案本来就该走的八个动作。

动作	纸质件怎么做	音视频怎么做
立卷归档	按年度、门类、期限、部门组卷，编卷号	一样组卷编号，只是不装进实体盒子
进库房	档案室某架某层某位	数字库房某区，三份副本分别在机房、异地、蓝光库
装盒贴标	装盒、打盒标签、贴条码	不装盒，但一样有卷号和位置，可扫码定位
盘点	扫条码核在不在架上	重算哈希核有没有被改过，探活三份副本还在不在
四性检测	主要看实体状况	真实性靠哈希、完整性看转写稿在不在、可用性看格式能不能打开
借阅	登记、取出、归还	批准后在线看，打申请人姓名水印，不给下载原件
到期鉴定	到期提醒，鉴定决定续存还是销毁	完全一样
销毁	碎纸，留销毁清册	三份副本一起彻底擦除，擦除过程留痕，一样留清册

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案1,284

科技档案936

会计档案12,650

照片档案3,027

录音录像档案148

实物档案42

人事档案806

电子档案5,913

AI加工

AI加工中心

语音转写

AI-OCR引擎

档案管理

立卷归档

按「年度 + 门类 + 保管期限 + 责任部门」自动组卷，录音录像和纸质件用同一套规则

年度2026

门类全部

状态全部

搜卷号、卷名

+ 新建卷

按规则批量立卷

导出卷内目录

卷号	卷名	件数	保管期限	责任部门	立卷日期	状态	操作
KL01-LX-2026-Y30-J001	2026 年安全生产类 录音录像 第 1 卷	14 件	30 年	安全环保部	2026-04-02	已立卷	看卷内目录
KL01-LX-2026-YJ-J002	2026 年工程施工过程记录 第 1 卷	9 件	永久	基建部	2026-04-08	待审核	看卷内目录
KL01-ZP-2026-Y30-J011	2026 年活动照片 第 2 卷	132 件	30 年	党群工作部	2026-03-30	已立卷	看卷内目录
—	待整理：2026 年一季度零散音视频	43 件	—	多部门	—	待整理	按规则自动立卷

立卷归档。按"年度 + 门类 + 保管期限 + 责任部门"四个条件自动组卷。最后一行 43 件待整理的零散音视频，点一下按规则自动分成 6 个卷，等档案员审核。

示例数据，不涉及真实客户信息

14

存放：数字件也要有"在哪儿"

纸质件存在几架几层几位，数字件也一样要有位置和副本数——出了事要说得清它在哪、还有几份。



虚拟库房。三个数字库房区加一个实体库房，容量和占用一目了然。音视频在 A 区，三份副本分别在机房、异地和蓝光归档库。下面是最近上架记录，每一件都标着存在哪、有几份副本。

纸质件还是要装盒贴标



装盒与标签。列表最后一行是"2026 年录音录像 第 1 卷 (148 件)"，位置写的是"数字库房 A 区 (不装盒)"——数字件不装盒，但一样编卷号、一样有位置。下面那张是盒标签的实际版式，含起止档号、件数、存放位置、到期日和条码，扫一下就能调出这一盒的卷内目录，盘点时不用开盒。

四性检测：能不能当凭证用，看的就是这四项

电子档案要能当凭证用，国标要求持续做四性检测。系统按批次跑，结果可以随时导出成报告。

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案 1,284

科技档案 936

会计档案 12,650

照片档案 3,027

录音录像档案 148

实物档案 42

人事档案 806

电子档案 5,913

AI 加工

AI 加工中心

语音转写

档案管理

四性检测

真实性、完整性、可用性、安全性。这是电子档案能当凭证用的前提，国标要求必须做

本次受检 24,806 件

四项全通过 24,801 件

真实性存疑 1 件

完整性缺项 2 件

格式待转换 2 件

这四项各查什么

真实性

归档时算哈希存证，之后每次校验重算比对。对不上就说明被动过，要查操作日志

完整性

该有的元数据、该有的附件、该有的签章、音视频该有的转写稿，缺一项报一项

可用性

格式能不能正常打开，淘汰格式有没有转成长期保存格式

安全性

权限配得对不对、加密有没有、三份副本在不在、操作有没有全程留痕

结果 全部 门类 全部

搜档号、题名

发起一次全量检测

导出检测报告

档号	题名	真实性	完整性	可用性	安全性	检测时间	操作
KL01-LX-2026-Y30-0087	三季度安全生产交底会 会议录像	✓	✓	✓	✓	2026-09-14	—
KL01-WS-2026-Y30-0142	三季度安全生产交底会 会议纪要	✓	✓	✓	✓	2026-09-14	—
KL01-LX-2026-Y30-0078	设备开箱验收 (塔吊 QTZ80) 完整性不通过: 转写稿缺失 (音轨采样率不支持, 加工失败)	✓	✗	✓	✓	2026-09-14	去处理
KL01-ZP-2016-Y30-0204	2016 年工程进度照片 可用性不通过: BMP 属于淘汰格式待转换, 哈希与归档时不一致	✓	✓	✗	✓	2026-09-14	去处理

四性检测。这一批 24,806 件里有 5 件没全通过，不通过的原因写得很具体：转写稿缺失、BMP 属于淘汰格式待转换、哈希与归档时不一致。

关于那条"哈希不一致"

截图里"2015 年归档邮件"真实性一栏是叉。查操作日志之后的结论是：2023 年做存储迁移时重新打包造成的，不是被人动过内容，已重新存证。

真实运行的系统不可能全是绿色对钩。重要的不是"从不报警"，而是报了警能查清楚、查清楚后有处理记录。一个从来不报错的四性检测，多半是没在真跑。

借阅：批下来也只能在线看

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案 1,284

科技档案 936

会计档案 12,650

照片档案 3,027

录音录像档案 148

实物档案 42

人事档案 806

电子档案 5,913

档案管理

借阅管理

批下来只能在线看，画面打申请人姓名水印，不给下原件。谁在什么时候看过哪一份，全程留痕

状态 全部 密级 全部

搜申请人、档号、题名

导出借阅台账

申请单号	申请人	要看的档案	事由	申请日期	可看到	状态	操作
JY-2026-0331	周立 财务部	三季度安全生产交底会 会议录像 KL01-LX-2026-Y30-0087 密级 内部	审计要核对培训费报销的真实性	2026-09-15	2026-10-15	待审批	审批
JY-2026-0328	陈静 人力资源部	安全交底培训签到表 KL01-WS-2026-Y10-0091 密级 内部	核对新入场人员三级教育记录	2026-09-12	2026-09-26	已批准	在线查看
JY-2026-0319	外部 市应急管理局	消防演练全过程记录 KL01-LX-2026-Y30-0079 密级 内部	安全生产标准化复评现场核查	2026-09-02	2026-09-09	已归还	—
JY-2026-0311	吴强 经营管理部	钢材供应商比价会议录音 KL01-LX-2026-Y30-0085 密级 秘密	合同纠纷取证	2026-08-28	—	已驳回	—

借阅管理。四条申请分别处于待审批、已批准、已归还、已驳回。最后一条因为密级是秘密、事由是合同纠纷取证，被驳回了。

批准之后申请人只能在线看，画面打着姓名和时间水印，下载按钮是灰的；什么时候打开过、看了多久都记在利用记录里。

示例数据，不涉及真实客户信息

16

到期：到期不等于删

这是档案口客户最在意、也最容易被软件做错的一件事。系统的口径是：**保管期限到了只触发提醒，不触发删除。**

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案 1,284

科技档案 936

会计档案 12,650

照片档案 3,027

录音录像档案 148

实物档案 42

人事档案 806

电子档案 5,913

AI 加工

AI 加工中心

语音转写

AI-OCR 引擎

档案管理

立卷归档

档案管理

鉴定与销毁

到期不等于该删。先鉴定，三条都排除了才进销毁流程，而且要双签、留清册

鉴定要回答的三个问题

还有没有查考价值

这份材料以后还会不会有人要看、能不能证明什么事

有没有官司或审计牵扯

正在打的官司、正在做的审计、监管正在查的事，牵扯到就不能销毁

上级有没有要求续存

行业主管部门、集团总部有没有下文要求继续保存

三条都是「否」，才能提同意销毁。同意销毁的必须档案室和分管领导双签，销毁前生成销毁清册——清册本身永久保存，哪一年销毁了什么、谁批的，永远查得到。数字件销毁是三份副本一起彻底擦除，擦除过程也留痕。

状态 全部 门类 全部 搜鉴定单号、题名

+ 发起鉴定

销毁清册

鉴定单号	鉴定对象	鉴定意见	签批	经办	状态	操作
JD-2026-0021	2016 年三季度安全例会录音等 12 件 录音录像档案 到期 2026-12-31	继续保存（续存 10 年）	待分管领导签批	李文娟	鉴定中	看鉴定单
JD-2026-0020	2016 年度考勤与办公用品单据 86 件 文书档案 到期 2026-12-31	同意销毁	档案室已签 / 分管领导已签	李文娟	待销毁	执行销毁
JD-2026-0019	2015 年度施工日志 42 件 科技档案·基建 到期 2025-12-31	转为永久保存	已签批	刘伟	已续存	看鉴定单
JD-2025-0016	2014 年度零星报销单据 214 件 会计档案 到期 2024-12-31	同意销毁	已签批	周立	已销毁	看鉴定单

鉴定与销毁。四条鉴定单分别处于鉴定中、待销毁、已续存、已销毁四种状态。"签批"一列写明谁签的、什么时候签的——**同意销毁的必须档案室和分管领导都签，缺一个走不下去。** 另有一张"到期预警"页列出 90 天内到期的 17 件，系统提前提醒，不会到点自动清理。

鉴定要回答的三个问题

还有没有查考价值——这份材料以后还会不会有人要看、能不能证明什么事

有没有官司或审计牵扯——正在打的官司、正在做的审计、监管正在查的事，牵扯到就不能销毁

上级有没有要求续存——行业主管部门、集团总部有没有下文要求继续保存

三条都是"否"才能提同意销毁。同意销毁的必须**档案室和分管领导双签**，**销毁前生成销毁清册，清册本身永久保存。**

销毁的是材料，不是记录

这一点值得单独说。系统销毁的是材料本体——纸质件碎纸、数字件三份副本一起彻底擦除，擦除过程留痕。 但"**哪一年销毁了什么、谁批的、依据是什么**"这条记录永久保存。

以后有人问"2016 年的考勤表哪儿去了"，档案室能拿出销毁清册 XH-2026-0007， 上面写着鉴定意见、双签人、销毁日期。这和"文件找不到了"是两回事。

示例数据，不涉及真实客户信息

17

第五章

AI 用在哪、数据去哪、花多少钱

客户在这一章一定会问三个问题：准不准、数据出不出机房、贵不贵。我们逐个回答。

一、准不准



语音转写。右上是专有名词表——把单位名、人名、专业词、规范号加进去， 准确率能提 3 到 5 个点。下面是最近五次转写的实际准确率。

材料类型	准确率	说明
会议录音录像（室内、有麦）	92 — 96%	正常会议场景，可直接用
谈话记录（无麦、多人插话）	88%	低置信度句子会标出来，人工订正
现场演练（噪音大、距离远）	86%	同上，建议现场配麦
印刷体文件 OCR	99%	红头文件、合同、报表
录像关键帧 OCR（投影、白板）	95%	投影反光和拍摄角度会影响
手写签名与批注 OCR	82%	一律标成待人工核对，不直接进检索库

最后一行我们不藏着：手写只有 82%。做法是把手写部分一律标成"待人工核对"， 档案员核完才进检索库。宁可慢一步，不让错的内容混进能被检索到的正式记录。

二、数据去哪

环节	在哪运行	数据怎么走
档案库、数据库、全文索引	客户机房 / 私有云	始终不出客户环境
抽关键帧、格式转换	客户机房	本机 FFmpeg 处理，不出网
语音转写	云端 阿里百炼	音频传到云端，转写完即时删除，不留存不训练
OCR 识字	云端 智谱 GLM-4V	图片传到云端，识别完即时删除，不留存不训练

也就是说：**档案本体和索引永远在客户自己的机房里**，只有待处理的那份音频或图片走一次云端。

涉密材料怎么办

两种做法，按客户的保密要求选：

- 按门类关闭 AI 加工**——在门类模板里把语音转写和 OCR 关掉，这一类材料只走归档流程，不做 AI 处理，一个字节都不出网。密级为"秘密""机密"的可以默认关闭。
- 整体改为本地部署模型**——转写和 OCR 都换成部署在客户机房的模型，全流程不出网。代价是需要额外配一台带 GPU 的服务器。

这两个开关在系统里是配置项，不需要改代码，也可以只对部分门类生效。

三、贵不贵



AI-OCR 引擎。本月识别 12,486 页，花费 10.20 元。

场景	用量	费用
一场 46 分钟的会议录像（转写 + 抽帧 + 识字）	1 份	约 1.2 元
一个月的录音录像加工	744 分钟	28.40 元
一个月的 OCR 识别	12,486 页	10.20 元
中等规模单位全年 AI 费用	—	数百元至数千元

这个量级通常低于客户预期。真正的投入在部署和实施，不在 AI 调用。

第六章

和你现在的系统怎么接

档案系统不产生业务数据，它的原材料全在业务系统里。接不进去，就只能靠人一份份导。



业务系统取数。每个系统标着怎么接的、抽什么过来、多久跑一次、最近一次什么时候。SAP 那一行是失败状态，原因写在下面的取数日志里：对方接口返回 401，服务账号口令过期。

已经接过的系统

业务系统	怎么接的	抽什么过来
泛微 OA	走对方接口	公文、会议纪要、用印审批，归到文书档案
用友 U9	直连数据库	凭证、账簿、报表、发票影像，归到会计档案
SAP S/4HANA	OData 接口	物料主数据、采购订单、供应商资料
汇联易	走对方接口	报销单、对公支付、借还款、合同

两种接法，各自的前提

接法	什么时候用	需要客户提供
走对方接口	对方有现成的开放接口，或者愿意开发。首选这种，对业务系统零侵入	接口文档、测试环境、一个服务账号
直连数据库	对方没有接口，或者接口不覆盖要归档的数据。只读，不写	只读账号、表结构说明、网络策略放行

关于取数失败

真实项目中最常见的取数问题不是技术问题，而是对方系统的服务账号口令到期了没人通知。

系统把失败原因写清楚，连续失败三次给档案员发提醒。它不会默默停在那里，让人三个月后才发现数据断了。

历史资料怎么进来

新产生的材料靠对接自动流进来，**存量的老资料要专门搬一次**。这通常是实施阶段工作量最大的一块。

档 数字化档案管理系统

门类口径 机关 / 事业单位 全宗 华兴建设 李文娟

首页

档案总览

档案采集

音视频上传

历史资料批量导入

扫描仪归档

业务系统取数

档案门类·机关口径

文书档案 1,284

科技档案 936

会计档案 12,650

照片档案 3,027

录音录像档案 148

实物档案 42

人事档案 806

电子档案 5,913

AI 加工

AI 加工中心

语音转写

AI-OCR 引擎

档案管理

档案采集

历史资料批量导入

一个文件夹几千份老资料，配一张 Excel 清单一起拖进来，系统按清单逐条建档、编档号、排进加工队列

今年已导入 10,095 件 本月 1,327 件 进行中 1 批 待处理失败 5 条

状态 全部 归到门类 全部 搜批次号、批次名

+ 新建导入 导出导入台账

批次号	这一批是什么	总数	进去了	没进去	状态	操作
DR-2026-0042	2016-2020 年会议录音（移交盘） 归到 录音录像档案 李文娟 2026-09-12 14:22	1,284	1,279	5	已完成 5 条：Excel 第 118、226、301、477、883 行的文件名在文件夹里找不到	看失败清单
DR-2026-0041	2021-2025 年工程竣工资料扫描件 归到 科技档案·基建 李文娟 2026-09-08 09:15	8,642	8,642	0	已完成	看导入报告
DR-2026-0040	历年荣誉证书与奖牌照片 归到 实物档案 赵倩 2026-08-30 16:40	126	126	0	已完成	看导入报告
DR-2026-0043	2026 年一季度零散音视频 归到 录音录像档案 李文娟 2026-09-16 10:05	43	—	0	进行中 38%	看导入报告

历史资料批量导入。四个批次里，2016–2020 年的会议录音那批 1,284 件进了 1,279 件， 5 件失败。点"看失败清单"能看到具体是 Excel 第几行、为什么没进去。

步骤	做什么
整理清单	一行一件，必填四列：文件名、题名、形成日期、形成部门。保管期限和密级不填就按门类模板的默认值走
先校验再导入	清单缺列、文件对不上、文件名重复，逐行报错到第几行， 改完再导，不会导到一半卡住
分批提交	大批量分批走，断了从断点接着导
出导入报告	成功多少、失败多少、每一条失败的具体原因
失败单独重导	改好清单或补上文件，只重导那几条，不用整批重来

截图里那五条失败的原因很典型：三条是清单里写的文件名在文件夹里找不到， 一条是文件名含中文空格对不上， 一条是文件本身损坏读不出时长。 这些问题在任何一次历史资料搬迁里都会出现，系统能做的是把它们一条条摆出来，而不是给一个"失败 5 条"了事。

第七章

对照国标，我们做到哪一步

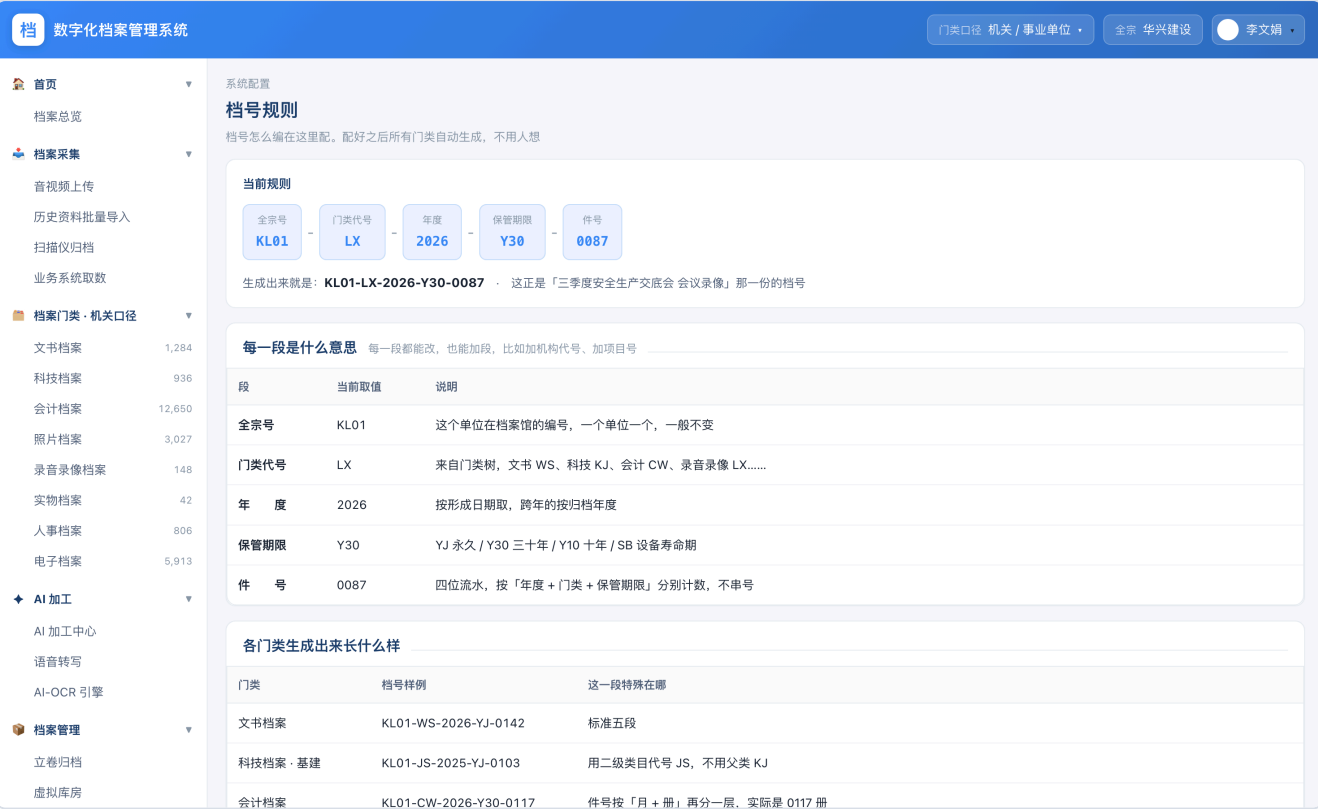
这一章是给档案专业人员和评审专家看的。逐条对照，不含糊。

涉及的主要标准

标准	管什么
DA/T 78—2019 录音录像档案管理规范	录音录像档案的收集、整理、著录、归档与管理，是本系统声像部分的主要依据
DA/T 42—2009 企业档案工作规范	企业档案的分类门类、业务工作方法，是企业口径十大类的依据
GB/T 18894 电子文件归档与电子档案管理规范	电子文件归档、四性检测、长期保存格式
国家档案局第 8 号令	机关文件材料归档范围和文书档案保管期限
DA/T 62—2017 / DA/T 63—2017	档案数字化规范、档案元数据方案
财会〔2015〕79 号	会计档案管理办法，本系统会计档案门类的依据

逐条对照

标准要求	系统里怎么做的	在哪一页能看到
按门类分类管理	两套国标口径内置，可切换，门类树可配置	系统配置 → 门类与著录模板
档号编制规则	全宗号-门类-年度-保管期限-件号，五段可配	系统配置 → 档号规则
著录项与元数据	每个门类一套著录模板，区分必填与自动读取	系统配置 → 门类与著录模板
保管期限划定	按门类配默认期限，支持导入客户自己的期限表	系统配置 → 保管期限规则
以件（案卷）为管理单位	一次活动的全部录音录像 + 情况说明 = 一件	门类模板 → 归档规矩
四性检测	真实性靠哈希存证比对，完整性查必备项，可用性查格式，安全性查权限与副本	档案管理 → 四性检测
长期保存格式转换	统一转 MP4/H.264，原件另存一份不动	门类模板 → 归档规矩
多套备份	本地 + 异地 + 蓝光三副本，盘点时逐一探活	档案管理 → 虚拟库房 / 盘点
利用登记与留痕	借阅两级审批，在线查看打水印，利用记录留档	档案管理 → 借阅管理
鉴定与销毁	鉴定三问、双签、销毁清册永久保存	档案管理 → 鉴定与销毁



档号规则。五段拆开摆出来，每一段的取值和含义写清楚。下面列了各门类实际生成出来的样子，包括几处特殊情况：科技档案用二级类目代号（JS）而不是父类（KJ），电子档案的保管期限段写 GEN 表示跟随原门类。

有一件事必须说在前面

系统里那张保管期限表，正式上线前必须换掉

系统内置了一套通用的保管期限默认值（文书永久/30 年/10 年、基建永久、会计按财会 79 号文……），目的是让系统装完就能跑起来。

但它不是依据。每个单位都有自己那张经本单位档案部门审定、报同级档案行政管理部门备案的《归档范围和文书档案保管期限表》，那张才是依据。

实施阶段的第一件事就是把客户那张表导进来替换默认值。系统支持按门类、按二级类目细配，改动留版本，能查是谁在什么时候改的。改了规则之后，存量档案的到期日默认立刻回算，不会出现"列表里还是老日期、要靠人再算一遍"的中间状态。

我们没做的事

为了不让客户产生错误预期，这几件事目前不在系统范围内，需要另行评估：

- 与国家或地方档案馆的移交接口——各地接口标准不一，需按当地档案馆的技术要求单独对接
- 电子签章与可信时间戳的第三方认证——系统留有接口，但 CA 服务需客户自行采购
- 人脸识别——技术上支持，但涉及个人信息处理，默认关闭，需客户书面确认后才开启
- 实体库房的温湿度、消防等物联监测——不属于本系统，可与第三方系统对接展示

第八章

怎么落地

这一章讲实施：装在哪、分几步、要多久、需要你配合什么。

装在哪

部署方式	适用情况	说明
客户自有机房	有机房、有运维、对数据出网敏感	全套装在客户内网。默认方案
客户私有云	已有阿里云/华为云等私有云资源	同上，只是承载环境不同
完全离线	涉密单位，网络物理隔离	AI 改为本地部署模型，需额外配一台 GPU 服务器

分几步

- 1

调研与基础配置

确定门类口径（机关还是企业）、导入客户自己的保管期限表、配档号规则、建组织架构和角色权限。这一步决定后面所有材料怎么归位，最不能省。
- 2

系统部署与验收

装环境、部署系统、配三副本存储、开通 AI 接口（或部署本地模型）。部署完逐页走查验收。
- 3

业务系统对接

按优先级逐个接：OA、财务、ERP、费控。每接一个要对方提供接口文档或只读账号，这一步的进度往往取决于对方系统厂商的配合速度。
- 4

历史资料迁移

整理清单、校验、分批导入、处理失败条目。工作量取决于存量规模和原始资料的规整程度，通常是整个项目里耗时最长的一块。
- 5

培训与试运行

档案员培训、业务部门培训，试运行一个归档周期，边跑边调。

需要你配合什么

事项	谁提供	什么时候要
经审定的归档范围和保管期限表	档案室	第一步，最关键
组织架构与人员名单	人力资源部 / 信息中心	第一步
服务器资源与网络策略	信息中心	第二步
业务系统接口文档或只读账号	信息中心 + 原厂商	第三步
历史资料及其清单	各业务部门 + 档案室	第四步

第一项是整个项目的地基。期限表没定下来，材料进了系统也不知道该留到什么时候，后面回算会很麻烦。我们建议在合同签订前就启动这一项。

如果你已经在用电子会计档案系统

这是常见情况，也是最省事的一种。会计档案那一类**原样保留，数据不动、流程不变**，这套系统在它旁边把其余七个门类补齐，并且让它们和会计档案连上。

原有部分	怎么处理
会计档案的数据	不迁移、不重建，原样保留
会计档案的业务流程	不改，档案员的操作习惯不变
已经接好的财务系统对接	继续用，不重新对接
档号规则、保管期限、权限体系	扩展到新门类，保持一致
全文检索	扩展索引范围，一次检索覆盖全部门类

试点还是全铺

我们建议**先从声像档案试点**，理由有三条：

- 声像是目前最没人管、最容易丢的一类，做出来效果最明显
- 存量规模通常不大（几百到几千件），迁移压力小，见效快
- "搜一句话跳到第几分几秒"这个能力，业务部门用一次就会主动往里传材料—— **这比行政推动有效得多**

声像跑顺了再铺文书、科技、实物，最后把人事和电子档案接进来。每铺一类都是加一套门类模板和一条对接，不动已经在跑的部分。

关于工期

基础配置和系统部署通常以周计，业务系统对接和历史资料迁移以月计，但具体周期**高度取决于两件事**：客户那张保管期限表定下来需要多久，以及存量资料的规整程度（有清单的和没清单的，工作量差好几倍）。

我们不在白皮书里承诺具体天数。实施方案会在调研之后给出分阶段的时间表。

第九章

客户常问的十个问题

1 我们的录音录像涉密，数据能不出机房吗？

能。两种做法：按门类把 AI 加工关掉，这一类只走归档流程，一个字节不出网；或者整体改成本地部署的模型，全流程不出网，代价是要配一台带 GPU 的服务器。两个开关都是配置项，也可以只对密级为“秘密”“机密”的材料生效。

2 转写准确率不够怎么办？

正常会议场景在 92—96%。低于这个数的通常是现场噪音大、无麦、多人插话。三个办法：把单位名、人名、专业词加进专有名词表（能提 3—5 个点）；低置信度的句子系统会标出来，档案员对着录音订正；重要会议建议现场配麦。订正后的稿子会重新进检索库，不是改了没用。

3 手写的东西认不准，那不是白做了？

手写准确率约 82%，这个数我们不藏。系统的做法是把手写部分一律标成“待人工核对”，不直接进检索库。宁可慢一步，也不让不确定的内容混进能被检索到的正式记录。

4 保管期限到了，系统会自动删吗？

不会。到期只触发提醒，要走鉴定流程。鉴定要回答三个问题（还有没有查考价值、有没有官司审计牵扯、上级有没有要求续存），三条都否才能提销毁，而且要档案室和分管领导双签，销毁前生成清册，清册永久保存。

5 老资料几千几万份，怎么进去？

配一张 Excel 清单，和文件夹一起批量导入。先校验再导入，清单缺列、文件对不上、重名都会逐行报错到第几行。大批量分批走，断了从断点接着导，失败的可以单独重导，不用整批重来。这一块通常是实施里耗时最长的，但不需要客户逐条手工录。

6 能不能只上一部分门类，不全铺？

可以，而且我们建议这么做。先从声像档案试点，跑顺了再铺其他门类。每铺一类是加一套门类模板和一条对接，不影响已经在跑的部分。

7 我们已经有电子会计档案系统了，要推倒重来吗？

不用。会计档案那一类原样保留，数据不动、流程不变、已接好的财务对接继续用。这套系统把其余七类补齐，并且让它们和会计档案连上，一次检索能覆盖全部门类。

8 AI 要花多少钱？会不会越用越贵？

按量计费，用多少算多少。一场 46 分钟的会议录像约 1.2 元，一个月 744 分钟录音录像的转写费 28.40 元，一个月 12,486 页 OCR 费 10.20 元。中等规模单位全年在数百元到数千元量级。加工只在材料入库时做一次，**不是每次检索都要花钱**，所以不会随着使用频次上涨。

9 要不要买 GPU 服务器？

默认方案不需要——转写和 OCR 走云端，本机只做抽帧和索引，普通服务器就够。只有选择"AI 全部本地部署"（涉密场景）时才需要，这时要单配一台带 GPU 的服务器。

10 万一你们的服务停了，数据还在吗？

在。档案本体、数据库、全文索引全部在客户自己的机房里，不在我们这边。长期保存格式统一转为 MP4/H.264、PDF 这类通用格式，**不依赖我们的软件也能打开**。档案目录、著录项、卷内目录都支持导出成标准格式。这是档案系统的基本要求，也是我们坚持私有化部署的原因。

还有一个没人明着问、但一定会想的问题

"这套东西上了之后，我们档案室的人会不会被替掉？"——不会，而且相反。系统替掉的是**录入著录项、翻找材料、手工统计**这些事，留给档案员的是**鉴定、审批、判断该不该销毁**这些必须由人来负责的事。这些恰恰是档案工作真正有价值、也必须留痕问责的部分。

关于开灵科技

北京开灵科技有限公司，专注于企业财税数字化与档案数字化，为客户提供可私有化部署、源代码可交付的业务系统。

我们做的几件事

方向	内容
数字化档案管理	全门类档案管理，音视频 AI 加工与内容检索，四性检测与长期保存
电子会计档案	凭证、账簿、报表、发票、回单的自动归档与关联，与主流财务系统对接
发票与开票	数电发票开具、反向开票、乐企联用、进销项发票采集
AI-OCR 与智能取数	票据识别、单据结构化、业务系统取数与接口封装

资质

- ISO9001 质量管理体系认证
- ISO27001 信息安全管理体系统认证
- ISO20000 信息技术服务管理体系认证
- ISO14001 环境管理体系认证 / ISO45001 职业健康安全管理体系认证
- 公安部信息系统安全等级保护三级
- 多项软件著作权证书

交付方式上的三条承诺

- 1 私有化部署**
系统装在客户自己的机房或私有云，档案本体与索引不出客户环境。
- 2 数据可带走**
长期保存格式采用通用标准，档案目录与著录项支持标准格式导出，不依赖我方软件也能打开。
- 3 逐页验收**
部署完成后按页走查验收，不是交完就走。

联系我们

北京开灵科技有限公司

官网 www.kailingteck.com

如需安排系统演示或提供针对贵单位的实施方案，请联系与您对接的商务人员。