零七零六（苏州）科技有限公司 · 自查工具

企业知识库体检表

10 项检查，判断你的知识库能不能用：看不懂 / 追不到 / 不敢用

版本：v1.0 · 发布日期：2026-10-31 · 建议用法：打印 A4 双面，逐项勾选 · 配套阅读：官网博客《上传文档，不等于建好知识库》

本表可自由打印与内部传阅，不得删改后另行发布。体检结论需结合现场数据判断，本表只回答「你填的这十项里有哪些短板」，不构成诊断结论。

使用说明

这份表用来回答一个问题：**已经建好的知识库，到底能不能用。**

不讲理论，只做 10 项检查。每项只有三个结果：通过、不通过、待验证。

- **打印**：建议 A4 双面，现场对着系统逐项过一遍；
- **时间**：熟练后约 40 分钟；
- **参与人**：业务使用方 1 人 + IT/数据 1 人（**不要只让 IT 填**，否则「敢不敢用」那三项一定全勾通过）；
- **判定口径**：凡需要「再确认一下」的，一律记「待验证」，不要记「通过」。

> 判断标准不是「传了多少份文档」，而是三条：**答得对、找得到出处、过期的不出现。**

打分表

#	检查项	通过标准	结果
1	图表、公式、非标准排版的解析	抽 10 份最难的文档，答出来的内容能与原文对应上	☐通过 ☐不通过 ☐待验证
2	文档层级结构统一	核心文档能说清「这一节讲什么、适用于哪台设备」	☐通过 ☐不通过 ☐待验证
3	答案可定位到章节	每条答案都能指到原文位置（页码或章节号）	☐通过 ☐不通过 ☐待验证
4	更新有机制	有明确规则：谁改、多久同步一次	☐通过 ☐不通过 ☐待验证
5	旧版本不再被检索	问一个只有新版才有的问题，答案是新版内容	☐通过 ☐不通过 ☐待验证
6	每条知识有生效时间与负责人	随机抽 10 条，能说清谁负责、什么时候生效	☐通过 ☐不通过 ☐待验证
7	冲突处理	指出存在冲突并给出两个来源，而不是挑一个答	☐通过 ☐不通过 ☐待验证
8	拒答能力	问到没有答案的问题时，明确拒答或转人工	☐通过 ☐不通过 ☐待验证
9	权限分级	工艺参数与客户名单的可见范围可以分开控制	☐通过 ☐不通过 ☐待验证
10	有失败样例集	至少 20 条真实问题，覆盖正常/边界/冲突/无答案	☐通过 ☐不通过 ☐待验证

通过数：____ / 10

三项必做的现场验证动作

填表之外，务必做这三个动作 —— 它们比表格更能暴露问题。

动作一：问一个只有新版才有的问题

目的：验证版本治理。

做法：找一份最近改过的规程，问一个只有新版才写进去的细节。

不合格表现：答出旧版内容，且语气很自信。

这意味着：知识库是静态快照，模型会持续输出过期答案。这在设备台账、操作规程类场景里几乎是必现问题。

动作二：问一个跨文档冲突的问题

目的：验证冲突处理。

做法：找两份对同一参数给了不同值的文档，问「这个参数是多少」。

合格表现：指出「存在两个说法，分别来自 A 文档第 X 节与 B 文档第 Y 节」。

不合格表现：随机挑一个答，且不给来源。

这意味着：知识库在做「择一回答」，一线无法判断该信谁。

动作三：问一个没有答案的问题

目的：验证拒答能力（最容易被跳过，但最关键）。

做法：问一个业务上合理、但知识库里确实没有的问题。

合格表现：明确说「知识库里没有这个信息」，或转人工。

不合格表现：编一个听起来合理的答案。

这意味着：在制造业场景里，编一个错的参数比说「我不知道」糟糕得多 —— 这一项不过，整套系统不该上生产。

结果解读

通过项	结论	下一步
0–3 项	还停在「文件柜」	先补内容结构化与版本机制，再谈问答效果；此时接模型只会放大问题
4–7 项	能用，但不敢信	优先补「出处可追溯」与「拒答能力」两项 —— 它们决定一线敢不敢用
8–10 项	可以做场景扩展	接到具体流程（客服、工单、培训），并用失败样例集做上线前回归

> 注意：无论多少分，**「冲突处理」与「拒答能力」单项不过，都不建议直接接入对客流程**。

三个改进动作（按 ROI 排序）

1. **给每条知识补「生效时间 + 负责人」**：工作量最小、收益最大，直接解决「追不到」。
2. **建 20 条失败样例集**：从一线真实提问里收，不要自己编；它是后续所有验收的基线。
3. **把「出处」加进回答模板**：答不出出处就不许答 —— 一条规则能同时解决「不敢用」与「乱编」两类问题。

体检之后

体检的交付物是「一页报告 + 三个改进动作」，**不含系统改造**。

如果体检发现是结构性问题（解析质量差、无版本机制、无权限分级），改造要单独立项：先把问题说清，再谈怎么修。

本表可自由打印与内部传阅，不得删改后另行发布。体检结论需结合现场数据判断，本表只回答「你填的这十项里有哪些短板」，不构成诊断结论。