

机票行程单查重 · 使用文档

批量上传行程单 PDF，自动识别发票号码比对重复，标注重复于哪张文件，一键导出查重结果 Excel。

一、功能介绍

「查重」是机票行程单处理工具的第三个功能页签，解决"同一批行程单重复报销、重复入账"的问题：

- 自动识别发票号码：无需手工核对号码，系统自动从每个 PDF 提取行程单发票号码比对
- 重复溯源：不止标记"重复"，还列出与哪张文件重复，方便回溯排查
- 多页合并 PDF 支持：一个 PDF 含多张行程单时，全部发票号码参与比对，任一号重复即整文件标记
- 三态结果：唯一（首次出现）/ 重复（与之前文件撞号）/ 未识别（扫描件等无法提号）
- 批量高效：单次最多 300 个文件，超过 50 个自动分批查重，跨批结果统一汇总
- 一键导出：查重结果导出 Excel（文件名 / 发票号码 / 查重结果 / 重复于 4 列）
- 隐私安全：网站不保留任何形式的用户数据，用完即删

适用人群：财务差旅报销审核、行程单台账查重、归档前批量核验。

二、详细操作步骤

第 1 步：选择文件

三种方式任选：

- 点击上传：点击上传区 → 弹出文件选择框 → 多选 PDF 文件
- 拖拽上传：从资源管理器拖一批 PDF 到上传区
- 整个文件夹：点「选择整个文件夹」→ 文件夹内全部 PDF 自动加入（含子文件夹）

要求：PDF 电子客票行程单（有文字层），单文件 ≤ 20MB，单次 ≤ 300 个。

第 2 步：等待查重

松开鼠标后系统自动开始识别发票号码并比对，状态区显示"正在识别发票号码..."。超过 50 个时自动分批处理（每批 50 个），跨批结果统一汇总。

第 3 步：查看结果

查重完成后，下方显示：

- 统计栏：共 N 个文件 / 唯一行程单 N 个文件 / 重复 N 个文件 / 未识别 N 个
- 结果表：文件名 / 发票号码 / 查重结果 / 重复于 4 列；重复行红色标注，未识别行灰色斜体
- 重复判定规则：同一发票号码第 2 次及以后出现标记「重复」，首次出现保留「唯一」

第 4 步：导出结果

点「导出查重结果」→ 下载 Excel 文件（如 `行程单查重结果\_20261009-153000.xlsx`），4 列数据可直接用于报销审核留档。

重新查重

上传区常驻显示，选择新文件即可开始新一轮查重（旧结果自动替换）。

三、常见问题解答

Q1：怎么判断重复？

系统自动识别每个行程单的发票号码（右上角红色 20 位号码），同一号码与之前文件相同即标记「重复」，并列出具体的重复文件；号码首次出现的文件保留为「唯一」。

Q2：一个 PDF 里有多张行程单怎么算？

多页 PDF 会提取全部发票号码（号码列分号分隔展示）。只要其中任一号码与之前文件重复，该文件即标记「重复」。

Q3：能处理多少文件？

单次最多 300 个 PDF（选文件夹时自动加入全部；超过 50 个自动分批查重，跨批结果统一汇总）。识别在服务器独立进程进行，完成后即删，不保存任何数据。

Q4：「未识别」是什么意思？

该文件无法解析出发票号码：可能是扫描件（图片式 PDF）、损坏文件或不是行程单的 PDF。这类文件不计入重复，建议人工确认。

Q5：文件会被保存吗？

不会。网站不保留任何形式的用户数据，用完即删。

Q6：下载的文件在哪里？

在浏览器的默认下载文件夹（通常是「下载」目录）。点击浏览器右上角的下载图标可直接找到，也可按 Ctrl + J 打开下载列表定位。

如果有其他问题或需求，可以联系技术处理，技术微信：ljq101450