



简明指南 · 代理 · 个人数据脱敏

还在用马克笔涂黑文件？ Redact Agent 标记，由您决定。

任何文件发出之前，第三方的姓名、地址和标识信息都必须去除。本简明指南说明 Redact Agent 如何在 PDF 和扫描件中标记个人数据，您如何检查预览，以及如何生成清理后的副本。适用于公共机关、法务部门和数据保护部门。

核心要点

13 个场景和 16 个可选数据类别，白名单和您自定义的词条，每个区域均附有原因的预览。被脱敏的文字会从副本中删除，原件保持不变。

01 · 问题所在

文件必须发出， 第三方的姓名不能外泄。

文件、专家报告和往来信函中，包含与申请毫无关系的人员的姓名、地址、电话号码和银行详情。发出之前，需要把页面打印出来，用笔涂黑，再重新扫描。这既费时，又无法撤回一处遗漏。

成因

黑框不等于脱敏

在 PDF 编辑器中画上的黑色方块，往往只是盖在文字上面。文字底下仍可复制，元数据或附件也可能在不知不觉中一并发出。

扫描件首先是图像

扫描页面包含的是像素，而不是文字。文本搜索找不到其中的姓名，因此每一页都必须人工阅读。

每种情形都有自己的规则

在数据主体查询中，申请人必须保持可见；在公开发布中，所有姓名都必须消失。每次都从头判断，会导致脱敏标准不一致。

如何察觉

- 文件被打印出来，用笔涂黑，再重新扫描。
- 没有人能确定黑框底下是否仍有文字。
- 查询、法院和公开发布各自遵循着各自不成文的规则。
- 已脱敏的副本与原件并排放置，没有清楚的标记。
- 文件较大时，脱敏要等上好几天，因为每一页都要手工完成。

02 · PAPEROFFICE 的解决方式

让它标记、检查， 然后放行副本。

Redact Agent 读取 PDF（包括扫描件），并标记所选场景中的数据类别。在预览中，您对每一个区域作出决定。随后，清理后的副本会作为新文档生成。

1

为情形选择场景

13 个场景设定相应规则，例如数据主体查询、对外共享、法律程序或病历。您也可以自行选择类别。

2

读取并标记

文字识别同样能读取扫描件。AI 会标记所选类别中的发现内容，从姓名和地址到银行详情、证件号码和健康数据。

模块: AI-OCR

3

在预览中决定

点击区域即可进行脱敏或取消，也可以拖动一个矩形框住遗漏的部分。悬停时会显示类别和原因。白名单和您自定义的词条适用于每一页。

4

真正删除文字

框内的文字会从 PDF 中删除，之后再次检查。如果底下仍有残留，副本将由脱敏后的页面图像生成。元数据和附件会被清除。

5

保存并发送副本

副本带有后缀 _redacted，保存在所选 Workspace 中，原件保持不变。通过电子邮件发送 会连同副本打开安全发送功能。

模块: 安全邮件发送

03 · 应用中的操作

从已标记的文档 到清理后的副本。

您可以在预览中处理单个文件，也可以用已保存的 Redact Agent 处理整批文件。两者使用相同的场景、类别和名单。

1

通过预览开始脱敏

在文档列表中选择文件，打开 **操作** > **代理与 workflows** > **Redact** 并选择 **通过预览编辑**。若要向第三方发布，**外部共享** 场景较为适用。

2

维护名单

在 **请勿删除（白名单）** 下，输入应保持可见的名称，例如您的机关名称。在 **额外隐藏（黑名单）** 下，添加您自己的词条，例如档案编号或项目名称。

3

审核并应用

在预览中点击区域或拖动矩形框，然后点击 **应用隐藏**。副本将保存到所选 Workspace，或下载到您的设备。

4

可读而非涂黑

在 **输出** 下，选择搭配 Ultra AI 模型的假名化。姓名和标识符会被替换为一致的虚构值，对应关系显示在对话框中。

5

用代理处理批量文件

在 **Document AI** > **Redact Agent** 下，保存一个包含场景、名单和目标 Workspace 的代理。对于大量 PDF，选择 **操作** > **代理与 workflows** > **Redact** 和 **使用 Redact Agent 自动执行**；处理在服务器端运行，无需预览。

实践经验

照片、印章和签名用矩形框处理 在预览中，请仔细检查每一页上的护照照片、印章或签名。用鼠标拖动矩形框覆盖这些区域，它们就会和已标记的文字一样从副本中删除。

04 · 案例

一个县政府机关。 一份查阅档案的申请。



一个县政府机关收到一份查阅某建设项目档案的申请。档案中包括专家报告、往来信函以及邻居提交的扫描异议书。档案管理员需要放行这份档案，但要让邻居的姓名、地址和电话号码无法辨认。

使用前

- 整份档案被打印出来，逐页阅读。
- 邻居的姓名和地址用马克笔涂黑。
- 页面被重新扫描，使底下的文字无法辨认。
- 完成的副本作为未加保护的邮件附件发送给申请人。

使用 PAPEROFFICE 后

- 档案以 PDF 形式保存在 Workspace 中，包括扫描的异议书。
- 使用 外部共享 场景，Redact Agent 标记姓名、地址和电话号码。
- 档案管理员在预览中检查，并用矩形框覆盖一个签名。
- 副本通过带密码的安全发送发出，原件仍保存在档案中。

这是一个用于说明的虚构示例。请您自行确认，或与您的法律或数据保护顾问一同确认，哪些信息必须脱敏，以及脱敏是否足以满足您的目的。

05 · 检查清单

发放文件前的 十个要点。

请与相关部门和数据保护负责人一起逐一核对这些要点。每一点都可以用是或否回答，并能显示您需要改进之处。

- ☐ 对于每种情形，已明确适用哪个场景：查询、共享、法院或公开发布。
- ☐ 必须保持可见的姓名已列入代理的白名单。
- ☐ 每个单独文件在放行之前都经过预览检查。
- ☐ 只放行带有后缀 _redacted 的副本。
- ☐ 对于批量处理，每种情形都已保存一个搭配 Ultra AI 模型的 Redact Agent。
- ☐ 放行之前，有人会在所有 Workspace 中搜索相关姓名，包括扫描件。
- ☐ 您自己的词条，例如档案编号，已列入始终脱敏的词条清单。
- ☐ 包含照片、印章和签名的页面，在预览中会单独检查。
- ☐ 清理后的副本保存在单独的 Workspace 中，并设有独立的组权限。
- ☐ 副本通过带密码的安全发送发出，而不是作为开放附件发送。

06 · 常见问题

档案管理员最先会问的问题。

被脱敏的文字真的消失了吗？

是的。框内的文字会从 PDF 中删除，然后 PaperOffice 检查底下是否仍有文字残留。如有残留，副本将由脱敏后的页面图像生成。元数据、附件、表单字段和链接都会被清除。

原件会被修改吗？

不会。副本是一份新文档，名称带有后缀 _redacted，保存在所选 Workspace 中或作为下载文件提供。原件保留在原处，保持不变。

假名化是什么意思？

PaperOffice 不再使用黑框，而是插入一致的虚构值，文档仍可阅读。此功能搭配 Ultra AI 模型使用；对应关系显示在对话框中，绝不会出现在 PDF 里。健康数据和访问凭据仍保持涂黑。

扫描页面如何处理？

文字识别读取扫描件，AI 标记的类别与数字 PDF 中相同。应用脱敏时，图像区域会被覆盖，PDF 随后重新生成。

Redact Agent 会标记哪些数据？

您可以从 16 个类别中选择，包括姓名、地址、电话号码、电子邮件地址、银行详情、税号和证件号码、出生日期以及健康数据。场景会预先勾选相应类别。在预览中，您可以用矩形框覆盖签名和照片。

能否将脱敏功能集成到我们自己的系统中？

可以。通过 REST API，POST /documents/redact-agent-run 可启动已保存的 Redact Agent；通过 MCP，工具 po_documents_redact_agent_run 可完成同样的操作。结果与应用中生成的清理副本相同。

07 · 如何开始

用预览功能脱敏您的下一份文件。

在 app.paperoffice.ai 创建账户，上传一份文件，并通过 [操作](#) > [代理与工作流](#) > [Redact](#) 使用 [通过预览编辑](#) 进行脱敏。费用按量计算，详情请参见[定价页面](#)。

按情形定义场景

与数据保护负责人商定，哪种场景适用于查询、共享和公开发布，以及哪些姓名应保持可见。

保存代理

在 [Document AI](#) > [Redact Agent](#) 下，为每种情形保存一个 Redact Agent，并设置场景、名单和目标 Workspace。

从测试开始

对一份包含扫描件的文件进行脱敏，然后在副本中尝试选中并复制框下的文字。

本主题相关的产品数据表

AI-OCR

paperoffice.ai/downloads/paperoffice-ai-datasheet-ocr-zh.pdf

智能搜索

paperoffice.ai/downloads/paperoffice-ai-datasheet-smart-search-zh.pdf

Workspaces

paperoffice.ai/downloads/paperoffice-ai-datasheet-workspaces-zh.pdf

安全邮件发送

paperoffice.ai/downloads/paperoffice-ai-datasheet-audit-safe-email-delivery-zh.pdf

本主题更多内容

个人数据脱敏

paperoffice.ai/zh/ai-pii-redaction-wenshu-tuoming-hua-he-shuju-tuoming-hua



本主题更多内容

paperoffice.ai/zh/ai-pii-redaction-wenshu-tuoming-hua-he-shuju-tuoming-hua

价格

paperoffice.ai/en/pricing

联系我们

paperoffice.ai/zh/lianxi