

关于双层 PDF

什么是双层 PDF 文件？

PDF 文件是一种具有多层结构的格式文件，其特点是：文件既可以是文本型的（比如由 word 生成的文件），也可以是图像型的（比如由扫描生成的文件）；双层 PDF 文件是指文件内容既包含文本层，也包含图像层，且其位置上下一一相对应。

使用双层 PDF 文件有什么好处？

PDF 文件格式已成为国际标准，越来越多的行业 and 单位正在使用它。其特点是具有多层结构，包括图层和文字层，这样，既可以 100% 保留原始版面效果（包括公章、签名），又可以通过下层的文字信息支持选择、复制、全文检索等功能。因此，双层 PDF 同时兼顾视觉效果和检索方便性，极大地方便了电子文件的管理。

双层 PDF 文件加工是什么意思？

是将 tif、jpg、bmp、gif、png、单层 PDF（图像）等图像类文件经加工后输出为可全文检索的双层（图层和文字层）PDF 文件。

双层 PDF 是这样的 PDF 文件：用户在阅读 PDF 文件时看到的是扫描图像，可以 100% 保留原始版面效果（包括公章、签名），在需要的时候，又可以通过下层的文字信息支持选择、复制、检索等功能。

word excel ppt txt 在转换 pdf 时里面的字体资源都是内嵌到到

pdf 文件中去的可以对文件进行复制、全文检索但是在数字化过程中，纸质文档或是扫描成图片或是扫描成 pdf 但是对文档的搜索复制都是不能实现的， 要想实现此功能就需要进行 ocr 识别将图片或是单层 pdf 文件中的字体资源提取出来转成双层的 pdf 实现文件的复制、选择和检索功能 这就是双层 pdf 的优势。