



Harness the power of PDF

What is iText pdfSweep?

pdfSweep is an iText 7 add-on for Java and C# (.NET) that removes (redacts) information from a PDF document in a reliable and secure way. pdfSweep intervenes as you edit a PDF document with iText 7's document stamping and watermarking tools. After adding a digital "blackout bar"

over the sensitive text, image or part of an image, pdfSweep changes the document's rendering instructions causing the hidden content of your digital document to become impossible to extract. This works for both text and images, affording you full information security.



Smart tip: You can use the `tentativeCleanUp` method to add redaction annotations to a document, allowing an end-user to choose which redactions to keep or delete.

How does iText pdfSweep work?

Looking at the advantages of pdfSweep and the data security it offers, you may find it surprising that it only takes five lines of code to integrate pdfSweep into your document workflow.

```
try (PdfDocument pdf = new PdfDocument(new PdfReader(SRC), new PdfWriter(new File(SRC, "redact")))) {
    final ICleanupStrategy cleanupStrategy = new RegexBasedCleanupStrategy(
        Pattern.compile("Alice", Pattern.CASE_INSENSITIVE)).setRedactionColor(
            ColorConstants.PINK);
    final PdfAutoSweep autoSweep = new PdfAutoSweep(cleanupStrategy);
    autoSweep.cleanUp(pdf);
}
```

Java code example

```
PdfDocument pdf = new PdfDocument(new PdfReader(SRC), new PdfWriter(DEST));
ICleanupStrategy cleanupStrategy = new RegexBasedCleanupStrategy(new Regex(
    @"Alice", RegexOptions.IgnoreCase)).SetRedactionColor(ColorConstants.PINK);
PdfAutoSweep autoSweep = new PdfAutoSweep(cleanupStrategy);
autoSweep.CleanUp(pdf);
pdf.Close();
```

C# code example



Smart tip: pdfSweep can be combined with pdfOCR to redact sensitive content from scanned documents or images.



Harness the power of PDF

Features



Automatic removal of words and phrases

Remove text from a document based on patterns like regular expressions.



Customized removal areas

Offers you the ability to remove content as necessary, just like a digital black bar.



Secure and reliable removal

As well as the visual appearance that is rendered when viewing or printing the PDF document, pdfSweep also takes care of the underlying rendering instructions and data structures to ensure the removed information is not retrievable.



Flexible options

Use recurring data or data fields to automate redaction throughout any volume of documents, with a set of predefined patterns for common data such as social security numbers, account numbers, ID numbers etc... Define custom redaction areas using coordinates to redact any content within.



Partial removal of text and images

When content is partially covered by a redaction area, it is only partially removed, allowing you to remove selected parts of text and images.



As pdfSweep is an iText 7 add-on, you also get all the benefits and features of iText 7 Core.

We're also developers!

In our 20 years of code, we know how important it is to have good documentation, and good processes in place. You can always find comprehensive documentation and code examples online. Commercial customers can also count on our world-class in-house customer support, offering direct access to expert assistance from our development team.

Get iText pdfSweep

Scan the QR-code or go to the link below:

itextpdf.com/itext7/pdfSweep

