

# Handling Skewed Documents

Understanding the Power of Solid Framework

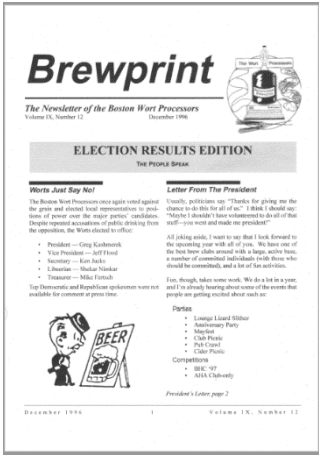
# What do we mean by “Skewed”?

Only applies to scanned documents.

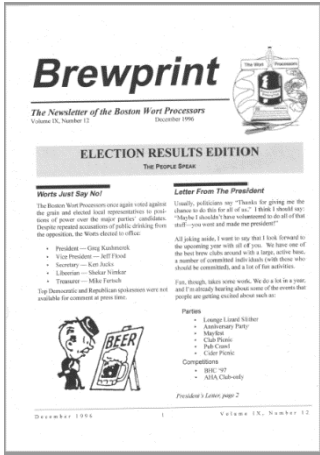
Caused by the original document being placed on the scanner at a slight angle and so the scanned image is also at an angle.

# Creation of the Sample File

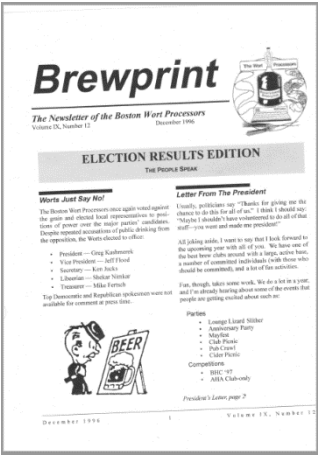
A single page was physically scanned five times, each with a different amount of skew, and then combined into a PDF.



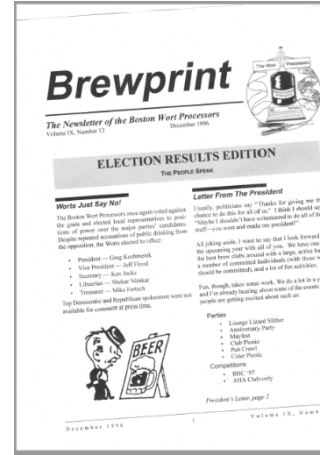
No skew



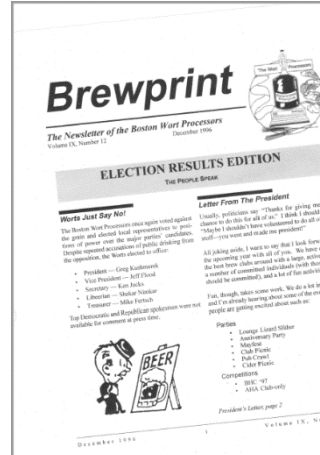
1°



2°



4°

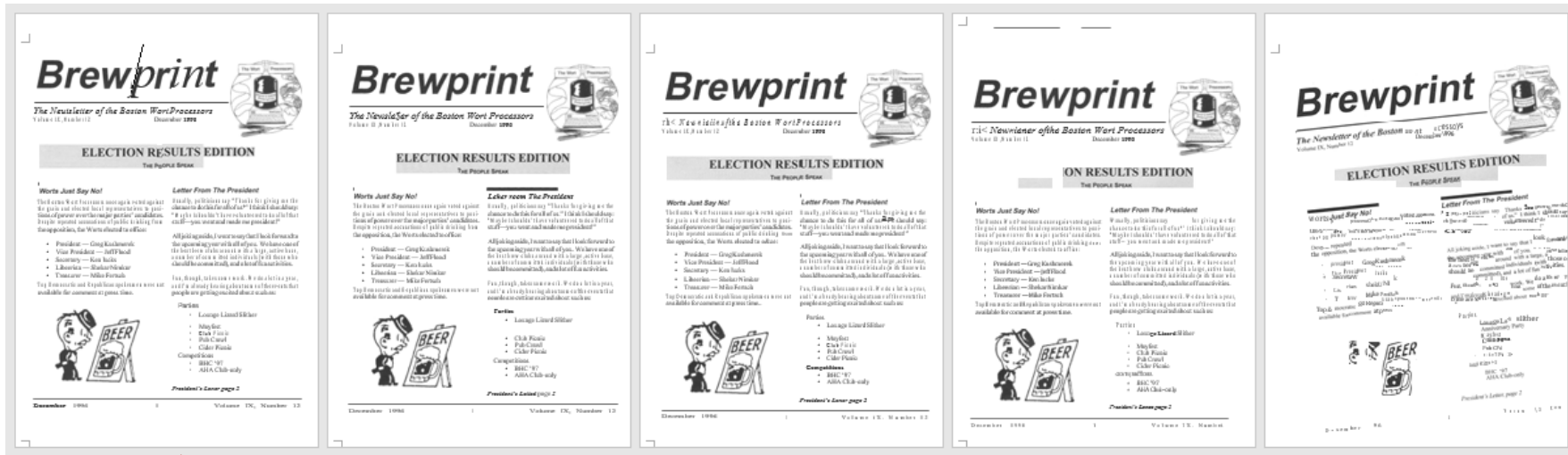


6°

# Solid Framework and Document Reconstruction

Solid Framework detects the presence of skewed text, and *automatically* corrects it before performing OCR.

# Reconstructing the Sample File back to Word



More accurate

Less accurate

The Skew here is so large that it cannot be corrected

Therefore, without any user intervention, deskewing allowed much of the text to be recovered:

Fun, though, takes some work. We do a lot in a year,  
and I'm already hearing about some of the events that  
people are getting excited about such as:



Fun, though, takes some work. We do a lot in a year,  
and I'm already hearing about some of the events that  
people are getting excited about such as:

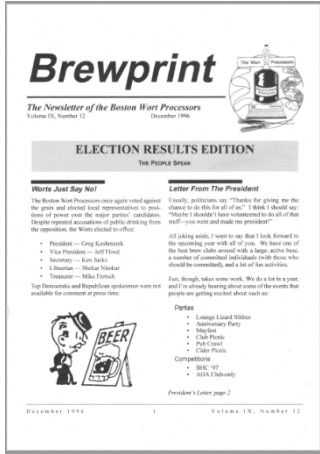
# Getting Information about the amount of “Skew”

Skew angle can be found programmatically

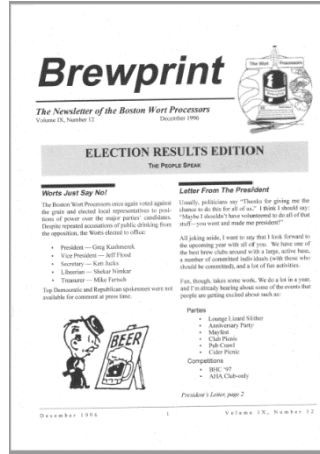
```
using (var doc = new PdfDocument(sourcePath))
{
    doc.Open();
    var pgs = doc.GetPages();
    foreach (var p in pgs)
    {
        var angle = p.Deskew();
        System.Diagnostics.Debug.WriteLine("Page angle was " + angle);
    }
}
```

**Important:** Do not call PdfDocument.RecognizeDocument before calling Deskew().

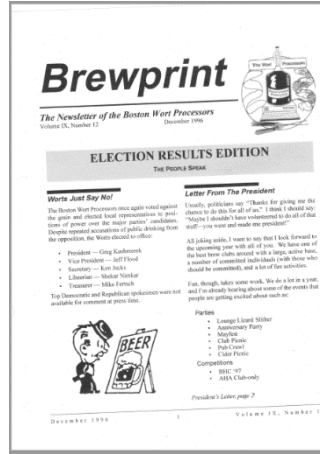
RecognizeDocument causes automatic Deskewing to occur, so there is nothing left to Deskew, and the returned value will always be 0.



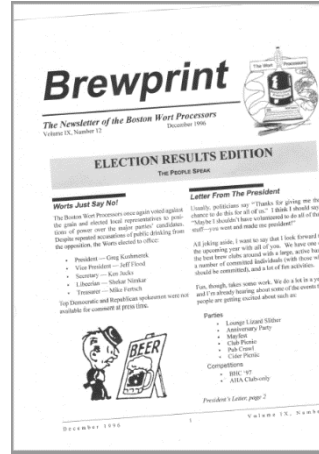
0.0°



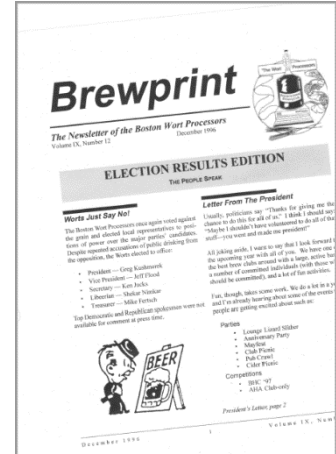
0.9°



2.2°



3.9°



0.0°

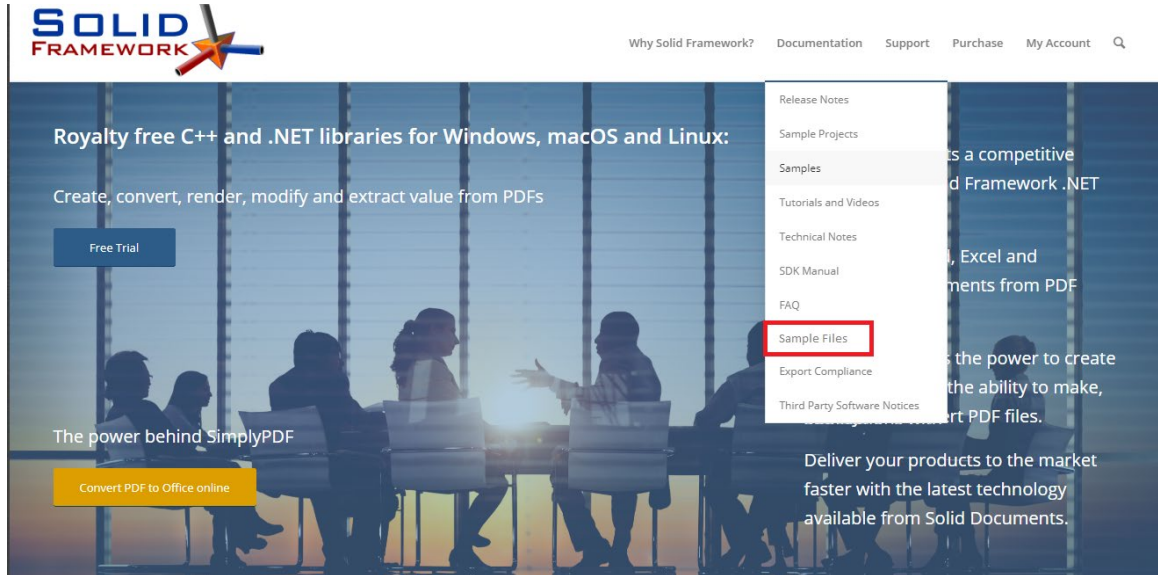
Skew angle is returned as a Double

The skew here is so extreme that Deskewing cannot occur



# Try it yourself

The PDF used for this demonstration can be downloaded from the Solid Framework web site.



The screenshot shows the Solid Framework website. The logo is in the top left, and the navigation menu is in the top right. A dropdown menu is open, with 'Sample Files' highlighted by a red box. The main content area features a large image of people in a meeting, with text describing the company's services and a 'Free Trial' button.

**SOLID FRAMEWORK**

Why Solid Framework? Documentation Support Purchase My Account

Royalty free C++ and .NET libraries for Windows, macOS and Linux:

Create, convert, render, modify and extract value from PDFs

Free Trial

The power behind SimplyPDF

Convert PDF to Office online

Deliver your products to the market faster with the latest technology available from Solid Documents.

- Release Notes
- Sample Projects
- Sample Files**
- Tutorials and Videos
- Technical Notes
- SDK Manual
- FAQ
- Export Compliance
- Third Party Software Notices

For more Samples and Tutorials, and for  
the latest version of Solid Framework  
please visit

[www.solidframework.net](http://www.solidframework.net)