Legislature committees have agendas and those are shown on the legislature website as PDF files. In order to process these files, Scout2 must first convert each to text.

Scout2 uses PDFSharp to translate PDF documents to text.

PDFSharp is open source, available on SourceForge, and comes with two zip files.

* PDFToText\_bin.zip, containing two DLLs: itextsharp.dll and SharpZipLib.dll. These have been placed in D:/CCHR/Projects/Scout2/PdfToText. References to them have been added to the project.
* PDFToText\_src.zip, containing a source-code project. In particular, PDFParser.cs is copied to D:/CCHR/Projects/Scout2/Scout2/Pdf and made part of the Scout2 project. Also, two lines from the Project.cs file formed the start of LegSiteController. ConvertPdfToText.
  + PDFParser pdfParser = new PDFParser();
  + pdfParser.ExtractText(file, Path.GetFileNameWithoutExtension(file)+".txt");

In addition, PdfSharpExtensions.cs has code taken from stackoverflow.com at

<https://stackoverflow.com/questions/10141143/c-sharp-extract-text-from-pdf-using-pdfsharp>

This provided the basis for the rest of LegSiteController. ConvertPdfToText.