

Formatting PDFs for EDGAR

Objective: To help 17(H) filers properly format PDF files to an EDGAR acceptable format.

Key Principles:

There are two key principles to understand for PDF files to comply with EDGAR standards:

1. Files CANNOT contain:
 - Active Content (Actions, embedded JavaScript, etc.),
 - Active External References (Destinations, Hyperlinks, etc.) Any Uniform Resource Locator (URL) must be in plain text,
 - Document passwords or security controls, or
 - No Embedded files (text, graphic, or audio)
2. Files need to be machine readable (See page 6, for more information about [Readable vs Unreadable](#) PDF Files)

Basic Guidance:

In most cases, EDGAR accepts documents that are electronically "printed" to a PDF from a source application, without performing any additional steps.

To "Electronically Print" your document to a PDF:

Select the "File" menu, then select "Print" (not the same as "Save to PDF"), then select "Adobe PDF" as the printer.

**This works because "printing" the document removes the active content (embedded fonts, scripts, links, etc.) that cause a problem for EDGAR. "Saving" as a PDF, may embed extra or active content within the PDF file, causing a problem in EDGAR.

Skip to guidance for a Specific
Error Message

[Invalid Filename](#)

[Java.lang.NullPointerException](#)

[External Reference Error](#)

[Invalid PDF Tag \(PDF Form Data FOUND\)](#)

Software Requirement:

The steps shown in this document require Adobe Acrobat software (or the equivalent), installed on your computer, the software allows you to create and modify PDF files. ** If you do not have Adobe Acrobat software (or the equivalent), you may be still be able to [solve the "external reference error" using the Google Chrome browser.](#)

How to format / troubleshoot different types of documents for EDGAR:

The formatting steps will vary based on the specific document, and how the document was created. Please refer to the following three document categories to find guidance and troubleshooting for your specific document.

Document Category 1	Document Category 2	Document Category 3
Documents created from another application (and "electronically printed" or saved as an electronic PDF file) Document Category 1 Formatting Steps (Page 3)	Documents that were physically printed, and scanned WITHOUT Optical Character Recognition (OCR), or Text Recognition Document Category 2 Formatting Steps (Page 4)	Documents that are a mix of category 1 and 2 (i.e. a signature page is scanned and added to an existing PDF) Document Category 3 Formatting Steps Page 5

** The size of the PDF file (kilobytes/ megabytes) can help give you an indication of the type of document you are working with. Also see the guidance on Page 6 about the readable and unreadable files. As a rule of thumb and dependent on the specific content:

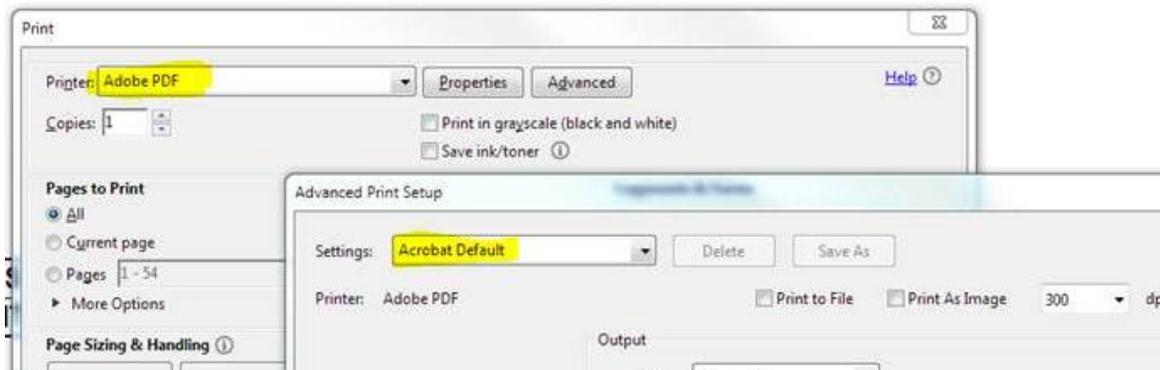
PDF pages that have been printed electronically may range between 15 kilobytes – 120 kilobytes per page.

PDF pages that have been scanned or printed as images may range between 350 kilobytes – 2 Megabytes per page.

For Document Category 1, make sure to follow the following steps to prepare them for EDGAR. You may have to perform these steps, if the file was created by the "Save as Adobe PDF" option within the respective application.

Document Category 1
Documents created from another application (and "electronically printed" or saved as an electronic PDF file)

1. From respective application select the File menu → Print – this not the same as the "Save to PDF" option
2. Select "Adobe PDF" as the printer
3. You will be prompted to save the file
4. If printing from within the Adobe Acrobat application, then make sure that you click on the "Advanced" button and that "Adobe Default" is selected in the "Settings" dropdown (See below for screenshot). Also, make sure that print to file and print as image are NOT selected.



5. Review the file size of the printed file. A large file size may be an indication that there are images within the file that will need to be further processed. As a rule of thumb, printed PDF files will be approximately 10-15 Kilobytes (Kb) per page, so if your document is 15 pages, total document size will be approximately 150kb- 200kb.

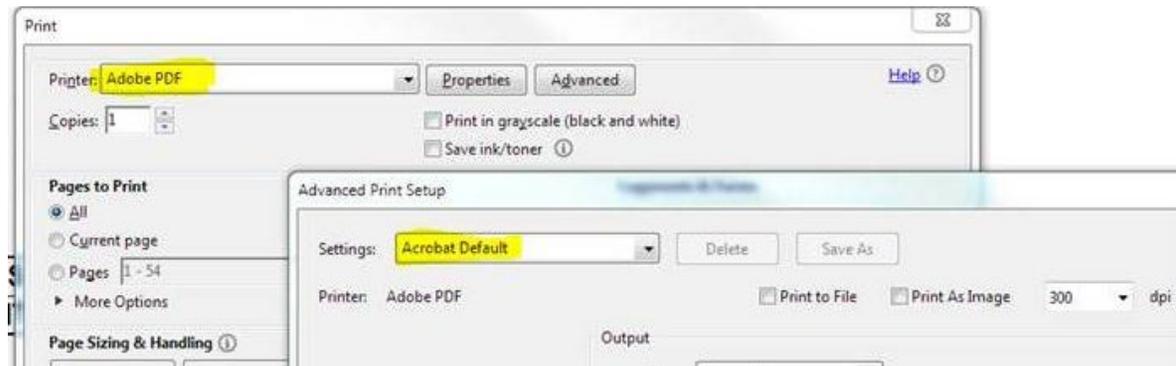
If the resulting file size is significantly higher (for example, 15 pages with a file size close to or higher than 1 Megabyte (Mb)). Refer to [Document Category 3](#), as there may be at least one "scanned page"

Document Category 2 files are not machine readable, and will cause them to be rejected by EDGAR. Follow these steps to prepare them for EDGAR. The scanning process reduces resolution, and converts all of the text on a page to an image.

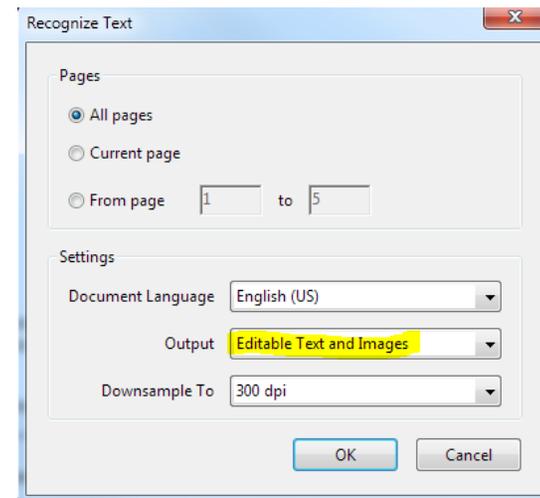
Document Category 2
Documents that were physically printed, and scanned **WITHOUT** Optical Character Recognition (OCR), or Text Recognition

1. (Optional) – Instead of trying to make the document machine readable, go back to the application where you created the document (before it was printed to a hard copy), and print the file electronically to a PDF. (Refer to the steps for Document Category 1 for those instructions), otherwise move to Step #2.
2. Open the PDF in Adobe Acrobat (Different versions of Adobe software ("Pro," "DC," etc.) may have different terminology, but should have the same functions/capabilities). Adobe “Reader” does not have this functionality.

3. Select File -> Print, then then make sure that you click on the "Advanced" button and that "Adobe Default" is selected in the “Settings” dropdown (Screenshot on right). Also, make sure that print to file and print as image are NOT selected. This step will ensure that there are no embedded items in the document.



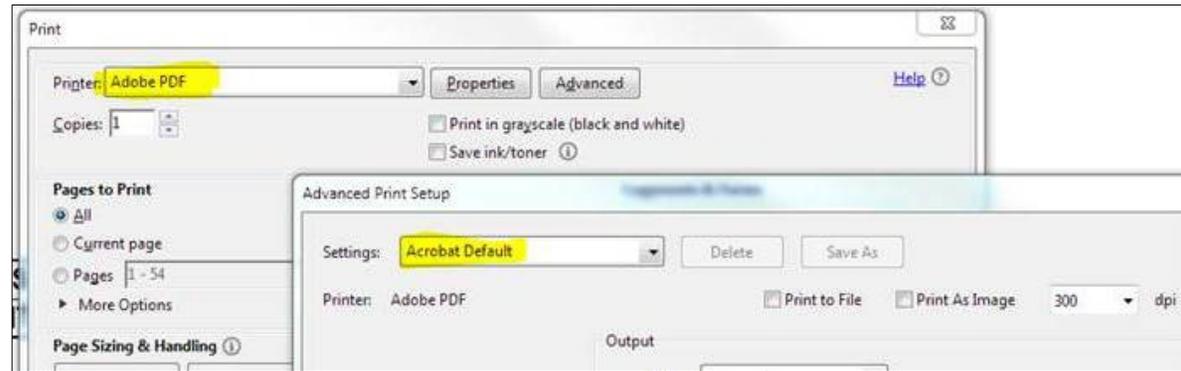
4. Click “Ok” and click “Print”
5. Save the file under a different name, so it does not overwrite over the original file (File size may be quite large)
6. Open up the newly reformatted file
7. View->Tools->Text Recognition (might say "OCR", "Recognize Text" or "Enhance Scans" in different versions of Adobe) – You may also have to specify "Editable Text and Images" for output options in the settings (See the image to the right)
8. After the text recognition tool is done running through the document, save the document with a new name.



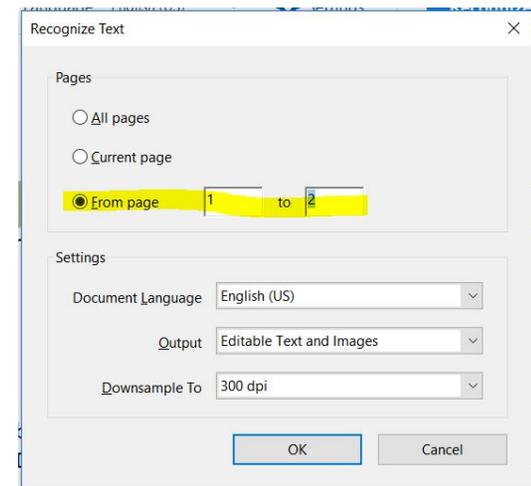
Follow these steps to prepare Document Category 3 files for EDGAR

Document Category 3
Documents that are a mix of category 1 and 2 (i.e. a signature page is scanned and added to an existing PDF)

- (Optional) – Instead of formatting as one document, it may be easier to separate the different types of pages. I.e. If you have one signature page that is scanned, then you could potentially remove that page (if not needed for filing) or split that page into a new document, and then process each document separately.
- Open the PDF in Adobe Acrobat (Different versions of Adobe software ("Pro," "DC," etc.) may have different terminology, but should have the same functions/capabilities). Adobe “Reader” does not have this functionality.
- Select File -> Print, then then make sure that you click on the "Advanced" button and that "Adobe Default" is selected in the “Settings” dropdown (See screenshot on right). Also, make sure that "print to file" and "print as image" are NOT selected. This step will ensure that there are no embedded items in the document.

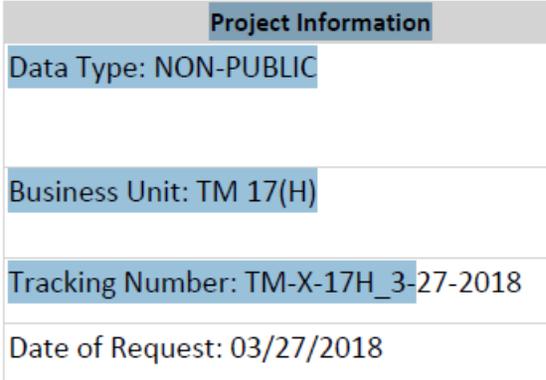
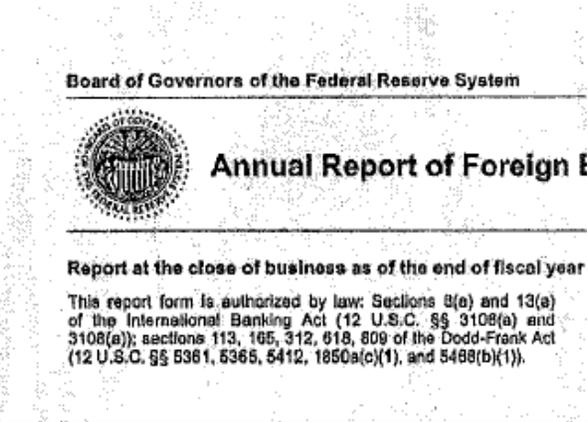


- Click “Ok” and click “Print”
- Save the file under a different name, so it does not overwrite over the original file (File size may be quite large)
- Open up the newly reformatted file
- View->Tools->Text Recognition (might say "OCR", "Recognize Text" or "Enhance Scans" in different versions of Adobe) – **Select only the pages that are not currently machine readable/where text is not already recognized. If you perform this step on the pages that are already machine readable, it may cause EDGAR to reject the document. (see screenshot on right)**
- After the text recognition tool is done running through the selected pages, save the document with a new name.



Readable vs Unreadable PDF Files

The following examples are a few common indicators to help determine the type of document, you likely will not have to confirm every property to identify the document type.

Readable PDF File property examples (may not be exhaustive)	Unreadable PDF File property examples (may not be exhaustive)
<ul style="list-style-type: none"> Text can be selected Whitespace is white with no noise marks Cursor shows as a "text selector" image like  Can use "<Ctrl><F>" to search and find a word (e.g. "Project") that you know is in the document. Text is clear and sharp, as seen in the following example Images or logos in readable documents may not be "selectable", so focus on the text portion of the document <div style="text-align: right; margin-top: 20px;"><i>[for Non</i></div> <div style="margin-top: 10px;">  </div>	<ul style="list-style-type: none"> Text is not able to be selected or searched Only "areas" are able to be selected (commonly these are scanned documents where the text is an image). Original documents may have been printed, and then scanned by a copy machine. Paper is not clear, may be off white in appearance (appears as though it was scanned or faxed) Cursor shows "crosshair" image like  Text may appear grainy as shown in the following example File is an unusually large file size <div style="text-align: center; margin-top: 20px;">  </div>

Form 17(H) Document Error Messages

If you receive an error message when validating a document, it will show up as a red “x” in the errors column, as shown below. If you click on the red 'x', there is a little more information available about what is causing the error. See below for a few instances.

Common Error Messages	Explanation and How to Resolve
Invalid file names	<p>File name Suggestions: file names can be no longer than 32 characters in length, should start with a letter (a-z), not contain spaces, may contain digits (0-9), one period (.), one hyphen (-), and one underscore (_) character.</p> <p>Section 5.1.1, “File Naming Standards” of the EDGAR Filer Manual - Volume II https://www.sec.gov/info/edgar/edmanuals.htm</p>
<p>Java.lang.NullPointerException</p>	<p>This occurs when the PDF contains scripts or embedded objects (which could also be an image file). The solution depends on whether the document is Readable (See Document Category 1) or Unreadable (See Document Category 2) to resolve.</p> <p>Per the Filer Manual Volume II, the following restrictions are enforced for all PDF documents:</p> <ul style="list-style-type: none"> • No Active Content (Actions, embedded JavaScript, etc.) is allowed • No External References (Destinations, Hyperlinks, etc.) are allowed. Any Uniform Resource Locator (URL) must be in plain text. • No document passwords or security controls are allowed

Common Error Messages	Explanation and How to Resolve
<p>External Reference (Active Link is in the document)</p> <p>ERROR_FILE_4584960856509785708_errors.txt</p> <p>Error Count: 1</p> <p>In linksave.pdf: found an invalid PDF tag (EXTERNAL REFERENCE URI FOUND).</p>	<p>This occurs when you have a readable PDF, that was likely “Saved” as a PDF from the source program, and has an active hyperlink in the document.</p> <p>See Document Category 1 details to resolve</p> <p>Solve Using Google Chrome (if you do not have Adobe Acrobat Installed)</p> <ol style="list-style-type: none"> 1. In Windows Explorer, right click on the file then select Open With > Google Chrome 2. Press Ctrl + P to open the Print dialog 3. Change the destination printer to "Save as PDF" 4. Click on Save 5. Type a new File name and then click "Save" 6. Close the Google Chrome tab
<p>PDF Form Data FOUND</p> <p>ERROR_FILE_3473634743613333166_errors.txt</p> <p>Error Count: 1</p> <p>In pages.pdf: found an invalid PDF tag (PDF Form Data FOUND).</p>	<p>This occurs when the PDF contains scripts or embedded objects (could be an image file). The solution depends on whether the document is Readable (See Document Category 1) or Unreadable (See Document Category 2) to resolve.</p> <p>Per the Filer Manual Volume II, the following restrictions are enforced for all PDF documents:</p> <ul style="list-style-type: none"> • No Active Content (Actions, embedded JavaScript, etc.) is allowed • No External References (Destinations, Hyperlinks, etc.) are allowed. Any Uniform Resource Locator (URL) must be in plain text. • No document passwords or security controls are allowed

Form 17(H) document filing guidelines and EDGAR document considerations

Unnecessary Documents - In some cases, firms have struggled to upload documents that are unnecessary for the 17(H) Filing, some of those documents are listed below with a brief explanation of why those documents do not need to be included within the 17(H) Filing:

- FINRA Focus Reports – the SEC already receives the FOCUS reports from FINRA directly
- Documents already filed with the SEC under additional SEC rules, including 10-K, 10-Q, etc. **
- Publicly available annual reports that are posted to a company's website **
- Broker-Dealer Annual Audits (X-17A-5) already filed with the SEC Division of Trading and Markets
- Executive management org chart – Rule 17(H) requires an “Org Chart” showing corporate entities and their relationships, not the names/titles of individuals that are leading the company.
- FOIA exemption request – Rule 17(H) is confidential and information submitted under the rule cannot be requested under FOIA

** If desirable to include with your 17(H) file as a reference, you may upload the first page of these documents, as applicable, or reference them on a cover page that these materials are already submitted/ posted to a public website, and should be read in conjunction with the 17(H) Filing.

Security - The SEC takes security very seriously and given the importance of EDGAR and the amount of documents received by the system, the SEC's approach is to restrict the types of files that could create a security issue within EDGAR.

Document Accessibility - Because of the volume of documents and the regulatory mandate EDGAR also requires that documents are machine readable (the text is recognizable).

Acceptable Files - **Besides PDF Files**, other acceptable file types include HTML (*.htm), XML (*.xml), XBRL and TXT (*.txt) files, each of which may have specific requirements. *.txt files cannot contain quote (") or dash (-) characters, and may be limited in the number of characters per row.