

Tech Tips from Mr G

What You Need to Know About Scanning Documents at the Liverpool Public Library



I think that the scanner is one of the best inventions of the personal computer era. With it, I'm able to take photos and documents that would've taken many file drawers to store, and put them into a nice compact format that I can store on my computer and reproduce almost perfectly later on. But there are a few things you need to know before you get started with our scanners, and that's what this handout is all about.

It's a Fancy Photocopier

That's right, at its heart, a scanner is nothing more than a fancy photocopying machine. The difference between the two is that the photocopier doesn't allow you to store the image to print out later on. Sure, you can make 87 copies of a document in one shot, but if you need an 88th later on, you need to come back with the original and copy it again. With a scanner, once you've scanned that document onto your computer or thumb drive, you can print off as many copies as you want whenever you want.

This makes it perfect for scanning in drawings, photographs, music scores, articles from magazines, handwritten letters, and other things that you just want a static copy of to refer to or print out later...just as if you were photocopying them. This is what most people want a scanner to be able to do for them, and doing this is relatively easy.

They can also be used for something called OCR (Optical Character Recognition) scanning. That's a little trickier.



These are the things that scanning is really easy for.

Types of Scanners and Scanning

Here at the Liverpool Public Library, we have five different scanners that can do three different types of scanning. I'll talk about the different types of scanning first. They are:

Image Scanning:	This is the type of scanning you do when you want to make copies of photographs or drawings. The files are saved in either jpg , png , or tif format. These files are later editable by programs like Photoshop and iPhoto .
Plain PDF (Document) Scanning:	This is the type of scanning you do if you want to scan magazine articles, old letters, music scores, medical forms, receipts, resumes, or anything else that you might simply want a stored photocopy of for future use. These files are saved in pdf format, and the text in them is not editable or searchable.
OCR Scanning:	This is the type of scanning you use when you either have a document that you want read into the computer and put into a text file to edit later , or scan into searchable pdf format.

Our five scanners are as follows:

Dell 4 Epson:	This can do image scanning, PDF scanning, and OCR scanning. If you use the Epson scanning software, the default setting is for image scanning into the jpg format, but that can be changed. If you scan to the pdf format, each page is saved as a separate document. This is the best machine to use for single page OCR scanning, and it's done through the ABBYY program.
Mac 5 Epson:	This can do image scanning, PDF scanning, and OCR scanning. Image and PDF scanning are done through Photoshop , and OCR is done through OmniPage Pro X . If you do pdf or OCR scanning, each page is saved as a separate file, but they can be combined into one document by opening them up in Preview or Microsoft Word .
Mac 5 Addesso:	This does slide and negative scanning into the jpg , png , and tif formats. You'll need to get the slide and negative holders from the Circulation Desk.
Mac Lite ScanSnap:	This can scan multipage, double-sided documents into pdf format. It can also do multipage OCR scanning.
Kodak Photo Scanner:	This is the machine to use if you have a shoebox full of photos that you'd like to scan. It saves to the jpg format, and you will need to have a thumb drive or CD with you to store the pictures on.

Searchable PDFs

While a plain pdf file is basically just a stored photocopy of the original document, a **searchable pdf** is a pdf file that you can search through for certain terms. Lawyers and researchers love this kind of file because it spares them the effort of having to read through an entire document that may be hundreds of pages long.

This is one of the best uses of OCR scanning, and the **Dell 4** and **Mac Lite** are the best machines for this.

OCR Scanning for Editing: It's Not As Easy As You Think

A lot of people come in with a text document, thinking that an OCR scan of it means that it will be scanned in exactly as it was on the original, and that making the few edits they need to do to it will be a piece of cake. They especially want to do this with resumes. This may or may not be true, depending on the scanning software you're using.

This is because OCR is an art, and not a science, and documents with complex formatting may not be handled correctly by the scanning software. As a result, **you may be better off just typing it in from scratch**. Really. Unless you are an expert at some of the more arcane features of Microsoft Word, by the time you're done fussing around with the text boxes and trying to get things moved around to the proper places, you could've typed the original all over again at least twice. And even people who are experts at Microsoft Word will usually just type the document in all over again.

However, if you insist on trying this, the best machine for this type of scanning is **Dell 4**.

After the Scan is Over

While we can easily get you set up to do your scanning, we simply cannot take the time, nor do we have the expertise to help you edit the resulting files, whether that be in **Photoshop**, **Microsoft Word**, or some other program.

The Takeaway

The takeaway from this is that while scanning is a great tool to be able to use, you really need to know which form of scanning is best for what you want to do with the finished document...and whether or not scanning it in is a good idea in the first place.

I hope this little handout has helped to make that clear.

keg 2014.05.19

Tech Tips from Mr G is a publication of the Liverpool Public Library.

You can find more of them at www.LPL.org/computers-and-computing/tech-tips

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/> or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

