

Paper Documents to e-Documents: An Empirical Approach towards Digitization for the Novice

Md. Rafiqul Alam*, Sourav Das**

Abstract

This paper deals with the process of digitization of paper documents using mainly available free software. It is necessary not only for library professional but also for the others. Because in the present era of IT any one may have require to digitized their valuable paper documents for preservation digitally. In this paper the digitization process has been discussed in a very lucid manner so that a novice person can do the digitization job using freely available software with ease.

Keywords: Digital Library; E-Document; E-Library; PDF-Searchable; Scanner; Storage Device.

Introduction

E-documents, an abbreviated term for electronic documents, are documents that exist only in electronic form such as data stored on a computer, network, backup, archive or other storage media [1]. Due to the rapid advancement of Information and Communication Technology it becomes very popular to us. When this type of documents are managed and maintained by a library then we called that library as an e-library or sometimes digital or virtual library. Remote access via network or internet to this type of library is possible from any part of the globe.

What is Digitization?

According to Wikipedia Digitizing or digitization is the representation of an object, image, sound, document or a signal (usually an analog signal) by a discrete set of its points or samples. The result is called digital representation or, more specifically, a digital image, for the object, and digital form, for the signal. Strictly speaking, digitizing simply means capturing an analog signal in digital form [2]. For a document the term means to trace the document image or capture the "corners" where the lines end or change direction.

Why Digitization?

There are so many reasons why digitization is necessary in today's context. Some of reasons are

- Digitally accessible
- Save storage space
- Data preservation

Besides these known benefits, digitization will critically provide you with a great head start in dealing with your documents and plans. It will not only allow you the flexibility and ability to be operable on existing and future digital platforms, it will also allow your digitized information to transform into useful, portable format. It is now commonly adapted universally as it can make a huge difference in many aspects, from information sharing to workflow processes [3].

The Differences between E-Documents and Paper Documents:

1. Variety of electronic documents is larger than paper documents.
2. In terms of accessibility and quick relevant content searching ability e-documents are more efficient than paper documents.
3. Electronic documents are more persistent and more difficult to destroy than paper documents. Throwing away or shredding makes paper documents disappear.
4. Changes or rectification of errors to an electronic document are faster and easier than paper documents.

Author's Affiliation: *Librarian, P.D. Women's College, Jalpaiguri-735101(W.B.) **Asst. Librarian, Siliguri Institute of Technology, Siliguri-734009(W.B.).

Reprint's Request: Md. Rafiqul Alam, Librarian, P.D. Women's College, Jalpaiguri-735101(W.B.).
E-mail: rafiqul_mlisc@yahoo.com

5. Electronic documents last longer than paper documents and always looks like new document while Paper documents deteriorate with time.
6. Electronic documents may be created by electronic means while paper documents are created by humans [4].

Equipments Needed for Digitization Work

Digital scanner, computer, USB pen drive or other storage device, PDF creator Software, PDF file joiner/splitter software, digital library software.

Steps of Digitization

Selection of Paper Documents

First of all we have to select the paper documents which to be digitized for the future preservation.

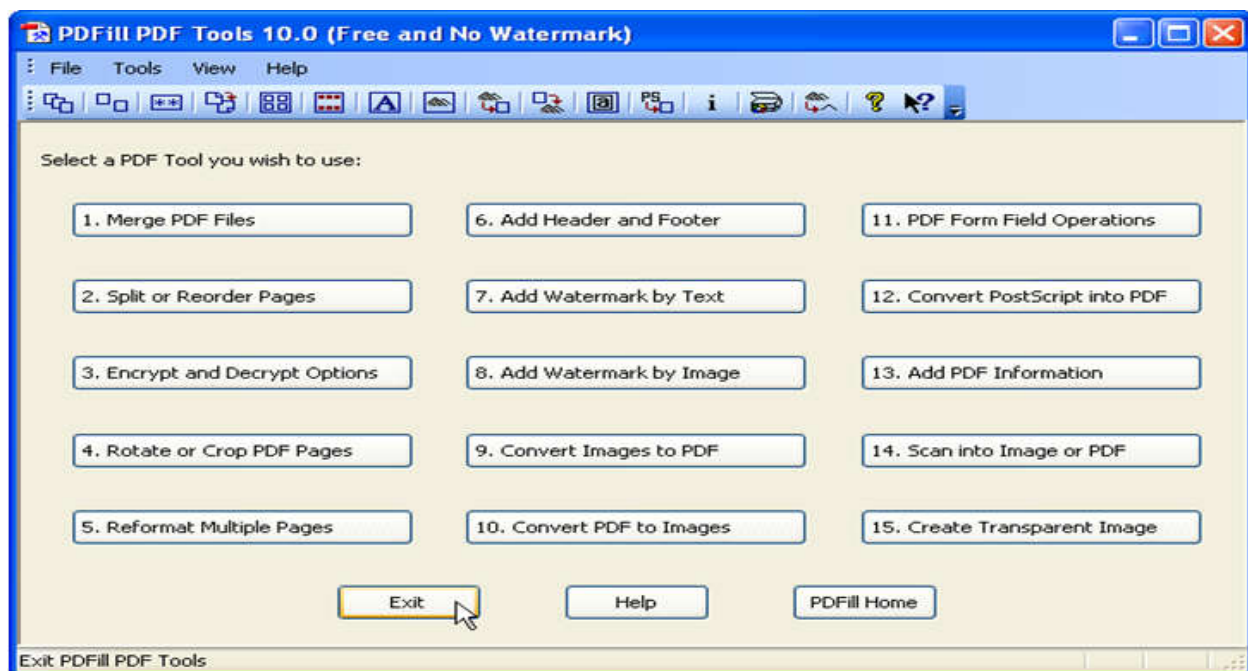
Selection of Scanner and Scanning of Documents

In this step the paper documents are to be scanned to convert it into a digital format. The paper

documents may be scanned in different file format for example .pdf, .png, .jpg, .bmp, etc. But my suggestion is, if the document is a black and white textual document then it should be scanned directly from the scanner in a PDF-Searchable file format because it occupy less space than that of other file format and it helps the user to find desired document with relevant content very quickly from the digital repository. If the document is a coloured document then only select .jpg or .bmp file format if necessary. But some scanner may not support PDF-Searchable file format. Hence at the time of selection of scanner it should be kept in mind that the scanner whether supports PDF-Searchable file format or not. The scanner should also have the facility of auto feeding of documents.

Software Selection and Joining of pdf Files

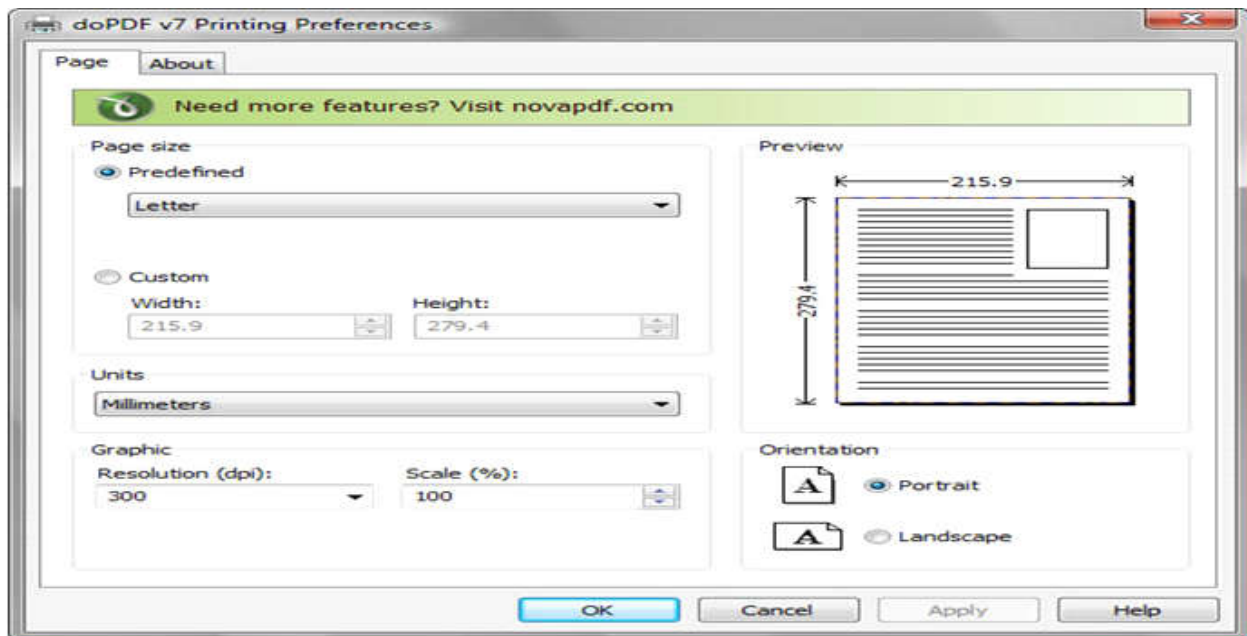
There are so many free software that can join, split, rotate and so many functions of pdf file as and when required. One of among them is PDFfill PDF Tools [5] that can be download from the site <http://www.pdfill.com/download.html>.



Another one is doPDF [6] that can convert any document into a pdf file format. Suppose if you want to convert a .jpg file extension document into a .pdf file extension. Then you simply select the print option. Then select do PDF as a printer and then the output will come as a pdf file format. DoPDF software can be download from the site <http://www.dopdf.com/download.php>

Building of Digital Library

Now the digitization process is complete. So the building of digital library is essential in this phase. Because digital library allow us to store, manage and retrieve the documents through search facility. Otherwise we could not be able to find out the required document at a right time. We can build



digital library by using DSpace or by Greenstone [7] digital library software. Greenstone can be

download from the site <http://www.greenstone.org/download>



Advantages of Digital Documents and of Digital Library

1. E-documents never become older and remain always as a new.
2. Multiple accesses.
3. Save space.
4. Allow information retrieval by any search term.
5. Easily accessible.
6. No physical boundary.

7. Round the clock availability.

8. The big advantage of e-documents in context of today's environmental issue is it encourages to 'going paperless, save trees'.

Limitations of Digital Documents and of Digital Library

1. Digitization violates the copy right law as the thought content of one author can be freely transfer by other without his acknowledgement.

2. As more and more computer are connected to the Internet its speed of access reasonably decreasing.
3. The infrastructure cost of digital library is generally very high.
4. Digital library will need high band for transfer of multimedia resources but the band width is decreasing day by day due to its over utilization.
5. With the much larger volume of digital information, finding the right material for a specific task becomes increasingly difficult.
6. Digital libraries cannot reproduce the environment of a traditional library.
7. Due to technological developments, a digital library can rapidly become out-of-date and its data may become inaccessible.^[8]

Conclusion

Document preservation in electronic format is the need of time and we cannot ignore this. Are e-documents really greener than paper documents? The debate is still going on. However, e-document encourages us to go paperless and discourage to use

printed paper document. By this way we can save trees and at the same time we can manage our valuable knowledge documents very efficiently.

References

1. <http://legalcareers.about.com/od/glossary/g/E-Documents.htm> [accessed on 15.12.2015]
2. <http://en.wikipedia.org/wiki/Digitizing> [accessed on 15.12.2015]
3. <http://www.weil-ebdm.com/digitization.htm> [accessed on 28.12.2015]
4. <http://www.seas.gwu.edu/~shmuel/WORK/Differences/Chapter%203%20-%20Sources.pdf> [accessed on 28.12.2015]
5. <http://www.pdf Fill.com> [accessed on 15.12.2015]
6. <http://www.dopdf.com> [accessed on 15.12.2015]
7. <http://www.greenstone.org/> [accessed on 15.01.2016]
8. <http://sscassrh.org/sri-sri-ayurveda-college/index.php/articles-by-doctors/item/320-digital-library-advantages-and-disadvantages> [accessed on 15.01.2016]

