# Swift Journals

*Original Research Article*

# The Development of Electronic Document Archiving System for Mindanao University of Science and Technology

**Ms. Love Jhoye M. Raboy, Mr. Dwiff C. Baconguis, Mr. Lee Jonuel C. Bagabuyo, Ms. Rose Glen P. Fuentes, Mr. Jefferson Y. Jamboy**

College of Industrial and Information Technology, Mindanao University of Science and Technology
Claro M. Recto Avenue, Lapasan, Cagayan de Oro City, 9000 Philippines

## ABSTRACT

**Manual handling of the documents is hard to manage due to these factors such as possible misplaced of documents, lost documents and searching is time consuming especially when the documents stored increases in volume. The study aims to develop an Electronic Document Archiving System (EDAS). A tool that can help offices keep backup copies of their physical files or documents and provide easy search and retrieval of the documents. The EDAS produces a digitized copy from a hard copy document through capturing an image of it. Various Application Programming Interfaces (APIs) were used to link the input devices, which are the webcam and scanner, in consolidating the images into one document and for archiving the files. Specified document information and keyword tagging method were applied to act as a metadata of the files that are stored in the database. Metadata serves as a complete tag of each file generated, that can be used in the identification of files. A text search query was used to match the metadata of files in searching and retrieval of the file. The EDAS average searching time is 0.09 seconds while manual searching takes 20.59 seconds. The result of the study states that the overall system searching time is faster than the manual searching.**

**Keywords:** Electronic Document Archiving, digital document, digital tags

## INTRODUCTION

Manually handled documents, has an advantage and several disadvantages. The advantage is that documents are filed in alphabetical order according to name so that it can be retrieved faster and easier. The disadvantage is that it can be misfiled and mixed in other categories (Recolisado, 2014). Another disadvantage of this so-called flat-file management system is the lengthy and inefficient way that files are meticulously positioned for search and retrieval. This became a problem for companies and offices over the past years.

So in the past few decades, the emergence of Electronic Document Management Systems has been the solution for this particular problem. Electronic Document Archiving System (EDAS) is a rapidly developing technology and is considered as the solution for organizations that need a way to manage information efficiently. The automation of document management system can lead to a more precise sorting and searching of documents. It also prevents loss of data and lessens the amount of time in arranging the documents according to its category and type. In addition, it retrieves and searches specified documents faster, so less time will be spent in retrieving documents as they can be retrieved by the users without leaving their desks.

One of the challenges encountered at the Dean's office of the College of Industrial and Information Technology is the management of documents. Problems encountered were misplaced documents, lost documents and lengthy searching time for a specific document. The objective then of this study to develop an Electronic Document Archiving System (EDAS). EDAS is a tool that can help the office to keepbackup copies of their physical files or documents, provide easy search and retrieval of the documents, and only the authorized personnel can access the system. Documents scanned and captured will

---

*Corresponding Author:* Ms. Love Jhoye M. Raboy, College of Industrial and Information Technology, Mindanao University of Science and Technology Claro M. Recto Avenue, Lapasan, Cagayan de Oro City, 9000 Philippines
Email: Jhoye.m@gmail.com, ljraboy@must.edu.ph

## METHODOLOGY

Figure 1 illustrates the system flow; the user has five options on how to interact with the system, namely: searching, capturing images via webcam, scanning documents, uploading pictures, and lastly archiving documents. In the "SEARCH" function, the user can look for a particular document with different filters: year, semester, month, category, keywords, author, or a file name. For the "NEW DOCUMENT" function, the user can take pictures of the document through the use of a web camera, or scan documents by using a scanner, or upload existing pictures directly into the system. Lastly, there's the "ARCHIVE" function, where the user can record the documents stored in the system by compressing it into a ZIP file.

There are different outcomes to each action that a user employs. For example, when the user chooses web camera as a type of input, the user will capture pictures of the document and will be able to categorize the document, or create another category. After categorizing, it will be saved to a device's storage where the user wants it to be saved.
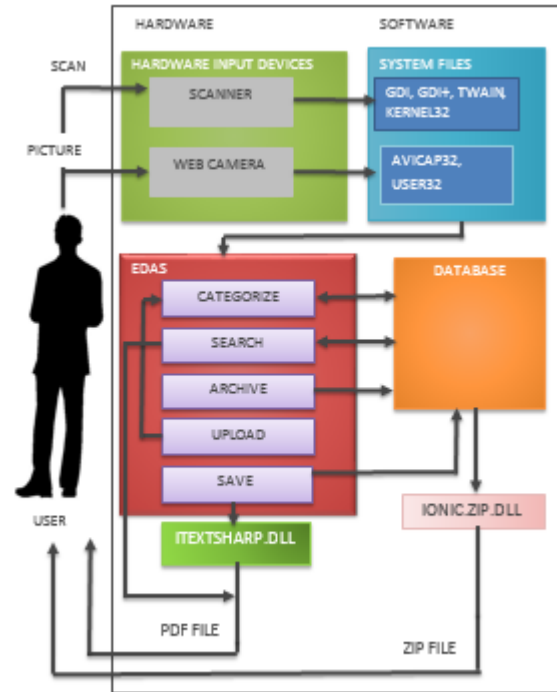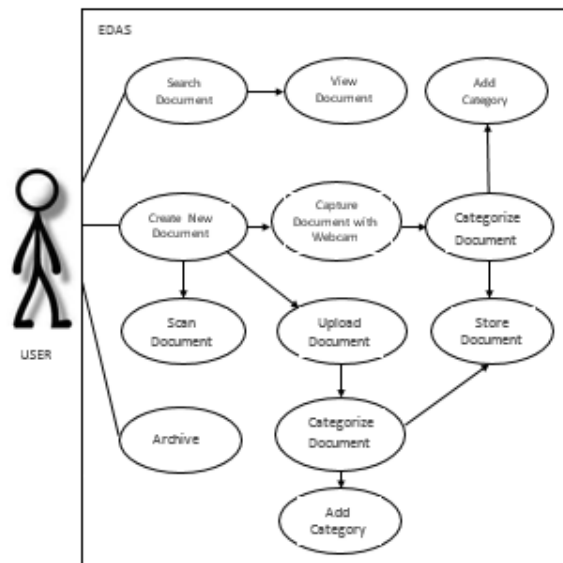


Figure 1 EDAS System Overview

### *System Architecture*

The EDAS System Architecture shows how the system interacts to the devices with the use of Dynamic-Link Library (dll) and how it is connected to the database. First, the user has the option to capture the images of the document using the scanner or the web camera. The hardware input devices use the system files or dlls to interact in the system and to access the full capability of the device. Then, the basic input of data functions as the metadata of the file that will be saved into the database. After getting the whole input, the user can now save the document. After saving, the system generates a file document (PDF file) that contains the images that were captured with the use of another type of dll. In retrieving the document, the information about the files that the user wants to retrieve are needed to search into the database. In this way, only the specified documents are identified and archived in a ZIP file with the use of a dll.



Figure 2 EDAS System Architecture

## RESULTS AND DISCUSSIONS

Figure 3 shows the main form. In this form, the user is able to choose the following actions such as searching, viewing, adding new documents, and archiving documents. The user can search documents by any of the following: by year, semester, month, category, author, keywords and file names.



Figure 3 the EDAS Main Form

### *Categorizing*

Figure 4 shows the New Document Form. This form has the segregated components and functionality that is only related to the basic input, choices, and webcam operations. In this form, the default webcam will be used; the category and page size (letter, A4 or legal) will then be chosen, and the author, input keywords, and determine the storage destination. After filling the fields, the user can now start capturing the document(s), then save it afterward.

Figure 7 Upload Form. The proponents designed the uploading of documents for the user to use alternative devices; the user only has to save the documents produced by the said device to the computer or any external storage. This form is designed to browse images as displayed above; the size of images being captured by different devices is a problem if it would save directly due to its different dimensions, so the solution of this



Figure 4 Category Form

### Capturing through Web Camera

Figure 5 shows the capture form. The purpose of this capture form is to capture documents in order to keep a back-up copy. Capturing document through webcam is faster than the scanner.
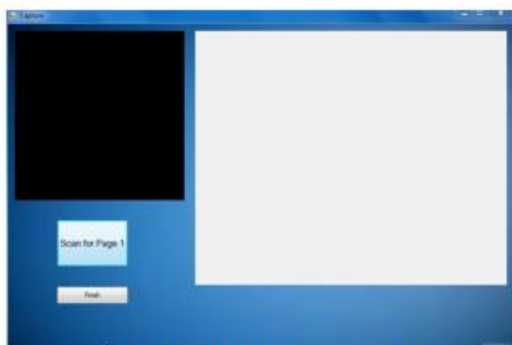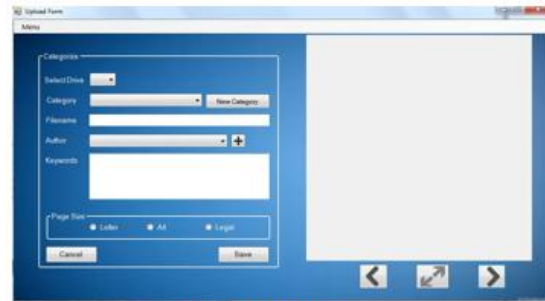


Figure 5 Web Camera Capture Form

### Capturing through Scanner

Figure 6 Scanner. This form has menu items in which the user only has to click the "select source" option to choose the scanner the user wants to use and "acquire" option to access the scanner's interface.



Figure 6 Scanner Form

### Input through uploading



Figure 7 Upload Image

problem is to resize the image based on what page size if it is a4, legal or letter and make sure that the user must save it in the same format as on how the user resized it.

### Archiving Functionality

In this system, archiving functions as the output that contains the files that the user wants to retrieve. The files are archived in a way that the documents needed are all consolidated into a ZIP file. There are choices on how the file will be archived namely by the combination of the following by year, by month, by category, and by a specific range of date. In by year, all the files created in a specific year will be archived; in by month, a specific year is needed and all the files on that specific month of that year will be archived; in by category a specific year and month is needed and in by range of date an author and category can be specified. The user can choose one of these choices and the files that are retrieved will be saved and compressed as a ZIP file to a specific drive or disk.

### Searching Functionality

There are choices on how the file will be searched, namely by the combination of the following fields: by year, semester, month, category, keywords, author, or filename. In by year, all files that are created on the specified year will be searched; in by semester, all files that are created during the specified semester will be searched; in by month, all files that are created on the specified month will be searched; in by category, all files that has the same category as specified by the user will be searched; in by keywords, all files with the same keywords specified by the user will be searched; in by author, the files with the same author specified by the user are searched; in by filename, all files with the same filenames specified by the user will be searched. The user can choose one or combine these choices to retrieve the specified document.
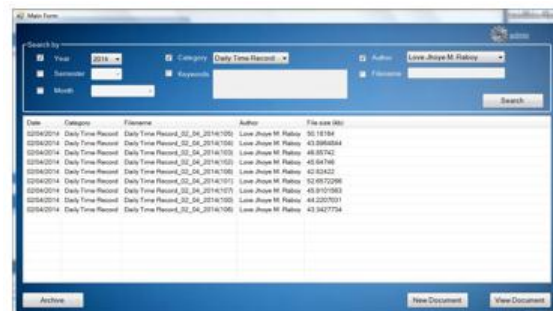


Figure 9 Searching Form

| Category | Manual Average Time | EDAS Average Time |
|---|---|---|
| Memo | 30.27 seconds | 0.10 seconds |
| DTR (Full-time Faculty) | 28.54 seconds | 0.09 seconds |
| DTR (Part-time Faculty) | 7.83 seconds | 0.08 seconds |
| Individual Teacher's load (Full-time Faculty) | 18.25 seconds | 0.11 seconds |
| Individual Teacher's load (Full-time Faculty) | 9.45 seconds | 0.08 seconds |
| Costing | 29.18 seconds | 0.09 seconds |
| Overall average Time: | 20.59 seconds | 0.09 seconds |

**Table 1 Manual Searching vs System Searching**

To test the searching time for the manual and system searching, the proponents sampled three documents to search per category. Table 1 shows that EDAS is faster in comparison with the Manual average time for searching for a specific documents. The system is faster because the system will only extract the information needed for a specified document in the database compared to the manual searching wherein the user has to go to the file cabinet and search through numerous folders and envelopes to locate a specified document.

### EDAS Organizing and Categorizing Capabilities

| Manual Categorizing and Organizing | System Categorizing and Organizing |
|---|---|
| 1. Identify category (memo, daily time record and etc) and sender if any. | 1. Identify the category, author, keywords, selection of storage (e.g. flash disks, external hard disks and etc) and the system will automatically generate the file names. |
| 2. The documents are kept in a folder marked with the category name and date but if the category is memo the folder will be labeled with the sender/author of the document. | 2. Save |
| 3. Sort by date. | 3. Then the system will automatically organize the documents. The system organizes the documents by year then by month and lastly the category. |
| 4. Then the folders will be kept in a porta file. The porta files are labeled with categories, year and semester. | |

**Table 2 Manual and EDAS Categorizing and Organizing**

Table 2 shows how the manual and the system's categorizing and organizing differences and similarity. For the manual categorizing, the user identifies the type of document received same for the system, while for the manual organizing the documents received are stored in a folder labeled with the category name and the author or sender's name. Lastly, the folder is stored in a porta file labeled with the category names, school year and semester. The system differs from the manual organizing, by organizing the documents by year, then by month, then by category and the user does not have to organize it manually. The system's organizing may not be specific like the manual organizing, but the lack of a more specific organizing are answered by a more extensive searching.

### EDAS Functionality Test

| Functionalities | Achieved |
|---|---|
| 1. To capture images using web camera | ✓ |
| 2. To scan documents using scanners | ✓ |
| 3. To upload images | ✓ |
| 4. Categorize documents | ✓ |
| 5. To search | ✓ |
| By year | ✓ |
| By month and year | ✓ |
| And other combination | ✓ |
| 6. To archive | ✓ |
| By year | ✓ |
| By month and other combinations | ✓ |

**Table 3 Functionalities of EDAS**

The proponents achieved the first functionality by successfully interfacing the web camera and grabbing the captured images into the system. The second functionality was achieved by interfacing the scanner into the system and the proponents were able to access the user interface of the scanner. The proponents were not able to grab the images scanned from the device. To address this problem the proponents came up with functionality number 3, this functionality is for devices that cannot be automatically grabbed by the system. With this functionality, the user will upload the already existing images into the system. Functionality number 4 was achieved by successfully categorizing each document, tagging keywords, and other categorizing options. Functionality number 5 was achieved by searching a document in many ways possible for the system. The system can now search by year, month, semester, author, keywords and file names. It can also search by combinations of year or month and other options. The archive functionality was achieved through successfully archiving documents by a given option. The archiving functionality can be done in many ways, it can be by year, month, category, author, and by a range of dates. And these options can be combined to give a more specified outcome.

### CONCLUSION

The proponents created a system that can fill the gaps of manual handling of documents. The proponents achieved the functionalities that are needed to comply. The system created by the proponents is suitable for the office that has troubles managing the documents.

The proponents achieved the stated objectives for the Electronic Document Archiving System. The proponents were able to develop a system that helps improve the operation of offices through categorizing, organizing and searching for a digitized hardcopy documents. The group was able to implement EDAS as a tool that can automate keeping and categorizing of documents. Implemented in a way that the proponents were able to link and get the data from the devices, create and consolidate the image into a single document, archive the retrieved files and search files with the information and keyword tagging used. Automation in terms of the technicality of the study that uses C# programming language and database. Lastly, the proponents were able to test the system to determine its performance in terms of the speed of searching for a specific document compared to manual searching. The EDAS average searching time is 0.09 seconds while manual searching takes 20.59 seconds. The result of the study states that the overall system searching time is faster than the manual searching.

## RECOMMENDATION

Referring to the results of the system dry run, the proponents can say that are some features that can be added to the system. It is highly recommended that the system will be modified further to cater the needs of firms and offices.

Recommendations are as follows:
- Provide security for folders and files.
- More flexible keyword and content-based searching capability.
- Automatically fetch the scanned images produced by the scanner.
- Integration of digital camera for better quality.
- Automatically disabling the built-in webcam for laptops and net books.
- Delete a specific image when capturing through the use of a web camera.

## BIBLIOGRAPHY

Internet:

Kawada, J. (Oct 25, 2007). "Scanned Document Management System". Retrieved Feb. 27, 2014. From https://www.google.com/patents/US20070250714

Combined Computer Resources, INC. (2009). "Municipal Records Management System". Retrieved Feb. 27, 2014. From WinOcular Products, WinOcular Innovations in Paperless Automation: http://www.winocular.com/winocularproducts/government/Municipal_Records_Management.htm

Microsoft (2014). "Windows Image Acquisition (WIA)". Retrieved Feb. 27, 2014. From Microsoft Library, Microsoft.com:

TWAIN Working Group (1992 - 2014). TWAIN. Retrieved Feb. 27, 2014. From http://www.twain.org/

Differencebetween.info (2012 - 2014). "Difference between BMP and JPG". Retrieved Feb. 27, 2014. From http://www.differencebetween.info/difference-between-bmp-and-jpg

WebFinance, Inc. (2014). "Portable Document File (PDF)". Retrieved Feb. 27, 2014. From http://www.businessdictionary.com/definition/Portable-Document-File-PDF.html

Alexander, R. (2002 - 2014)." Zip File Format". Retrieved Feb. 27, 2014. From http://www.rarlab.com/zip_file.htm

Wilson, L. (2012). "Using a webcam in VB6 with AVICAP32.DLL". Retrieved Feb, 27, 2014. From http://www.imajeenyus.com/computer/20120526_webcam_in_vb6/index.shtml

Microsoft (2014). "About GDI+ Retrieved". Feb. 27, 2014. From Microsoft Library, Microsoft.com: http://msdn.microsoft.com/enus/library/windows/desktop/ms533797(v=vs.85).aspx

Microsoft Corporation (n.d.)." gdi32.dll". Retrieved Feb 27, 2014. From http://www.paretologic.com/resources/paretolabs/dll/gdi32_dll.aspx

Kindig, A. (n.d.) "What is Kernel32.dll?". Retrieved Feb. 27, 2014. Fromhttp://www.ehow.com/about_5209211_kernel32_dll_.html

Jack Hubbert (2006). "Flat File". Retrieved July 2006. From http://searchsqlserver.techtarget.com/definition/flat-file

Corry Janssen (n.d). "Multifunctional Printer". n.d http://www.techopedia.com/definition/3611/multifunction-peripheral-mfp

Margaret Rouse (2010)." Scanner". Retrieved from May 2010. From http://whatis.techtarget.com/definition/scanner

Interview:

Interview with Beverly Recolizado, CIIT Dean's Secretary, Mindanao University of Science and Technology, Cagayan de Oro City. 15 February 2014

Interview with Mayette Cayetano, CIIT Dean's Secretary, Mindanao University of Science and Technology, Cagayan de Oro City. 15 February 2014

LociSolutions, "7 disadvantages of paper-based document management", June 26,2014. From http://www.locisolutions.com/blog/7-disadvantages-paper-based-document-management on March 10, 2015