Secure Digital Asset Repository

By

Nazurah Husna Binti Mohd Mahpos

Dissertation submitted in partial fulfilment of
the requirements for the
Bachelor of Technology (Hons)
(Business Information Systems)

MAY 2011

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

Secure Digital Asset Repository

By

Nazurah Husna Binti Mohd Mahpos

A project dissertation submitted to the
Business Information Systems Programme
University Teknologi PETRONAS
in partial fulfilment of the requirement for the
BACHELOR OF TECHNOLOGY (Hons)
(BUSINESS INFORMATION SYSTEMS)

Approved by:

....

(Mr. Abdullah Sani B Abd Rahman)
Project Supervisor

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
May 2011

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

NAZURAH HUSNA MOHD MAHPOS

SECURE DIGITAL ASSET REPOSITORY

ABSTRACT

As the industry evolves into technology driven businesses, an increasing number of companies are reaching a critical pain in needing to control and manage their vast amounts of digital media assets. Technically speaking, a digital asset is any form of media that has been turned into a binary source. Digital assets, which for textile mills include everything from artwork, logos, photos, text documents and even email, are proving to be valuable assets in terms of both productivity and company valuation. Many organizations thought that secure digital asset repository as part of their business-critical strategy for managing and sharing digital assets. They need dedicated solution to help overcome operational and organisational challenges unique to them. Secure digital asset repository allows company to create a custom-branded, password-protected area where company can exchange business files with clients easily, securely, and professionally. Whether company deal with files that are simply too large to transfer by email, need secure file transfer or just need a certain data being revealed, secure digital asset repository can provide a solution. This paper focuses several features in the secure digital asset repository application.

ACKNOWLEDGEMENT

This report, like everything else I have created in my life, is the result of a huge team effort. I extend my deepest gratitude and thanks to

Norkhalifah Binti Redzuan, my beloved mother who always gives incredible support, deep insights and keeping me focused. You are truly awesome!

Mr. Abdullah Sani Bin Abd Rahman, my great supervisor. You were there every step of the way with your academic insights, enthusiastic encouragement and authentic friendship. I admire your integrity and your professionalism.

And finally, thanks to all my friends for their thoughts, comments and understanding.

I really enjoy in doing this project with all of you around me. Thank you again.

LIST OF FIGURES

LIST OF SYMBOLS / ABBREVIATIONS

ACL	Access control list
PKI	Public key infrasturcture
CA	Certificate authorities
CRL	Certificate revocation lists
DAC	Discretionary access control
SDLC	System development life cycle
HCL	Hardware compatibility list
CPU	Central processing unit
TSA	Time Stamp Authority
TST	Time Stamp Token
ITA	International Time Authority
USNO	United States Naval Observatory

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1.1 Background

A Secure Digital Asset Repository application is simply a tool for organizing digital media assets for storage and retrieval. In asset repositories the content itself is physically stored inside a secure database. This results in a host of benefits, including security levels, replication, referential integrity and centralized data management. Also included is the comfort of full hierarchical storage management.

Solution based on the asset repository model are ideal when systematizing studios with industrial workflow, managing rights and permissions such as intellectual property of either your company or third party, and structuring global access by employees, contractors, suppliers, partners, and customers.

The process of secure digital asset repository insures that only approved brand elements are used in the proper context. It helps to build relationships by supporting the ability to share assets over an extranet with clients and suppliers. In addition, with the ability to allow clients or other users to observe creative work in progress, secure digital asset repository fosters communication and collaboration.

1.2 Problem Statement

Firstly, when most security professionals think of access control, they think mainly of the management of access: the creation and administration of access control lists (ACLs) and the formality of access control procedures. ACLs are indeed a way of managing and controlling access to information, but only if the information is in a computer and stored in files. If the information is being displayed or communicated, ACLs are of little help.

Controlling access while information is being displayed is very different problem from controlling access while information is being communicated over long distances. There are access controls that are specific to information states and access control that are general in nature. This is the kind of operational environment that must be taken into account when considering the use of general-purpose controls for information-security-specific purpose.

Secondly, there are a variety of functions that authorized individuals may need to perform. Some may need to read the information but not need to be able to change or destroy it. Others may need to append information without necessarily changing what is already there. The ability to change and ultimately to destroy information is given only to those persons in whom we have great faith.

Thus, corresponding to the level and duration of protection desired for each item of information, every person who is given authorized access to the information must be assigned a level of trustworthiness. The level of trustworthiness assigned to an individual depends, roughly, upon his or her need for access to the various levels of classified information and compartments.

1.3 Aims and Objectives

The aims and objectives for this project are:

- To provide flexibility distribution of the documents.
- To provide secured information management system.
- To be able to store and retrieve data.
- To allow people to spend less time locating assets and more time working on current project.
- To allow people access the system via personal computer and mobile phone.

1.4 Scope of Study

The main focuses for this project are the security of the information management systems and the access control the information. Firstly, the science behind authentication method is sophisticated, and there are different ways to use these methods. Digital signatures are a result of the computer and its influence in our daily lives. Digital signatures are perhaps the best known of the various scientific methods for authentication. Digital signatures can be generated very quickly and facilitate binding a document to its owner to help authenticate the message. Since digital signatures require effective –binding processes for a public key infrastructure (PKI) to be practical, a review of the requirements for certificate authorities (CAs) yield some interesting results.

Secondly, the concept of controlling access to information is a twofold issue. First, access to information can and should be controlled, and second, what someone can do with information once access has been achieved can in some cases be controlled. Thus, access control as a topic that will be discussed in two sections which are limit access and manage access.

CHAPTER II

LITERATURE REVIEW AND THEORY

2.1 Digital Signature

A digital signature is an electronic method of signing an electronic method of signing an electronic document that is reliable, convenient, and secure. An electronic document includes any instrument created or stored on a computer. These include e-mail, letters, contracts, and images. Many states have enacting legislation concerning the use of digital signatures. The most widely used type of electronic signature relies on public or private key encryption. A public or private key encryption system involves two mathematically related keys that are like large passwords for each user. The private key, known only by the "signer", can encrypt a message that only their public key can decrypt. Once the private or public keys are generated using a valid certificate authority, they remain associated with the person who generates the two keys. The certificate authority is a trusted entity that issues certificates to end entities and other CAs. CAs also issue certificate revocation lists (CRLs) periodically and post certificates and CRLs to a database or repository.

Signing a document is a unique two-step process. First, the message is hashed using a message digest function. This is done because the cryptographic algorithms used to encrypt messages are slow. A short representation of the message (message digest) can be created using a cryptographic algorithm known as a one-way hash function to speed the process. The hash function maps strings of bits to fixed-length strings of bits that make it computationally infeasible to find a given input or output that maps to a particular output or input [1].

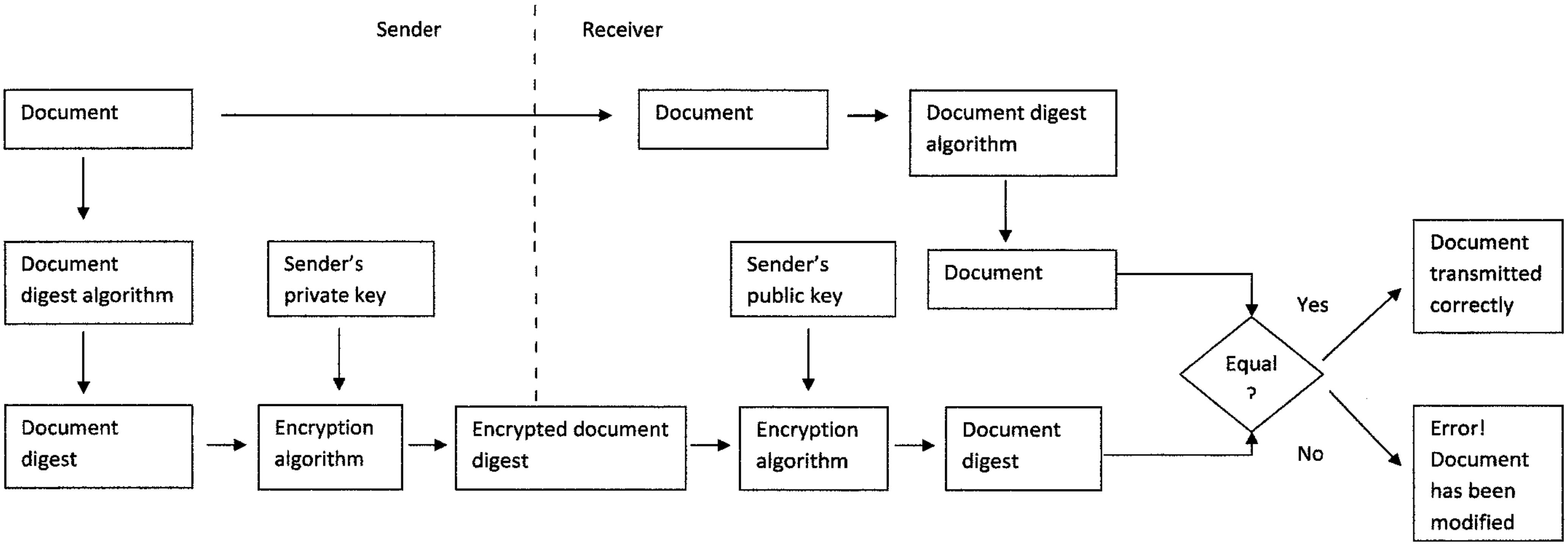


Figure 1: Process in getting digital signature

2.2 The validation process

Figure 2 shows the basic validation process. The relying party presumably knows that the signer is the subject of the certificate, checks the issuer name to determine the subject certificate was issued by the Sub CA, and checks that issuer name to determine the Sub CA certificate was issued by the Root CA. The relying party then verifies the Root CA self-signed certificate using the Root CA public key; verifies the Sub CA certificate using the Root CA public key, and ultimately verifies the subject certificate using the Sub CA public key. The validation of the subject certificate and the verification of the message digital signature provide a degree of authenticity to the relying party but do not necessarily impart data integrity. Because the message content and the system clock are both under the control of the signer, there is a risk that the signer can reset the system clock, change the message content, and sign the modified message. The relying party would then be unable to distinguish between the original message and the modified message, as both display the same time stamp and valid digital signatures. Genuine data integrity requires a cryptographic-based integrity mechanism (e.g. digital signature) and an independent verifiable time stamp.

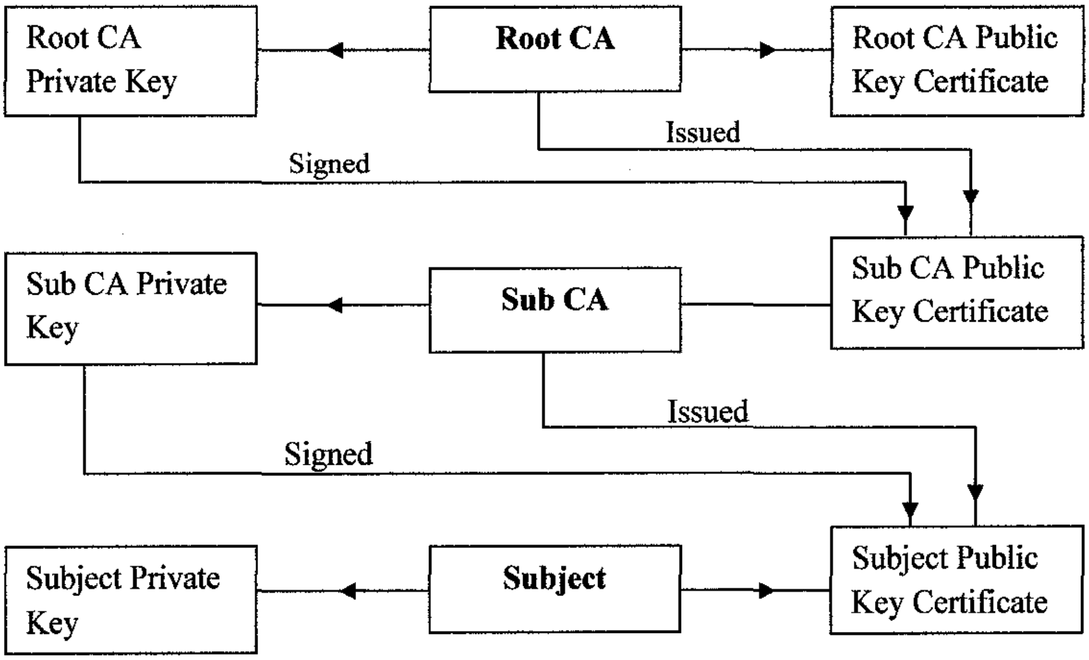


Figure 2: Issuance of Certificate

2.3 Timestamp

A timestamp is a sequence of characters, denoting the date and time at which a certain event occurred. A timestamp is the time at which an event is recorded by a computer, not the time of the event itself. In many cases, the difference may be inconsequential: the time at which an event is recorded by a timestamp example, entered into a log file should be very close to the time of the occurrence of the event recorded. This data usually presented in a consistent format, allowing for easy comparison of two different records and tracking progress over time; the practice of recording timestamps in a consistent manner along with the actual data is called time stamping. Timestamps are typically used for logging events, in which case each event in a log is marked with a timestamp. In file systems, timestamp may mean the stored date and time of creation or modification of a file. [2]

Time stamping is the process of securely keeping track of the creation and modification time of a document. Here security means that once the document has been recorded, no one can be able to change it, provided that the time stamper's integrity is never compromised. The technique is based on hash functions and digital signature. First a hash is calculated from the data which is a hash is a sort of digital fingerprint of the original data: a string of bits that is different for each set of data. If the original data is changed, hash will also change. Anyone trusting the time stamper can then verify that the document had not been posed after the date that the time stamper vouches and also it can no longer be repudiated that the requester of the time stamp was in possession of the original data at the time given by the time stamp. [3]

2.3.1 Trusted Time Stamps

The American National Standard X9.95-2005 Trusted Time Stamps describes the roles, responsibilities and requirements for users of trusted time stamps—time source entities, time stamp authorities, time stamp requestors, and time stamp relying parties. The standard also specifies data objects, processing flows, error handling, and message formats as well as defines technology methods for digital signature, message authentication code, linked-token and transient key. In addition, the standard offers a comprehensive set of time stamp control objectives to validate a trusted time stamp system for use by a professional audit practitioner. It also provides sample time stamp policy and time stamp practice statements. The benefit of trusted time stamp technology is in its ability to verify managed data integrity against a reliable time source provable to any third party.

2.3.2 Hash-Based Time Stamp Tokens

Figure 3 shows the time source entities for a Time Stamp Authority (TSA); and the transaction schema for the issuance of a Time Stamp Token (TST) by a TSA to a Requestor (R). The TSA system clock is calibrated to a regional master clock or a national measurement institute (NMI) whose clock is calibrated to the international time authority (ITA) Bureau International des Poids et Mesures (BIPM) located in France. In the United States, the national measurement institutes (NMI) are the National Institute of Standards and Technology (NIST) and the United States Naval Observatory (USNO). The requestor (R) inputs the data into a HASH algorithm which produces a hash value; the hash value is sent to the RSA in a Request message. The TSA returns a Time Stamp Token (TST) to the requestor, consisting of the (i) original hash value, (ii) a timestamp from the TSA trusted clock, and (iii) a digital signature generated by the TSA using its private key. A relying party, having received the original data and the TST from the Requestor (R), can verify the TSA digital signature using the TSA public key certificate. The relying party also regenerates a hash value from the original data and validates that it matches the hash value in the TST. The validated TST provides a high degree of assurance of the data integrity at the time indicated per the TST timestamp.

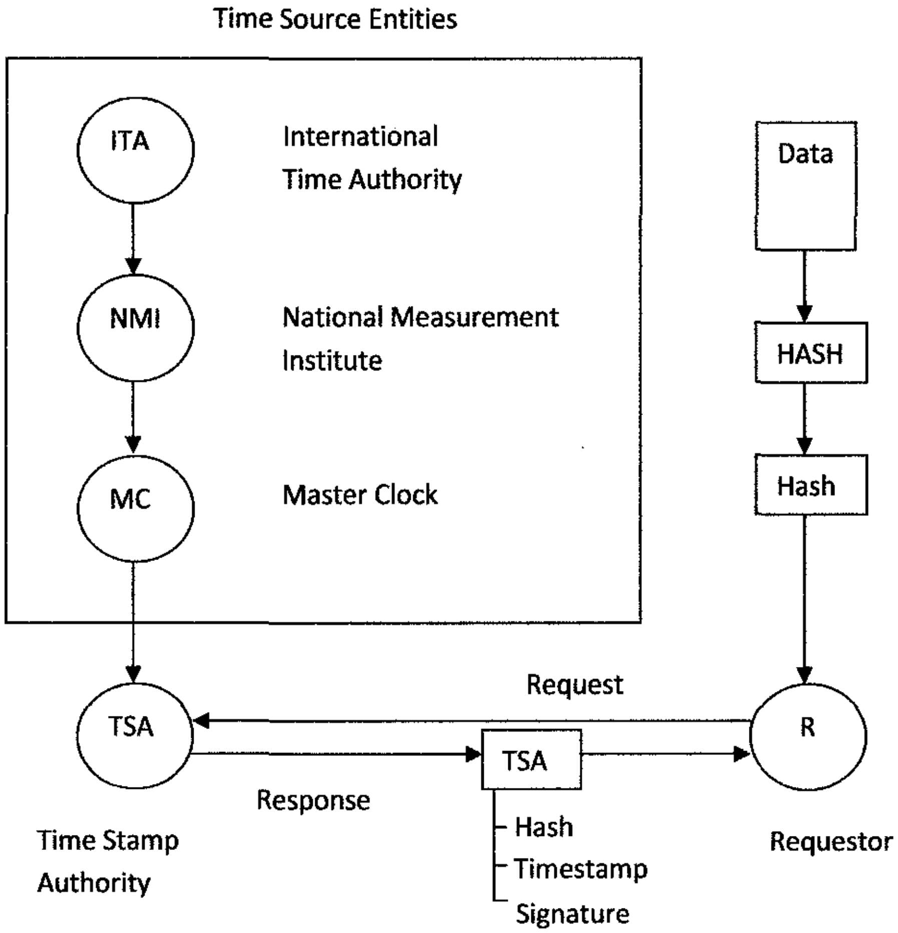


Figure 3: Time Stamp Token with digital signature

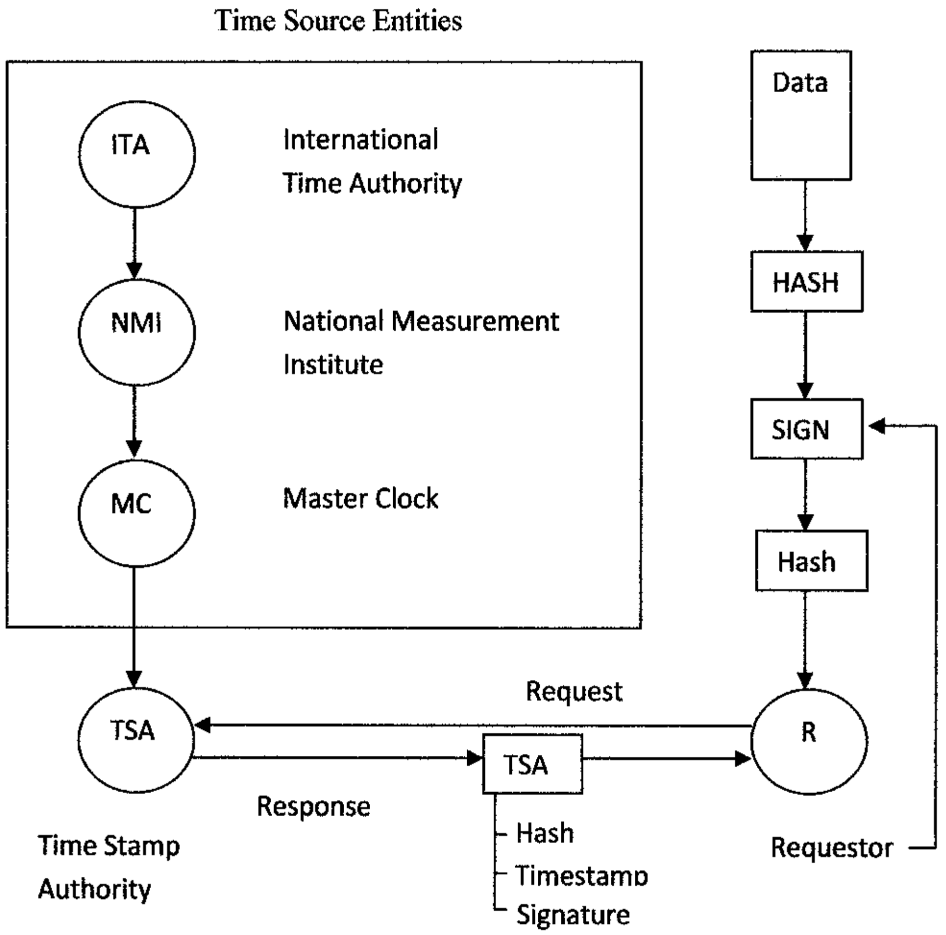


Figure 4: Time Stamp Token with digital signature with the requestor submitting a digital signature of the data in the Request message instead of hash

2.4 Signature-Based Time Stamp Tokens

The trusted time stamp method described in Figure 3 can be combined with a requestor's digital signature to provide an even higher degree of assurance of the data integrity and authenticity. Figure 4 shows the same time source entities and TSA with the requestor submitting a digital signature of the data in the Request message instead of a hash. The TSA generates and returns the TST to the Requestor as normal, containing the digital signature. In this case, the TSA is not cognizant and does not need to know whether a hash or digital signature was sent in the Request message, and incurs no liability of the Requestor's digital signature. The inclusion of the Requestor digital signature in the TST provides the Relying Party with a higher degree of assurance of the data integrity and that the data was digitally signed by the Requestor at the time indicated per the TST timestamp. The risk scenario described above is not theory, and in fact has several real world occurrences. [4]

2.5 Access Control

Managing access controls implies specifying what kind of access is allowed to each user and then enforcing those limitations. The kind of access can range from none at all to full unfettered access with full privileges. There are three schemas that provide a framework for managing access which are discretionary access controls, mandatory access controls, and role based access controls. Each of these stems from a different philosophy in the management of access limitations but can complement each other in complex situations. Access control lists (ACLs) are a key tool in executing the management aspects of each.

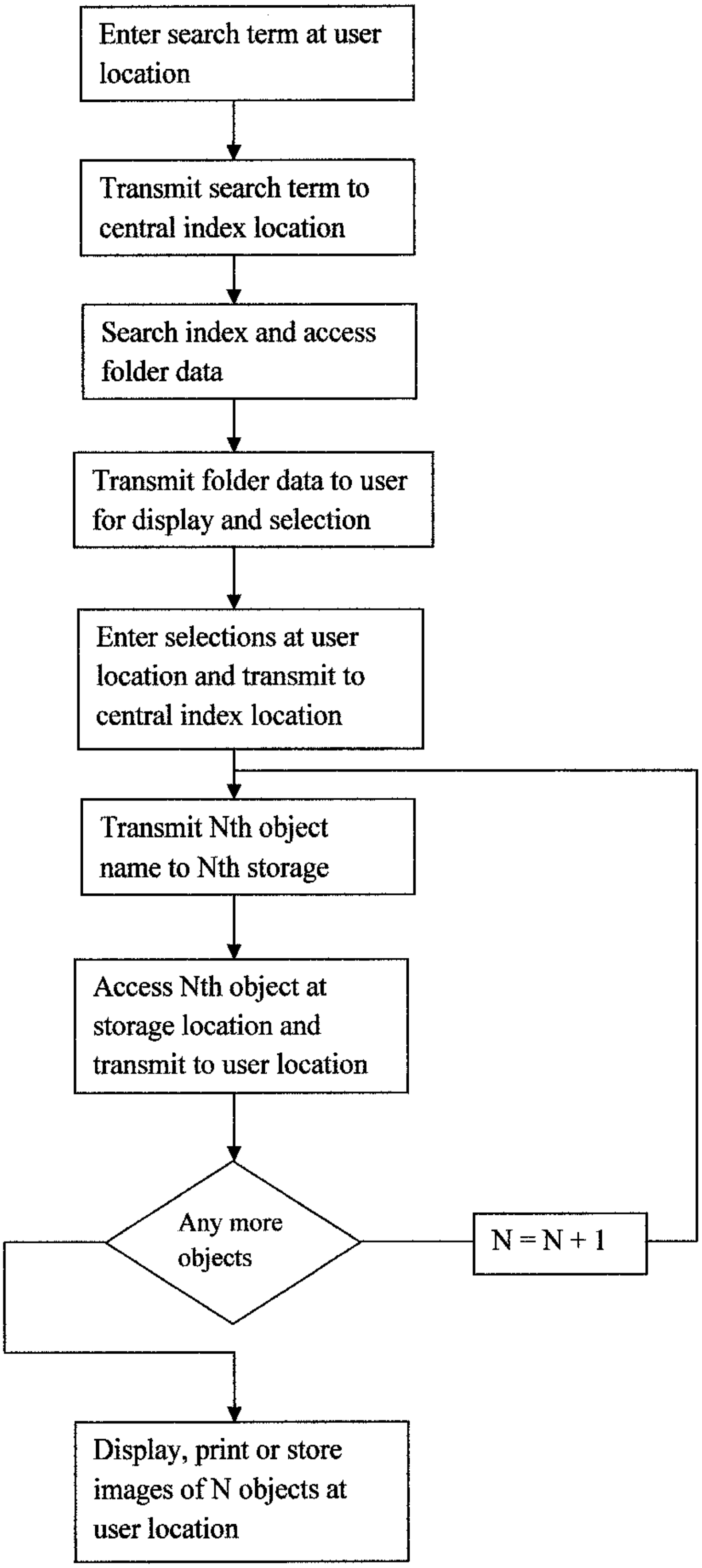
The philosophy underlying discretionary access control (DAC) is that the owner or administrator of the information has the knowledge, skill, and ability to limit access appropriately, to control who can see or work with information. A person running a working group could provide working group members with access to shared files but deny access to anyone else. Alternatively, and more frequently, global access is provided by default.

In single-user situations where physical access to a system is carefully controlled, the identity of the user may not be ascertained directly. In the more common situation in which many users have access to a system, it is important to unambiguously identify specific users. There are three ways this can be accomplished which are passwords, token and biometrics [5].

2.6 Storage and Retrieval

A system is described in which a user enters a search term at a user location that search term being associated with a plurality of image objects which are stored at a plurality of geographically diverse locations within the system. Figure 5 shows that the search term is then transmitted to a central index which includes the object name and location and associated data for each of the distributed objects associated with the search term. Optionally, the associated data may be transmitted from the central location back to the user location for display to the user, who can then select which of the several objects or perhaps all of the objects to be accessed from the diverse storage location. The user can enter his selection at the user location and that selection will be transmitted back to the central index location. [6]

Figure 5: Retrieval Process


Enter search term at user location
Transmit search term to central index location
Search index and access folder data
Transmit folder data to user for display and selection
Enter selections at user location and transmit to central index location
Transmit Nth object name to Nth storage
Access Nth object at storage location and transmit to user location
Any more objects
N = N + 1
Display, print or store images of N objects at user location

CHAPTER III

METHODOLOGY

3.1 Introduction

This chapter will cover the details explanation of methodology that is being used to make this project complete and working well. The method is use to achieve the objective of the project that will accomplish a perfect result. In order to evaluate this project, the methodology based on System Development Life Cycle (SDLC), generally three major steps, which is planning, implementing and analysis.

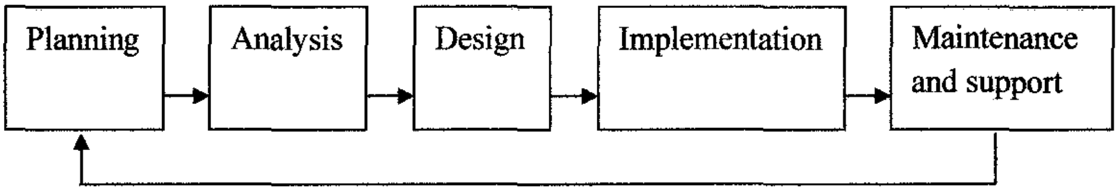


Figure 6: SDLC Phase

This final year project used three major steps to implement the project which are planning, implementing and analysis.

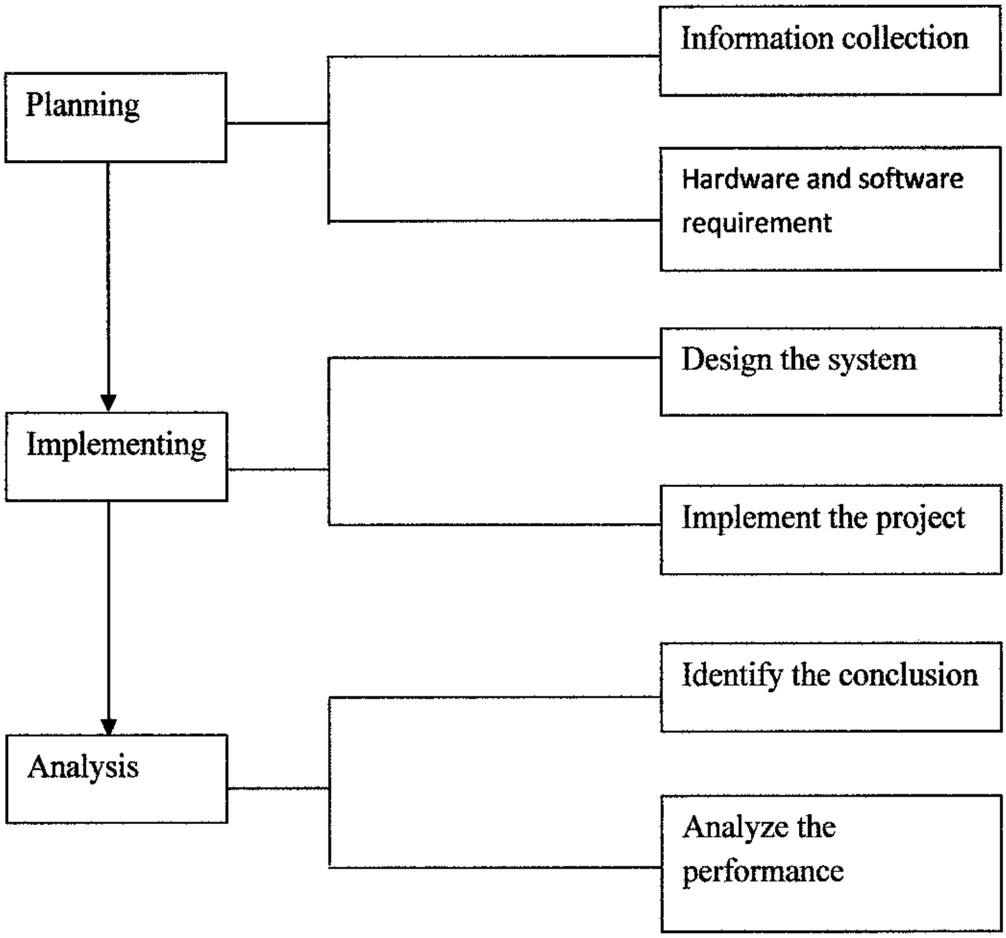


Figure 7: Steps of Methodology

3.2 Planning

To identify all the information and requirement such as hardware and software, planning must be done in the proper manner. The planning phase has two main elements which are information collection and the requirements of hardware and software.

3.3 Information collection

At this point, the supervisor and student have reach agreement on the evaluation work plan, establishing a clear and mutual understanding of how the project is to be carried out and what is to be achieved. Information obtained from journal, text book and research paper. The information collected is then analyzed and distilled into credible, reliable and useful results for presentation.

3.3.1 Hardware and software requirement

Hardware requirement

The most common set of requirements defined by any operating system is the physical computer resources also known as hardware. A hardware requirements list is often accompanied by a hardware compatibility list (HCL), especially in case of operating systems. An HCL lists tested, compatible, and sometimes incompatible hardware devices for a particular operating system or application. The following sub-sections discuss the various aspects of hardware requirements for this project.

Architecture

All computer operating systems are designed for particular computer architecture. Most software applications are limited to particular operating systems running on particular architectures. Although architecture-independent operating systems and application exist, most need to be recompiled to run on a new architecture.

Processing power

The power of the central processing unit (CPU) is a fundamental system requirement for any software. Most software running on x86 architecture define processing power as the model and the clock speed of the CPU. Many other features of a CPU that influence its speed and power, like bus speed, cache, and MIPS are often ignored.

Memory

All software, when run, resides in the random access memory (RAM) of a computer. Memory requirements are defined after considering demands of the application, operating system, supporting software and files, and other running process. Optimal performance of other unrelated software running on a multi-tasking computer system is also considered when defining this requirement.

Secondary storage

Hard-disk requirements vary, depending on the size of software installation, temporary files created and maintained while installing or running the software, and possible use of swap space.

Display adapter

Software requiring a better than average computer graphics display, like graphic editors and high-end games, often define high-end display adapters in the system requirements.

Peripherals

Some software applications need to make extensive and special use of some peripherals, demanding the higher performance or functionality of such peripherals. Such peripherals include CD-ROM drives, keyboards, pointing devices, network devices, etc.

Software requirement

Software requirements deal with defining software resource requirements and pre-requisites that need to be installed on a computer to provide optimal functioning of an application. These requirements or pre-requisites are generally not included in the software installation package and need to be installed separately before the software is installed.

Platform

In computing, a platform describes some sort of framework, either in hardware or software, which allows software to run. Typical platforms include a computer's architecture, operating system, or programming languages and their runtime libraries. For this project, the platform that will be used is .NET, JAVA and PHP.

3.4 Analysis

Requirements analysis is done in order to understand the problem for which the system is to solve. For example, the problem had been focused more on security of the digital asset and the flexibility of user excess to the digital asset. Understanding the requirements of the system is a major task. The emphasis in requirements analysis is on identifying what is needed from the system. There are two major activities in this phase which are problem understanding; means has to understand the problem and its context. Such analysis typically requires a thorough understanding of the existing system, the parts of which must be automated. Once the problem is analyzed and the essentials understood, the requirements must be specified in the requirement specification document. The requirements documents must specify all functional and performance requirements, the formats of inputs, outputs and any required standards, and all design constraints that exist due to political, economic environmental, and security reasons. The phase ends with validation of requirements specified in the document. The basic purpose of validation is to make sure that the requirements specified in the document, actually reflect the actual requirements or needs, and that all requirements are specified. Validation is often done through requirement review in which a group of people including me and my supervisor critically review the requirements specification.

3.5 Design

The design stage takes as its initial input the requirements identified in the approved requirements document. For each requirement, a set of one or more design elements will be produces as a result of interviews, discussion and prototype efforts. Design elements describe the desired software features in detail, and generally include functional hierarchy diagrams, screen layout diagrams, tables of system rules, system process diagrams, pseudo code, and a complete entity-relationship diagram with a full data dictionary. These design elements are intended to describe the software in sufficient detail that skilled programmers may develop the software with minimal additional input.

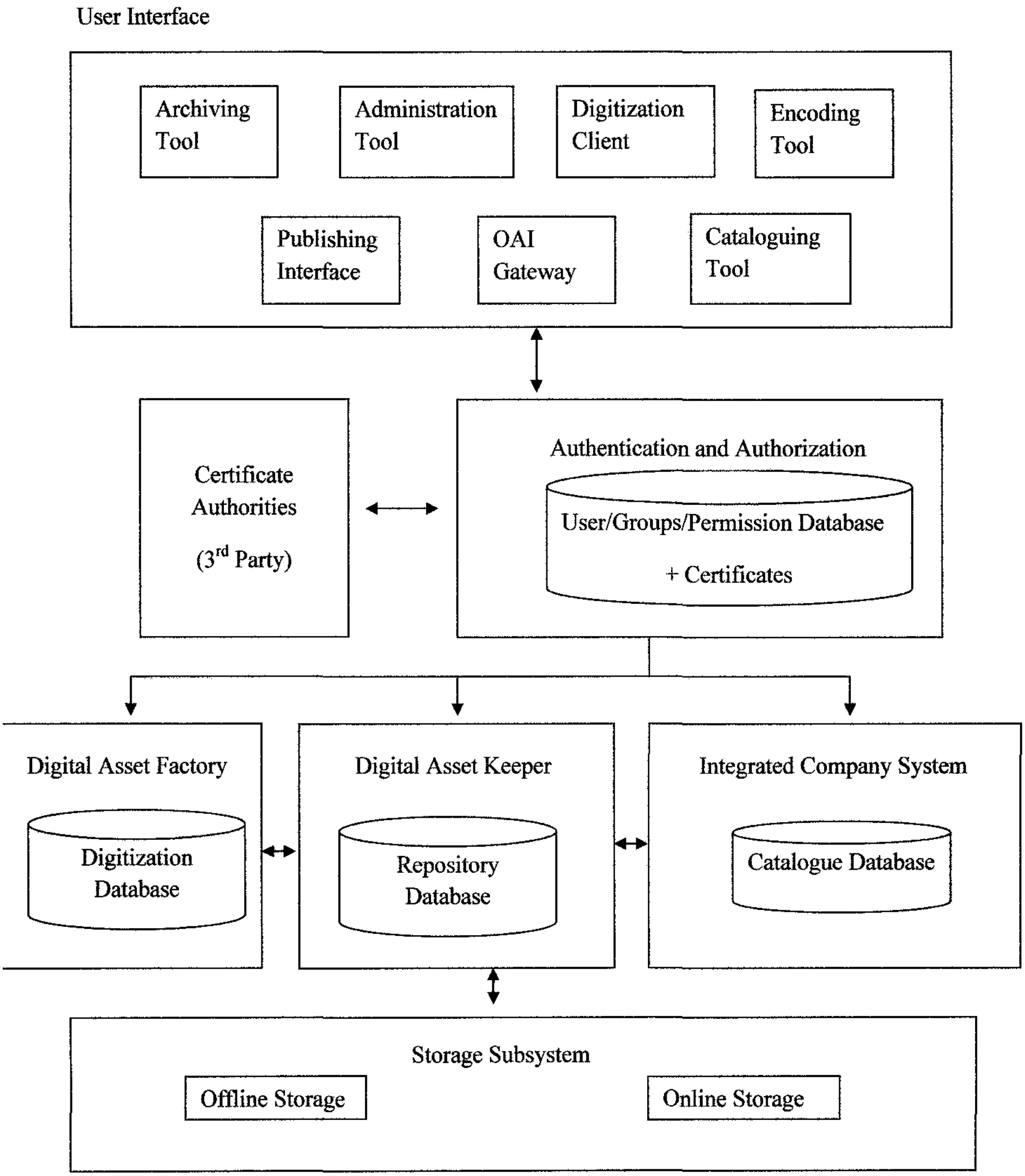


Figure 8: System architecture of SDAR

The architecture of Secure Digital Asset Repository (SDAR) is depicted in Figure 8.

The system core consists of two fundamental modules:

- The Digital Assets Keeper (DAK) which acts as a repository for digital assets and,
- The Digital Assets Factory (DAF) which is responsible for the automation of the digitization workflow.

Both systems interact with the digital objects storage system. The storage system is used to store digital files either for online access and publishing purpose, or offline for long-term preservation. The system contains a set of user interfaces that interact with the system components through Application Program Interface (APIs). The user interfaces provide tools for the automation of the digitization process, metadata entry, searching and browsing the repository content, and tools for the interoperability with other repositories. An authentication and authorization system controls the access to the repository contents and functionalities based on the user identity. The repository is integrated with the Integrated Company System . The system is implemented in HTML, JavaScript, PHP, XML and .NET. The web-based components are implemented as ASPX pages running on Microsoft IIS web server. The repository APIs are implemented as Web services. SQL server database is used as the main repository database.

3.5.1 Digital Assets Keeper – DAK

The DAK acts as a repository for digital material in a digital format. All metadata related to a digital object is stored in the DAK repository database. Within the SDAR data model, the system holds several categories of metadata describing assets and its digital reproductions.

1 Digital Content Metadata

This includes metadata describing a Digital Object. SDAR supports a variety of digital objects' formats including JPG, TIFF, PDF, Text and others. Metadata such as image resolution, dimensions or profile are extracted from digital files automatically and stored in DAK. New formats can be introduced into the system and appropriate tools can be integrated to deal with the new file formats.

2 Publishing Metadata

Encoded objects for publishing are stored on online storage. The publishing metadata includes the path of the published Digital Object on the server, the date of publishing as well as the category of targeted users e.g. students, researchers, clients, suppliers, etc.

3 Authentication and Authorization Metadata

SDAR users are identified by a username and password. Further, user groups are defined where a user can belong to one or more groups. Permissions are given to each user or group. Anonymous access to the repository is also allowed, the access right of an anonymous user are defined by the permissions assigned to a special group Guest. However, group Guest has its disadvantage where they have expiry date of accessing the system. To authenticate the document upload by user, the user will be given a digital certificate which is a special kind of machine-readable document issued by a trusted Certificate Authority (3^{rd} party) to an individual or organization which is unique to them. A certificate authority is an authority in a network that issues and manages security credentials and public keys for message encryption. As part of a public key infrastructure (PKI), a CA checks with a registration authority (RA) to verify information provided by the requestor of a digital certificate. If the RA verifies the requestor's information, the CA can then issue a certificate. Digital certificate typically contain the Owner's public key, expiration date of the public key, the expiration date of the certificate, name of the issuer (CA), the owner's name, serial number and the Digital signature of the issuer (CA). Other than that, time stamp is also being used for logging event. In this system, timestamp may mean the stored date and time of creation or modification of a file.

3.5.2 Digital Asset Factory – DAF

The DAF governs the digitization process of the company or organization collection at the digital lab. DAF realizes one of the main goals of SDAR which is the automation of the digitization process. This supports the digitization of company assets including textual material, slides, maps and others. It provides the digital lab operators with tools for entering a digitization job metadata, keeping track of digitization status, applying validation tests on digitized material, archiving the digitized material for long term preservation and retrieving the archived material when needed. The system supports different workflows for different workflows for different types of material.

The asset passes by the general phases depicted in Figure 9:

- Scanning the material.
- Processing the scanned files to enhance the quality.
- Encoding the digitized material by generating a version suitable for publishing.
- Archiving the output of each step of the digitization. Two offline backups are taken for a file, one on CD and the other on tape. Encoded versions are moved to online storage for publishing.

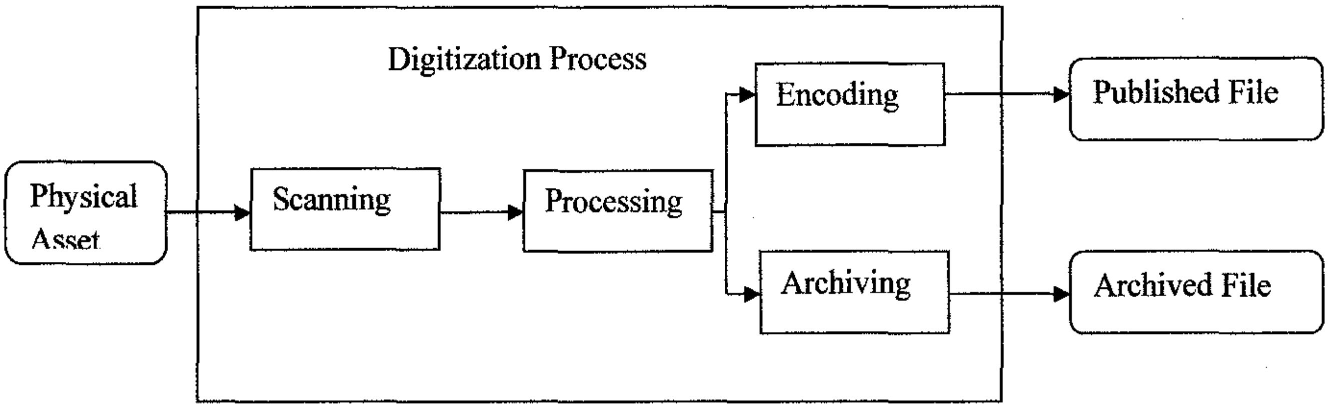


Figure 9: Digitization Phases

The main goals of DAF are:

- To provide a database system to keep track of the digitization process through the scanning, processing, archiving and publishing.
- To keep track of digitized materials; unifying the naming conventions and exhaustively checking the produced folders and files for consistency.
- To provide timely reports to various levels of management describing the workflow on a daily, weekly or longer basis and to allow online queries about the current status of a certain asset.
- To apply necessary encodings on the scanned materials to be suitable for electronic publishing.
- To manage the archiving and retrieval of the digitized material.

The SDAR system deals with three types of users which are digitization operators, administrators which are divided into cataloguers and reviewers and the end user which categorized into registered user and special user. Each type of users is provided with tools to make use of the system functionalities.

1. Administration Tool

The Administration Tool is one of the DAF Web-based tools used by the operator in the digital lab. The tool is used to initiate a new job by entering minimal descriptive metadata for the material to be digitized. If the material is catalogued in the company catalogue, the Integrated Company System id – an item barcode, for example – is used to retrieve metadata from the company catalogue. This id is also used to link the record in SDAR to the one in the library catalogue for future synchronization. If the material is not previously catalogued, the operator enters the minimal metadata that can be deduced from the physical item in hand. The tool uses this metadata to derive a unique folder name for the scanned files.

2. Digitization Client

The digitization client is a DAF application used by the operator in the digital lab. The tool creates structured folders for new digitization jobs and after the completion of digitization phase, the digitization client tool is used to perform the following:

- Validate the files; generate warnings if any inconsistencies are detected.
- Update the job status in the database by setting the username for the operator who performed the job, the job completion date and the count of produced files.
- Move the folders and files to the queue of the next digitization phase on a storage server. Before moving any folder, a lock is acquired on the folder and sub files to avoid concurrent access to the folder while moving.

The digitization client is used by the operator through the three main digitization phases; scanning, processing and reprocessing.

Scanning

Physical assets submitted to the lab for digitization are placed in a scanning queue. The operator retrieves a job from the queue and uses the digitization client to create the folder structure where scanned files are to be stored. Mainly, the digitization folder contains three subfolders for three types of files: the original scanned files, the processed files and the encoded output. The encoded output, the folder structure and the scanning resolution varies according to the material type; text, image, audio or video. When the scanning is done, the digitization client places the produced files in the processing queue.

Processing

The operators use the digitization client to retrieve a job processing queue. A combination of manual and automated image processing tools is used to enhance the quality of the scanned images. After the job is done, the digitization client places the job directly to the archiving queue for other types of material.

Reprocessing

The system supports a special workflow for reprocessing a digitized material. Reprocessing may be needed to generate new publishing format of the digitized material. Reprocessing begins by searching and retrieving the files to be reprocessed from the archive. The files are then placed in the appropriate digitization queue. The reprocessed files go through the normal digitization steps described before until they reach the archiving phase. Only altered files are re-archived, changes in files are detected using checksums that are calculated before and after the reprocessing. The archiving information of a new file version is recorded in the repository database and a link is made to the parent version archiving location so that file versions may be tracked in the database from the most recent to the base version.

3. Archiving Tool

In the current version, a digital object is represented by one or more files with different formats and resolutions, these files are stored for online access or on offline storage for long-term preservation. Typically, the preserved material is the scanned originals and the processed version with high resolution. Lower versions derived for publishing purpose are saved on online storage for ease of access; this includes low resolution JPG, PDF and DJVU. Files stored offline are archived on two medias; CDs and tapes. Unique labels are generated, printed and attached to the media for future retrieval. The system keeps track of different versions of a file by linking a newer version to its older one.

The archiving tool is one of the DAF Windows-based applications used by the lab operators and administrator and offers the following functionalities:

- Checking files and folders consistency.
- Preparing the folders for archiving by compressing the subfolders and files, grouping them into bundles that fit into the media capacity (CD or tape), generating the media label, printing the label.
- The tool generates checksums for the archived files to detect changes in case of downloading and reprocessing a file.
- A search facility enables the user to retrieve an archived folder by locating the folder, uncompressing the subfolders and files and copying the uncompressed files and folders to a destination specified by the user.
- Managing the space on the storage server hard drives, the tool generates warnings when storage level exceeds a predefined value for each drive.
- The tool updates the DAK database by recording the archiving information related to a digital file.

4. Encoding Tool

In the encoding step, a final product is generated for publishing. For images, slides and maps, different JPG resolutions are generated. For audio and video, different qualities are generated to accommodate for different network connections' speed.

5. Cataloguing Tool

The cataloguing tool is a Web-based application used by the administrator to add and edit metadata in the DAK subsystem. Using the cataloguing tool, the administrator enriches the digital repository records by adding metadata. The administrator can also create new records for digital objects and upload their corresponding files. The repository is preloaded with controlled vocabularies lists. The tool allows defining configurable templates, importing metadata from external sources and automatic extraction of digital content metadata.

6. Publishing Interface

The publishing interface is a Web-based interface related to the DAK that provides access to the repository of digital objects through search and browsing facilities. The repository publishing interface offers the following functions:

- Browse the repository contents by collection, categories, subject, creator and title.
- For textual material, a search in the full text can be conducted.
- For images, different levels of zooming are available.
- Display brief and full record information with links to the digital objects.
- Hyperlinked data fields that can invoke searchers e.g. by Keywords, Subjects and Creator.

7. OAI Gateway

DAR OAI Gateway implements the Open Archive Protocol to provide access to the repository contents across an organization's architecture. The Gateway receives XML requests and translates them to the equivalent database queries. When the request result sets are retrieved, the gateway translates them into XML and responds to the requesting application.

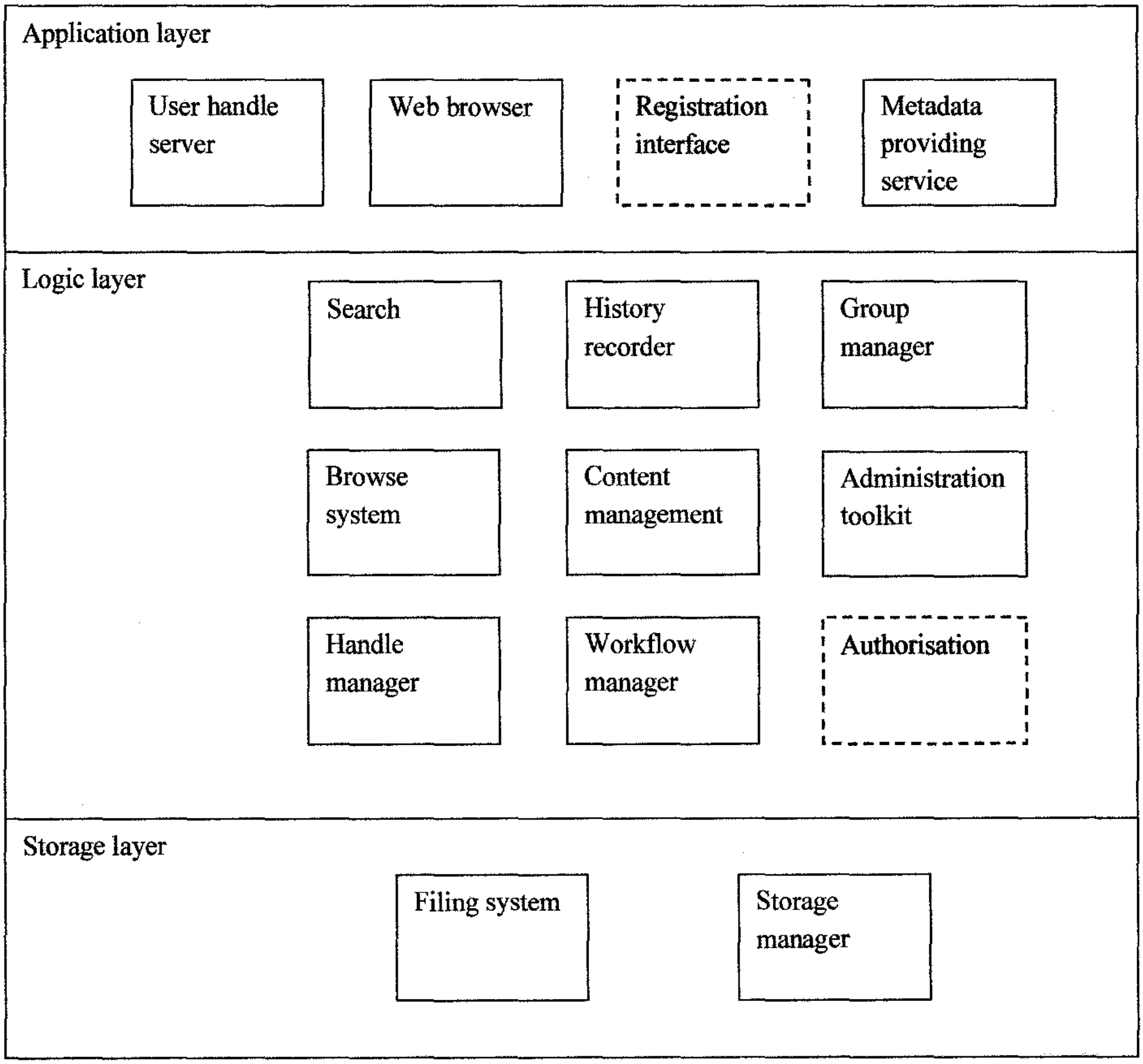


Figure 10: Functional layer in the Secure Digital Asset Repository system

Based on the Figure 10 above, we focus more on the discussion about the registration process that related to a strict authorisation since the system is emphasize on the security access control.

Registration

The system (organization) holds registration information both on the user's computer and on system server. The system uses this information to authenticate user and to assure the user seeking authentication are accessing SDAR server. SDAR uses public key encryption as defined before to accomplish this. The SDAR server has a private key, and distributes that key's public counterpart with each time the user want to access the system. As part of registration, the user selects a desired username and password. The SDAR locally generates public and private keys. The private key and hash of the password are stored as securely as possible on the user's computer. Then a 256-bit advanced encryption standard-encrypted session is established with the SDAR server. The client creates a session key using its random number generator. The SDAR server verifies that the selected username is unique and that follows SDAR's naming rules. The server stores the username and a hash of the hash of the user's password in its database. The server now forms and signs an identity certificate for the username that binds the username, its verification key and the key identifier.

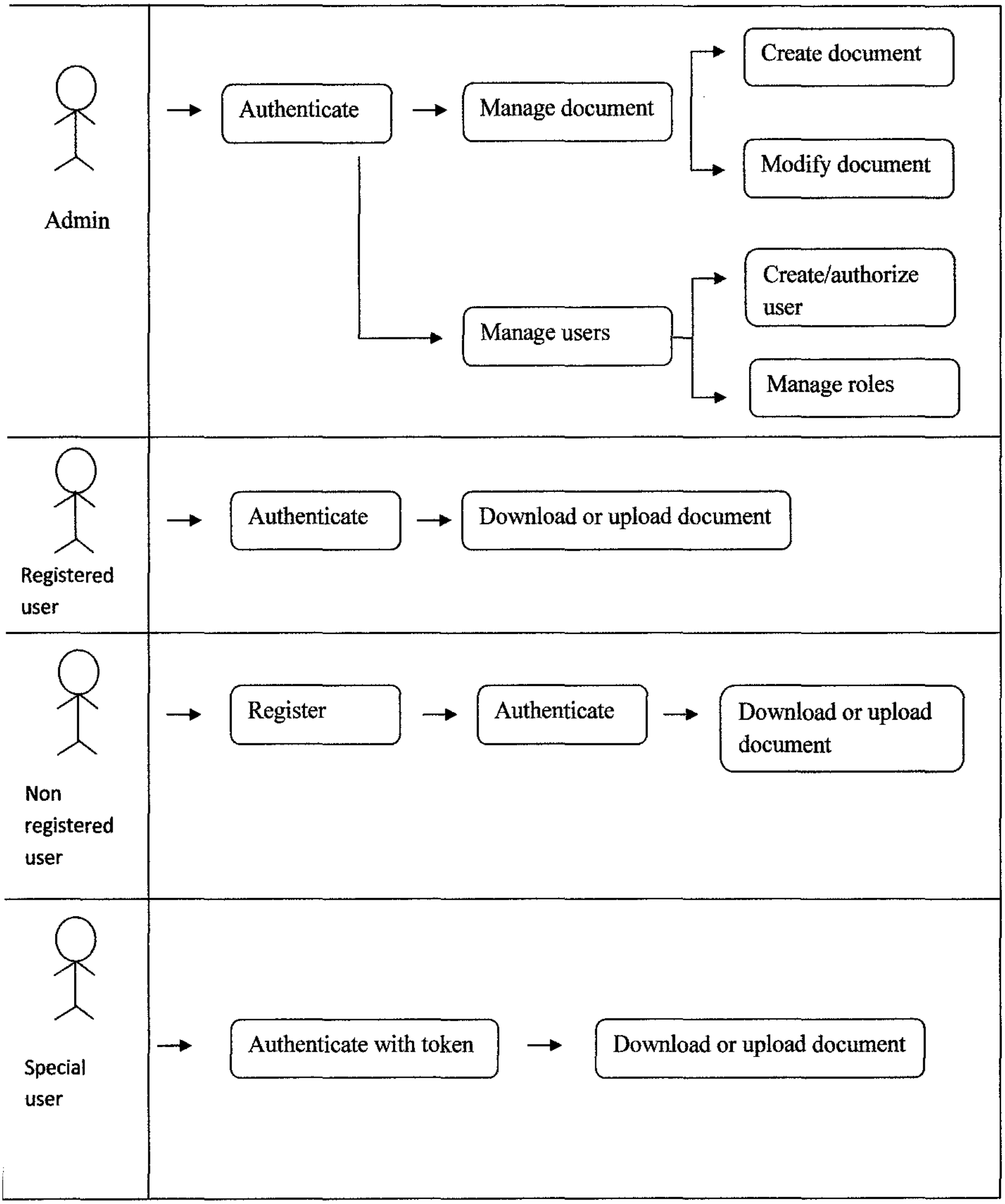


Figure 11: Process involve in Secure Digital Asset Repository system

3.6 Features of the System

3.6.1 Digital Signatures

Digital Signatures are industry-standard mathematical functions that provide a mechanism for uniquely identifying pieces of data and revealing if the data has been modified. For every asset that is loaded into the system, Secure Digital Asset Repository computes multiple digital signatures and stores these signatures with the asset as long as the asset remains in the system. The digital signatures of assets play a key role in proving that the assets remain unchanged as they are managed. Using multiple signature functions yields a more robust system which is protected from possible weaknesses that may arise from single signature functions.

3.6.2 Import Audit Log

As part of the Secure Digital Asset Repository secure workflow, the system can produce hard-copy import audit logs that contain the assets' thumbnails and digital signatures for storage outside of the system. This is particularly useful in the event that the system itself is challenged. On demand, Secure Digital Asset Repository can recompute any particular asset's digital signatures, which can then be compared against the original audit logs, thereby proving that the asset coming out of the system is mathematically identical to the asset that went into the system.

3.6.3 Import Stations

In an idealized secure environment, Secure Digital Asset Repository recommends locked-down image import stations that have secure connections to the Secure Digital Asset Repository server. However, the system is not limited to this deployment scenario. Regardless of the workstation that is performing an import, all communication between the workstation and Secure Digital Asset Repository is securely encrypted.

3.6.4 Organization

Secure Digital Asset Repository allows the organization of assets according to desired conventions. The system supports an arbitrary number of categories that can be layered in complex hierarchies. Each category can be grouped under more than one parent, allowing multiple logical organizations. Categories may be divisions, clients, cases, photographers, crime scenes, or any other logical grouping required.

3.6.5 Access Control

Secure Digital Asset Repository provides an extensive fine-grained security model. This model allows system administrators to control access at both the user and the group level. The security administrator can enumerate the specific users and/or groups that are allowed to perform various operations -- view, organize, and download -- on the assets. Access to any assets in the system is not granted without permission from the security administrator.

3.6.6 Secure Backups

Backup is a crucial component in providing reliability in an asset management system. Executing a backup plan for such a system is often complicated because of the enormous volume of data involved (often terabytes of data). Secure Digital Asset Repository provides multiple methods for backing up data. At the database level, Secure Digital Asset Repository supports full database replication to a backup database system that may be onsite or located remotely. Secure Digital Asset Repository can also produce per-import, daily or periodic, 'chunked' backups, which are sized to the desired media type (CD, DVD, etc). All backups from Secure Digital Asset Repository are encrypted to ensure that, in the event of a data restoration, the integrity of the assets is maintained. Secure Digital Asset Repository allows both partial and full restorations of data. Partial restoration is useful for fast recovery of more recent content, while full restoration is used in the event of a catastrophic loss.

3.6.7 Indexing and Searching

Secure Digital Asset Repository maintains an index of all assets that are imported into the system. This index allows fast and efficient searching across multi-terabyte asset databases based on multiple attributes of the asset including (but not limited to) caption, filename, and orientation.

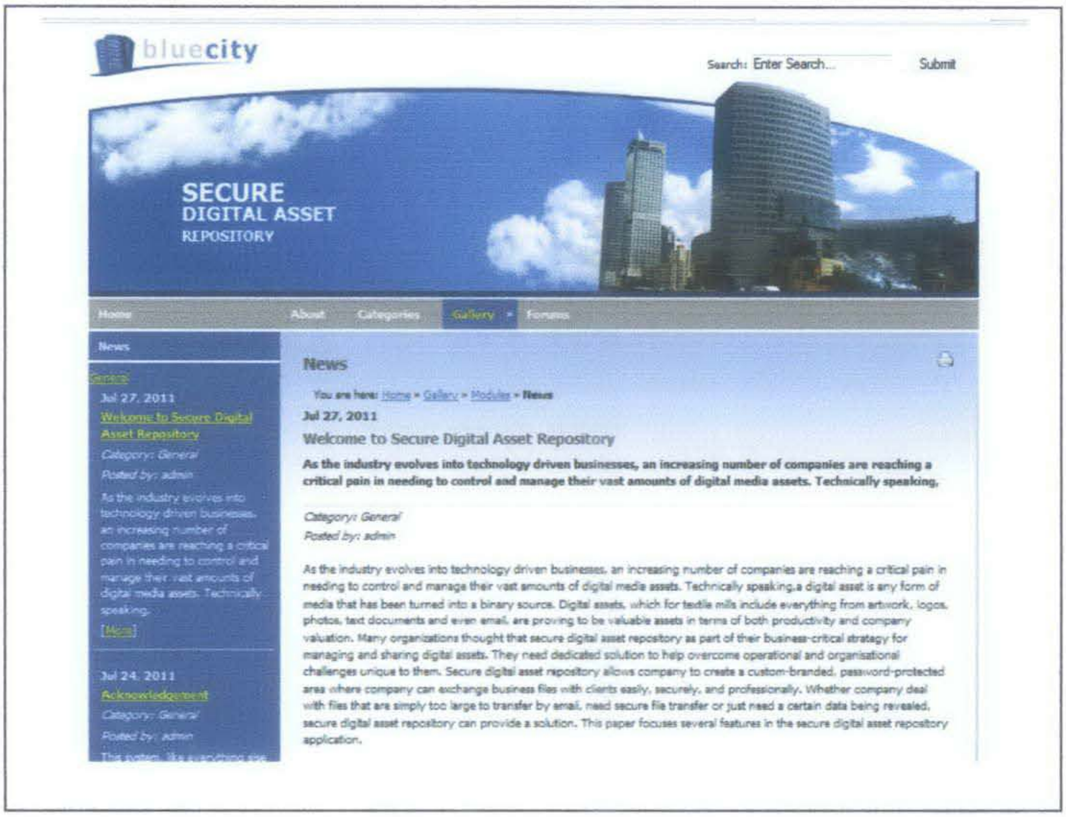
3.6.8 Review and Request Management

Most organizations that manage asset collections must also fulfil output requests for those assets. Secure Digital Asset Repository provides several mechanisms for managing the review and request process. Users of the system can browse the catalogue of assets and build multiple personal galleries of images for later reference. Additionally, users can, given appropriate permissions, email expiring claim tickets to personal galleries. Recipients of these claim tickets can be granted visitor access to the gallery. For each asset in a gallery, Secure Digital Asset Repository provides and preserves a discussion thread where gallery owners and viewers can converse about the asset. From a gallery, users can choose to download or request various versions of the assets. Secure Digital Asset Repository allows users to choose the desired output size and format and can automatically fulfil the requests. If a user does not have permission to download assets, a request for an administrator to output the assets may be submitted. Secure Digital Asset Repository also provides a management console for reviewing pending output requests so that designated users can track requests. Secure Digital Asset Repository provides an extensive feature-set that automates and assists in many aspects of an asset management organization. From import to output fulfilment and management, the system ensures asset integrity and security with multiple mathematically proven techniques. Since security measures are also included in the backup system, organizations can be assured their assets are protected and the authenticity of the assets will withstand even the strongest scrutiny.

CHAPTER IV

RESULTS AND DISCUSSION

4.1 User Interface (Application Layer)



General

Jul 27, 2011

Welcome to Secure Digital Asset Repository

Category: General

Posted by: admin

As the industry evolves into technology driven businesses, an increasing number of companies are reaching a critical pain in needing to control and manage their vast amounts of digital media assets. Technically speaking.

[More]

Jul 24, 2011

Acknowledgement

Category: General

Posted by: admin

This system, like everything else

News

You are here: Home » Gallery » Modules » **News**

Jul 27, 2011

Welcome to Secure Digital Asset Repository

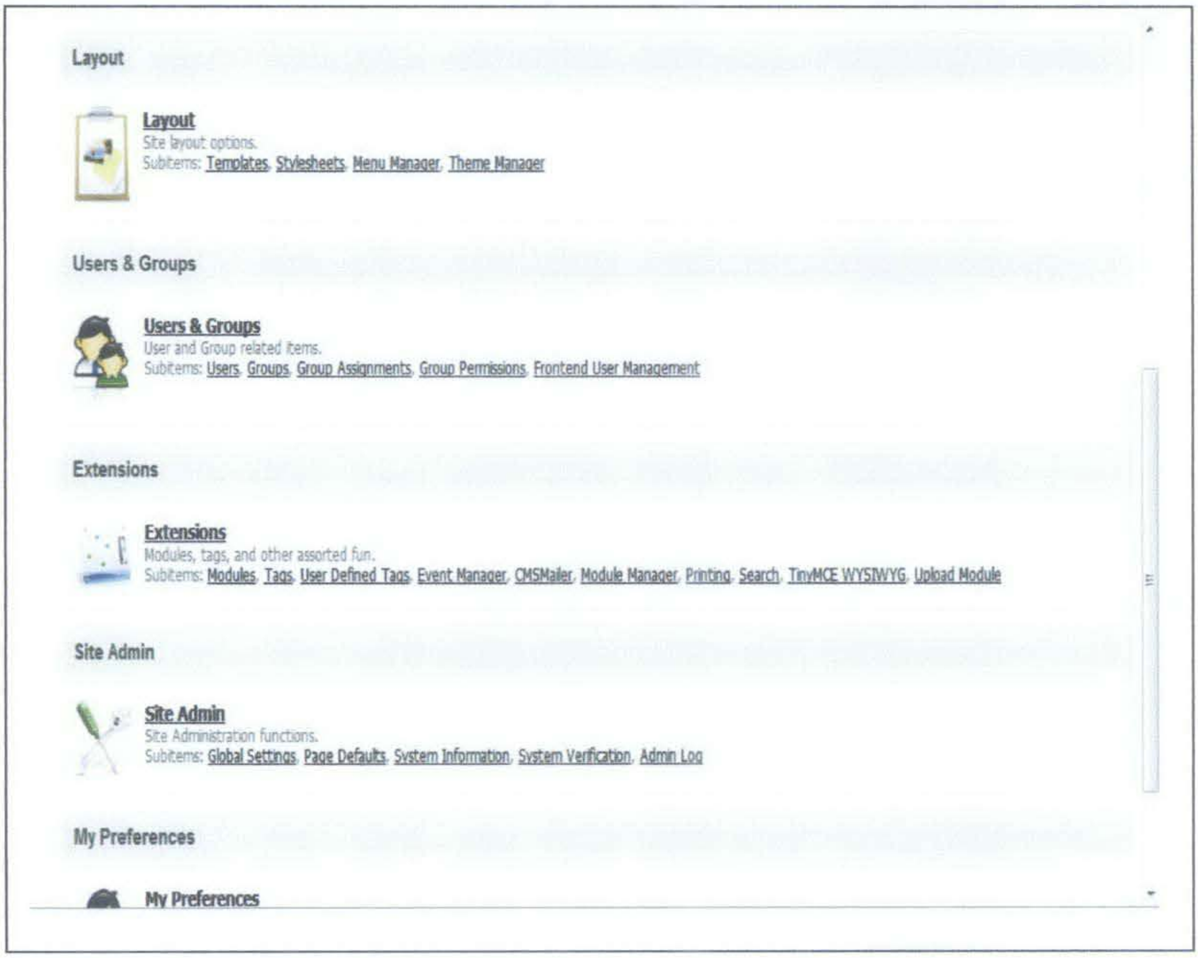
As the industry evolves into technology driven businesses, an increasing number of companies are reaching a critical pain in needing to control and manage their vast amounts of digital media assets. Technically speaking,

Category: General

Posted by: admin

As the industry evolves into technology driven businesses, an increasing number of companies are reaching a critical pain in needing to control and manage their vast amounts of digital media assets. Technically speaking,a digital asset is any form of media that has been turned into a binary source. Digital assets, which for textile mills include everything from artwork, logos, photos, text documents and even email, are proving to be valuable assets in terms of both productivity and company valuation. Many organizations thought that secure digital asset repository as part of their business-critical strategy for managing and sharing digital assets. They need dedicated solution to help overcome operational and organisational challenges unique to them. Secure digital asset repository allows company to create a custom-branded, password-protected area where company can exchange business files with clients easily, securely, and professionally. Whether company deal with files that are simply too large to transfer by email, need secure file transfer or just need a certain data being revealed, secure digital asset repository can provide a solution. This paper focuses several features in the secure digital asset repository application.

4.2 **Administration interface** (Logic layer)



Layout

Layout
Site layout options.
Subitems: Templates, Stylesheets, Menu Manager, Theme Manager

Users & Groups

Users & Groups
User and Group related items.
Subitems: Users, Groups, Group Assignments, Group Permissions, Frontend User Management

Extensions

Extensions
Modules, tags, and other assorted fun.
Subitems: Modules, Tags, User Defined Tags, Event Manager, CMSMailer, Module Manager, Printing, Search, TinyMCE WYSIWYG, Upload Module

Site Admin

Site Admin
Site Administration functions.
Subitems: Global Settings, Page Defaults, System Information, System Verification, Admin Log

My Preferences

My Preferences

4.3 Database interface (Storage layer)

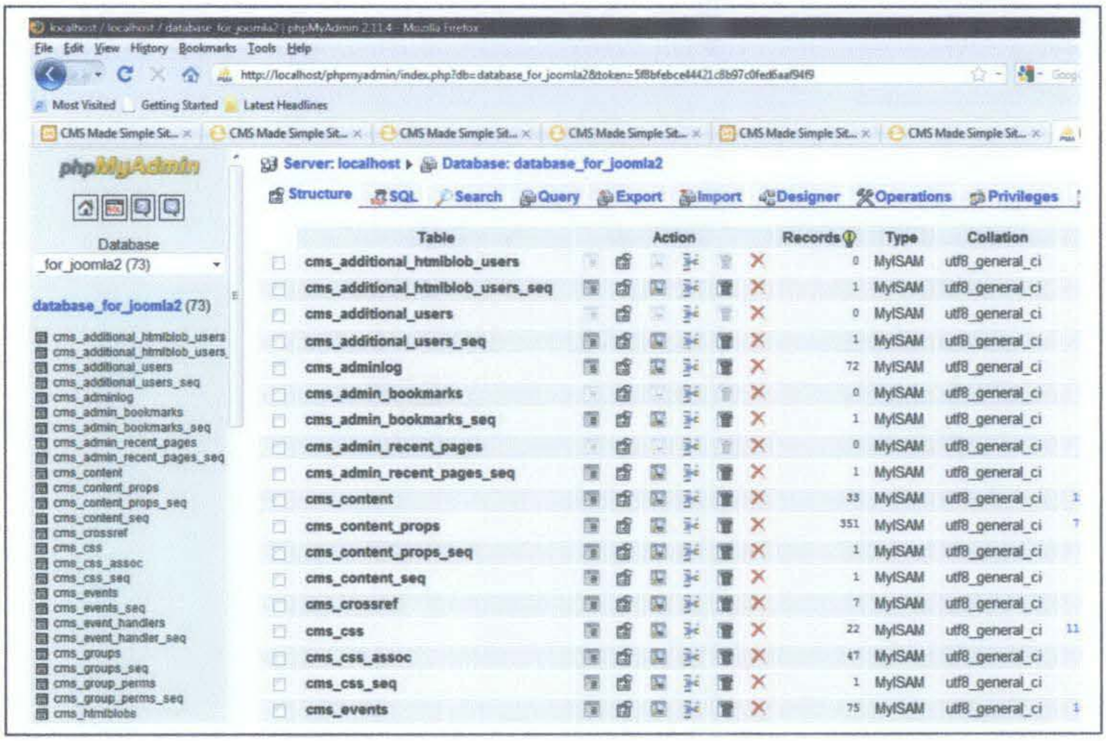
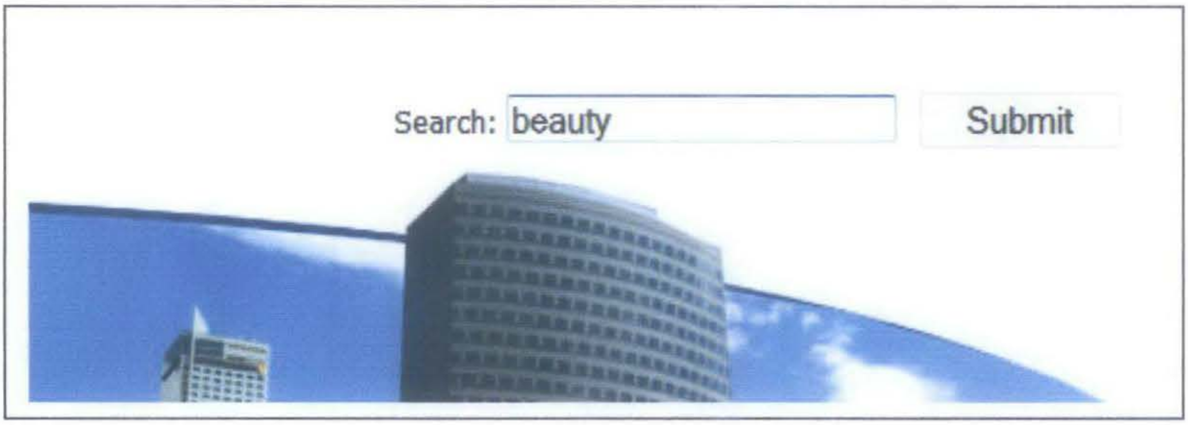


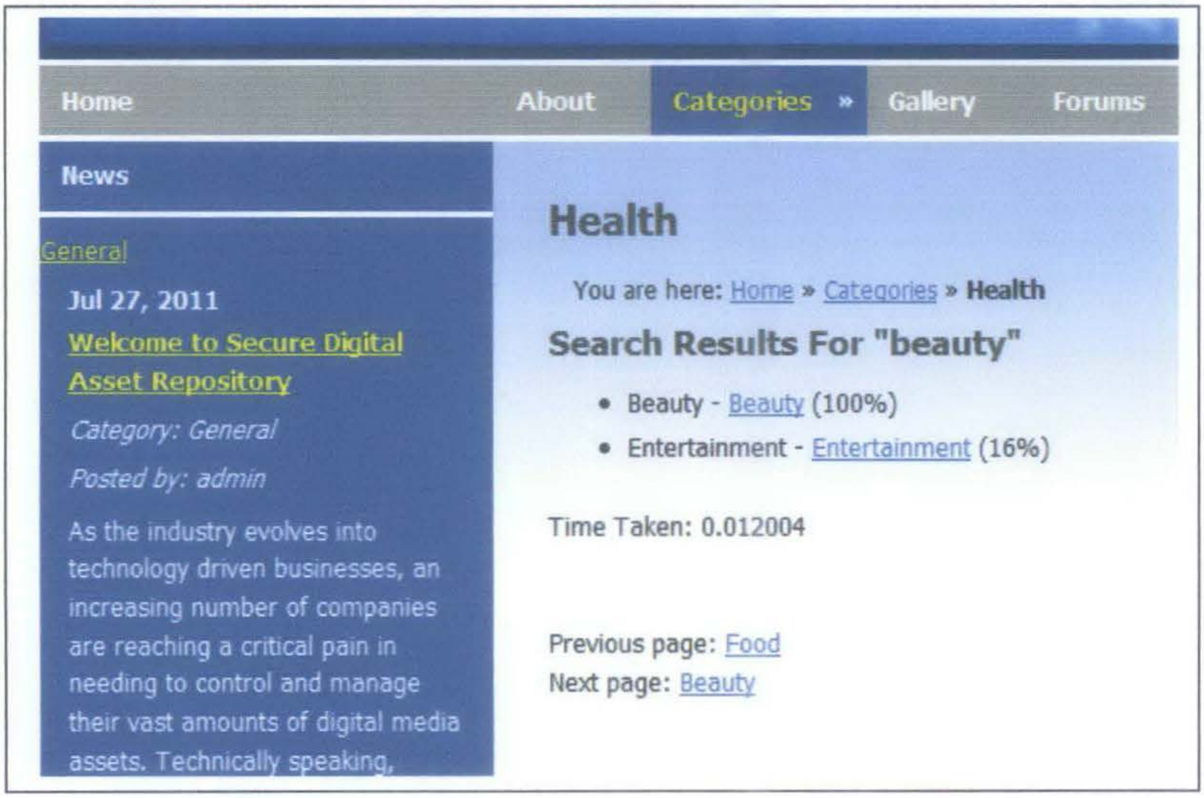
	Table	Action	Records	Type	Collation	
□	cms_additional_htmlblob_users	✕	0	MyISAM	utf8_general_ci	
□	cms_additional_htmlblob_users_seq	✕	1	MyISAM	utf8_general_ci	
□	cms_additional_users	✕	0	MyISAM	utf8_general_ci	
□	cms_additional_users_seq	✕	1	MyISAM	utf8_general_ci	
□	cms_adminlog	✕	72	MyISAM	utf8_general_ci	
□	cms_admin_bookmarks	✕	0	MyISAM	utf8_general_ci	
□	cms_admin_bookmarks_seq	✕	1	MyISAM	utf8_general_ci	
□	cms_admin_recent_pages	✕	0	MyISAM	utf8_general_ci	
□	cms_admin_recent_pages_seq	✕	1	MyISAM	utf8_general_ci	
□	cms_content	✕	33	MyISAM	utf8_general_ci	1
□	cms_content_props	✕	351	MyISAM	utf8_general_ci	7
□	cms_content_props_seq	✕	1	MyISAM	utf8_general_ci	
□	cms_content_seq	✕	1	MyISAM	utf8_general_ci	
□	cms_crossref	✕	7	MyISAM	utf8_general_ci	
□	cms_css	✕	22	MyISAM	utf8_general_ci	11
□	cms_css_assoc	✕	47	MyISAM	utf8_general_ci	
□	cms_css_seq	✕	1	MyISAM	utf8_general_ci	
□	cms_events	✕	75	MyISAM	utf8_general_ci	1

4.4 Specialty of the System

4.4.1 Search features

Searching is fast and easy. With Secure Digital Asset Repository Systems, user can find the documents they want in seconds.





Jul 27, 2011

Welcome to Secure Digital Asset Repository

Category: General

Posted by: admin

As the industry evolves into technology driven businesses, an increasing number of companies are reaching a critical pain in needing to control and manage their vast amounts of digital media assets. Technically speaking,

Health

You are here: Home » Categories » **Health**

Search Results For "beauty"

- Beauty - Beauty (100%)
- Entertainment - Entertainment (16%)

Time Taken: 0.012004

Previous page: Food
Next page: Beauty

4.4.2 Token-based Authentication features

SDAR provide the ability to authenticate users via any of the leading password tokens. It authorizes others which are non-user to fetch or upload document from the system. Other than that, it has start and stop dates whereby for those temporary users, once the contract is over, the token is disabled. The advantage of the temporary user is that they do not have to register in order to enter the system. However, the disadvantage of these users is they cannot enter the system once the token date is expired.

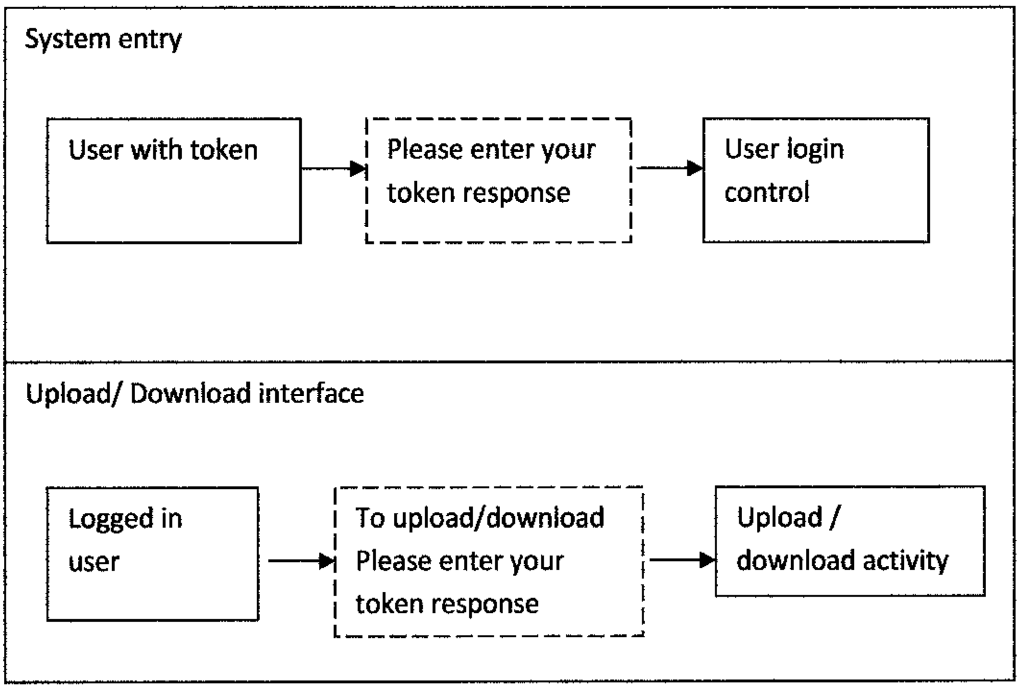
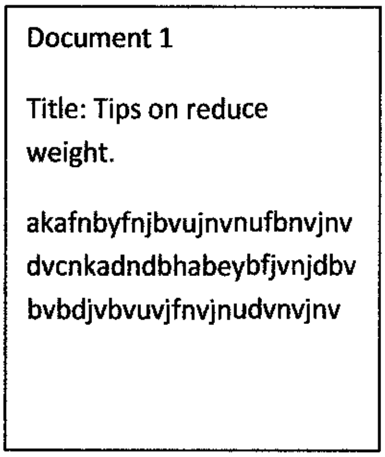

System entry
User with token
Please enter your token response
User login control
Upload/ Download interface
Logged in user
To upload/download Please enter your token response
Upload / download activity

Figure 11: Process of the special user with the token in SDAR

4.4.3 Document censorship features

The objective of this feature is to prevent a third-party from modifying a document uploaded in the system. Our system should make it extremely difficult for user or third-party to make changes to or force the deletion of published materials.

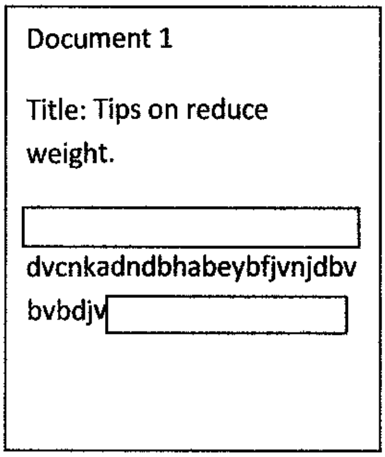


Document 1

Title: Tips on reduce weight.

akafnbyfnjbvujnvnufbnvjnv
dvcnkadndbhabeybfjvnjdbv
bvbdjvbvuvjfnvjnudvnvjnv

Without censorship feature



Document 1

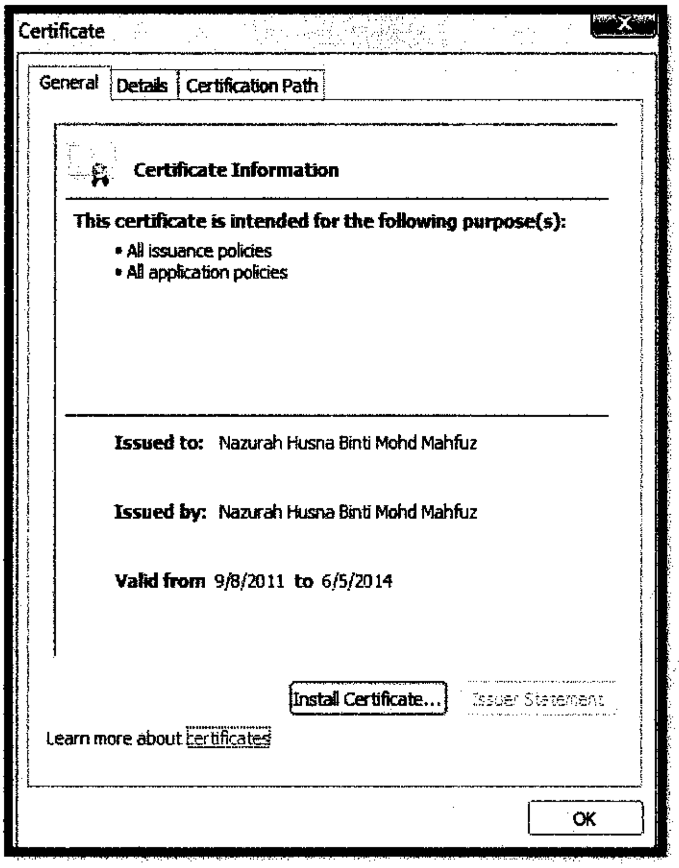
Title: Tips on reduce weight.

dvcnkadndbhabeybfjvnjdbv
bvbdjv

With censorship feature

4.4.4 Certification features

Electronic certificate will be given to the author or the up loader of the document. This certificate will ensure there is no copyright issue being arise after they up load their document in the system.



Certificate Information

This certificate is intended for the following purpose(s):

- All issuance policies
- All application policies

Issued to: Nazurah Husna Binti Mohd Mahfuz

Issued by: Nazurah Husna Binti Mohd Mahfuz

Valid from 9/8/2011 **to** 6/5/2014

Figure 12: A sample of certificate

CHAPTER V

CONCLUSION AND RECOMMENDATION

As a conclusion, we have presented in this paper the Secure Digital Asset Repository (SDAR) will be implemented at future. The system acts as a repository for digital assets owned by the system owner (company or institution) and associates the metadata with the content to provide efficient search and retrieval. SDAR supports different digital formats and incorporates in one integrated system the security, preservation and dissemination of material. The system addresses the main challenges faced by secure digital repositories; preservation of digital material and content dissemination. Future enhancements include; building a more sophisticated security system based on existing and emerging standards that are appropriate for the web services environment and designing and implementing a more sophisticated and efficient content versioning support. Therefore, with a good cooperation between my supervisor and myself, personally I hope this project may success and could beneficial to society.

REFERENCES

[1] Wikipedia, 2010. Retrieved March 1, 2011 from Wikipedia website <http://en.wikipedia.org/wiki/Digital Asset>

[2] Wikipedia, 2010. Retrieved March 1, 2011 from Wikipedia website <http://en.wikipedia.org/wiki/Access Control>

[3] Etter P.C. 2006, Digital Asset Management The Art of Archiving (3rd edition), L Australia, Spoon Press.

[4] Harvard University Office for Information Systems. Introduction to digital repository services. (2004).

<http://hul.harvard.edu/ois/systems/drs/introdrs.html>

[5] Chitre M., Freitag L., Sozer E., Shahabudeen S., Stojanovic M., and Potter J. 2006, “Managing Digital Repositories,” in *Proc. OCEANS 2006 - Asia Pacific*, Singapore: IEEE

[6] Journal. Retrieved April 18, 2011 from journal website

<http://airccse.org/journal/jcsit/1210ijcsit07.pdf.>

[7] Wikipedia. Retrieved April 20, 2011 from Wikipedia website

<http://en.wikipedia.org/wiki/Timestamp>

[8] Journal. Retrieved April 20, 2011 from journal website

<http://www.issa.org/Library/Journals/2006/January/Stapleton,TepplerDigital Signature are NotEnough.pdf>

[9] Journal. Retrieved August 6, 2011 from journal website

<http://www.issa.org/Library/Journals/1992/March/Parksetl.StorageAnd RetrievalSystems.pdf>

[10] Website. Retrieved August 7, 2011 from token authentication website

http://www.afp548.com/filemgmt_data/files/cryptocard_MW07.pdf

[11] Website. Retrieved August 9, 2011 from digital asset repository website

http://www.bibalex.org/isis/UploadedFiles/Publications/DAR_Ext_View.pdf

APPENDIX A: GANTT CHART

cure Digital Asset Repository
ıiversiti Teknologi PETRONAS

BS	Tasks	Start	End	Duration (Days)	Complete %	Remarks
	Finding information:	18/6/2011	24/6/2011	7	100	
	Computer component					
	Finding information:	25/6/2011	1/7/2011	7	80	
	Application layer					
	Logic layer					
	Develop the system:	4/7/2011	21/8/2011	45	50	
	Find CMS available					
	Choose software					
	Edit the coding					