CLARMONT
A CLEAR VIEW OF TECHNOLOGY

CLARMONT I. S. Ltd.
Királyhágó street 2.
Budapest, Hungary-1126
Tel.: (+36) 1 225-3973
Fax: (+36) 1 225-3974
sales@clarmont.hu

http://www.clarmont.hu

Document Management System

" Master your documents "

Introduction

Implementing a Document management system is much more complex than it may appear at first glance. It is not enough to just solve the problems of **storage** and retrieval, since basic requirement also include **document authentication**, system **optimization**, and ensuring constant data access speeds regardless of the amount of data stored. To mention a few more selected topics (there are many more), the system must also enable a wide range of **access** to the data with the proper security, provide data **encryption and a full range of logging functions**.

The SANity™ DOC system has all of these functions, which we have grouped into 3 main fields. These follow the logical structure of a **document management** system and refer to document entry/upload, storage, and finally subsequent **access/retrieval**.

Adding a Document

Documents are entered into the system using a **custom developed client**. Use of this client is fully governed by the systems access control system and can be enhanced with several types of authentication systems.

During the systems customization phase, the types of documents are defined. These documents undergo conversion when entering the system in order to provide for **read-only access**, and allow the content of these documents to be accessible to a global full-text search feature. Scanned documents can be processed using OCR, and the original **scanned images** can also be attached and stored in the system.

Features and Benefits

- Fully configurable security and access control system
- Allows storage of Audio/video documents
- Intra/Internet interface to document management, and query/retrieval
- Transaction based, fault-tolerant, clusterable system – scalable performance and load balancing
- Standard J2EE technologies
- Multilingual ready
- Easy to use
- 24/7 document availability
- Guaranteed manufacturer and platform independence
- Document authentication and management of digital signatures
- Original document (or image) indexed and storage
- PDF generation and storage for read-only presentation
- Text format generation and storage enables full text search
- OCR functionality can be integrated
- Document Version control
- Meta-data registered and stored for each document
- Data Storage:
  - RAID based HD storage
  - Optical media based storage (BD DVD Storage)

Document entry actions

Meta-data generation and storage

A set of meta-data is attached to each document entered into the system. This **meta data** can be information pertaining to the document (source, external reference ID, **digital signature**, time-stamp etc). The range of meta-data also needs to be defined during the implementation.

The meta-data results in a sort of envelope, or registration card for the document to be archived.

Text file generation

Independent of the format of the document being entered into the system, you will need to be able to find it. In other words, we must ensure that your document can be easily located. To aid in this, a **text extract is created** for each document processed (or a text summary must be entered for non-text documents such as audio/video files). This text "extract" is then **stored in the database** and allows for comprehensive full-text searches against the content stored in the SANity™ DOC system.

PDF file generation

In a system such as this, there must be no possibility for unauthorized users to modify a stored document in any way. Consequently, files **cannot be overwritten**. Additionally, we provide an solution wherein he system **generates read-only, graphical (PDF) versions of the files**. These PDF format, read-only files **guarantee the authenticity of the document and the security of the original document**. At the same time, this provides a platform independent format that can be viewed anywhere.

Document Authentication

By authentication, we mean that we guaranteed that the content and format of the document is identical to the original. On one hand, this can be solved by implementing a **centralized document filing infrastructure**, in which case the person(s) performing the data entry provide the guarantee. For individual document filing, the system log in process, or rather the **access control system** can guarantee this. The best form of document authentication is that of the use of a digital signature, the use of which practically ensures that undetected data modification cannot occur, and that the document's **authenticity** can easily and accurately be checked.

Version Control and Management

Experience shows that some **documents will require version control**. This means that multiple version of a given document (e.g. A contract) may exist over time, possibly with different authors, and multiple changes to the contents. SANity DOC can **handle** this multiplicity of documents, and deal with them as version of a single file. Depending on the access rights of a user they can view and compare **different versions** or alternatively just view the most recent (current) version.

SANity™ DOC

Document Storage

Document Database

All formats of each document (original, text, pdf) and the associated **meta-data are all** stored.

In the case of larger source files and graphical format files, the original files are stored on some physical data storage medium, but always with **online, or near-line** availability. (archival of documents to tape or CD does not in any way constitute a workable solution). The **text extracts**, meta-data (including location of original files) and system information are stored in the document database.

Physical Data Storage

A document management system requires a high capacity storage solution. This must provide suitable **data security** (ensuring that data will not become lost or corrupted), and allow for establishment of **fault-tolerant** (and consequently **disaster tolerant**) solutions.

There are currently only a couple solutions that satisfy these requirements:

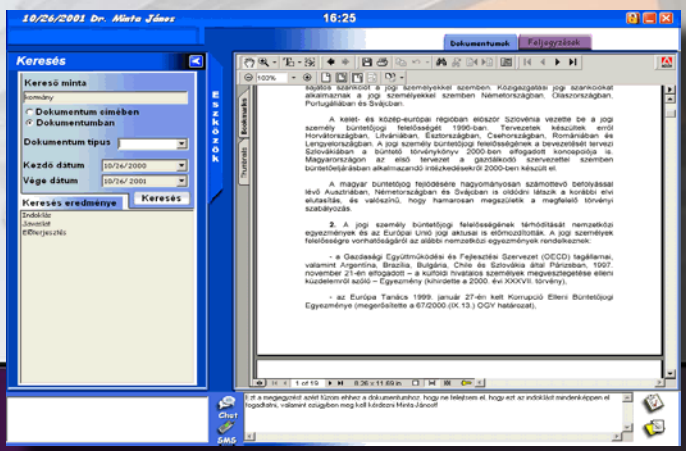
- *RAID based HD storage systems*
- *Optical media based storage systems*

Recommendation: DVD based Storage systems provide **cost-effective**, reliable and long-term storage for the SANity archive family of products.

Logging

A basic element of system security is a comprehensive **logging functionality**. SANity™ DOC provides a complete, event-based logging capability. This means that logging can be configure to include the most basic information (who-when-what), but can also be set to provide detailed system level event **(every event)** information.

Optionally, the log information pertaining to a specific document can be attached to that document, providing an access history similar to that in a library (when checked-out/ viewed, by who, how often etc.)



Document access

Retrieval

Document retrieval can be accomplished using a **custom client**, or via the **Intra/Internet** using a thin client. In both cases the use of these clients is dependent on the necessary permissions, and in the latter case also requires an Internet browser. The advantage of the WEB based solution is that the client is **platform independent** and does not require client-side installation.

Search

There are two ways in which a search of the document database can occur.

If **searching for a specific document**, the search will look through the stored **meta-data** (including predefined **document identification** elements) using the parameters you specify. These search parameters could include date, author, filename, document title, external filing ID, category etc. The more data you can provide, the more accurate the results. The other possibility is to **search for a phrase - in any document**. In this case the search operates on the stored text summaries, and uses a special algorithm developed for this task.

Security and Access Control

The security system of SANity DOC is highly **configurable at multiple levels**. You can define users, user groups, roles, role groups, permissions, permission groups which can then be linked together as desired. In practice this means that you can define (even on a document level) who has what type of access: read, save, write, edit etc. These **permissions can be assigned at any level** – for instance a user can be a member of several user groups, each with multiple role groups, each of with its own set of permissions. The security system can be **modified at any time**, allowing highly flexible access to the archived data.

Data security

In addition to the basic **(user/password)** security, the security of your data can be ensured on multiple levels depending on your requirements.

- *Biometric and/or chip card (authentication)*
- *Encrypted data transfer between central system and clients*
- *Encryption of stored data*

  For encryption **SSA, RSA protocols** can be used as **desired by the customer.**

SANity™ DOC