



# Large Scale Document Management System: Creating Effective Public Sector Knowledge Management System

Christian A. Bolu, PhD

Director, ICT/Innovation Centre, University of Nigeria Nsukka

Department of Mechanical Engineering, University of Nigeria, 410101, Enugu State, Nigeria

christian.bolu@unn.edu.ng

## ARTICLE INFO

### Article History:

Submitted 15 February 2010

## ABSTRACT

The digital age has redefined the production process and utilisation of documents globally. In the information age, the process of input, delivery, storage, receipt, and categorization of data is critical. The public sector has to rely more and more on automated, reliable solutions in order to keep their information safe and readily accessible for effective governance.

A document management system is a computer system used to track and store electronic documents and/or images of paper documents. The term has some overlap with the concepts of content management systems often viewed as a component of enterprise content management systems and related to digital asset management, document imaging, workflow systems and records management systems.

This paper examines an ongoing document management implementation case study in a public sector of digital assets of over twelve million pages, scalable to billions of pages, highlighting the taxonomy, content and knowledge management creation using an enterprise content management system and discusses the role in national development and growth.

## Keywords

Content management  
Knowledge management  
Document management  
Taxonomy  
e-Business

## TRACTS

Optimization Theory and Application  
Project and Engineering Management

## 1.0 Introduction

In the information age, the process of input, delivery, storage, receipt, and categorization of data is critical. The public sector has to rely more and more on automated, reliable solutions in order to keep their information safe and readily accessible. These content when managed by workflow in a collaborative environment supports knowledge management for effective e-Governance.

| |

**A document management system[1]** (DMS) is a computer system (or set of computer programs) used to track and store electronic documents and/or images of paper documents. The term has some overlap with the concepts of content management systems. It is often viewed as a component of enterprise content management (ECM) systems and related to digital asset management, document imaging, workflow systems and records management systems.

Digitisation of existing paper document is in two parts - scanning and ‘rasterising’ the records to become searchable, and then warehousing them using an enterprise software solution.

The benefits gained by implementing Document Management System include:

- Secure storage of documents
- Easy retrieval of documents
- Excellent search capabilities to cut down on time searching for documents
- A controlled environment for updates to document
- Complex security rules to control access
- Increased visibility of key documents
- One source and one set of rules for managing documents
- Reduced time and effort spent on document management
- Ability to maintain document history to meet legal requirements

## 2.0 Objectives

This paper examines an ongoing implementation case study in a public sector of digitisation of over twelve million pages, highlighting the taxonomy, content management system and the knowledge management implementation using an enterprise content management system. Discussion on the benefits in the public sector for national development and e-Governance is made.

## 3.0 Content and Knowledge Management

**A content management system[2]** (CMS) is a collection of procedures used to manage work flow in a collaborative environment. These procedures can be manual or computer-based. The procedures are designed to:

- . Allow for a large number of people to contribute to and share stored data
- . Control access to data, based on user roles. User roles define what information each user can view or edit
- . Aid in easy storage and retrieval of data
- . Reduce repetitive duplicate input
- . Improve the ease of report writing
- . Improve communication between users

In a CMS , data can be defined as almost anything - documents, movies, pictures, phone numbers,

scientific data, etc. CMS's are frequently used for storing, controlling, revising, semantically enriching, and publishing documentation.

**Knowledge management**[3] (**KM**) comprises a range of strategies and practices used in an organisation to identify, create, represent, distribute, and enable adoption of insights and experiences. Such insights and experiences comprise knowledge, either embodied in individuals or embedded in organisational processes or practice. SAP[4] Netweaver Knowledge Management[5] is an umbrella term for the management of unstructured information, that is, documents of all types. It converts unstructured information into organized knowledge – an essential function in the age of global e-business

**Taxonomy** is the practice and science of classification. Taxonomies[6] are enterprise-wide navigation structure which can searched for information according to categories. In these categories, information from different sources can be classified according to theme, organisational criteria or other criteria.

## **4.0 Document Management and Effective National Development**

Document management are essential for the effective and productive functioning of public organizations. Records document the decisions and activities of governments and public institutions, and serve as a benchmark by which future activities and decisions are measured. They document fundamental rights and obligations, and differentiate the rule of law from the actions of arbitrary states.

Without good document management there can be no rule of law and no accountability, public officials are forced to take decisions on an ad hoc basis without the benefit of an institutional memory, fraud cannot be proven, meaningful audits cannot be carried out, and government actions are not open to review. In national development process, the effectiveness of development projects could suffer as there will be no means of verifying that the development project falls within acceptable legal, financial and cultural boundaries of a client government. There will be no means to verify that funds for development are used as intended.

## **5.0 Methodology**

In the design and implementation of public-domain content for effective utilization and collaboration, the following issues are to be considered.

### a. National case

- *What national documents do we want to manage?*
  - o Documents for the rule of law - legislative records, court records, police and prison records.
  - o Documents to demonstrate accountability to its citizens, - policy files, budget papers, accounting records, procurement records, personnel records, tax records, customs records, and electoral registers, property and fixed assets registers
  - o Documents to protection entitlements - pension records, social security records, land registration records, and birth/death records.

- o Documents in providing services for its citizens - hospital records, school records, and environmental protection monitoring records.
- o Documents for government's relationship with other countries - foreign relations and international obligations, treaties, correspondence with national and international bodies, loan agreements, etc.

b. Case under study

- o Intellectual Properties – University theses, publications, patents, copyrights, inventions, personnel records, physical planning drawings, accounting documents, etc
- *How do documents fit into the overall national business process?* What is the Business Process Flow? Are documents created or required at certain steps in the national business process? Which business objects are documents associated with? What are the documents and what is their significance?
- *How do we want to search for documents?* What are the attributes of the document? Standard attributes – Description, Owner, Responsible Lab/Office Additional Attributes – Application, Release; Full Text Search
- *Define Lifecycle of Documents.* What are the steps in the lifecycle of the document? In Work, Pending Approval, Approved, Released – No more change-Released version remain as history
- *What is the change control process?* Are updates controlled through a change control process? Changes of document through Engineering Change Management - Capture reason for change, element of workflow, and digital signature for release; Provides a complete history of when and why a document was updated.
- *Is there a formal approval process?* Before a document is officially released, does it go through a formal approval process? Facilitated through a workflow process; Might require digital signature; Formal approval results to a released version of document with record of the approver; Further change to document must be by a new version
- *What are the security requirements?* What roles in the national business are allowed to change each document? Consider status for changes – *In Work* seen by select group and *Release* seen by all.
- *What type of application files will be stored?* What output file of a specific application is

stored? Microsoft Word, Adobe Acrobat, Autodesk AutoCAD; Application configured to behave in a certain manner when associated file is launched for display or change.

- *How are versions and revisions used in the business?* What do the term Version and Revision mean to the Business; Version is a separate instance of a document that has its own status such as *In Work* or *Released*. It is a snapshot in time; Revision level is assigned to a document version and is associated with a release state. Represents a major change; For each document you can store multiple versions. With each version, you can assign a revision identifier.
- *Do you need to support searching and maintenance in multiple languages?* Maintain some attributes in multiple languages? Attribute *Description* could be maintained in English and French; Capability to maintain entry, display and search attributes in multiple languages.
- *What is the volume and size of documents to be stored?* How large is document to be stored? Infrastructure requirements to be considered; Content Server to be sized appropriately, say 100 TB; Size of each file help in Network Sizing, say 4Mbps; Document consumers may be in a number of different geographic locations – Require Cache Server
- *Location of Creators vs Consumers;* What are the different locations of creators and Consumers? Creator is someone who generates and stores document in the system; Consumer is someone who searches and displays documents; For large number of Creators install Content Server at that location; For large number of Consumers install a Cache Server; This helps reduce the impact on the performance of WAN.
- *Are there document retention requirements?* How long should document be stored or made available based on national business and legal requirement? How to handle document when retention period expires, say archived or deleted; What does the national & international law require?
- *Do documents need to be converted to a neutral format for long term retention?* What neutral format should be used – PDF or TIF? Conversion can be carried out automatically by DMS Conversion Server when the status of released is reached.

## 6.0 Results

The following steps were taken:

- a. Over 12 million pages of paper documents are being scanned with fast scanners, raterised/OCR to become searchable and bookmarked.

- b. Development of the taxonomy and the unstructured documents classified
- c. Using the Enterprise Content management system configured the following
  - Search design
  - Change control process
  - Workflow
  - Roles for security implementation
  - Application types
  - Versioning and revision structure
  - Language requirement
- d. Creation of Content, Cache and Conversion Servers
- e. Deployment of the collaboration rooms, etc in the Knowledge Management Enterprise Portal
- f. User management and training

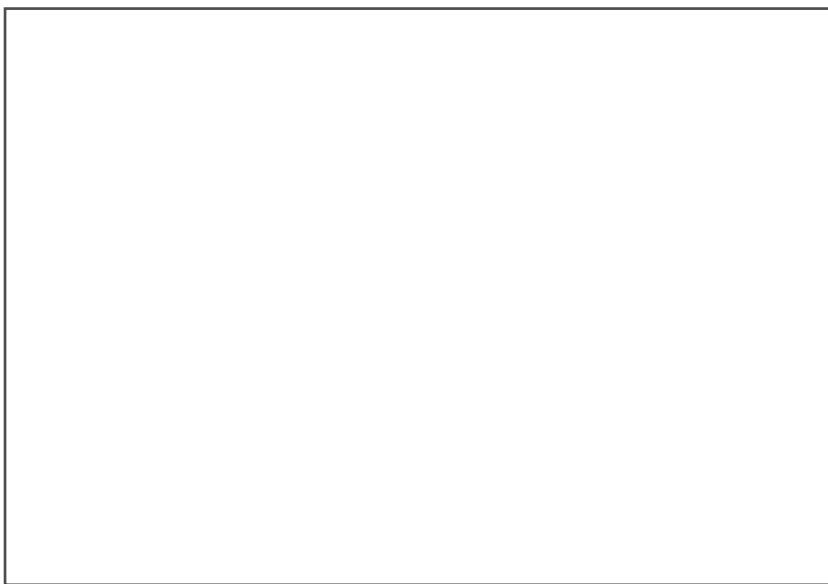


Fig 1: Document Creation Screen in SAP DMS

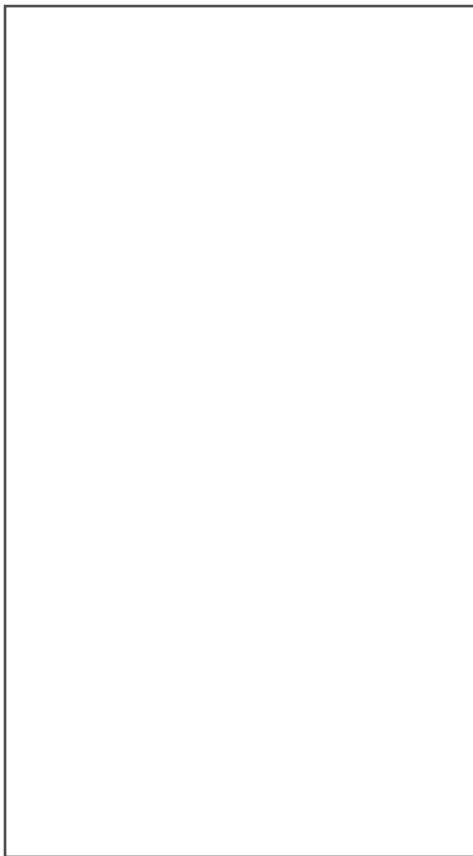


Fig 2: Document Type creation in SAP DMS

## 7.0 Discussion

The World Bank structures its assistance according to the Comprehensive Development Framework (CDF), a paradigm for cooperative development aid planned and organized by the client countries in consultation with development partners.

The Four Pillars of the Comprehensive Development Framework[7] are:

- good governance
- equitable judicial system
- accountable financial system, and
- enforceable civil rights.

All of the elements for effective development depend upon an effective document management infrastructure. Without a document management infrastructure, governments and organizations are incapable of effectively managing current operations, and have no ability to use the experience of the past for guidance. Records are inextricably entwined with increased transparency, accountability and good governance.

Lack of good document management system is directly linked to the persistence of corruption and fraud. Experts in financial management and control recognize that well-managed record systems are vital to the success of most anti-corruption strategies. Records provide verifiable evidence of fraud and can lead investigators to the root of corruption. Well-managed records can act as a cost

effective restraint. On the whole, prevention is much cheaper than prosecution.

In many developing countries, such as Nigeria and Ghana, document management problem is a massive one. Existing record keeping systems - if they exist at all - are inadequate and unable to cope with the growing mass of unmanaged papers. Administrators find it ever more difficult to retrieve the information they need to formulate, implement, and monitor policy and to manage key personnel and financial resources. National archival institutions - when they exist at all - are becoming marginalized - given the role of maintaining the records of the colonial period only. Countries emerging from the colonial period to statehood have done so without adequate systems to keep the governments functioning - either no adequate record keeping systems were ever established or they were not maintained. This situation impedes the capacity of these countries to carry out economic and administrative reform programs aimed at achieving efficiency, accountability, and enhanced services to citizens. Moreover, the decline, and in some cases total collapse, of record keeping systems makes it virtually impossible to determine responsibility for actions and to hold individuals accountable.

The symptoms[8] of this problem are many:

- There is a low awareness of the role of records management in supporting organizational efficiency and accountability.
- There is an absence of legislation to enable modern records management practice.
- There is an absence of core competencies.
- Overcrowded and unsuitable storage of paper and electronic records;
- Absence of purpose built record centres such as Content and Cache Servers
- Absence of a dedicated budget for records management
- Poor security and confidentiality controls
- Absence of vital records, disaster recovery and preparedness plans
- Limited capacity to manage electronic records.

## 8.0 Conclusion

Proper document management requires trained staff, adequate and continuous funding, appropriate environmental conditions and physical security. Appropriate document management structures and governmental legislation and/or regulation are needed. A document management system should have realistic targets and project design. This can be achieved by a scalable DMS implementation.

Computerized systems must be adopted appropriately, with regard for local capacity, with concern for legal requirements for evidence. They must fit business requirements. Long range planning for systems support and upgrades is also needed to sustain efforts. There must be well organized, accurate and easily accessible source data, a reliable power supply, realistic back-up and storage procedures, and adequate communications and sustainable technical support.

## 9.0 References

1. Anderson, E. Greenspun, P. & Grumet, A, 2005. *Software Engineering for Internet*

*Applications*, Massachusetts Institute of Technology.

2. Bolu, C. A., 2010. Technical Reports on Document Management. [Booklet], University of Nigeria
3. Jay, R., 2008. The complete Reference – SAP Netweaver Portal Technology, 1<sup>st</sup> ed., McGraw Hill
4. SAP, 2006. SAP Netweaver Portal Training Manual [Manual] SAP AG
5. Stajda, E, 2009. *Effective Document Management with SAP DMS*, Galileo Press
6. Wikipedia, 2010. *Document Management System*, [Online] (Updated 5 March 2010) Available at [http://en.wikipedia.org/wiki/Document\\_management\\_system](http://en.wikipedia.org/wiki/Document_management_system) [Accessed 6 March 2010]
7. World Bank Group, 2010. *Why Records Management?* [Online] Available at <http://go.worldbank.org/889BWHZPL0> [Accessed 15 February 2010]



- 
- [1] Wikipedia, 2010
  - [2] Wikipedia, 2010, SAP Netweaver Portal training manual
  - [3] Wikipedia, 2010
  - [4] Systems Applications Product AG of Germany
  - [5] SAP Implementation and operation of Netweaver Portal
  - [6] SAP Implementation and operation of Netweaver Portal
  - [7] World Bank Group Report, 2009
  - [8] World Bank Group Report, 2009