ISSN:2320-0790

An International Journal of Advanced Computer Technology

# Rules to be followed when executing Electronic Archiving Projects to Build E-Government

Ibrahim M. Osman[1] and Ashraf Gasim Elsid Abdalla[2]

[1]Computer Science and information Technology, Omdurman Islamic University
[2]School of Electronic Engineering. The Future University

Corresponding author: Ibrahim Mohamed Osman Ibrahim

**ABSTRACT:** This is study aim to highlight the risks may be occur to the digital information of the current archiving system at National Information Centre (N I C) in Sudan, if the appropriate rules and strategies not followed in implementing Electronic-Government (E.Gov) programs and other risks which may threat the information in the future, the matter which may lead to fail the project. Thus, following analyzing data gathered from the field study in the N I C. The result revealed that information might be subject to loss due to natural disaster or human fault. The study compared with between policies in N I C related to E.Gov and international adopted policies. The gathered results has been analyzed, the outcome is to design Data Center of standard specifications which conform with the requirement of E.Gov by upgrading the current design from Tire2 to Tire 4 and establish two Backup Data Centers, provided that one of those center to be outside the N I C to protect the information in case of natural or intended disasters, and adopting daily information backup strategy to protect it from loss instead of weekly backup. However, this also is not sufficient because in future more complicated risks require to following digital storing strategy to protect the information against loss and damage by replacing the current strategy (technology preservation) to strategy of (technology simulation), the concern is not only limited in backup but also to other aspects, the digital files itself need to be saved in fixed medium as far as possible SAN, in addition to maintain the programs required to read such files and decode it. The research conclude that the E. Gov project in Sudan it depend on implementing National integrated archive connecting all ministries and governmental authorities through national network to provide data and information can be utilized to implement E.Gov gateways.

**KEYWORDS**: E-Government, Data Center, Storage, Backup, Antivirus

## I. INTRODUCTION

The improvement of E.Gov programs in the developing countries is a necessity, due to urgent requirements of these countries to develop their administration systems by relying on information technology applications to reach integrated information community. International experience in this filed express the usefulness of information technology applications in tasks and different duties, which it has major role in overtaking several problems faced by the conventional systems particularly the time and effort by the applicants due to complication of executive procedures followed. Locally the administration system adopted by the governmental institutions suffered of several routine obstacles and issues in addition to financial and administration corruption, and production of huge amount of paper documents which need a lot of efforts to be organized, recorded and archived along with wasting huge amount of money, note that such paper documents will worthless after short term of period. The administration system in Sudan required the citizen to have several identification documents to execute his transactions in addition to routine complicated procedures. The process of implementing E.Gov gateways in Sudan, which aimed to connect the all concerned governmental institutions site together, and make it available to the citizens will be considered as great step, this process, as the researcher though, will not be successful unless a qualified electronic archiving system has been implemented in conform with rules and strategies along with electronic administration program and to facilitate the communication medium to all community individuals.

## II. MATERIAL AND METHODS:

To implement qualified electronic archiving systems which contribute in the success of E.Gov project, there are some strategies to be followed. In this part we will take into

account such strategies to be compared by the current situation in Sudan, the strategies will be as follow:

**First: Prepare appropriate strategy to choose the design of Data Center:**

There are several recognized designs in the implementing of Data Centers, can be summarized as follow: Data Center types are frequently classified by tier, with the breakdown looking like this [1]:

- Tier 1 data center (99.671% minimum uptime)
- Tier 2 data center (99.741% minimum uptime)
- Tier 3 data center (99.982% minimum uptime)
- Tier 4 data center (99.995% minimum uptime)

The first thing you'll probably notice about this table is that the minimum uptime is over 99% for every tier, with differences of less than one percent. Of course, these fractional slices of time do add up. A Tier I data center can be down 28.8 hours in the course of a year, while a Tier IV data center has an annual maximum downtime of 0.4 hours. A Tier II data center can be down 22 hours a year, while a Tier III data center can be down 1.6 hours. The other major difference between the tiers is in redundancy requirements. Tier I data centers have no redundancy, and when they're down—for maintenance, for instance—they're completely down. A Tier IV data center is fully redundant, with multiple power paths. Tier IV data centers are able to sustain a 96-hour power outage; with a Tier I data center, everything will likely go down when the power goes out

How much does this matter? What is the significance of downtime for specific businesses?

The answer to these questions depends on the business. A recent survey from Ponemon Institute finds the total cost of data outages has increased more than 40% in the last three years. Every business needs to know how costly downtime will be for its specific bottom line, and the tier system is helpful in determining the type of data center offering that will provide just the right benefits. Tiers are an invaluable assistance for any business in thinking about exactly how much uptime it must have and exactly how much downtime will cost.

**Second: Prepare appropriate strategy for digital storing:**

In the projects of E.Gov a strategy of digital storing should be prepared before implementing the program, there are two types of storing strategies[2]:

1. **Organizational strategy**: concerned with administrational aspect for implementing the technical strategy a part of that adopting administrational structure with available budget and trained professional.
2. **Technical Strategy**: concerned with technical aspects to be done in order to ensure operation-ability and availability of digital items even in case of future technological changes whether for the programs or media, there are three types of technical strategy for digital storing:

a) **Technology Preservation:**
Pursuant to this strategy the hard and software will be kept which will allow using the future digital resources taking into account transfer of information from a media to another of same type every interval in order to protect the information against loss. In case of medium damage a refreshing process will be taken.

b) **Technology Simulation:**
Information will be kept in its original format and create an environment simulate such environment used in creation and operation of digital resources after aging of the original work platform, accordingly this strategy will provide future and operation system simulate the original platform used in the project. This strategy also required activation of data every time. Although this strategy similar to the previous one as it adopt same technology used in creation and operation of digital project but it also different as it not required saving the application program, operation system and hardware[3]

**c) Information Migration:**
The digital items i.e. specifications, programs and other programs or computer generations to another will be transferred in order save such items in the light of rapid technological changes, so that this kind of strategies required transferring information every interval to another medium work on the new generation of computer, accordingly we are not reserve the medium itself in this case but it will be updated every interval of time, also the files may copied to work with new versions of the same program in backward compatibility, taking into account that some program allow the new versions to read files of previous one or two versions of the same program.

**Three: Choose the storage medium:**

Digital storing medium are different than old storing medium which has long lifetime because it made of material not easy damaged like magnetic tapes, the lifetime of the most of digital storing medium is short, even the medium which considered as long lifetime than tapes like Compact Discs (CDs) its subject to damage if not be under care. The storing space vary pursuant to several considerations i.e. nature of stored material, the matter which lead to use the file compacting, but the size of file used in the project of E-Government will stay large and need high storing space, hence the ideal storing of digital items should be of wide space, low cost and high speed in information storing and reading beside the long lifetime,

accordingly Storage Area Network "SAN" technology has been adopted in the project of E-Government.

**(SAN)Technology "Storage Area Network" [4] :**

This is the most resilient Storage solution. It consists of the three parts viz the host, fabric (network of Fiber Channel Switches) and the Storage array. The servers and clients in the host level can access the Storage via the Fabric which is generally a multi-path system providing a high level of redundancy. The ISCI protocols and the fiber channel connections make the SAN a very high performing and scalable storage environment.

**Four: Adopting the standardization for files format:**

In the digital projects, the form of non standard files should be avoided, because it's subject to change and disappearing and stopping possibility of the company or developing new forms. This idea depend on that the program of widely used in the market are less subject to change other than such another program, there are two methods to have digital items:

a) Convert the content to digital by using digital devices like scanners, digital cameras or others in order to create matrix of pixels.

b) Create the content in digital forms "born digital" by using the keyboard, pens or drawing by mouse to create digital content, it's possible to create digital content for text, images, sound, animation and 3D content, accordingly the result of digitalization will create files of different shapes suit the several contents of the digital project, [6]so the management of such files will be easy on the long term if such files conform the standard forms. The American Standard Code for Information Interchange (ASSCII) or Rich Text Format (RTF) make the reading possible for the long term due to availability of the required programs, the text files maybe stored in standard form to process words i.e. SGML (ISO Standard 8879) or Portable Document Format (PDF) its popular format and widely spread and recently it become ISO Standard 32000

Usually the images will be available on the format specified for images, the most common type of standard image files are:

**Tagged Image File Format (TIFF)**, high definition and quality type of files, but it take space, so it's suitable for archive storing for the main copies of digital images, but it's difficult to transfer files in this format through network unless it was high speed network, this format will not normally use the compression technology.

**Joint Photographic Experts Formats (JPEG),** this format use compression technology which leads to loss some information due size reduction (lossy compression), it suitable for transfer the file through network but not suitable for archive storing.

**Graphic Interchange Format (GIF)**, file format suitable for transfer image through internet and used file compression technology also.

**Portable Network Graphics (PNG),** its big size files and uses the compression technology in good manner and high definition better than JPEG and GIF because the compression used without losing information (lossless compression).

The standard formats for audio-video format are the following:

**WAV**, used to operate the multimedia in Windows.

**Motion Pictures Expert Group (MPEG),** format used in movies compression.

**MP3** of **MPEG** for multimedia standard the main feature of this format it has small size and good sound quality.

**Real Audio,** common files format because it uses free program downloadable from internet which is audio player software. There are other formats i.e. Real Video, Quick Time, Digital Video Standard and Audio Video Interleave (AVI).

**Five: Create Metadata required to represent the data:**

Metadata by definition[6] is coded structured data describes specification of information entities to identify, explore and manage the information. The good storage of digital resource need active metadata system to describe, access control and ensure saving the same, the metadata elements required to support the strategy of storing followed in the digital project (Technology preservation, simulation or information migration) which will make detailed information to describe used file formats, software environment and the hardware, it may contain information regarding title management and access control[7], the importance of metadata will mainly consist of the following:

- Facilitate exploring of resources by identifying its entity and existence.
- Provide digital identities to differentiate among entities.
- Ensure availability of resources through metadata of historical saving and storing of digital resources.

**Six: Prepare appropriate strategy for information security:**

The document of E.Gov should be secured against loss by applying security and protection strategy, which can be executed by the following:

1- **Safety Principles measures for (Data Center Server Room), by using the following:**

- Fire Alarm: important device protect against fire.
- Fire Distinguishers: redundancy number of fire distinguishers should be available to avoid risks.
- Concealing all electrical wires going through archive hall inside metal pipes.
- Secure the gate of data center server rooms.

- Comply with standard specifications for electrical connection from power supply to data center server rooms.

2- **Safety Principles measures for Archived data (DB Security) against outsiders:**

- Prepare passwords for computers used in archiving system and for the storing medium of capability to create passwords for authorized and should be in combination difficult to guess.
- Use application program which allow making passwords for electronic files.
- Use antivirus.

## III. RESULTS:

In this study, data were collected from the N I C in Sudan as an entity entrusted with the Council of Ministers implementation of E.Gov project, it has been the questionnaire design based on rules and strategies universally adopted in the implementation of E.Gov projects to be compared with the results based on the rules and strategies applicable to the center through the analysis of information that will be obtained them.

**Table 1**: show the Educating level, Specialist, Standard Data center availability for saving data, Which data center type used and Availability of back-up data center for E.Gov in N I C

| Parameters | | | | | Total |
|---|---|---|---|---|---|
| **Education level** | B Sc 17 (57%) | M Sc 11 (37%) | PhD 2 (6%) | | 30 (100%) |
| **Specialist** | Computer Sciences 9 (30%) | programmer 3 (10%) | IT 5 (17%) | Computer Engineering 13 (43%) | 30 (100%) |
| **Standard Data center availability for saving data of E/Gov** | Yes 30 (100%) | | No 0 (0%) | | 30 (100%) |
| **Which data center type used** | Tire 1 0 (0%) | Tire 2 27 (90%) | Tire 3 3 (10%) | Tire 4 0 (0%) | 30 (100%) |
| **Availability of back-up data center** | Yes 0 (0%) | | No 30 (100%) | | 30 (100%) |

From the above table (75%) of the community are Bsc holder and most of them are graduate of computer engineering and telecommunication (47%).
The design of Data Center currently used in the N I C is type Tire2, there is no back-up data center to store the electronic documents in order to protect it against loss in case of a disaster occurred in the main data center.

**Table 2**: show Availability of special way for saving document, Technology preservation (TP) or Technology simulation (TS) or information migration (IM), type of strategies used and storage type used for E.Gov in N I C

| Parameters | | | | Total |
|---|---|---|---|---|
| Availability of special way for saving document of E/Gov | Yes 30 (100%) | No 0 (0%) | | 30 (100%) |
| Which type of strategies used | TP 30 (100%) | TS 0 (0%) | IM 0 (0%) | 30 (100%) |
| Which type of storage used | SAN 30 (100%) | NAS 0 (0%) | | 30 (100%) |

The community state (100%) that there is prepared strategy to save and store electronic document related to E.Gov which is strategy of Technology preservation with using of SAN storage medium.

**Table 3**: show Dose the center following standard procedure for types for saving the documents, which types of text used, which types of Video used, which types of Audiovisual used, Dose the center following standard procedure for metadata

| Parameters | | | | | Total |
|---|---|---|---|---|---|
| Dose the center following standard procedure for types for saving the documents | Yes 30 (100%) | | No 0 (0%) | | 30 (100%) |
| Which types of text used (More than one choice) | ASSCII 30(100%) | RTF 0 (0%) | ASGML 0 (0%) | PDF 30(100%) | 30 (100%) |
| Which types of Image used (More than one choice) | TIFF 30(100%) | JPEG 0(0%) | GIF 30(100%) | PNG 30(100%) | 30 (100%) |
| Which types of Audiovisual used (More than one choice) | WAV 30(100%) | MPEG 0(0%) | MP3 30(100%) | AVI 30(100%) | 30 (100%) |
| Dose the center following standard procedure for metadata | Yes 0 (0%) | | No 30(100%) | | 30(100%) |

The community state (100%) that the N IC is complying by the standardization regarding the format of digital documents of E.Gov, the format used for text are ASSCII and PDF (100%), the file used for video (TIFF, GIF, PNG) 100% and for audiovisual file format are (WAV, MP3, AVI) of 100% . The community also state (100%) that the NIC is not complying with standardization for metadata rules to write the digital documents used in E-Gov.

**Table 4**: show Dose the center following information security for saving the data, which type of producer to back-up for data base, Dose the center used antivirus for saving DB, Dose the center following procedure for saving the location of data center

| **Parameters** | | | | **Total** |
|---|---|---|---|---|
| Dose the center following information security for saving the data | Yes 30(100%) | No 0 (0%) | | 30(100%) |
| Which type of producer to back-up for data base | Daily 2(7%) | Weekly 28(93%) | Monthly 0 (0%) | 30(100%) |
| Dose the center used antivirus for saving DB | Yes 30(100%) | No 0 (0%) | | 30(100%) |
| Dose the center following procedure for saving the location of data center | Yes 30(100%) | No 0 (0%) | | 30(100%) |

100% of the community state that the NIC is complying of applying information security regarding storing the digital documents for E.Gov, 93% state that the backup is done in weekly basis. And 100% of the community state that the center is using antivirus to protect the database against malware, 100% of community state that the center is applying safety and security for the database server rooms.

**Table 5**: show Did the connection between data center and all governmental agencies in public net work has playing a key role in successful of E/ Gov, Coverage ratio, types of documents "Identity Documents(ID) , Bank Accounts(BA), Academic Documents(AD), Criminal Records(CR), Health Records(HR) ,How these documents are transfer to be use in E/ Gov, If the answer of Take a copy regularly did the data center Have prevents these documents, What types of surfaces are produced by public net work for governmental agencies

| Parameters | | | | | | Total |
|---|---|---|---|---|---|---|
| Did the connection between data center and all governmental agencies in public net work has playing a key role in successful of E.Gov | Yes 30(100%) | | No 0 (0%) | | | 30(100%) |
| Dose the public net work covering all governmental agencies | 100% 0 (0%) | 75% 10(33%) | 50% 20(67%) | 25% 0 (0%) | 0% 0 (0%) | 30(100%) |
| Which types of documents are available by public net work from governmental agencies to E.Gov (More than one choice) | ID 30(100%) | BA 0(0%) | AD 30(100%) | CR 30(100%) | HR 0(0%) | 30(100%) |
| How these documents are transfer to be use in E.Gov | Save in data center 20(67%) | | Take copy regularly 10(33%) | | | 30(100%) |
| If the answer of Take a copy regularly did the data center Have prevents these documents | Yes 30(100%) | | No 0(0%) | | | 30(100%) |
| What types of surfaces are produced by public net work for governmental agencies (More than one choice) | Mailing 30(100%) | VOIP 15(50%) | VC 15(50%) | Internet 30(100%) | Training 30(100%) | 30(100%) |

100% of the community confirms the vital and important role to connect all governmental institutions ministries in the national network in order to success the project of E.Gov in Sudan, the coverage of governmental institutions and ministries inside the national network is 50%, but 100% of the community states that networking provide the N I C the identity, academic and criminal record. There are some documents stored outside the center 33% a copy of which will be taken periodically, and 100% of the community states that the networking provides services to all governmental institutions and ministries at 100% i.e. post, internet and training. Other services provided to governmental institution and ministries of 50% i.e. VOIP & VC.

## IV. DISCUSSION:

Its easily to state that the successfulness of any E.Gov project depend on electronic archiving system, so that the rules of implementation projects of electronic archiving should followed in order to build E.Gov, such approach can be achieved by preparing clear policy to manage the electronic archive in order to make the digital data ready anytime for use and backup in day to day operation, stored under perfect environment to become a historical memory handed to the next generation in order to build projects of E.Gov, there is some weaknesses which threaten the project of E.Gov in Sudan:

1. The design of current Data Center is Tire 2, its not suitable for E.Gov project as its lack of several technical specifications, no electricity system and high-performance adapting, with possibility of power cutoff from Data Center for period of time reach 22 hours in the year the matter threaten the transactions done through E.Gov which it based on continuity.
2. No backup Data Center to store the documents of E.Gov in case of fault happen to the Data Center the matter which lead to the fail of E.Gov project.
3. the strategy of technology preservation for storing and saving the digital documents its not suitable for the project of E.Gov, the most drawback is to keep the old hardware in a manner which may lead constitute semi-museum which contain old technology which threaten the project in the future.
4. Adopting of American Standard Code for Information Interchange (ASSCII) for text files, which is not widely spread will threat the future of files stored under this format.
5. Adopting (TIFF & GIF) format for image files, its main disadvantages is use type of documents compression leads to loss some information in order to reduce the size (Lossy Compression).
6. Non compliance with metadata rules regarding writing detailed information for description of stored files format as documents for the E.Gov.
7. Adopting weekly backup policy the matter which threats the project in case of fault to the database before copy.
8. Some governmental institutions and ministries are out of the network the matter which fails the project of E.Gov.
9. Some ministries and governmental institution storing their data in internal data centers the matter which threat the documents in case of neglecting the technical related aspects.

## V. RECOMMENDATIONS:

The electronic archive suppose providing several supports along with software and suitable hardware to open and read the files, this support improving continuously as the new version will supersede the old ones in the meantime the programs and devices will changed the matter which make the use of electronic archive not so easy if there's no continuous efforts to abreast the change and technological improvement, because the integrated archive system will provide the citizen with sufficient information to process their administrational transactions in addition to open wide service of browsing, the citizen can depend on the national archive after himself and download the required documents to complete personal transactions i.e. issuing passport, pay bills and fines, … etc to achieve the end, a series of electronic archiving should be followed which will have huge benefit to the E-Government summarized as follow:

A. Electronic archiving for civil records and unifies the identification documents of the citizen in one ID card through completing the unified national identification number.
B. Electronic archiving for bank accounts and spread the credit cards issuance.
C. Electronic for academic documents for several educational levels.
D. Electronic archiving for real and private properties.
E. Electronic archiving to the health records and activation of family health cards.
F. Electronic archiving to the criminal records.
G. Electronic archiving to the statistical records for several ministries.

In order to achieve these goals by manner which makes the E.Gov in Sudan succeed the current weaknesses in N I C should be remedied through the following:

1. Change the design of Data Center to Tire 4 as it's the ideal design for the project of E.Gov for the following reasons:
   - ✓ 99.995 uptime
   - ✓ 2.4 Minutes Downtime Per Year
   - ✓ Fully Redundant in Power and Cooling
   - ✓ 96 Hour Power Outage Production
2. Establish backup Data Center in order to guarantee the continuity of E.Gov by providing

services to candidates in case of fault happened to the main Data Center provided that the design should comply with the following:

- ✓ Create two Backup Data Center, Classification type Tire 4
- ✓ One Backup Data Center must place it outside the scope of the main Data Center to works in the event of a disaster in the city.

3. Adopt Technology Simulation (TS) strategy instead of Technology Preservation (TP) in order to store and save the digital documents, its works in environment simulate the platform of creation and operation such digital resources after aging the original work environment, this strategy provide hardware and software simulate the original used in the digital project the matter which make it the suitable strategy in the E.Gov project.
4. Adopting the format of Portable Document Format in order to save the E.Gov documents, its famous and poplar format, recently became ISO Standard ISO32000 instead of using (ASSCII).
5. Adopting Portable Network Graphics (PNG) to store the documents of E.Gov as its big size files and use the compression technology in better way with high definition than JPEG & GIF, since it use the compression without losing information (lossless compression).
6. Adopting metadata to identify the stored documents in the database of E.Gov in N I C, as its coded structured data describe specified information entities assist in identifying, exploring and manage the information. The perfect storing of digital required active system for metadata to describe such resources.
7. Adopting daily backup policy for the database instead of current weekly backup.
8. Connect all ministries and governmental institutions within the national network to fulfill the implementation of E.Gov which is depend on collect and manage data from all relevant authorities within E-Government.
9. Not allowing any ministry or institution to save the document related to E.Gov in their internal data center, the data should be saved in Data Center located in the N I C since it's the implementing authority for the project of E.Gov in Sudan.

## VI. REFERENCES

[1] http://www.vxchnge.com/blog/types-tiers-data-centers-primer (retrieved 2006).

[2] sameh zeinhom Abd Elgwad, ''Digital libraries and archives: planning, construction and management '', Volume 2 . NAS Publishers, Cairo, 2006.

[3] Chowdhury, G. G. & Chowdhury, Sudatta, ''Introduction to digital Libraries''. Facet publishing, London,2003.

[4] https://www.quora.com/What-is-the-difference-between-SAN-DFS-and-NAS (retrieved 2015).

[5] Hendley, Tony. Comparison of methods & costs of digital preservation. Report 106-British Library Research and Innovation – available at: http://www.ukoln.ac.uk/services/elib/papers/tavistock/hendley/hendley. html(2007).

[6] Mohammed Amin Marghalani and Susan Mustafa '' ,metadata in the websites of university libraries Saudi Arabia, analytical study, information studies, p. 16 , 2008.

[7] William Y. Arms,'' Digital Libraries'' MIT Press publisher, New York , 2000.