APPLYING DIGITAL FORENSICS TO THE OFF-LINE TRANSFER OF ELECTRONIC RECORDS

Ji-Hye Park
National Archives of Korea

Abstract

Beginning in 2015, the National Archives of Korea (NAK) plans to make off-line transfers of over three terabytes of standard electronic records a year. However, many have pointed out that compared to the online transfer procedure, which has a carefully planned and designed record control system, the off-line counterpart is much more prone to various security risks, including forgery and falsification risks. The goal of this study is to find ways that will ensure the security of the off-line procedure by applying digital forensics, which is emerging as a promising field of digital preservation and conservation. The 2013 Records Preservation Technology R & D Project has developed to a new digital forensic transfer tool. The tool was subjected to a standardization process and is being reviewed for policy feasibility in 2014. Two government agencies in Korea have tried out the new digital forensic transfer tool by applying it to the transfer of electronic records produced in 2004. The test has helped strengthen the tool and the related procedure.

1. Introduction

Since 2004, the Korean government has been producing electronic records using a standardized system. Beginning in 2015, it will transfer over three terabytes of such records each year, using a Record Management System (RMS) and a Archive Management System (AMS). In designing the procedure for the transfer of electronic records, the Korean government standardized the data interchange specification between RMS and AMS and the technical specification for online transfer. In an effort to ensure the efficiency of the online transfer process and the integrity of the transferred records, the RMS was given functions to verify users, generate time logs on the transfer details, encrypt of transferred files, screen the compatibility of files, and transmit large-sized records with log management functions.[1]

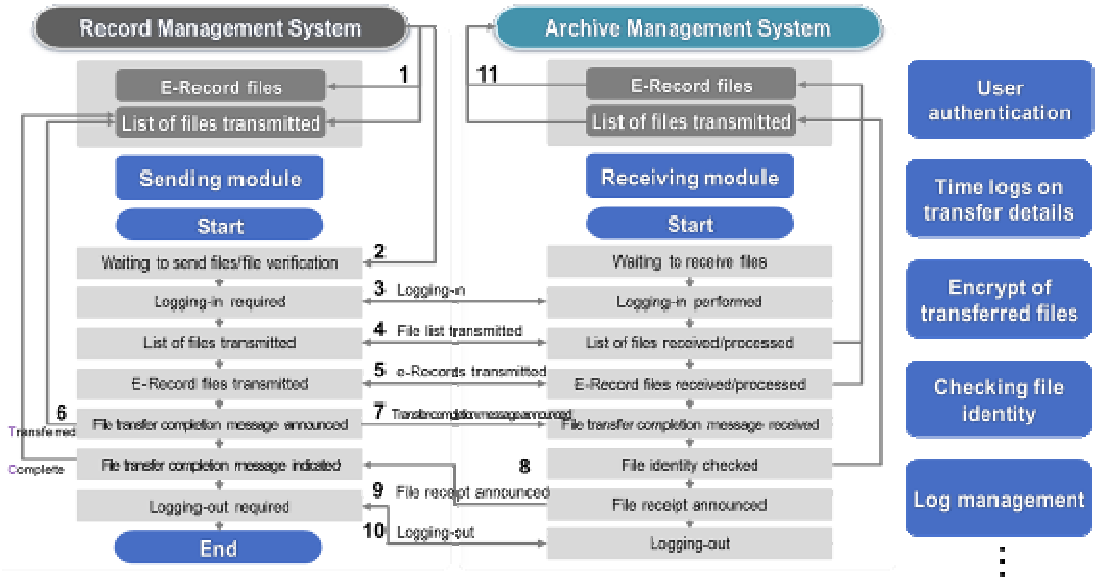

Figure 1. Concept and functions of the procedure for the online transfer of electronic records

Contrary to intended function, off-line transfer is often preferred over the online method as numerous government agencies and organizations are incapable of adopting the RMS, while some only resort to closed networks. There are also audiovisual materials and presidential records that are too large for online transfer. However, off-line transfer also increases security

risks, such as forgery and falsification, theft and loss, and damage of the transferred media that requires repair.

Electronic records can be easily deleted and falsified, and often raise questions of integrity and reliability in forensic settings. The off-line transfer of these records necessarily involves the media migration. Copy of these records, however, carries risks of damage to the meta-data information and/or the bit streams of the original data, making it nearly impossible to verify and certify that no acts of forgery or falsification have been attempted.

In accordance, the NAK began its search for ways to ensure the security of off-line transfer of electronic records using digital forensics. Digital forensics refers to a series of techniques and practices relating to the collection and analysis of digital information using legitimate procedures and science so that such information can be treated as admissible evidence in the court of law. Until now, digital forensics has been almost an exclusive purview of prosecutors and police. Nevertheless, it has much use for the management of electronic records, as digital forensics can help ascertain the integrity of given records and ensure the chain of custody in all the steps of the process, from production to the final receipt of those records, including collection, transfer, analysis, conservation, and presentation to the court of law.

An Inter-Institutional Model for Stewardship (2012), a project that involved the University of Virginia, Stanford University, the University of Hull, and Yale University, presents a workflow that applies different reaching modes to different types of storage media, checks for viruses, generates technical meta-data like checksums, and identifies software and file formatting issues using DROID.[2] BitCurator Project (2013), a joint study of the University of North Carolina and the University of Maryland, involves creating a Linux-based virtual environment, and performs disk imaging, identification of sensitive information, data classification, and meta-data extraction using a variety of open-source tools, such as Guymager(a multi-process tool for generating forensic images), GTK Hash (a tool for generating encrypted hashes), and DFXML (a tool for processing digital forensic extensible markup languages).[3] The British National Archives has also developed a guideline (2013) for the digital transfer of electronic records, involving the use of such hardware media as USB sticks, DVDs, and encrypted hard disks, as well as the use of encrypted electronic assemblies.[4]

The NAK conducted a study on the development of a digital forensic tool for the transfer of electronic records that is safe against security risks, such as the possibilities of forgery and falsification, loss, and theft (2013). The tool was given a reinforced process of collecting user verification and system information so as to ensure the continuity of custody over electronic records in off-line transit. The applied password algorithms are meant to prove the integrity of the transfer process. In addition, other technical details were added to enable the tool to generate image files and hashes and to verify files.[5] In other words, the study aimed to produce a tool that can facilitate the off-line transfer of records for those that cannot be transferred online because of network-related or size reasons. Chapter 2 discusses the main findings of the study in detail.

2. Function analysis and design of the digital forensic transfer tools

2.1. Function analysis of the digital forensic transfer tools in existing studies

As previously mentioned, numerous studies precede NAK's study worldwide, involving the application of digital forensics to the control and transfer of electronic records. These projects mostly relied on Encase or FTK—tools that investigators use to collect digital evidence—or some other tools developed by the researchers themselves. These tools, however, fail to verify users and also to conduct analysis of a given system information. User verification and signature authorization can be crucial to ensuring the reliability of given electronic records. It may also be equally important to collect and analyze system information in order to ensure proper control of the records' histories. Therefore, NAK decided to develop a new digital forensic tool that can verify users, collect time logs and other system information on given records so as to ensure a greater reliability of the transfer process.

Table 1. Survey of the literature: Analysis of digital forensic transfer tools in existing studies

Tool (version)	Platform	Imaging	Hash	Time Stamp	User Log-in	System Inspection	Encrypt	Project	Open Source
FTK Imager (3.1.2)	Windows	O	O	O	X	X	O	AIMS	X
FTK (4.2.1)	Windows	X	X	X	X	X	O	AIMS	X
EnCase Imager (7.06)	Windows	O	O	O	X	X	O	-	X
Guymager (0.7.1)	Unix	O	O	O	X	X	X	BitCurator	O
Tableau Imager (1.2)	Windows	O	O	O	X	X	X	-	X

2.2. Design of the digital forensic transfer tools

A digital forensic transfer tool applies digital forensic techniques to the off-line transfer process of electronic records so as to ensure the integrity of the transferred records, as well as the security of the transfer process itself. Figure 2 illustrates the off-line transfer procedure using the digital forensic transfer tool developed by the NAK. Along with the online RMS that seeks to ensure the security and efficiency of the online transfer process and the integrity of electronic records, the digital forensic transfer tool likewise verifies users, generates time logs on transfer details, encrypts electronic record files, tests file compatibility using the hash function, and keeps logs, thus avoiding the many shortcomings and deficiencies of existing off-line processes.

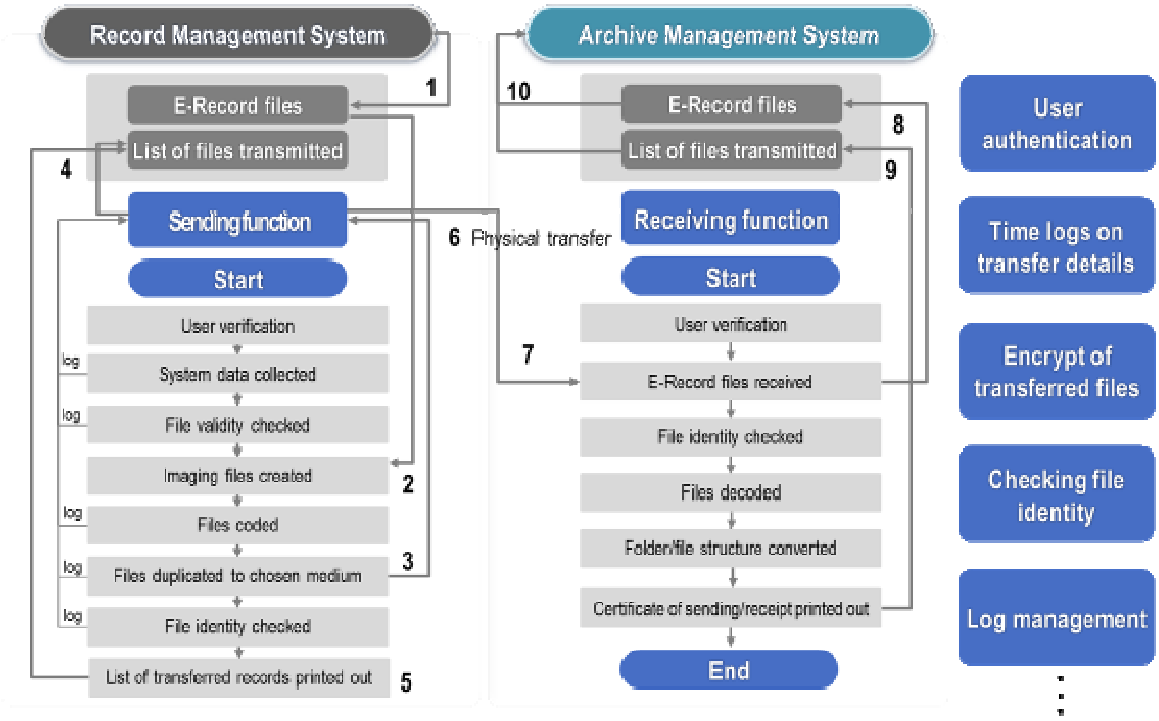

Figure 2. Concept and functions of the off-line transfer procedure of electronic records

2.3. Implementation of the digital forensic transfer tool

2.3.1. User verification and authorization

A major problem of the off-line transfer process is that it can leave the information blank on the persons actually involved in the transmission of given electronic records. The NAK tool allows the user to enter the information of persons involved in the transfer, and to test the match between the entered information and the approved information as part of verifying those persons' identities. This allows the user to block unauthorized third parties from accessing the transfer records, and to also ascertain by whom the given electronic records were collected, handled, and transferred.

2.3.2. Collecting time logs and other system information

The actual sources and recipients of electronic records are an unidentifiable number of computers used at a variety of institutions and agencies, where it is virtually impossible to keep exact track of who accesses the Internet and the given records. However, in order to ensure the integrity of transferred files, it is crucial to keep track of the times at which given records changed their status. These time data are central to the verification of transferred records, and therefore must be kept precisely. Users of Windows operating systems can change the time data on their computers arbitrarily, thus leaving incorrect time information on the transferred records. While one may automatically update the time information via the Internet, there are many government computers that are not connected to the Internet. The NAK tool, therefore, requires the user to enter the time at the very beginning of using the tool. Moreover, the tool collects and keeps information on the operating systems of the computers to which files are to be transmitted, the names of system users, and whether or not the recipient computers are connected to the Internet.

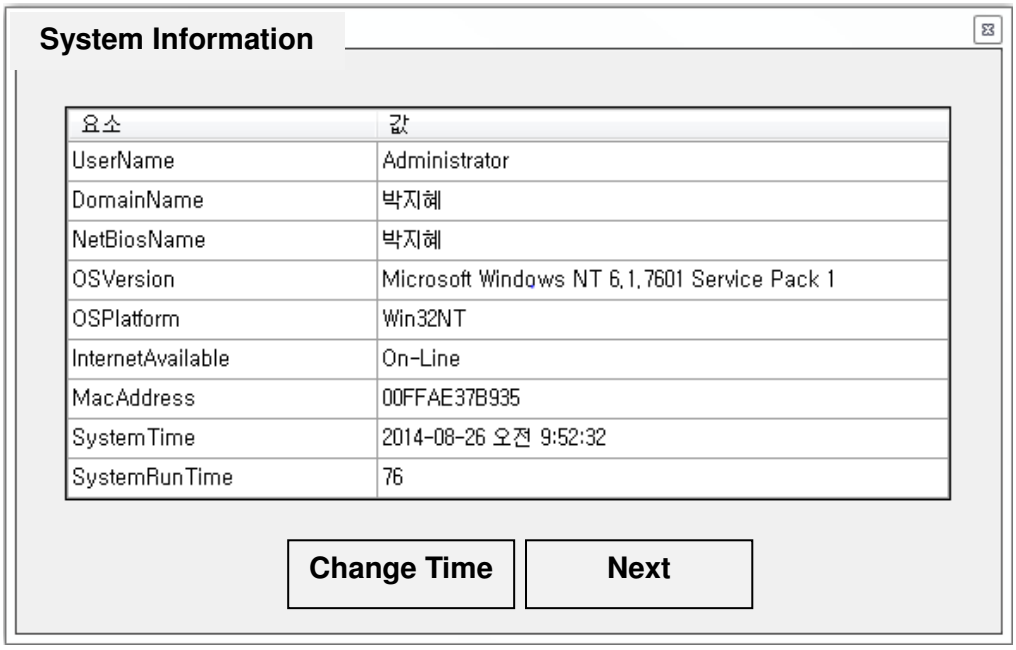

Figure 3. Time logs and other collected system information

2.3.3. Verifying the validity of electronic records

Electronic records come in diverse formats. In transferring these records of diverse formats, it is crucial to ensure the validity of those formats using an automated tool for format distinction. With the NAK tool, the validity of the transferred records can be validated by selecting “File Signature Analysis” in the “File” menu. A file signature refers to a series of bytes found at the front of a given digital file that is often used to identify the specific type or format of the file. The NAK tool includes a database consisting of file signatures of different formats that have been preauthorized. The transferred electronic records are then compared against these signatures to ensure their validity.

2.3.4. Imaging electronic records

You can save a given electronic record on the medium of your choice through duplication, reproduction, or imaging. Until now, the most preferred way of saving transferred electronic records was duplication. Duplication involves the simple relocation of original files or directors, using the copy-and-paste (Ctrl+C and Ctrl+V) function. However, this method can only be used on logical data that can be viewed and confirmed through the browser function of the computer. Also, it does not allow the recovery of deleted or removed data. Therefore, digital forensics resorts to imaging instead of duplication. Imaging involves creating copies of the original data, and saving the data of all sectors in the form of files. With imaging, not only browsable files, directories, and their structures can be saved, but also the relative positions of the given data as well as deleted files saved in nonallocated areas that have not been written over can be all retrieved and recovered in the analysis process. Applying the imaging function to the digital forensic transfer tool also facilitates the processes of hash extraction and encrypt.

2.3.5. Encrypt of electronic record files

It is necessary to encrypt the imaging files to protect electronic records against possible loss or theft during the transfer process. The NAK tool requires the user to submit a code consisting of six or more digits and letters. The tool then combines the code with the user’s ID and password, and applies the hash value as a key to perform encryption. In the process of relocating electronic records from a medium to the AMS, the tool enters the code provided by the user and thereby decodes the encrypted image-formatted files. The encrypt algorithm

used for the NAK tool is Advanced Encryption Standard (AES)-256.

2.3.6. Checking identity of files

It is absolutely essential to ensure the identity of the transferred records with the original ones. In order to identify the transferred files, the NAK tool generates hash values on the encrypted image-formatted files upon sending and receiving them, and compares the values. The hash algorithm applied is Secure Hash Algorithm (SHA)-2.

Figure 4. Automatic imaging, decrypt, file/folder structuralization, and hash verification

2.3.7. Issuing certificates for the sending and receipt of electronic records

The NAK tool also prints out detailed certificates of electronic records that have been sent and received. At the end of sending and receiving a given record, the sender and the recipient exchange their manual signature to confirm the completion of the transaction, and keep record of it as part of the history of the entire off-line transfer process.

Certificate of Sending and Receiving Electronic Records

[Transaction Details]

· Number of files: 27
· Total file size: 54,333,320 bytes
· Image name: EncryptedDigitalRecord_test2_2014116145955
· Number of images: 1
· Transfer start time: 15:01:07, 16-JAN-2014
· Transfer end time: 15:07:43, 16-JAN-2014

[Process Details]

· 15:01:07, 16-JAN-2014: Transfer began under supervision of “test 2.”
· 15:01:07, 16-JAN-2014: System information were saved.
· 15:02:43, 16-JAN-2014: 20 folders were selected.
· 15:02:43, 16-JAN-2014: Logs of selected folders were saved.
· 15:02:44, 16-JAN-2014: Validity of selected files was reviewed.
· 15:07:10, 16-JAN-2014: 27 files were selected.
· 15:07:11, 16-JAN-2014: Logs of selected files were saved.
· 15:07:36, 16-JAN-2014: Key for encryption was entered.
· 15:07:39, 16-JAN-2014: 1 image file was created.
· 15:07:41, 16-JAN-2014: 1 image file was encrypted.
· 15:07:43, 16-JAN-2014: 1 image file was transferred to pre-defined medium.
· 15:07:43, 16-JAN-2014: Hash value of duplicated image file was verified.
· 15:07:43, 16-JAN-2014: Log of duplicated image file was saved.
· 15:07:43, 16-JAN-2014: Transfer of electronic records was successfully completed.

[Hash Value Verification]

· Original record hash value:
B34E013945ACBD9711C98C85050B8C7F15B59B3D
· Transferred record hash value:
B34E013945ACBD9711C98C85050B8C7F15B59B3D

Transfer of aforementioned details has been successfully completed.
16 January 2014

Sender: (Signature)
Recipient: (Signature)

Figure 5. Certificate of sending and receiving electronic records using the digital forensic transfer tool

2.4. Example of transferring electronic records off-line using the digital forensic transfer tool

The NAK tested its digital forensic transfer tool by applying it to the online and off-line transfer of 1,852 files (total of 30 gigabytes) from the Ministry of Security and Public Administration and the Fair Trade Commission, using a software program for the transmission of large-sized files. The records, which were of the standard format and size defined by the RMS, were transferred and subjected to the process of meta-data analysis, electronic signature validation, and virus screening before they were saved on the chosen storage. The test demonstrated that each and every function of the tool was working as intended. The tool, however, has much more functions other than duplication—such as imaging file creation, encryption, and hash comparison—and the transfer process therefore, took twice as long with the tool than without it.

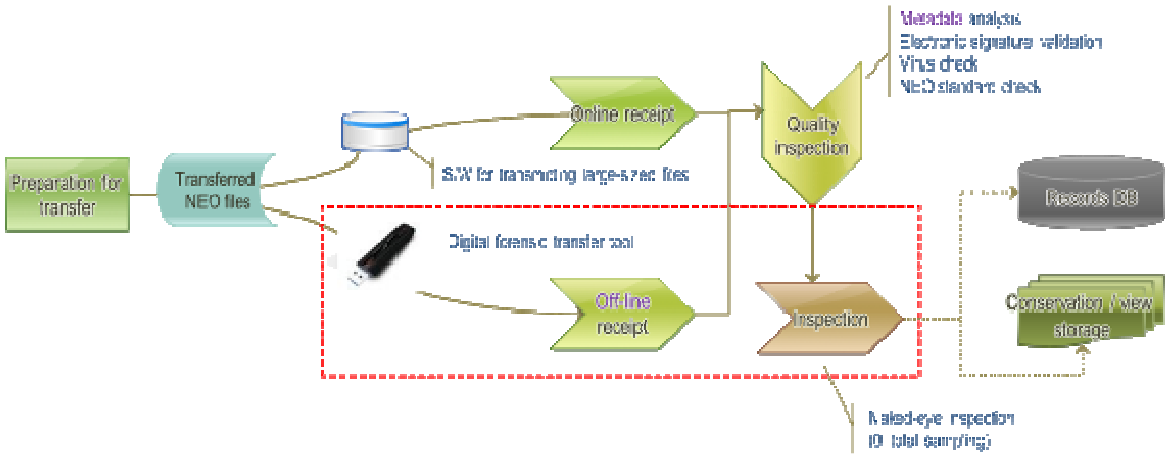

Figure 6. Process of trial online and off-line transfer using the digital forensic transfer tool

3. Research on the preservation of electronic records

Supported by new policies and institutional measures since 2007, such as the Public Records Management Act and the Guideline for the Management of the NAK R & D Projects, the NAK has been investing KRW 1 billion each year into developing five new technologies for the conservation and restoration of paper records, long-term conservation of electronic records, conservation and restoration of audiovisual records, conservation and restoration of public administration relics, and maintaining optimal conservation conditions for records in general. The research and development of the digital forensic transfer tool forms part of these larger projects. It is therefore proper that we lay out some of the main accomplishments of the NAK's R & D projects in this regard until now.

3.1. Digital Format Registry(DFR)

Digital formats are dependent upon specific software and hardware technologies. Electronic records of a given format may not be retrievable or playable once the technology supporting that format disappears. Migration, emulation, format standardization, and other common methods of conservation all have benefits and advantages, but the use of a single method is not suited to the long-term conservation of electronic records. Therefore, a core task in ensuring the long-term conservation of electronic records is developing a system that can collect and manage information on the evolving digital format technologies. While researchers in the United Kingdom and the United States have developed such systems, there is a critical dearth of efforts in Korea to keep track of the key evolving format technologies, including word-processing programs like Hangul and Hunminjeongeum.

The NAK has thus launched a project for researching and developing a DFR and a related database on evolving format technologies. Based on the findings and outcomes of this project, the NAK has been collecting and preserving technical information on the formats of electronic records and verifying structural errors in file formats since 2014. Using the registry, one can now keep track of and manage technical information on file formats favored by the Korean government agencies and unique to Korea, such as Hangul and Hunminjeongeum. The registry is also being upgraded so that it can identify and verify structural errors in file formats because of its extensions and signature information. The registry identifies formats based on certain bytes found at the beginning and end of a given electronic file. To identify and verify structural errors in these formats, the registry parses the structure of a given electronic file and obtains structural data on possible damage, passwords, empty files, presence of information linked to images contained within, and errors in the images themselves. The automatic verification function of this registry is expected to maximize the efficiency of the transfer process in terms of both time and money required. The NAK plans

to apply this registry to a greater number of its records over time.

3.2. Disaster Recovery System(DRS)

The Public Records Management Act requires institutions keeping permanent records, such as the NAK, to establish an emergency recovery system, including electronic procedures for recovering data, recording media, and systems that are lost in the event of an emergency. In 2008, the NAK embarked on a research project for the development of a standard model for emergency recovery systems for electronic records. From 2009 to 2012, their findings were applied to the development of an electronic emergency recovery system. To this end, the NAK established a DRS (in the class of a warm site) using resources in Daejeon and Seongnam, which are about 120 kilometers away from each other. The recovery system provides a dual-protection and recovery for online and off-line media alike. The system is designed to recover lost data and information within a fixed period of time, using records and equipment stored at a remote location, in the events of disasters or other emergencies that may interrupt the workings of the peacetime computer center.

In designing the system, the NAK researchers distinguished between electronic records to be recovered and details of system management to be backed up and recovered. They also distinguished between the emergency recovery process and the peacetime recovery process. The long-term electronic signature verification data can now be duplicated and recovered between Daejeon and Seongnam. The emergency system operation and process manuals will also support the recovery and normalization of records within fixed periods of time. Upon detecting errors in records, users of the system can activate both online and off-line media to recover the damaged records and verify their identity with original records after recovery.

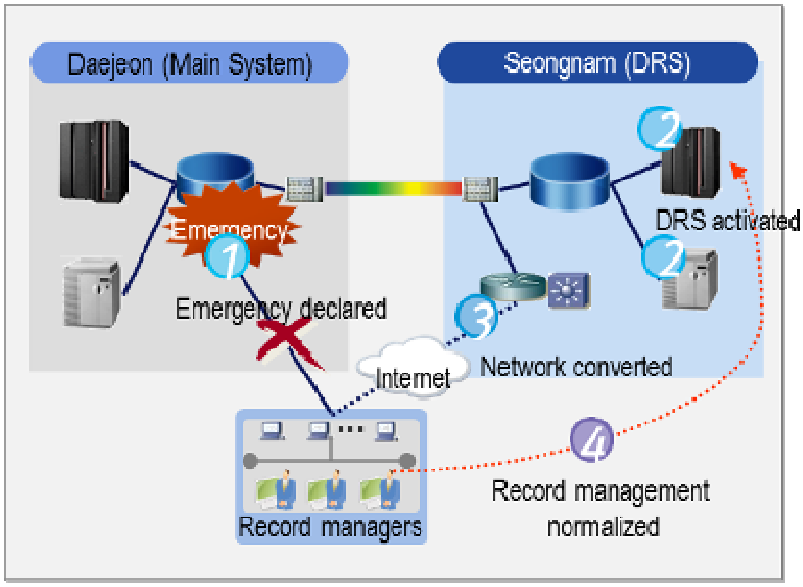

Figure 7. Process of recovering electronic records in the event of an emergency

3.3. Optical media condition assessment

Electronic media are harder to preserve than their non-electronic counterparts. Therefore, it is essential to keep monitoring electronic records regularly and frequently. The electronic media into which records are written are prone to deterioration by heat, humidity, light, and other such factors, which may eventually render the records unreadable over time. ISO/IEC 29121:2009 for instance, requires that the condition and state of a DVD be distinguishable by reference to the PI error value. A PI error value is the maximum number of PI errors detected in eight sequential ECC blocks. According to the International Organization for

Standardization (ISO), the integrity of data cannot be ensured on a DVD with a PI error value of 280 or higher. With this, the data have to be relocated somewhere else.[6] Thus, the NAK has developed a technique for assessing the condition of given optic media by gauging the number of PI errors based on the ISO standard.

Table 2. DVD data migration method according to the measurement results

Stage	Condition	DVD-R, DVD-RW, +R, +RW
1	Current condition kept	< 200
2	To be relocated	200 – 280
3	To be relocated immediately	> 280

References

[1] National Archives of Korea (2013). “Technical Specification for Online Transmission of Digital Records (v1.1),” NAK/TS, p. 7–17.
[2] University of Hull, Stanford University, University of Virginia, Yale University, “AIMS Born-Digital Collections; An Inter-Institutional Model for Stewardship,” 2012.
[3] http://www.bitcurator.net.
[4] http://www.nationalarchives.gov.uk/information-management/our-services/digital-records-transfer.htm.
[5] National Archives of Korea (2013). “Development of Digital Forensic-based Electronic Records Migration Tools.”
[6] ISO/IEC 29121 (2009). “Information technology - Digitally recorded media for information interchange and storage - Data migration method for DVD-R, DVD-RW, DVD-RAM, +R, and +RW disks.”