

## 전자기록의 장기보존전략을 위한 의사결정 절차에 관한 연구

최주호<sup>○</sup>, 차현철<sup>\*</sup>

<sup>○</sup>주식회사 이산티에스

<sup>\*</sup>동양대학교 컴퓨터정보전학과

e-mail: iq2chun@esants.co.kr<sup>○</sup>, hccha@dyu.ac.kr<sup>\*</sup>

## A Study on the establishment of criteria for selecting long-term preservation strategy of electronic records

Joo-Ho Choi<sup>○</sup>, Hyun-Chul, Cha<sup>\*</sup>

<sup>○</sup>EsanTS Co. Ltd.,

<sup>\*</sup>Dept. of Computer Information Warfare, Dong-Yang University

### ● 요약 ●

본 논문에서는 파일 포맷에 대한 위험요소 평가를 기반으로 해당 포맷에 대한 장기보존 방안을 제시하는 절차 및 시스템 구축에 필요한 아키텍처를 정의하였다. 파일포맷이 사양화된 경우 이를 재현할 수 있는 장기보존전략으로 마이그레이션과 에뮬레이션 방안이 제시되고 있는데 개별 포맷에 대한 위험도 평가를 통해 가장 적합한 보존전략을 선정하여 전자기록의 재현성을 부여하는 연구를 진행하였다.

**키워드:** 위험평가(Risk Evaluation), 마이그레이션(Migration), 에뮬레이션(Emulation)

### I. Introduction

전자기록에 대한 위험평가는 전자기록에 적용된 파일포맷의 위험도를 평가하는 것으로 본 연구에서는 파일포맷의 위험도 평가를 위해 앞서 제시된 21개의 위험요소 항목을 대상으로 포맷의 위험도를 평가한 후 장기보존 전략을 선정하는 프로세스 및 시제품 구현을 위한 아키텍처에 대한 연구를 수행하였다.

국가 공공기관에서 운영하는 시스템에 저장된 파일포맷의 유형별 분포 현황 및 특징을 분석한 결과 문서, 이미지, 소리, 동영상 등 4개 유형의 파일 포맷이 81%를 점유하고 있으며 총 파일포맷 수는 약 261개로 조사되었다. 그러나 파일 포맷별로 다양한 포맷 버전이 생산되므로 향후 관리되어야 할 파일 포맷의 대상 수는 대폭 증가할 것으로 예상된다.

### II. Preliminaries

#### 1. Related works

##### 1.1 국내 공공기관 파일포맷 보유 현황 분석<sup>1)</sup>

1) 대상 : 중앙부처(23), 지자체(4), 교육청(2), 공공기관 (56)

### III. The Proposed Scheme

#### 1. 장기보존 의사결정 프로세스 및 아키텍처

##### 1.1 의사결정 프로세스

파일포맷에 대한 위험요소 평가를 수행한 후 그 결과를 바탕으로 해당 포맷에 대한 장기보존 방안을 제시하기 위한 전체 프로세스를

<Fig 1>에 제시하였다. 평가를 위한 전자기록의 External Signature 와 Internal Signature 추출 정보와 기술정보 Registry에 등록된 정보를 바탕으로 포맷에 대한 위험평가를 수행하여 위험평가 결과와 위험도를 평가한 후 장기보존전략을 선정하는 절차이다.

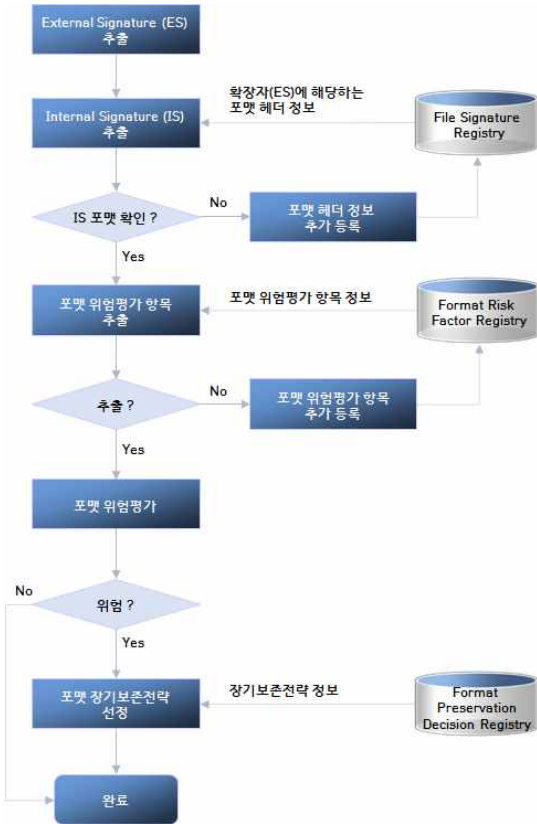


Fig. 1. 장기보존 의사결정 프로세스

### 1.2 아키텍처 정의

위험평가 시스템을 구성하는 모듈은 크게 Service, Evaluator, Manager로 구성되는데 Service 모듈은 사용자 인터페이스를 담당하고 Evaluator는 본 연구에서 제시한 알고리즘을 기반으로 평가를 담당하며 Manager는 평가를 위해 필요한 5개의 레지스트리를 관리하도록 구성하였다. Manager를 통해 관리하는 영역은 포맷의 기술정보(FSR), 포맷의 위험요소 값과 지수 및 가중치(FRR, RIF, RWR)와 장기보존전략(FPR)을 관리하도록 설계하였다.

### 1.3 시제품 구현

본 연구에서는 제시한 알고리즘과 절차를 기반으로 위험평가 서비스를 수행하도록 시제품을 개발하였고, 시제품을 구동하여 7개의 전자기록을 등록한 후 평가를 수행한 최종 결과 화면을 <Fig 2>에 제시하였다.

번호	종류	파일명	파일 크기	위험요소	평가결과	보존방안	문자코드	
1	Text (Word Processing)	안우영(국무총리)서명	MSIP Document File	5	7.5	안전	가용부치	완료
2	Text (Word Processing)	MS09KMT1000L	HandeJungam	200	2.0	위험	가용부치	완료
3	Page Description	공민_001_2008_2008_Annex_L2012...	Alphabet PDF 1.2 - Portable Document Format	1.3	0.5	안전	가용부치	완료
4	Audio	DaesilEonam-BesulkiSundae_김용...	Windows Media Audio	9.2	7	안전	가용부치	완료
5	Audio	나을과-이-나을-그림(mp3)	MPEG-12 Audio Layer 3	2.0	1.0	안전	가용부치	완료
6	Image (Raster)	197가평출안안 and	Adobe Photoshop	1.1	4.2	보통	가용부치	완료
7	Text	백기서(문신)가 서명	Adobe Flash	3.0	3.2	위험	가용부치	완료

Fig. 2. 위험평가 결과

## IV. Conclusions

본 연구에서 전자기록의 장기보존 의사결정 지원을 위한 위험평가 알고리즘을 개발하여 다양한 포맷을 대상으로 위험 평가를 수행한 결과 다음과 같은 결론 및 향후 연구가 필요하다고 판단되었다. 우선 다양한 분야에서 포맷 정보의 표준화가 수행되고 있으므로 포맷 유형별 표준화 연구에 대한 지속적인 관심과 연구가 필요하고, 제시한 21개의 위험요소 항목은 해외 사례를 기반으로 추출된 위험 요소 항목이므로 향후 위험요소 항목에 대한 국내 상황을 반영하는 심도 깊은 연구가 필요하다.

## Acknowledgment

본 논문은 “2016년 안전행정부 국가기록원의 기록물 보존기술 연구개발 사업”의 일환으로 진행한 “전자기록 기술정보기반 장기보존 의사결정 지원체계에 관한 연구” 과제 결과를 토대로 작성됨

## References

- [1] Adrian Brown, Head of Digital Preservation, “Automatic Format Identification Using PRONOM and DROID”, The National Archives 2006.
- [2] Myongji University, “A study on the reproduction technology and the prototype for the electronic records of administrative agency”, National Archives of Korea (11-1312125-000014-01) 2013.
- [3] David Levy, “AONS As Built Specification”, National Library of Australia, August 2007.