

N a t i o n a l A r c h i v e s S t a n d a r d

I 전자기록물 보존포맷 선정기준(v1.0)

Selection Criteria for Preservation Format of Digital Records

Version 1.0

- 제 정 자 : 행정안전부 국가기록원장
- 제 정 일 : 2022년 10월 17일(국가기록원 고시 제2022-5호)
- 심 의 : 국가기록관리위원회, 기록·정보정책전문위원회
- 원안작성 :
 - 이철편(국가기록원 서기관)
 - 임형렬(국가기록원 공업연구사)
- 검 토 :

<ul style="list-style-type: none"> • 김언희(부산진구청) • 백영미(국가철도공단) • 이인동(한국가스안전공사) • 전세웅(대전지방보훈청) • 한승우(국방과학연구소) • 김명옥(국가기록원) 	<ul style="list-style-type: none"> • 김정훈(청주시청) • 성민결(강원지방병무청) • 이수진(한국남동발전(주)) • 하성혜(합동참모본부) • 황인규(충청남도태안교육지원청) • 송헌규(국가기록원)
--	--
- 관 리 :
 - 국가기록원 정책기획과
- 자 문 :
 - 양동민(전북대학교)
 - 남동선(주)팬소프트 대표)

(1) 이 표준의 열람은 국가기록원 홈페이지를 이용하시고, 의견 또는 질문은 아래 전화로 연락 주십시오.

- 표준열람 : 국가기록원(<http://www.archives.go.kr>)→기록관리업무
→기록관리표준→표준화현황
- 행정안전부 국가기록원 기록정책부 디지털혁신과(042-481-1759)
기록정책부 정책기획과(042-481-6231)

(2) 이 표준은 「저작권법」 제24조의 2(공공저작물의 자유이용)에 따라 저작권자인 국가기록원의 허락 없이 자유롭게 이용할 수 있습니다. 다만, 저작물을 이용하는 자는 그 출처를 명시하여야 하며, 영리를 목적으로 하는 무단 복제 및 배포는 금지합니다.

Copyright© National Archives of Korea(2022). All Rights Reserved.



목 차

머리말	ii
1 적용범위	1
2 적용근거	1
2.1 법적 근거	1
2.2 인용표준	2
2.3 다른 표준과의 연계	2
3 용어정의	2
4 일반사항	3
4.1 보존포맷의 다양화 필요성	3
4.2 보존포맷 선정기준 필요성	4
5 보존포맷 선정을 위한 공통기준	5
5.1 일반사항	5
5.2 공통기준의 구성	5
5.3 공통기준에 따른 보존포맷 평가방법	8
6 보존포맷 선정을 위한 고유기준	12
6.1 일반사항	12
6.2 고유기준의 구성	13
6.3 고유기준에 따른 보존포맷 평가방법	18
7 전자기록물의 보존포맷 선정방법	22
7.1 권고포맷 선정방법	22
7.2 보존포맷 선정방법	22
부속서 (참고)	24
참고문헌	43

머리말

이 표준은 장기간에 걸쳐 전자기록물의 진본성을 보호하고, 지속적으로 재현 가능하도록 하는데 필요한 전자기록물의 보존포맷 선정을 위한 기준을 제공하기 위해 제정되었다.

이 표준은 「공공기록물 관리에 관한 법률」 제20조(전자기록물의 관리) 및 동법 시행령 제36조(기록관 및 특수기록관의 전자기록물 보존)에서 규정한 중앙기록물관리기관의 장이 정하는 '보존포맷'의 선정기준에 해당한다.

기술 환경의 발전에 따라 파일 포맷의 종류나 특성이 다양해지고 이에 따라 생산되는 전자기록물이 다양해지고 있으나, 기술의 변화에 종속되지 않으면서 모든 유형의 전자기록물 파일에 적용가능한 보존포맷 선정을 위한 공통 기준과 기록물 각 유형별 고유기준을 제공한다. 문서유형 이외의 다른 유형의 기록물에 적합한 고유기준은 지속적으로 표준의 '부(part)'를 추가하여 제공할 예정이다.

중앙기록물관리기관의 장은 이 표준에서 제시하는 기준에 따라 후보가 되는 포맷들을 주기적으로 평가하여 각 기록물 유형별 권고포맷(보존포맷, 수용가능포맷)을 정보통신망 등을 통해 게시한다. 다만, 이 표준은 현 시점에서 문서유형 전자기록물에 적용가능한 권고포맷(보존포맷과 수용가능포맷)을 부속서로 제공한다.

NAK 30:2008(v1.0) 「전자기록물 문서보존포맷 기술규격」은 문서유형 전자기록물 보존포맷 중 PDF/A-1b 기반의 포맷에 해당하는 기술규격으로 이 표준이 제시하는 보존포맷의 기준을 충족하므로 내용을 일부 수정하여 유지한다. 더불어 각 공공기관이나 영구기록물관리기관이 선택 가능한 보존포맷의 범위를 확대하고 유연화하려는 보존정책의 일환으로 단일한 보존포맷을 선정하여 제공하기 보다는 보존포맷 선정을 위한 기준을 제공하기 위해 이 표준을 제정한다.

이 표준에서 제시하는 보존포맷 선정을 위한 공통기준, 문서유형 전자기록물에 적합한 고유기준은 국가기록원이 '19년~'20년도에 수행한 아래 연구개발사업의 결과물을 기반으로 한다.

- '19년 데이터세트 유형 전자기록의 장기보존기술 연구
- '20년 문서유형 보존포맷 및 장기보존패키지 다양화 연구

이 표준은 기록·정보정책전문위원회 및 국가기록관리위원회 심의를 거쳐 제정되었으며, 국가기록원이 유지·관리한다. 이 표준은 관련 법령의 개정, 관계 기관 및 이해당사자의 요청 등 개정사유가 발생할 경우 그 필요성 및 타당성을 검토한 후 개정안을 마련하고 의견수렴 및 심의 절차를 거쳐 개정한다.

전자기록물 보존포맷 선정기준

1 적용범위

이 표준은 「공공기록물 관리에 관한 법률」 및 동법 시행령 규정에 따른 보존포맷의 선정기준을 제공하며, 공공기록물법의 적용을 받는 모든 공공기관 및 기록물관리기관에 적용된다.

공공기관은 이 표준에서 제공하는 선정기준에 따라 관할 영구기록물관리기관과 협의를 통해 자체적으로 적합한 보존포맷을 선정하여 적용하거나, 영구기록물관리기관이 관할 공공기관에 적용되는 보존포맷을 결정하여 일괄 적용할 수 있다.

비고 「공공기록물 관리에 관한 법률」 시행령 개정으로, '문서보존포맷'은 '보존포맷'으로, '장기보존포맷'은 '장기보존패키지'로 변경됨에 따라, NAK 30 2008(v1.0) 「전자기록물 문서보존포맷 기술규격」을 「문서유형 전자기록물 보존포맷 기술규격: PDF/A-1b 기반의 포맷」으로 변경한다.

2 적용근거

2.1 법적 근거

이 표준의 구체적인 법적 근거는 다음과 같다.

- 「공공기록물 관리에 관한 법률」 제20조(전자기록물의 관리)
- 「공공기록물 관리에 관한 법률」 시행령 제36조(기록관 및 특수기록관의 전자기록물 보존)
- 「공공기록물 관리에 관한 법률」 시행령 제46조(영구기록물관리기관의 전자기록물 보존 및 관리)

2.2 인용표준

다음의 인용표준은 이 표준의 적용을 위해 필수적이다. 발행연도가 표기된 인용표준은 인용된 판만을 적용한다. 발행연도가 표기되지 않은 인용표준은 최신판(모든 개정내용을 포함)을 적용한다.

- ISO 15489-1(2016) Information and documentation - Records management - Part 1: Concepts and principles
- ISO 30300(2020) Information and documentation - Records management - Core concepts and vocabulary

2.3 다른 표준과의 연계

이 표준의 적용을 위해 필요하거나 직접적으로 연관이 있는 표준은 다음과 같다. 발행연도가 표기되지 않은 표준은 최신판(모든 개정내용을 포함)을 적용한다.

- NAK 30 2022(v1.1) 문서유형 전자기록물 보존포맷 기술규격: PDF/A-1b 기반의 포맷

3 용어정의

이 표준의 목적을 위하여 다음의 용어와 정의를 적용한다.

3.1 기록관리 메타데이터(Metadata for managing records)

시간이 지나가도 영역(domain) 안과 영역 간에 기록의 생산, 관리 및 이용을 가능하게 하는, 구조화되었거나 반구조화된 정보

[ISO 15489-1(2016), ISO 23081-2(2009)]

3.2 보존포맷(Preservation Format)

전자기록물 생산 당시의 내용과 외형 등 주요 특성을 유지함으로써 시간과

기술의 변화와 상관없이 해당 기록물을 재현할 수 있는 포맷

비고 전자기록물의 진본성을 유지하며 장기간 보존하고, 지속적으로 재현할 수 있기 위하여, 전자기록물의 내용, 구조, 맥락, 외관, 기능을 보존할 수 있도록 설계된 포맷을 의미한다.

3.3 이용가능성(Useability)

이해관계자들이 합당하다고 여기는 시간 안에 위치를 찾을 수 있고, 검색할 수 있으며, 보여줄 수 있고, 해석할 수 있는 기록의 속성

[ISO 30300(2020), ISO 15489-1(2016) 참조하여 개작]

3.4 장기보존(Long-term Preservation)

오랜 시간이 경과된 후에도 전자기록물에 접근할 수 있고 진본의 상태를 유지하여 증거로서 인정받을 수 있도록 하는 보존행위

3.5 진본성(Authenticity)

기록이 표방하는바 그대로의 것인지, 그것을 생산했거나 보낸 것으로 되어 있는 바로 그 행위 주체가 생산했거나 보냈는지, 그리고 명시된 시점에 생산되었거나 보내졌는지를 증명할 수 있는 기록의 품질

[ISO 30300(2020), ISO 15489-1(2016) 참조하여 개작]

3.6 필수보존속성(SP: Significant Properties)

시간이 지나도 접근가능하고 의미 있는 상태를 유지하기 위해서 보존되어야 하는 전자기록물의 속성

4 일반사항

4.1 보존포맷의 다양화 필요성

전자기록물의 장기보존을 위한 보존포맷으로 PDF/A-1b 기반의 포맷(NAK 30 2022(v1.1))이 유일한 포맷으로 제공되었으나, 이는 공문서(문서형)에 한정된 단일한 보존포맷으로, 기록물이 생산되는 환경과 기술의 발전으로 공문서도 다양한 보존포맷의 적용이 가능하며, 일반적인 공문서 이외의 전자기록물 유

형에 따라 각각 적합한 보존포맷이 필요하다. 따라서 파일포맷의 기술 변화에 대응하기 위해서는 특정 포맷을 지정하기보다, 환경이나 기술의 변화에도 적용이 가능한 전자기록물 파일의 보존포맷을 선정하는 기준을 마련하여, 각 공공기관이나 영구기록물관리기관이 자체적으로 적용하도록 하는 체계로의 전환이 필요하다.

또한, 행정환경에서 생산되는 문서유형 전자기록물에는 텍스트형 파일포맷(hwp, doc, odt, epub, txt 등)이 가장 많이 사용되고 있으나, 스프레드시트형과 프리젠테이션형 포맷들도 많이 활용되고 있다. 그러나 현행 PDF/A-1b 기반의 단일한 보존포맷으로는 텍스트형, 스프레드시트형, 프리젠테이션형 전자기록물의 내용, 맥락, 표현, 기능, 구조의 주요 속성들을 보존하는 것이 어려우므로, PDF/A-1b 일괄 변환 과정에서 '필수보존속성(Significant Properties)'들이 유실되는 문제를 해결하기 위해 적용가능한 전자기록물의 보존포맷을 확대하고 다양화할 필요가 있다.

4.2 보존포맷 선정기준 필요성

이 표준에서는 모든 유형의 전자기록물에 적용되는 공통기준과 문서유형 전자기록물의 고유한 특성을 보존할 수 있는 고유기준을 제공하고자 한다. 다시 말해, 이 표준에서는 보존포맷 선정을 위한 공통기준과 그 평가방법, 문서유형 전자기록물의 고유기준과 그 평가방법을 규정하고, 이러한 평가과정을 거쳐 보존포맷을 선정할 수 있는 방법을 제시한다.

보존포맷으로 선정될 수 있는 파일로 기록물이 생산된 경우 추후 보존포맷으로 다시 변환할 필요가 없으므로, 각급기관에서는 '보존포맷' 또는 최소한 '수용가능포맷'으로 기록물이 생산되도록 독려하고 안내할 필요가 있다.

5 보존포맷 선정을 위한 공통기준

5.1 일반사항

전자기록물을 장기보존하고 활용하기 위해서는 먼저 현재 포맷이 보존포맷으로 적합한지 검토하고, 적합하지 않으면 이 표준에서 제시하는 보존포맷 중 기관의 환경에 적합한 보존포맷으로 변환하여 전자기록물의 진본성, 이용가능성이 보호될 수 있도록 해야 한다.

5.2 공통기준의 구성

기록의 유형과 관계없이 보존포맷을 선정하는데 필요한 공통기준은 상위기준 4개와 세부기준 10개(예외적 기준 적용 시 11개)로 구성하며, 그 내용은 표 1과 같다.

5.2.1 공통기준의 정의

표 1 - 전자기록물 보존포맷 선정 공통기준의 정의

공통기준	내용
개방성	해당 포맷 및 구동 SW의 소유권에 대한 특정업체의 독점여부 및 해당 포맷의 '표준(standard)'과 '공개코드(Open Source)'의 존재여부를 판단하는 기준 · 세부기준 : 공개가용성, 공표
상호운용성	해당 포맷과 외부의 다양한 요인(HW, 구동SW, OS 등)과의 독립성, 호환성 및 변환가능성을 판단하는 기준 · 세부기준 : 독립성, 호환성, 변환가능성
채택	해당 포맷의 사용 범위를 판단하는 기준 · 세부기준 : 편재성, 편중성(예외적 적용)
기능성	해당 포맷이 지원하는 다양한 기능(오류감지, 검색기능 등)을 판단하는 기준 · 세부기준 : 자체 문서화(Self-Documentation), 기계가독성, 보호메커니즘, 검색기능

(1) 개방성

○ 공개가용성

- 해당 포맷 및 구동 SW가 특정 기업에 '독점'되어 있는지 여부
 - 해당 포맷을 구동시킬 수 있는 다른 SW 존재여부
 - 해당 포맷 사용에 대한 제한여부(라이선스, 구독, 특허료 등)
 - 기본 도구(메모장, 그림판 등) 사용을 통한 분석가능 여부

○ 공표

- 해당 포맷에 대한 '표준(Standard)'의 존재여부
 - 공개적인 참조 및 이용 여부
 - 체계적이고 권위 있는 기관에 의한 표준화 여부
- 해당 포맷의 '공개코드(Open Source)' 존재여부
 - 라이선스 형태(저작권 등)

(2) 상호운용성

○ 독립성

- (OS 관점) 해당 포맷을 구동할 때 특정 OS에 의해 받는 영향도
- (HW 관점) 해당 포맷을 구동할 때 HW에 의해 받는 영향도
- (SW 관점) 해당 포맷 또는 구동 SW가 특정 기술/표준/부가 SW에 의해 받는 영향도

○ 호환성

- 해당 포맷의 현재 구동 SW 지원 여부 및 이전/이후 구동 SW 버전과의 호환 가능 여부(동일한 SW에 한함)
 - ※ 동일한 SW : 같은 제조사, 계열사, 인수회사 등
- 구동하는 SW의 배포(Release) 주기에 따른 형식 및 사양의 업데이트 여부
- 해당 포맷의 버전 업데이트 개발 로드맵 또는 계획의 존재 여부

○ 변환가능성

- 해당 포맷이 정보 손실 없이 다른 포맷으로 변환 가능 여부 및 변환 가능한 포맷의 다양성
- 해당 포맷을 활용하기 위한 목적으로 SW, 서비스 및 도구 등에 의해서 다른 포맷으로 쉽게 변환되고 재사용할 수 있는지 여부

(3) 채택

○ 편재성(Ubiquity)

- 해당 포맷에 대한 수요와 공급의 법칙이 잘 확립되어 있고 광범위하게 사용되는 포맷인지의 여부
 - OS에서 별도의 응용 SW의 설치 없이 해당 포맷을 인식하고 내용 확인이

가능한지 여부

- 브라우저에서 별도의 확장 응용 SW의 설치 없이 해당 포맷을 인식하고 내용 확인이 가능한지 여부
- 해당 포맷이 표준화 단체에 의해 표준화 과정을 거쳐 저명한 컨소시엄과 그룹에 의해 채택되어 전 세계에서 사용하고 있는지 여부
- 해당 포맷이 시장을 선도하는 포맷인지 여부
- 해당 포맷을 제작/조작/렌더링(Rendering) 할 수 있는 많은 경쟁 제품의 존재 여부

○ 편중성(Local Prominence)

- 해당 포맷이 도서관, 기록원 및 기타 기록유산기관이 공식적으로 채택한 보존포맷인지의 여부
- 해당 포맷이 특정 전문 분야의 커뮤니티에서 채택되어 사용되고 있는 포맷인지 여부

(4) 기능성

○ 자체문서화(Self-Documentation)

- 해당 포맷의 자동 생성 메타데이터 기능 제공 여부
- 해당 포맷의 사용자 정의 메타데이터 기능 지원 여부
- 해당 포맷으로부터 메타데이터 추출 가능 여부

○ 기계가독성

- 해당 포맷이 데이터 분석, 인공지능 학습 등의 목적으로 컴퓨터가 처리할 수 있는 형태로 설계되었는지에 대한 여부를 판단하는 기준

○ 보호메커니즘

- 해당 포맷이 복사 방지, 디지털 서명, 인쇄 방지 및 콘텐츠 추출 보호와 같은 기술보호메커니즘이 적용되는 포맷인지의 여부
- 해당 포맷의 오류 감지, 수정 기능의 수용 여부
- 해당 포맷의 우발적인 손상에 대한 탄력성 여부

○ 검색기능

- 문서내용에 대한 유니코드(ISO/IEC 10646:2021) 텍스트 추출 가능 여부

5.2.2 공통기준의 가중치

보존포맷을 선정하기 위한 공통기준의 평가는 세부기준에 따라 평가요소별로 쌍대비교(Pairwise Comparison)를 통해 부여한 가중치를 적용하였다. 가중치 부여에는 AHP(Analytic Hierarchy Process) 기법이 사용되었는데, AHP

는 의사결정의 계층을 구성하는 요소들의 쌍대비교(Pairwise Comparison)를 통해 중요도를 산출하는 분석 방법으로 사회과학 분야에서 널리 활용되는 기법이다.

공통기준의 가중치 도출은 표준화작업반과 관련분야 전문가 설문 결과를 활용하였다. 표준을 적용하는 기관은 해당 기관의 정책적 판단에 따라 가중치나 배점을 달리 적용하여 사용할 수 있다.

표 2 - 전자기록물 보존포맷 선정 공통기준의 가중치

공통기준	세부기준	가중치(%)
개방성	공개가용성	14.3
	공표	9.8
상호운용성	독립성	10.9
	호환성	10.3
	변환가능성	11.2
채택	편재성, 편중성(예외적 적용)	10.2
기능성	자체 문서화(Self-Documentation)	6.3
	기계가독성	11.0
	보호메커니즘	7.7
	검색기능	8.3

5.3 공통기준에 따른 보존포맷 평가방법

5.1의 공통기준은 어떤 포맷이 보존포맷으로 적정한가를 평가하는 평가요소로 활용하며, Y/N로 형태로 응답하고 Y=1, N=0을 부여한다. 아래 ‘채택’ 공통기준의 ‘편재성’과 ‘편중성’은 서로 상충되는 개념을 가진 평가기준이므로 이 두 개의 기준은 분리하여 평가한다. 편재성은 기본적인 평가기준으로 하고, 편중성은 특별한 조건을 충족하는 경우에만 예외적으로 적용하는 것으로 한다.

표 3 - 전자기록물 보존포맷 공통기준 평가표

공통 기준	세부 기준	평가항목		배점	Y/N
개방성	공개 가용성	1	특정 기업 외 해당 포맷을 구동시킬 수 있는 다른 SW가 있는가?	2.1	

		2	해당 포맷 사용에 대한 제한여부(라이선스, 구독, 특허료 등)	무료 Read인가?	2.1	
		3		무료 Write인가?	2.1	
		4	기본 도구(메모장, 그림판 등) 사용을 통한 분석 가능 여부	기본 도구를 통해 해당 포맷을 구성하는 콘텐츠 전체를 해석할 수 있는가?	2	
		5		텍스트 콘텐츠가 표준 문자 인코딩(UTF-8, 유니코드, 아스키 코드 등)으로 되어 있는가?	2	
		6		압축되어 있는 경우 신뢰성 있는 압축(zip, gzip, lzw 등)으로 되어 있는가?	2	
		7		멀티미디어 콘텐츠가 공개 포맷(jpeg, gif, mpeg 등)으로 되어 있는가?	2	
		공표	8	해당 포맷의 '표준' 존재 여부	해당 포맷의 표준을 인터넷 등을 통해 공개적으로 참조 및 이용이 가능한가?	2.5
	9		해당 포맷의 표준을 인터넷 등을 통해 공개적으로 참조 및 이용할 때 무료인가?		2.5	
	10		체계적이고 권위있는 기관에 의해 표준화 과정을 거쳤는가?		2.4	
	11		해당 포맷이 오픈소스 라이선스인가?		2.4	
	상호 운용성	독립성	12	OS 관점	해당 포맷을 구동할 수 있는 OS의 개수가 다수인가?	2.8
13			HW 관점	해당 포맷을 특별한 HW없이 구동할 수 있는가?	2.7	
				해당 포맷을 개인용 컴퓨터 수준의 HW에서 구동할 수 있는가?	2.7	
15			특정 기술, 표준, 부가SW	해당 포맷 또는 구동 SW에 특수 코덱 및 특수 플레이어와 같은 특정 기술이나 부가 SW 등의 영향이 없는가?	2.7	
호환성		16	해당 포맷이 현재 구동 SW에서 지원하는가? (동일한 SW(같은 제조사, 계열사, 인수회사 등)에 한함)	2.6		
		17	해당 포맷이 이전/이후 구동 SW 버전과 호환이 가능한가? (동일한 SW(같은 제조사, 계열사, 인수회사 등)에 한함)	2.6		
		18	해당 포맷은 구동하는 SW의 배포(Release) 주기에 따라 형식이나 사양이 자주 업데이트되는가? (현재	2.6		

			가장 대표성 있는 구동 SW)				
		19	해당 포맷의 버전 업데이트 개발 로드맵 또는 계획이 존재하는가?		2.5		
	변환 가능성	20	보존, 추후 안정적인 마이그레이션 보장 가능성	해당 포맷이 정보의 손실없이 다른 포맷으로 변환 가능한가?	3.8		
		21		변환 가능한 포맷이 다양한가?	3.8		
	22	해당 포맷을 활용하기 쉬운 포맷으로 변환가능 여부	해당 포맷이 SW, 서비스 및 툴과 상호운용되어 새로운 목적으로 콘텐츠를 조작하고 재사용할 수 있는가?	3.6			
채택	편재성	23	OS에서 별도의 응용 SW 설치 없이 해당 포맷을 인식하고 내용을 확인할 수 있는가?		2.1		
		24	브라우저(Microsoft Edge, Internet Explorer, Chrome, Firefox 등)에서 별도의 확장 응용 SW 설치 없이 해당 포맷을 인식하고 내용을 확인할 수 있는가?		2.1		
		25	해당 포맷이 표준화 단체에 의해 표준화 과정을 거쳐 저명한 컨소시엄과 그룹에 의해 채택되어 전 세계에서 사용하는가?		2		
		26	해당 포맷이 시장을 선도하는가?		2		
		27	해당 포맷을 제작/조작/렌더링하는 많은 경쟁 제품의 존재하는가?		2		
기능성	자체 문서화	28	해당 포맷이 자동 생성 메타데이터 기능을 제공하는가?		2.1		
		29	해당 포맷이 사용자 지정 메타데이터 기능을 제공하는가?		2.1		
		30	해당 포맷으로부터 메타데이터를 추출할 수 있는 기능을 지원하는가?		2.1		
	기계 가독성	31	해당 포맷이 데이터 분석, 인공지능 학습 등의 목적으로 컴퓨터가 처리할 수 있는 형태(기계가독성을 위한 요소(태깅 등) 포함, 개방형 파일 포맷, 표준화된 파일 포맷 등)로 설계되었는가?		11		
		보호 메커니즘	32	해당 포맷이 암호 보호, 복사 방지, 디지털 서명, 인쇄 방지 및 콘텐츠 추출 보호와 같은 기술보호메커니즘이 적용되어 있지 않은가?		2.6	
			33	해당 포맷이 오류 감지, 수정 메커니즘을 수용하는가?		2.6	
		34	해당 포맷이 우발적인 손상에 대한 탄력성이 있는가?		2.5		
	검색 기능	35	해당 포맷이 이용자가 원하는 문서내용에 대한 검색 기능을 제공하는가?		8.3		
			합계(점수)			100	

표 4 - 전자기록물 보존포맷 공통기준 평가표(예외적 적용)

공통 기준	세부 기준	평가항목		배점	Y/N
채택	편중성	1	해당 포맷이 국립도서관, 기록원 및 기타 기록유산 기관이 공식적으로 채택한 보존포맷인가?		
		2	해당 포맷이 특정 전문 분야의 커뮤니티에서 채택되어 사용되고 있는가?		

※ 공통기준 '채택'의 '편중성'과 별도로 기관이 별도 배점 부여하여 평가 가능

6 보존포맷 선정을 위한 고유기준

6.1 일반사항

오랜 시간 접근할 수 있고 의미를 담고 있는 상태를 유지할 수 있도록 보존되어야 하는 전자기록물의 중요한 특성을 ‘필수보존속성(Significant Properties, 이하 SP)’이라 한다. 영국 국립기록원(TNA)의 InSPECT 프로젝트에서 제시된 SP는 ‘시간이 지나도 접근 가능하면서 의미있는 상태를 유지하기 위해서 보존되어야 하는 디지털 기록물의 속성’으로 정의된다.¹⁾

SP는 전자기록물에 내재되어 있는 고유의 속성으로 전자기록물의 각 유형별로 어떤 SP가 있는가를 분석하고, 이러한 SP를 보존한다면 기록의 4대 속성을 유지한 상태로 전자기록물을 보존함과 동시에 향후 유형별 전자기록물의 장기보존 전략 수립에도 활용할 수 있다. 각국 국가 아카이브 등에서 연구된 결과로 현재까지 공통적으로 도출된 SP는 다음과 같다.

표 5 - 필수보존속성(Significant Properties)의 5가지 범주

범주	의미
외관 (Appearance)	· 기록의 외형적인 모습 (예. 폰트, 색상)
기능 (Behavior)	· 기록과 연결되어있는 외부와의 상호작용에 의한 기능
내용 (Content)	· 기록 내 모든 데이터 및 수식
맥락 (Context)	· 기록의 메타데이터 (예. 작성자, 작성일)
구조 (Structure)	· 기록의 구조정보

문서유형 전자기록물이 가지는 고유의 속성을 분석하기 위해 MS office, 한컴 오피스, ODF 등의 표준을 분석한 결과 텍스트, 스프레드시트, 프리젠테이션형 전자문서에서 도출될 수 있는 SP는 다음과 같다.

1) Knight, Gareth. (2008). Framework for the definition of significant properties. The National Archives, InSPECT Project Document.

표 6 - 전자문서 유형별 필수보존속성(Significant Properties)

범주	텍스트	프리젠테이션	스프레드시트
외관	글꼴, 단락, 페이지, 개체 서식, 화면 보기	양식, 서식, 화면 보기	양식, 서식, 화면 보기, 데이터 정렬/필터
기능	참조, 변경내용 추적, 매크로	매크로, 애니메이션, 데이터 참조	수식, 데이터 유효성 검사, 데이터 참조
내용	개체 삽입	개체 삽입	개체 삽입
맥락	설명 메타데이터	설명 메타데이터	설명 메타데이터
구조	단락, 페이지, 참조	슬라이드 구조	시트 구조

6.2 고유기준의 구성

6.2.1 텍스트형

6.1의 SP에 해당하는 전자기록물의 고유기준을 각 유형별로 설명하면 아래 표와 같다.

표 7 - 텍스트형 보존포맷 선정 고유기준

고유기준	내용
글꼴	<ul style="list-style-type: none"> 보존포맷은 문서 가독성을 제고하는 모든 서식을 생산 당시의 모습으로 보존해야 한다는 기준 텍스트 문서를 구성하는 글자, 단락, 페이지 요소, 표, 이미지 등 각 개체는 이용자 가독성을 제고하는 다양한 서식을 활용하므로 서식은 보존되어야 함 글꼴의 경우 폰트 임베딩(Embedding)이 적용되어야 함
단락	
페이지	
개체	
화면 보기	<ul style="list-style-type: none"> 보존포맷은 이용자가 정의한 기준에 따라 배열된 데이터의 가시성을 잘 드러내며 보존해야 한다는 기준 텍스트 문서가 화면에 표시되는 정도를 추후에도 동일하게 보일 수 있도록 보존할 수 있어야 함
참조	<ul style="list-style-type: none"> 보존포맷은 문서와 연결된 외부 객체 및 데이터와의 관계를 드러내는 참조성을 보존해야 한다는 기준 텍스트 문서와 연결된 외부 객체와 데이터를 추후에 활용할 수 있도록 보존할 수 있어야 함.

변경내용 추적	<ul style="list-style-type: none"> · 보존포맷은 이용자가 추적한 문서의 변경내용을 드러내는 부분을 보존해야 한다는 기준 - 이용자가 텍스트 문서의 변경내용을 추적할 수 있도록 하는 기능을 사용하여 변경한 경우 추후에 사용할 수 있도록 보존할 수 있어야 함
매크로	<ul style="list-style-type: none"> · 보존포맷은 여러 트리거를 통해 유발되는 자동기능을 보존해야 한다는 기준 - 텍스트 문서에서의 단축키와 같은 자동기능을 추후에 사용할 수 있도록 보존할 수 있어야 함
개체 삽입	<ul style="list-style-type: none"> · 보존포맷은 문서가 내포한 콘텐츠의 이질성을 보존해야 한다는 기준 - 텍스트 문서는 다양한 콘텐츠를 내포할 수 있기 때문에, 보존포맷은 유형에 구애받지 않고 모든 콘텐츠를 보존할 수 있어야 함

6.2.2 프리젠테이션형

6.1의 SP에 해당하는 프리젠테이션형 전자기록물의 고유기준을 설명하면 아래 표와 같다.

표 8 - 프리젠테이션 보존포맷 선정 고유기준

고유기준	내용
양식 ²⁾	<ul style="list-style-type: none"> · 보존포맷은 전자문서의 템플릿을 구성하는 모든 요소(예, 배경 이미지, 글꼴, 색상, 레이아웃 등)를 보존해야 한다는 기준 - 프리젠테이션 문서에 적용된 템플릿은 다양한 요소로 구성되어 있으므로, 다른 기준보다 더욱 안정적으로 보존되어야 함
서식	<ul style="list-style-type: none"> · 보존포맷은 전자문서 가독성을 제고하는 모든 서식을 생산 당시의 모습으로 보존해야 한다는 기준 - 프리젠테이션 문서를 구성하는 슬라이드 내 텍스트, 도형, 이미지 등 개체는 이용자 가독성을 제고하는 다양한 서식을 활용하므로 서식은 보존되어야 함
화면 보기	<ul style="list-style-type: none"> · 보존포맷은 전자문서의 가시성을 확보하고자 생산 당시의 화면 비율로 보존해야 한다는 기준 - 프리젠테이션 문서에 대한 최소한의 가시성을 확보하려면 문서를 생산할 당시에 적용된 화면 비율로 보존되어야 함

2) 양식(template)은 배경 이미지, 글꼴, 색상, 레이아웃 등 전자문서를 구성하는 다양한 요소에 적용되는 광범위한 기준

애니메이션	<ul style="list-style-type: none"> • 보존포맷은 프리젠테이션 문서에서 사용된 애니메이션을 보존해야 한다는 기준 - 프리젠테이션 문서에 사용된 애니메이션 기능은 이용자에게 시각 효과를 제공하므로 추후 문서 활용 시 문서 이해도 제고에 도움을 줄 수 있기 때문에 보존되어야 함
매크로	<ul style="list-style-type: none"> • 보존포맷은 프리젠테이션 문서에서 사용된 매크로를 보존해야 한다는 기준 - 프리젠테이션 문서에서 사용된 매크로는 외부 객체의 연결 및 특수 함수의 수행 과정을 추후에 활용할 수 있도록 보존할 수 있어야 함
데이터 참조	<ul style="list-style-type: none"> • 보존포맷은 전자문서에 외부 객체와의 연결을 보존해야 한다는 기준 - 프리젠테이션과 연결된 외부 객체와 데이터를 추후에 활용할 수 있도록 연결을 보존할 수 있어야 함
개체 삽입	<ul style="list-style-type: none"> • 보존포맷은 전자문서가 내포한 이질적인 콘텐츠를 보존해야 한다는 기준 - 프리젠테이션은 다양하고 이질적인 콘텐츠를 내포할 수 있기 때문에, 보존포맷은 유형에 구애받지 않고 모든 콘텐츠를 보존할 수 있어야 함
슬라이드 구조	<ul style="list-style-type: none"> • 보존포맷은 전자문서가 생산 당시의 구조를 변형 없이 보존해야 한다는 기준 - 슬라이드, 레이아웃 등 프리젠테이션의 구조가 생산 당시의 모습으로 보존될 수 있어야 함

6.2.3 스프레드시트형

6.1의 SP에 해당하는 스프레드시트형 전자기록물의 고유기준을 설명하면 아래 표와 같다. 피벗 테이블, 함수와 같이 스프레드시트의 특징적인 요소를 보존함으로써 전자기록물의 진본성을 보장하는 장기보존을 위해 고유기준에 대한 평가가 필요하다.

표 9 - 스프레드시트형 보존포맷 선정 고유기준

고유기준	내용
양식	<ul style="list-style-type: none"> • 보존포맷은 전자문서의 템플릿을 구성하는 모든 요소(예. 배경 이미지, 글꼴, 색상, 레이아웃 등)를 보존해야 한다는 기준 - 스프레드시트 문서에 적용된 템플릿은 다양한 요소로 구성되어 있으므로, 다른 기준보다 더욱 안정적으로 보존되어야 함

서식	<ul style="list-style-type: none"> · 보존포맷은 전자문서 가독성을 제고하는 모든 서식을 생산 당시의 모습으로 보존해야 한다는 기준 - 스프레드시트 문서를 구성하는 시트 내 텍스트, 데이터, 도형, 이미지 등 각 개체는 이용자 가독성을 제고하는 다양한 서식을 활용하므로 서식은 보존되어야 함
화면 보기	<ul style="list-style-type: none"> · 보존포맷은 전자문서의 가시성을 확보하기 위해 생산 당시의 화면 비율로 보존해야 한다는 기준 - 스프레드시트 문서에 대한 최소한의 가시성을 확보하려면 문서를 생산할 당시에 적용된 화면 비율로 보존되어야 함
데이터 정렬/필터	<ul style="list-style-type: none"> · 보존포맷은 이용자가 정의한 기준에 따라 배열된 데이터의 모습을 보존해야 한다는 기준 - 스프레드시트 문서의 표시 기준, 데이터의 중요도와 표현 방식을 확보하기 위해서는 배열된 데이터의 모습을 보존해야 함
수식	<ul style="list-style-type: none"> · 보존포맷은 전자문서에 사용된 함수, 매크로 등 수식을 보존해야 한다는 기준 - 스프레드시트에 사용되는 다양한 함수와 매크로 등을 수용 및 보존할 수 있어야 함
데이터 유효성 검사	<ul style="list-style-type: none"> · 보존포맷은 전자문서에 입력 가능하며, 유효한 데이터 유형을 판별할 수 있어야 한다는 기준 - 사용자가 사전에 정의한 데이터 유형만을 스프레드시트에 입력할 수 있어야 하며, 정의된 데이터의 유형을 수용 및 보존해야 함
데이터 참조	<ul style="list-style-type: none"> · 보존포맷은 전자문서에 참조된 외부 객체와의 연결을 보존해야 한다는 기준 - 스프레드시트와 연결된 외부 객체와 데이터를 추후에 활용할 수 있도록 연결을 보존해야 한다는 기준
개체 삽입	<ul style="list-style-type: none"> · 보존포맷은 전자문서가 내포한 이질적인 콘텐츠를 보존해야 한다는 기준 - 스프레드시트는 다양하고 이질적인 콘텐츠를 내포할 수 있기 때문에, 보존포맷은 유형에 구애받지 않고 모든 콘텐츠를 보존할 수 있어야 함
시트 구조	<ul style="list-style-type: none"> · 보존포맷은 전자문서가 생산된 당시의 구조를 변형 없이 보존해야 한다는 기준 - 행, 열, 셀 등 스프레드시트의 구조가 생산 당시의 모습으로 보존될 수 있어야 함

6.2.4 보존포맷 선정 고유기준 가중치 부여

각 유형별 고유기준의 평가는 SP 각 평가요소별로 쌍대비교를 통해 가중치를 부여하였다. 고유기준의 가중치 도출은 표준화작업반과 관련분야 전문가 설문 결과를 활용하였다. 표준을 적용하는 기관은 문서 생산현황이나 문서작성의 특수성 등을 참고하여 가중치나 배점을 달리 적용하여 사용할 수 있다.

표 10 - 전자문서 유형 고유기준 가중치 및 중요도 순위

텍스트형		프리젠테이션형		스프레드시트형	
고유기준	가중치(%)	고유기준	가중치(%)	고유기준	가중치(%)
단락	20	개체 삽입	21	수식	19
개체 삽입	15	슬라이드구조	18	시트 구조	15
페이지	14	서식	16	양식	13
글꼴	13	양식	13	개체 삽입	12
개체서식	12	데이터 참조	10	데이터 정렬/필터	11
참조	9	애니메이션	10	서식	11
변경내용추적	7	매크로	8	데이터 유효성검사	9
매크로	6	화면 보기	4	데이터 참조	7
화면보기	4			화면 보기	3

위의 표 10과 같이 가중치를 부여한 결과 각 유형별 고유기준 중 중요도가 높은 요소는 각 전자문서 유형의 특징적인 구조, 기능(예. 텍스트형의 '단락 서식', 프리젠테이션형의 '슬라이드 구조', 스프레드시트형의 '수식', '시트 구조' 등)과 관련됨을 알 수 있다. 또한 공통적으로 '화면 보기' 기준이 가장 낮은 중요도를 보이고 있어 '화면 보기' 기준은 유사한 기준(예. 텍스트형의 '페이지 서식', 프리젠테이션형의 '슬라이드 구조', 스프레드시트형의 '시트 구조')에 포함시켰다.

6.3 고유기준에 따른 보존포맷 평가방법

6.3.1 텍스트형

텍스트형 기록물의 고유기준 8개 요소를 대상으로 평가문항과 각 평가문항별 배점을 부여하면 다음과 같다. 문항별 배점은 8개 각 요소의 가중치를 문항 수로 나누어 부여한다. 예를 들어, '단락의 중요도'는 20점이고 '단락의 평가문항'은 2개이므로, 각 문항마다 10점을 배점한다. 텍스트형 기록물의 고유기준을 대상으로 보존포맷을 평가하기 위한 평가표는 다음과 같다.

표 11 - 텍스트형 전자기록물의 보존포맷 평가표

고유기준	평가문항		배점	Y/N
글꼴	1	보존포맷은 텍스트 문서의 텍스트 서식(예. 폰트, 자간, 크기, 굵기, 기울임꼴, 글자색 등) 보존이 가능한가? ※ 글꼴의 경우 폰트 임베딩(Embedding)이 적용되어야 함	13	
단락	2	보존포맷은 텍스트 문서 단락 서식(예. 단락 정렬, 맞춤, 간격 등) 보존이 가능한가?	10	
	3	보존포맷은 텍스트 문서 단락목록(예. 글머리 기호) 보존이 가능한가?	10	
개체 서식	4	보존포맷은 텍스트 문서 개체 중 테이블(표) 서식(예. 테이블 크기, 색상, 간격, 테두리 등) 보존이 가능한가?	4	
	5	보존포맷은 텍스트 문서에 입력한 개체 중 그림 서식(예. 그림 크기, 색상, 테두리 등) 보존이 가능한가?	4	
	6	보존포맷은 텍스트 문서의 테이블 및 그림 외 개체 서식(예. 이미지, 도형, 수식 등의 크기·정렬·색상) 보존이 가능한가?	4	
참조	7	보존포맷은 텍스트 문서가 참조한 내·외부 개체와 연결(linkage) 보존이 가능한가?	9	
변경내용 추적	8	보존포맷은 텍스트 문서의 변경내용 추적 보존이 가능한가?	7	
매크로	9	보존포맷은 텍스트 문서에 사용된 매크로 보존이 가능한가?	6	

개체 삽입	10	보존포맷은 텍스트 문서에 입력한 개체 중 테이블(표) 보존이 가능한가?	5	
	11	보존포맷은 텍스트 문서에 입력한 개체 중 그림 보존이 가능한가?	5	
	12	보존포맷은 텍스트 문서의 입력한 개체 중 테이블 및 그림 외 개체 보존이 가능한가?	5	
페이지	13	보존포맷은 텍스트 문서의 페이지 요소(머리말, 꼬리말, 쪽 번호, 테두리 및 배경 등) 보존이 가능한가?	6	
	14	보존포맷은 텍스트 문서 페이지의 크기, 방향(landscape, portrait), 여백 등의 보존이 가능한가?	6	
	15	보존포맷은 텍스트 문서 생산 당시의 화면 비율로 보존이 가능한가?	6	
합계(점수)			100	

6.3.2 프리젠테이션형

프리젠테이션형 기록물의 고유기준 7개 요소를 대상으로 평가문항과 각 평가문항별 배점을 부여하면 다음과 같다.

표 12 - 프리젠테이션 전자기록물의 보존포맷 평가표

고유기준	평가문항		배점	Y/N
양식	1	보존포맷은 프리젠테이션 문서에 적용된 테마(예. 배경 이미지 서식, 글꼴, 색상 등) 보존이 가능한가?	13	
서식	2	보존포맷은 프리젠테이션 문서의 텍스트 서식(예. 폰트, 자간, 크기, 굵기, 기울임꼴, 색상 등) 보존이 가능한가?	4	
	3	보존포맷은 프리젠테이션 문서 텍스트 정렬(예. 정렬, 맞춤, 간격, 행간 등) 보존이 가능한가?	4	
	4	보존포맷은 프리젠테이션 문서 개체 서식(예. 크기, 정렬, 색상 등) 보존이 가능한가?	4	
	5	보존포맷은 프리젠테이션 문서 슬라이드 주석(예. 머리말, 꼬리말, 쪽 번호 등) 보존이 가능한가?	4	
애니메이션	6	보존포맷은 프리젠테이션 문서 내 개체에 적용된 애니메이션 보존이 가능한가?	5	
	7	보존포맷은 프리젠테이션 문서 내 슬라이드 전환시 적용된 애니메이션 보존이 가능한가?	5	
매크로	8	보존포맷은 프리젠테이션 문서가 사용한 매크로 보존이 가능한가?	8	

데이터 참조	9	보존포맷은 프리젠테이션 문서가 참조한 외부 개체와 연결(예. 하이퍼링크) 보존이 가능한가?	10	
개체 삽입	10	보존포맷은 프리젠테이션 문서가 내포한 정적 개체(예. 텍스트, 이미지, 도형 등) 보존이 가능한가?	11	
	11	보존포맷은 프리젠테이션 문서가 내포한 미디어 파일 보존이 가능한가?	11	
슬라이드 구조	12	보존포맷은 프리젠테이션 문서 생산 당시의 구조로(예. 슬라이드 크기·비율, 레이아웃 위치 등) 보존이 가능한가?	7	
	13	보존포맷은 프리젠테이션 문서에 적용된 슬라이드 마스터 보존이 가능한가?	7	
	14	보존포맷은 프리젠테이션 문서 생산 당시의 화면 비율로 보존이 가능한가?	7	
합계(점수)			100	

6.3.3 스프레드시트형

스프레드시트형 기록물의 고유기준 8개 요소를 대상으로 평가문항과 각 평가문항별 배점을 부여하면 다음과 같다.

표 13 - 스프레드시트 전자기록물의 보존포맷 평가표

고유기준	평가문항		배점	Y/N
양식	1	보존포맷은 스프레드시트 문서에 적용된 템플릿(예. 배경화면 서식, 글꼴, 색상 등) 보존이 가능한가?	13	
서식	2	보존포맷은 스프레드시트 문서의 텍스트 서식(예. 폰트, 자간, 크기, 굵기, 기울임꼴, 색상 등) 보존이 가능한가?	2.75	
	3	보존포맷은 스프레드시트 문서의 텍스트 정렬(예. 정렬, 맞춤, 간격, 행간 등) 보존이 가능한가?	2.75	
	4	보존포맷은 스프레드시트 문서의 개체 서식(예. 크기, 정렬, 색상 등) 보존이 가능한가?	2.75	
	5	보존포맷은 스프레드시트 문서 시트 주석(예. 머리말, 꼬리말, 쪽 번호 등) 보존이 가능한가?	2.75	
데이터 정렬/필터	6	보존포맷은 스프레드시트 문서의 데이터가 특정 기준으로 정렬 된 모습 그대로 보존이 가능한가?	5.5	
	7	보존포맷은 스프레드시트 문서의 데이터가 특정 요소로 필터링 된 모습 그대로 보존이 가능한가?	5.5	

수식	8	보존포맷은 스프레드시트 문서에 사용된 함수 보존이 가능한가?	9.5	
	9	보존포맷은 스프레드시트 문서에 사용된 매크로 보존이 가능한가?	9.5	
데이터 유효성검사	10	보존포맷은 스프레드시트 문서에 입력 가능한 특정 데이터 유형 보존이 가능한가?	9	
데이터 참조	11	보존포맷은 스프레드시트 문서가 참조한 외부 개체와 연결(linkage) 보존이 가능한가?	7	
개체 삽입	12	보존포맷은 스프레드시트 문서가 내포한 도표 보존이 가능한가?	4	
	13	보존포맷은 스프레드시트 문서가 생성한 피벗 테이블 보존이 가능한가?	4	
	14	보존포맷은 스프레드시트 문서가 내포한 정적 개체(도형, 텍스트 상자, 그림 등) 보존이 가능한가?	4	
시트 구조	15	보존포맷은 스프레드시트 문서 생산 당시의 구조로(행 높이, 열 너비 등) 보존이 가능한가?	4.5	
	16	보존포맷은 스프레드시트 문서 내 셀의 그룹화 보존이 가능한가?	4.5	
	17	보존포맷은 스프레드시트 문서 생산 당시의 화면 비율로 보존이 가능한가?	4.5	
	18	보존포맷은 스프레드시트 문서 내 특정 행, 열을 고정시킨 상태 보존이 가능한가?	4.5	
합계(점수)			100	

7 전자기록물의 보존포맷 선정 방법

7.1 권고포맷 선정방법

5장 공통기준, 6장 고유기준을 적용하여 파일포맷을 평가하고 그 결과에 따라 아래와 같이 권고포맷을 선정한다.

- 권고포맷 후보 : 권고포맷을 선정하기 위한 후보로 공통기준 평가점수가 특정점수 이상인 파일포맷
- 권고포맷 : 권고포맷 후보 중 보존포맷 선정 고유기준 평가점수가 특정점수 이상인 파일포맷. 보존포맷과 수용가능포맷을 포함한다.

7.2 보존포맷 선정방법

권고포맷은 아래와 같이 보존포맷, 수용가능포맷으로 구분³⁾한다.

- 보존포맷 : 권고포맷 중 ISO, ECMA 등⁴⁾에 의해 국제표준으로 승인된 파일포맷. 가장 높은 수준의 안정성을 보장할 수 있는 포맷
- 수용가능포맷 : 권고포맷 중 국제표준화 과정을 거치지 않은 포맷

보존포맷 선정을 위한 평가과정에서 공통기준과 고유기준을 적용하되 보존에 적합한 포맷(권고포맷)으로 선정될 수 있는 기준점수는 각 기관의 정책 및 평가 시점에 따라 달라질 수 있다.

기준점수를 상향하면 엄격한 품질을 갖춘 소수의 보존포맷 선정이 가능하게 되어 안정적 보존포맷 운영과 신속한 의사결정이 가능한 반면, 기준점수를 하향하면 선택의 폭이 확대되어, 보다 기술 변화에 수용적인 보존포맷 운영과 유연한 의사결정이 가능할 것이다.

또한, 평가점수를 등급화하여 보존이나 이관을 위한 다양한 포맷 정책에도

3) 미국 NARA 홈페이지 참고 : 'NARA Transfer Guidance'의 Transfer Guidance Format Tables('Appendix A: Tables of File Formats' - ① Textual Data, ② Presentation Formats, ③ Structured Data Formats...)

4) International Organization for Standardization(국제표준화기구), European Computer Manufacturers Association(국제표준화기구)

반영할 수 있다.

보기 최상위 구간은 보존포맷으로, 차상위 구간은 수용가능포맷으로 하고, 수용가능포맷까지를 이관가능한 포맷으로 허용하는 등의 정책 결정이 가능하다.

포맷에 관련된 정책 결정을 위해서는 새로운 포맷의 등장, 기존 포맷의 단종과 소멸, 포맷의 표준화 동향, 포맷 버전의 변화, 포맷관련 기술 발전 등을 모니터링하면서 적절한 결정을 내릴 수 있도록 주기적으로 정보를 수집하고 현행화하여야 한다. 또한, 포맷을 평가하여 이관이나 보존에 적합한 포맷을 적시에 선정할 수 있도록 하는 업무절차 정립이 필요하다.

국가기록원이 적용한 보존포맷 선정 평가 기준점수와 이에 따른 문서유형 파일포맷의 평가결과는 부속서를 참고한다.

부속서 (참고)

보존포맷 선정 평가 기준점수

1. 문서유형 보존포맷 선정 평가 기준점수

공통기준 평가결과 75점 이상을 획득하면 권고포맷 후보로 한다. 이러한 후보 포맷 중 고유기준을 평가*하여 4레벨로 구분 후 1레벨과 2레벨을 권고포맷으로 한다.

* (고유기준 평가방법) 1레벨(75점 이상), 2레벨(50~75점 미만), 3레벨(25~50점 미만), 4레벨(25점 미만)

보존포맷은 권고포맷에서 국제표준으로 승인된 포맷, 수용가능포맷은 국제표준화 과정을 거치지 않은 포맷으로 한다. 이러한 절차를 그림으로 나타내면 아래 그림 1과 같다.

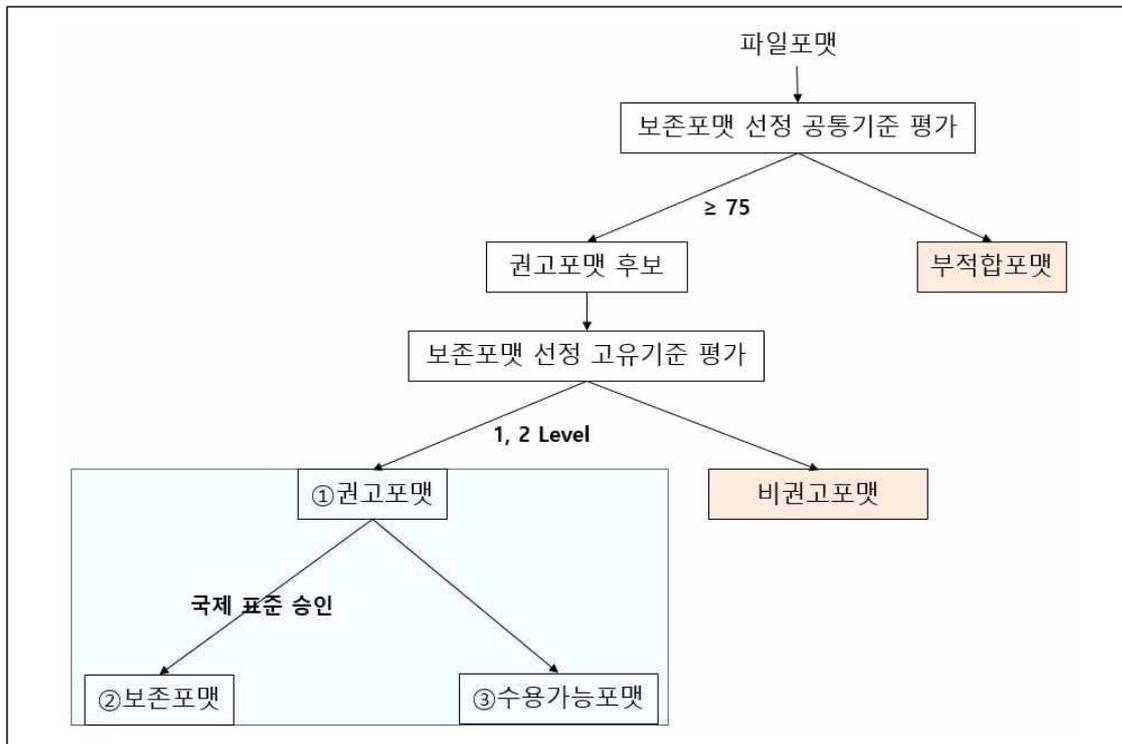


그림 1 - 권고포맷(보존포맷, 수용가능포맷) 선정 평가 예시

2. 문서유형 전자기록물 파일포맷 평가결과

2.1 텍스트형

공통기준 및 고유기준 평가를 포함하는 보존포맷 선정 방법에 따라 현재 텍스트형 포맷으로 가장 많이 사용되고 있는 HWP, HWPX, PDF, PDF/A-1⁵⁾, PDF/A-2⁶⁾, DOC, DOCX, TXT, ODT, EPUB 확장자를 가진 파일포맷을 평가한 결과는 아래 표 1과 같다. 공통기준 및 고유기준에 대한 평가점수는 표 2, 표 3을 각각 참조한다.

표 1 - 텍스트형 권고포맷

보존포맷	수용가능포맷
PDF, DOCX, PDF/A-1, PDF/A-2, ODT, EPUB	HWPX

5) PDF/A-1, Level A와 B, 그 수준을 따르는 포맷을 모두 포함

6) PDF/A-2, Level A와 B, 그 수준을 따르는 포맷을 모두 포함

표 2 - 텍스트형 보존포맷 공통기준 평가표

공통기준	세부기준	평가문항	배점	HWP	HWPX	PDF	PDF/A-1	PDF/A-2	DOC	DOCX	TXT	ODT	EPUB		
개방성	공개 가용성	1	특정 기업 외 해당 포맷을 구동시킬 수 있는 다른 SW가 있는가?	2.1	Y	Y	Y	Y	Y	Y	Y	Y	Y		
		2	해당 포맷 사용에 대한 제한여부(라이선스, 구독, 특허료 등)	무료 Read인가?	2.1	Y	Y	Y	Y	Y	Y	Y	Y	Y	
		3		무료 Write인가?	2.1	N	N	Y	Y	Y	N	N	Y	Y	Y
		4	기본 도구(메모장, 그림판 등) 사용을 통한 분석가능 여부	기본 도구를 통해 해당 포맷을 구성하는 콘텐츠 전체를 해석할 수 있는가?	2	N	N	N	N	N	N	Y	Y	Y	N
		5		텍스트 콘텐츠가 표준 문자 인코딩(UTF-8, 유니코드, 아스키 코드 등)으로 되어 있는가?	2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
		6		압축되어 있는 경우 표준 압축(zip, gzip, lzw 등)으로 되어 있는가?	2	N	Y	Y	N	N	N	Y	N	Y	Y
		7		멀티미디어 콘텐츠가 공개 포맷(jpeg, gif, mpeg 등)으로 되어 있는가?	2	N	Y	Y	N	Y	N	Y	N	Y	Y
	공표	8	해당 포맷의 '표준' 존재 여부	해당 포맷의 표준을 인터넷 등을 통해 공개적으로 참조 및 이용이 가능한가?	2.5	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
		9		해당 포맷의 표준을 인터넷 등을 통해 공개적으로 참조 및 이용할 때 무료인가?	2.5	Y	Y	N	N	N	Y	Y	Y	Y	Y
		10		체계적이고 권위 있는 기관에 의해 표준화 과정을 거쳤는가?	2.4	N	N	Y	Y	Y	N	Y	Y	Y	Y
		11		해당 포맷의 '공개코드' 존재 여부	해당 포맷이 오픈소스 라이선스인가?	2.4	N	N	Y	N	N	N	N	Y	Y
상호 운용성	독립성	12	OS 관점	해당 포맷을 구동할 수 있는 OS의 개수가 다수인가?	2.8	Y	Y	Y	Y	Y	Y	Y	Y		

공통기준	세부기준	평가문항		배점	HWP	HWPX	PDF	PDF/A-1	PDF/A-2	DOC	DOCX	TXT	ODT	EPUB	
		13	HW 관점	해당 포맷을 특별한 HW없이 구동할 수 있는가?	2.7	Y	Y	Y	Y	Y	Y	Y	Y	Y	
		14		해당 포맷을 개인용 컴퓨터 수준의 HW에서 구동할 수 있는가?	2.7	Y	Y	Y	Y	Y	Y	Y	Y	Y	
		15	특정 기술, 표준, 부가SW	해당 포맷 또는 구동 SW에 특수 코덱 및 특수 플레이어와 같은 특정 기술이나 부가 SW 등의 영향이 없는가?	2.7	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	호환성		16	해당 포맷이 현재 구동 SW에서 지원하는가? (동일한 SW(같은 제조사, 계열사, 인수회사 등)에 한함)		2.6	Y	Y	Y	Y	Y	Y	Y	Y	Y
			17	해당 포맷이 이전/이후 구동 SW 버전과 호환이 가능한가? (동일한 SW(같은 제조사, 계열사, 인수회사 등)에 한함)		2.6	Y	Y	Y	Y	Y	Y	Y	Y	Y
			18	해당 포맷은 구동하는 SW의 배포(Release) 주기에 따라 형식이나 사양이 자주 업데이트되는가? (현재 가장 대표성 있는 구동 SW)		2.6	Y	Y	Y	Y	Y	Y	Y	Y	Y
			19	해당 포맷의 버전 업데이트 개발 로드맵 또는 계획이 존재하는가?		2.5	Y	Y	Y	Y	Y	Y	Y	Y	Y
	변환 가능성		20	보존, 추후 안정적인 마이그레이션	해당 포맷이 정보의 손실없이 다른 포맷으로 변환 가능한가?	3.8	Y	Y	Y	Y	Y	Y	Y	Y	Y
			21	보장 가능성	변환 가능한 포맷이 다양한가?	3.8	Y	Y	Y	Y	Y	Y	Y	Y	Y
			22	해당 포맷을 활용하기 쉬운 포맷으로 변환가능 여부 (AIP → DIP)	해당 포맷이 SW, 서비스 및 툴과 상호 운용되어 새로운 목적으로 콘텐츠를 조작하고 재사용할 수 있는가?	3.6	Y	Y	Y	Y	Y	Y	Y	Y	Y
	채택	편재성	23	OS에서 별도의 응용 SW 설치 없이 해당 포맷을 인식하고 내용을 확인할 수 있는가?		2.1	N	Y	N	N	N	Y	Y	Y	N
			24	브라우저(Microsoft Edge, Internet Explorer, Chrome, Firefox 등)에서 별도의 확장 응용 SW 설치 없이 해당 포맷을 인식		2.1	N	N	Y	Y	Y	N	N	Y	N

공통기준	세부기준	평가문항	배점	HWP	HWPX	PDF	PDF/A-1	PDF/A-2	DOC	DOCX	TXT	ODT	EPUB		
		하고 내용을 확인할 수 있는가?													
	25	해당 포맷이 표준화 단체에 의해 표준화 과정을 거쳐 저명한 컨소시엄과 그룹에 의해 채택되어 전 세계에서 사용하는가?	2	N	N	Y	Y	Y	N	Y	Y	Y	Y		
	26	해당 포맷이 시장을 선도하는가?	2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		
	27	해당 포맷을 제작/조작/렌더링하는 많은 경쟁 제품의 존재하는가?	2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		
기능성	자체 문서화	28	해당 포맷이 자동 생성 메타데이터 기능을 제공하는가?	2.1	Y	Y	Y	Y	Y	Y	Y	Y	Y		
		29	해당 포맷이 사용자 지정 메타데이터 기능을 제공하는가?	2.1	N	N	Y	N	N	N	N	N	N	Y	
		30	해당 포맷으로부터 메타데이터를 추출할 수 있는 기능을 지원하는가?	2.1	N	N	Y	N	N	N	N	N	N	Y	
	기계 가독성	31	해당 포맷이 데이터 분석, 인공지능 학습 등의 목적으로 컴퓨터가 처리할 수 있는 형태(기계가독성을 위한 요소(태깅 등) 포함, 개방형 파일 포맷, 표준화된 파일 포맷 등)로 설계되었는가?	11	N	Y	Y	Y	Y	N	Y	N	Y	Y	
		보호 메커니즘	32	해당 포맷이 복사 방지, 디지털 서명, 인쇄 방지 및 콘텐츠 추출 보호와 같은 기술보호 메커니즘이 적용되어 있는가?	2.6	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
			33	해당 포맷이 오류 감지, 수정 메커니즘을 수용하는가?	2.6	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
	34		해당 포맷이 우발적인 손상에 대한 탄력성이 있는가?	2.5	Y	Y	Y	Y	Y	Y	Y	N	Y	N	
	검색기능	35	해당 포맷이 이용자가 원하는 문서내용에 대한 검색 기능을 제공하는가?	8.3	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
합계(점수)			100	65.7	82.8	91.3	82.8	84.8	65.7	89.3	73.1	93.7	90.3		

표 3 - 텍스트형 보존포맷 고유기준 평가점수

고유기준	평가문항	배점	HWPX	PDF	PDF/A-1	PDF/A-2	DOCX	ODT	EPUB
글꼴	1 보존포맷은 텍스트 문서의 텍스트 서식(예. 폰트, 자간, 크기, 굵기, 기울임꼴, 글자색 등) 보존이 가능한가? ※ 폰트의 경우 Embedding 되어 있어야 함	13	N	N	Y	Y	N	N	N
단락	2 보존포맷은 텍스트 문서 단락 서식(예. 단락 정렬·맞춤·간격 등) 보존이 가능한가?	9.5	Y	Y	Y	Y	Y	Y	Y
	3 보존포맷은 텍스트 문서 단락목록(예. 글머리 기호) 보존이 가능한가?	9.5	Y	Y	Y	Y	Y	Y	Y
개체 서식	4 보존포맷은 텍스트 문서 개체 중 테이블(표) 서식(예. 테이블 크기, 색상, 간격, 테두리 등) 보존이 가능한가?	4.3	Y	Y	Y	Y	Y	Y	Y
	5 보존포맷은 텍스트 문서에 입력한 개체 중 그림 서식(예. 그림 크기, 색상, 테두리 등) 보존이 가능한가?	4.3	Y	Y	Y	Y	Y	Y	Y
	6 보존포맷은 텍스트 문서의 테이블 및 그림 외 개체 서식(예. 이미지, 도형, 수식 등의 크기·정렬·색상) 보존이 가능한가?	4.3	Y	Y	Y	Y	Y	Y	Y
참조	7 보존포맷은 텍스트 문서가 참조한 내·외부 개체와 연결(linkage) 보존이 가능한가?	9	Y	Y	Y	Y	Y	Y	Y
변경내용 추적	8 보존포맷은 텍스트 문서의 변경내용 추적 보존이 가능한가?	7	Y	N	N	N	Y	Y	N
매크로	9 보존포맷은 텍스트 문서에 사용된 매크로 보존이 가능한가?	6	Y	N	N	N	Y	Y	N

고유기준	평가문항	배점	HWPX	PDF	PDF/A-1	PDF/A-2	DOCX	ODT	EPUB
개체 삽입	10	보존포맷은 텍스트 문서에 입력한 개체 중 테이블(표) 보존이 가능한가?	5	Y	Y	Y	Y	Y	Y
	11	보존포맷은 텍스트 문서에 입력한 개체 중 그림 보존이 가능한가?	5	Y	Y	Y	Y	Y	Y
	12	보존포맷은 텍스트 문서의 입력한 개체 중 테이블 및 그림 외 개체 보존이 가능한가?	5	Y	Y	Y	Y	N	Y
페이지	13	보존포맷은 텍스트 문서의 페이지 요소(머리말, 꼬리말, 쪽 번호, 테두리 및 배경 등) 보존이 가능한가?	6	Y	Y	Y	Y	Y	Y
	14	보존포맷은 텍스트 문서 페이지의 크기, 방향(landscape, portrait), 여백 등의 보존이 가능한가?	6	Y	Y	Y	Y	Y	Y
	15	보존포맷은 텍스트 문서 생산 당시의 화면 비율로 보존이 가능한가?	6	Y	Y	Y	Y	Y	Y
합계(점수)		100	87	74	87	87	87	82	74

2.2 프리젠테이션형

공통기준 및 고유기준 평가를 포함하는 보존포맷 선정 방법에 따라 현재 프리젠테이션형 포맷으로 가장 많이 사용되고 있는 PPT, PPTX, SHOW, PPS, PPTM, ODP, PDF, PDF/A-1, PDF/A-2 확장자를 가진 파일포맷을 평가한 결과는 아래 표 4와 같다. 공통기준 및 고유기준에 대한 평가점수는 표 5, 표 6을 각각 참조한다.

표 4 - 프리젠테이션형 권고포맷

보존포맷	수용가능포맷
PDF, PDF/A-1, PDF/A-2, PPTX, ODP	-

표 5 - 프리젠테이션형 보존포맷 공통기준 평가점수

공통기준	세부기준	평가문항	배점	PPT	PPTX	SHOW	PPS	PPTM	ODP	PDF	PDF/A-1	PDF/A-2		
개방성	공개 가용성	1	특정 기업 외 해당 포맷을 구동시킬 수 있는 다른 SW가 있는가?	2.1	Y	Y	N	Y	N	Y	Y	Y	Y	
		2	해당 포맷 사용에 대한 제한여부(라이	무료 Read인가?	2.1	Y	Y	Y	Y	Y	Y	Y	Y	Y
		3	센스, 구독, 특허로 등)	무료 Write인가?	2.1	N	N	N	N	N	Y	Y	Y	Y
		4	기본 도구(메모장, 그림판 등) 사용을 통한 분석가능 여부	기본 도구를 통해 해당 포맷을 구성하는 콘텐츠 전체를 해석할 수 있는가?	2	N	Y	Y	N	Y	Y	N	N	N
		5		텍스트 콘텐츠가 표준 문자 인코딩 (UTF-8, 유니코드, 아스키 코드 등)으로 되어 있는가?	2	Y	Y	Y	Y	Y	Y	Y	Y	Y
		6		압축되어 있는 경우 표준 압축(zip, gzip, lzw 등)으로 되어 있는가?	2	N	Y	Y	N	Y	Y	Y	N	N
		7		멀티미디어 콘텐츠가 공개 포맷(jpeg, gif, mpeg 등)으로 되어 있는가?	2	N	Y	Y	N	Y	Y	Y	N	Y
	공표	8	해당 포맷의 표준을 인터넷 등을 통해 공개적으로 참조 및 이용이 가능한가?	2.5	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
		9	해당 포맷의 '표준' 존재 여부	해당 포맷의 표준을 인터넷 등을 통해 공개적으로 참조 및 이용할 때 무료인가?	2.5	Y	Y	N	Y	Y	Y	N	N	N
		10	체계적이고 권위있는 기관에 의해 표준화 과정을 거쳤는가?	2.4	N	Y	N	N	Y	Y	Y	Y	Y	Y
		11	해당 포맷의 '공개 코드' 존재 여부	해당 포맷이 오픈소스 라이선스인가?	2.4	N	N	N	N	N	Y	Y	N	N

공통기준	세부기준	평가문항		배점	PPT	PPTX	SHOW	PPS	PPTM	ODP	PDF	PDF/A-1	PDF/A-2	
상호 운용성	독립성	12	OS 관점	해당 포맷을 구동할 수 있는 OS의 개수가 다수인가?	2.8	Y	Y	Y	Y	Y	Y	Y	Y	
		13	HW 관점	해당 포맷을 특별한 HW없이 구동할 수 있는가?	2.7	Y	Y	Y	Y	Y	Y	Y	Y	
		14		해당 포맷을 개인용 컴퓨터 수준의 HW에서 구동할 수 있는가?	2.7	Y	Y	Y	Y	Y	Y	Y	Y	
		15	특정 기술, 표준, 부가SW	해당 포맷 또는 구동 SW에 특수 코덱 및 특수 플레이어와 같은 특정 기술이나 부가 SW 등의 영향이 없는가?	2.7	Y	Y	Y	Y	Y	Y	Y	Y	
	호환성	16	해당 포맷이 현재 구동 SW에서 지원하는가? (동일한 SW(같은 제조사, 계열사, 인수회사 등)에 한함)		2.6	Y	Y	Y	Y	Y	Y	Y	Y	
		17	해당 포맷이 이전/이후 구동 SW 버전과 호환이 가능한가? (동일한 SW(같은 제조사, 계열사, 인수회사 등)에 한함)		2.6	Y	Y	Y	Y	Y	Y	Y	Y	
		18	해당 포맷은 구동하는 SW의 Release 주기(공개 주기)에 따라 형식이나 사양이 자주 업데이트되는가? (현재 가장 대표성 있는 구동 SW)		2.6	Y	Y	Y	Y	Y	Y	Y	Y	
		19	해당 포맷의 버전 업데이트 개발 로드맵 또는 계획이 존재하는가?		2.5	Y	Y	Y	Y	Y	Y	Y	Y	
	변환 가능성	20	보존, 추후 안정적인 마이그레이션	해당 포맷이 정보의 손실없이 다른 포맷으로 변환 가능한가?	3.8	Y	Y	Y	Y	N	Y	Y	Y	Y
		21	보장 가능성	변환 가능한 포맷이 다양한가?	3.8	Y	Y	Y	Y	N	Y	Y	Y	Y
		22	해당 포맷을 활용하기 쉬운 포맷으로 변환가능 여부 (AIP → DIP)	해당 포맷이 SW, 서비스 및 툴과 상호 운용되어 새로운 목적으로 콘텐츠를 조작하고 재사용할 수 있는가?	3.6	Y	Y	Y	Y	N	Y	Y	Y	Y
	채택	편재성	23	OS에서 별도의 응용 SW 설치 없이 해당 포맷을 인식하		2.1	N	N	N	N	N	N	N	N

공통기준	세부기준	평가문항	배점	PPT	PPTX	SHOW	PPS	PPTM	ODP	PDF	PDF/A-1	PDF/A-2	
		고 내용을 확인할 수 있는가?											
	24	브라우저(Microsoft Edge, Internet Explorer, Chrome, Firefox 등)에서 별도의 확장 응용 SW 설치 없이 해당 포맷을 인식하고 내용을 확인할 수 있는가?	2.1	N	N	N	N	N	N	Y	Y	Y	
	25	해당 포맷이 표준화 단체에 의해 표준화 과정을 거쳐 저명한 컨소시엄과 그룹에 의해 채택되어 전 세계에서 사용하는가?	2	N	Y	N	N	Y	Y	Y	Y	Y	
	26	해당 포맷이 시장을 선도하는가?	2	Y	Y	N	N	N	Y	Y	Y	Y	
	27	해당 포맷을 제작/조작/렌더링하는 많은 경쟁 제품의 존재하는가?	2	Y	Y	N	Y	N	Y	Y	Y	Y	
기능성	자체 문서화	28	해당 포맷이 자동 생성 메타데이터 기능을 제공하는가?	2.1	Y	Y	Y	Y	Y	Y	Y	Y	
		29	해당 포맷이 사용자 지정 메타데이터 기능을 제공하는가?	2.1	N	N	N	N	N	N	Y	N	N
		30	해당 포맷으로부터 메타데이터를 추출할 수 있는 기능을 지원하는가?	2.1	N	N	N	N	N	N	Y	N	N
	기계 가독성	31	해당 포맷이 데이터 분석, 인공지능 학습 등의 목적으로 컴퓨터가 처리할 수 있는 형태(기계가독성을 위한 요소(태깅 등) 포함, 개방형 파일 포맷, 표준화된 파일 포맷 등)로 설계되었는가?	11	N	Y	N	Y	Y	Y	Y	Y	Y
		보호 메커니즘	32	해당 포맷이 복사 방지, 디지털 서명, 인쇄 방지 및 콘텐츠 추출 보호와 같은 기술보호 메커니즘이 적용되어 있지 않은가?	2.6	Y	Y	Y	Y	Y	Y	Y	Y
	33		해당 포맷이 오류 감지, 수정 메커니즘을 수용하는가?	2.6	Y	Y	Y	Y	Y	Y	Y	Y	Y
	34		해당 포맷이 우발적인 손상에 대한 탄력성이 있는가?	2.5	Y	Y	Y	Y	Y	Y	Y	Y	Y
검색기능	35	해당 포맷이 이용자가 원하는 문서내용에 대한 검색 기능을 제공하는가?	8.3	Y	Y	Y	Y	Y	Y	Y	Y		
합계(점수)			100	65.7	87.1	60.6	74.7	69.8	91.6	93.4	82.8	84.8	

표 6 - 프리젠테이션형 보존포맷 고유기준 평가점수

고유기준	평가문항		배점	PPTX	ODP	PDF	PDF/A-1	PDF/A-2
양식	1	보존포맷은 프리젠테이션 문서에 적용된 테마(예. 배경 이미지 서식, 글꼴, 색상 등) 보존이 가능한가?	13	Y	Y	Y	Y	Y
서식	2	보존포맷은 프리젠테이션 문서의 텍스트 서식(예. 폰트, 자간, 크기, 굵기, 기울임꼴, 색상 등) 보존이 가능한가?	4	Y	Y	Y	Y	Y
	3	보존포맷은 프리젠테이션 문서 텍스트 정렬(예. 정렬, 맞춤, 간격, 행간 등) 보존이 가능한가?	4	Y	Y	Y	Y	Y
	4	보존포맷은 프리젠테이션 문서 개체 서식(예. 크기, 정렬, 색상 등) 보존이 가능한가?	4	Y	Y	Y	Y	Y
	5	보존포맷은 프리젠테이션 문서 슬라이드 주석(예. 머리말, 꼬리말, 쪽 번호 등) 보존이 가능한가?	4	Y	Y	Y	Y	Y
애니메이션	6	보존포맷은 프리젠테이션 문서 내 개체에 적용된 애니메이션 보존이 가능한가?	5	Y	Y	N	N	N
	7	보존포맷은 프리젠테이션 문서 내 슬라이드 전환 시 적용된 애니메이션 보존이 가능한가?	5	Y	Y	N	N	N
매크로	8	보존포맷은 프리젠테이션 문서가 사용한 매크로 보존이 가능한가?	8	N	Y	N	N	N
데이터 참조	9	보존포맷은 프리젠테이션 문서가 참조한 외부 개체와 연결(예. 하이퍼링크) 보존이 가능한가?	10	Y	Y	Y	Y	Y
개체 삽입	10	보존포맷은 프리젠테이션 문서가 내포한 정적 개체(예. 텍스트, 이미지, 도형 등) 보존이 가능한가?	11	Y	Y	Y	Y	Y
	11	보존포맷은 프리젠테이션 문서가 내포한 미디어 파일 보존이 가능한가?	11	Y	Y	Y	Y	Y

고유기준	평가문항		배점	PPTX	ODP	PDF	PDF/A-1	PDF/A-2
슬라이드 구조	12	보존포맷은 프리젠테이션 문서 생산 당시의 구조로(예. 슬라이드 크기·비율, 레이아웃 위치 등) 보존이 가능한가?	7	Y	Y	Y	Y	Y
	13	보존포맷은 프리젠테이션 문서에 적용된 슬라이드 마스터 보존이 가능한가?	7	Y	Y	Y	Y	Y
	14	보존포맷은 프리젠테이션 문서 생산 당시의 화면 비율로 보존이 가능한가?	7	Y	Y	Y	Y	Y
합계(점수)			100	92	100	71	71	71

2.3 스프레드시트형

공통기준 및 고유기준 평가를 포함하는 보존포맷 선정 방법에 따라 현재 스프레드시트형 포맷으로 가장 많이 사용되고 있는 XLS, XLSX, CELL, CSV, XLK, ODS 확장자를 가진 파일포맷을 평가한 결과는 아래 표 7과 같다. 공통기준 및 고유기준에 대한 평가점수는 표 8, 표 9를 각각 참조한다.

표 7 - 스프레드시트형 권고포맷

보존포맷	수용가능포맷
XLSX, ODS	-

표 8 - 스프레드시트형 보존포맷 공통기준 평가점수

공통기준	세부기준	평가문항		배점	XLS	XLSX	CELL	CSV	XLK	ODS	
개방성	공개 가용성	1	특정 기업 외 해당 포맷을 구동시킬 수 있는 다른 SW가 있는가?	2.1	Y	Y	N	Y	Y	Y	
		2	해당 포맷 사용에 대한 제한여부(라이선스, 구독, 특허료 등)	무료 Read인가?	2.1	Y	Y	Y	Y	Y	Y
		3		무료 Write인가?	2.1	N	N	N	Y	N	Y
		4	기본 도구(메모장, 그림판 등) 사용을 통한 분석가능 여부	기본 도구를 통해 해당 포맷을 구성하는 콘텐츠 전체를 해석할 수 있는가?	2	N	Y	Y	Y	Y	Y
		5		텍스트 콘텐츠가 표준 문자 인코딩(UTF-8, 유니코드, 아스키 코드 등)으로 되어 있는가?	2	Y	Y	Y	Y	Y	Y
		6		압축되어 있는 경우 표준 압축(zip, gzip, lzw 등)으로 되어 있는가?	2	N	Y	Y	N	Y	Y
		7		멀티미디어 콘텐츠가 공개 포맷(jpeg, gif, mpeg 등)으로 되어 있는가?	2	Y	Y	Y	N	Y	Y
	공표	8	해당 포맷의 '표준' 존재 여부	해당 포맷의 표준을 인터넷 등을 통해 공개적으로 참조 및 이용이 가능한가?	2.5	Y	Y	N	Y	Y	Y
		9		해당 포맷의 표준을 인터넷 등을 통해 공개적으로 참조 및 이용할 때 무료인가?	2.5	Y	Y	N	Y	Y	Y
		10		체계적이고 권위있는 기관에 의해 표준화 과정을 거치는가?	2.4	N	Y	N	Y	N	Y
		11		해당 포맷의 '공개코드' 존재 여부	해당 포맷이 오픈소스 라이선스인가?	2.4	N	N	N	Y	N
상호	독립성	12	OS 관점	해당 포맷을 구동할 수 있는 OS의 개수가 다수인가?	2.8	Y	Y	Y	Y	Y	

공통기준	세부기준	평가문항		배점	XLS	XLSX	CELL	CSV	XLK	ODS	
운용성	HW 관점	13	해당 포맷을 특별한 HW없이 구동할 수 있는가?	2.7	Y	Y	Y	Y	Y	Y	
		14	해당 포맷을 개인용 컴퓨터 수준의 HW에서 구동할 수 있는가?	2.7	Y	Y	Y	Y	Y	Y	
		15	특정 기술, 표준, 부가SW 해당 포맷 또는 구동 SW에 특수 코덱 및 특수 플레 이어와 같은 특정 기술이나 부가 SW 등의 영향이 없는가?	2.7	Y	Y	Y	Y	Y	Y	
	호환성	16	해당 포맷이 현재 구동 SW에서 지원하는가? (동일한 SW(같은 제조사, 계열사, 인수회사 등)에 한함)		2.6	Y	Y	Y	Y	Y	Y
		17	해당 포맷이 이전/이후 구동 SW 버전과 호환이 가능한가? (동일한 SW(같은 제조사, 계열사, 인수회사 등)에 한함)		2.6	Y	Y	Y	Y	Y	Y
		18	해당 포맷은 구동하는 SW의 배포(Release) 주기에 따라 형식이나 사양이 자주 업데이트되는가? (현재 가장 대표성 있는 구동 SW)		2.6	Y	Y	Y	Y	Y	Y
		19	해당 포맷의 버전 업데이트 개발 로드맵 또는 계획이 존재하는가?		2.5	Y	Y	Y	Y	Y	Y
	변환 가능성	20	해당 포맷이 정보의 손실없이 다른 포맷으로 변환 가능한가?	3.8	Y	Y	Y	Y	Y	N	Y
		21	보존, 추후 안정적인 마이그레이션 보장 가능성 변환 가능한 포맷이 다양한가?	3.8	Y	Y	Y	Y	Y	N	Y
		22	해당 포맷을 활용하기 쉬운 포맷으로 변환가능 여부 (AIP → DIP) 해당 포맷이 SW, 서비스 및 툴과 상호운용되어 새로운 목적으로 콘텐츠를 조작하고 재사용할 수 있는가?	3.6	Y	Y	Y	Y	Y	N	Y
	채택	편재성	23	OS에서 별도의 응용 SW 설치 없이 해당 포맷을 인식하고 내용을 확인할 수 있는가?	2.1	N	Y	Y	Y	Y	Y
24			브라우저(Microsoft Edge, Internet Explorer, Chrome, Firefox 등)에서 별도의 확장	2.1	N	N	N	N	Y	N	

공통기준	세부기준	평가문항	배점	XLS	XLSX	CELL	CSV	XLK	ODS
		응용 SW 설치 없이 해당 포맷을 인식하고 내용을 확인할 수 있는가?							
		25 해당 포맷이 표준화 단체에 의해 표준화 과정을 거쳐 저명한 컨소시엄과 그룹에 의해 채택되어 전 세계에서 사용하는가?	2	N	Y	N	Y	N	Y
		26 해당 포맷이 시장을 선도하는가?	2	Y	Y	N	Y	N	Y
		27 해당 포맷을 제작/조작/렌더링하는 많은 경쟁 제품의 존재하는가?	2	Y	Y	N	Y	Y	Y
기능성	자체 문서화	28 해당 포맷이 자동 생성 메타데이터 기능을 제공하는가?	2.1	Y	Y	Y	Y	Y	Y
		29 해당 포맷이 사용자 지정 메타데이터 기능을 제공하는가?	2.1	N	N	N	N	N	N
		30 해당 포맷으로부터 메타데이터를 추출할 수 있는 기능을 지원하는가?	2.1	N	N	N	N	N	N
	기계 가독성	31 해당 포맷이 데이터 분석, 인공지능 학습 등의 목적으로 컴퓨터가 처리할 수 있는 형태(기계가독성을 위한 요소(태깅 등) 포함, 개방형 파일 포맷, 표준화된 파일 포맷 등)로 설계되었는가?	11	N	Y	N	Y	N	Y
		보호 메커니즘	32 해당 포맷이 복사 방지, 디지털 서명, 인쇄 방지 및 콘텐츠 추출 보호와 같은 기술보호메커니즘이 적용되어 있지 않은가?	2.6	Y	Y	Y	Y	Y
	33 해당 포맷이 오류 감지, 수정 메커니즘을 수용하는가?		2.6	Y	Y	Y	Y	Y	Y
	34 해당 포맷이 우발적인 손상에 대한 탄력성이 있는가?		2.5	Y	Y	Y	Y	Y	Y
검색기능	35 해당 포맷이 이용자가 원하는 문서내용에 대한 검색 기능을 제공하는가?	8.3	Y	Y	Y	Y	Y	Y	
합계(점수)			100	67.7	89.2	62.8	89.7	62.5	93.8

표 9 - 스프레드시트형 보존포맷 고유기준 평가점수

고유기준	평가문항		배점	XLSX	CSV	ODS
양식	1	보존포맷은 스프레드시트 문서에 적용된 템플릿(예. 배경화면 서식, 글꼴, 색상 등) 보존이 가능한가?	13	Y	N	Y
서식	2	보존포맷은 스프레드시트 문서의 텍스트 서식(예. 폰트, 자간, 크기, 굵기, 기울임꼴, 색상 등) 보존이 가능한가?	2.75	Y	N	Y
	3	보존포맷은 스프레드시트 문서의 텍스트 정렬(예. 정렬, 맞춤, 간격, 행간 등) 보존이 가능한가?	2.75	Y	N	Y
	4	보존포맷은 스프레드시트 문서의 개체 서식(예. 크기, 정렬, 색상 등) 보존이 가능한가?	2.75	Y	N	Y
	5	보존포맷은 스프레드시트 문서 시트 주석(예. 머리말, 꼬리말, 쪽 번호 등) 보존이 가능한가?	2.75	Y	N	Y
데이터 정렬/필터	6	보존포맷은 스프레드시트 문서의 데이터가 특정 기준으로 정렬 된 모습 그대로 보존이 가능한가?	5.5	Y	Y	Y
	7	보존포맷은 스프레드시트 문서의 데이터가 특정 요소로 필터링 된 모습 그대로 보존이 가능한가?	5.5	Y	N	Y
수식	8	보존포맷은 스프레드시트 문서에 사용된 함수 보존이 가능한가?	9.5	Y	N	Y
	9	보존포맷은 스프레드시트 문서에 사용된 매크로 보존이 가능한가?	9.5	Y	N	Y
데이터 유효성 검사	10	보존포맷은 스프레드시트 문서에 입력 가능한 특정 데이터 유형 보존이 가능한가?	9	Y	N	Y
데이터 참조	11	보존포맷은 스프레드시트 문서가 참조한 외부 개체와 연결(linkage) 보존이 가능한가?	7	Y	N	Y

고유기준	평가문항		배점	XLSX	CSV	ODS
개체 삽입	12	보존포맷은 스프레드시트 문서가 내포한 도표 보존이 가능한가?	4	Y	N	Y
	13	보존포맷은 스프레드시트 문서가 생성한 피벗 테이블 보존이 가능한가?	4	Y	N	Y
	14	보존포맷은 스프레드시트 문서가 내포한 정적 개체(도형, 텍스트 상자, 그림 등) 보존이 가능한가?	4	Y	N	Y
시트 구조	15	보존포맷은 스프레드시트 문서 생산 당시의 구조로(행 높이, 열 너비 등) 보존이 가능한가?	4.5	Y	N	Y
	16	보존포맷은 스프레드시트 문서 내 셀의 그룹화 보존이 가능한가?	4.5	Y	N	Y
	17	보존포맷은 스프레드시트 문서 생산 당시의 화면 비율로 보존이 가능한가?	4.5	Y	N	Y
	18	보존포맷은 스프레드시트 문서 내 특정 행, 열을 고정시킨 상태 보존이 가능한가?	4.5	Y	N	Y
합계(점수)			100	100	5.5	100

참고문헌

- [1] 국가기록원, 전북대학교 산학협력단. 2019년 11월. 「데이터세트 유형 전자기록의 장기보존기술 연구」.
- [2] 국가기록원, 전북대학교 산학협력단. 2020년 11월. 「문서유형 보존포맷 및 장기보존패키지 다양화 연구」.