

ISSN 2765-3005

국립중앙도서관 이슈페이퍼

# ISSUE PAPER

뉴미디어 시대의  
시청각 자료 디지털 보존 전략

Vol. 5

2021. 8



문화체육관광부  
국립중앙도서관

국립중앙도서관 이슈페이퍼 제5호

발행일 2021년 8월 30일

발행처 국립중앙도서관

발행인 국립중앙도서관장 서혜란

주소 서울특별시 서초구 반포대로 201

전화 02-590-0799

팩스 02-590-0546

누리집 <http://www.nl.go.kr>

ISSN 2765-3005

- 본지에 실린 글의 내용은 집필자의 개인적인 견해이며, 국립중앙도서관의 공식적인 의견과 다를 수 있습니다.
- 본지의 저작권은 국립중앙도서관에 있으며, 사전 허락없이 무단으로 복제·변경·배포할 수 없습니다.

## 뉴미디어 시대의 시청각 자료 디지털 보존 전략

박소연, 이연수 학예연구사 국립중앙도서관 자료보존연구센터

### I. 뉴미디어 시대의 미디어 시장 환경 변화 02

### II. 도서관 시청각 자료 현황 03

1. 시청각 자료 유형 및 수집 현황
2. 시청각 자료 보존처리 및 디지털화

### III. 기술메타데이터 연구 개발 및 보존 전략 06

1. 시청각 자료 기술메타데이터 선정
2. 장기보존패키지의 메타데이터 구성
3. 시청각 기술메타데이터 구조 분석
4. 기술메타데이터 자동 추출 도구 개발
5. 장기보존패키지 적용 및 보존 전략

### IV. 맺음말 13

### 참고문헌 14

#### | 주요 키워드 |

시청각 자료, 장기보존, 기술메타데이터

# 뉴미디어 시대의 시청각 자료 디지털 보존 전략

박소연, 이연수 학예연구사 국립중앙도서관 자료보존연구센터

## 요약

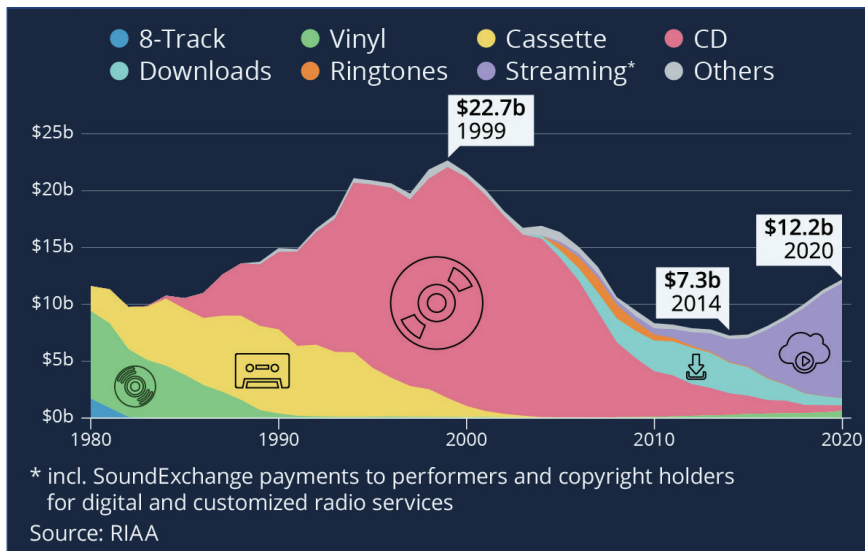
국립중앙도서관은 1950년대에 생산된 LP부터 최근의 디지털 파일까지 약 84만 점에 달하는 시청각 자료를 소장하고 있다. LP, VHS, 카세트테이프 등 다양한 매체를 소장하고 있으며, 각 매체의 정보를 판독하려면 재생장비를 사용해야 한다. 하지만 재생장비의 단종과 매체의 물리적 노화로 소장자료의 정보 또한 소실될 위기에 놓여 있다. 이에 국립중앙도서관은 시청각 자료의 장기보존과 미래활용을 위해 ① 디지털화 대상자료 선정 ② 원본자료의 보존상태 점검 ③ 원본자료 및 재생장비 세척공정 ④ 디지털화 순으로 네 단계 작업 공정을 거쳐 체계적으로 보존 처리함으로써 자료의 수명을 연장하고 있다. 하지만 IT 기술의 집약체라고 볼 수 있는 변환된 디지털 자료는 대용량 시청각 자료를 효율적으로 처리하기 위한 코덱, 컨테이너 포맷 등 디지털 자료의 특성에 맞는 보존체계 수립을 추가로 요구한다.

이와 같은 디지털 환경에서 요구되는 새로운 보존 과제를 해결하고자 국립중앙도서관은 시청각 자료 메타데이터 구조 연구 및 자동 추출 도구 개발 등으로 디지털 자료 장기보존 연구를 추진하고 있다. 또한 앞으로 다양한 시청각 자료를 관리하기 위해 고려해야 하는 환경적 요인도 분석해오고 있다. 이 글에서는 지금까지의 연구 결과와 함께 앞으로 소장 시청각 자료를 어떻게 보존하고 관리해야 하는지 방향성을 제시하고자 한다.

**주요 키워드** 시청각 자료, 장기보존, 기술메타데이터

## I. 뉴미디어 시대의 미디어 시장 환경 변화

뉴미디어<sup>1</sup>는 미디어 생태계를 기존 오프라인 매체에서 유튜브와 같은 온라인 매체로 사용 환경을 빠르게 변화시키고 있다. 불과 몇 년 전만 해도 영상이나 음원은 규격화된 저장매체에 담겨 유통되었다. 미국음반산업협회(RIAA)의 자료<sup>2</sup>에 따르면, 40년간 음악을 저장하는 매체는 비닐 레코드(vinyl record), 카세트테이프, CD, 스트리밍 플랫폼으로 변화하고 있다고 한다. 또한 영상 콘텐츠 산업 역시 VHS, DVD, 블루레이(Blu-ray), 온라인 스트리밍 플랫폼 등 다양한 매체로 매우 빠르게 이동하고 있다. 이러한 상황은 시청각 자료를 수집하고 서비스하는 도서관, 기록원, 박물관 등 아카이브 기관에 깊은 고민을 던지고 있다.



〈그림 1〉 음반 저장매체 유통 현황

출처: Statista(<https://www.statista.com/chart/17244/us-music-revenue-by-format>)

도서관은 납본된 다양한 자료를 보존하고 일반에게 서비스해야 하는 사회적 책무를 지고 있다. 구형자료를 보존하고 서비스 방안을 개발하려고 노력하지만 시청각 자료의 경우 단종된 장비가 많아서 정보 확인에 어려움이 있다. 동시에 시청각 디지털 자료는 생명 주기가 매우 짧고, 한때 유행했던 포맷이

1 20세기 후반부터 통용된 매체 연구 용어로 영화, 그림, 음악, 언어, 문자 등의 전통적인 전달 매체에 컴퓨터와 통신 기술, 스마트 모바일 기기, 인터넷 등의 높은 상호작용성이 더해져 만들어진 새로운 개념의 매체를 가리킨다. 대표적인 예로 문자 다중 방송, 양방향 케이블 TV, 인터넷 등이 있다. 뉴미디어를 통하여 원하는 콘텐츠를 언제 어디서나 어떤 기기로도 접근할 수 있으며, 콘텐츠에 대한 이용자의 자유로운 피드백을 허용하여 상호작용성이 높다. 뉴미디어는 전통적인 매체와 달리 디지털화된 콘텐츠를 가지며 빠른 시간에 콘텐츠를 많은 양 생성하는 특징도 있다. [위키백과(2021. 7. 12.). 뉴미디어. 출처: <https://ko.wikipedia.org/wiki/뉴미디어>]

2 RIAA(n.d.). U.S. sales database. Retrieved July 14, 2021, from <https://www.riaa.com/u-s-sales-database>

한순간에 구형이 되거나 코텍이 사라져 못 읽는 일도 생긴다. 만약 이들을 보존·관리하지 않는다면 결국 전국 도서관에 소장된 시청각 자료 수백만 점은 무용지물이 될 것이다. 공공도서관뿐만 아니라 민간기관에서도 이러한 문제를 심각하게 인식하지만 예산, 기술 지원 등 다양한 정책적 결정이 필요하기 때문에 쉽게 해결책을 찾기 어렵다. 이에 국립중앙도서관은 수집된 모든 시청각 자료를 미래에도 활용할 수 있도록 대응 방안을 지속적으로 연구하고 있다.

다음 장에서는 시청각 자료를 보존하기 위한 디지털화 작업 공정과 장기간 보존하기 위한 기술메타데이터 연구 사례를 소개한다.

## II. 도서관 시청각 자료 현황

### 1. 시청각 자료 유형 및 수집 현황

국립중앙도서관 장서개발 지침에 따르면 시청각 자료는 영상자료(DVD, 블루레이, 슬라이드, 비디오테이프, 비디오디스크 등), 음향자료(음반, 카세트테이프, 오디오테이프, CD 등), 실물 또는 모형자료(지구 표본 등)를 의미한다.<sup>3</sup>

현재 국립중앙도서관이 소장하고 있는 영상자료는 <표 1>과 같이 비디오테이프(VHS), 비디오카세트, 비디오디스크(V-CD), DVD, 블루레이(Blu-ray) 형태로 구성되며, DVD 396,678점(72.8%), 비디오테이프 136,776점(25.1%), 비디오디스크 5,320점(1%), 블루레이 4,802점(0.9%), 비디오카세트 1,456점(0.3%)으로 DVD가 가장 큰 비중을 차지한다. 반면 음향자료는 CD 매체가 53.1%(158,306점)를 차지하며, 카세트테이프 43.9%(130,854건), LP 3.1%(9,248점) 순으로 구성되어 있다.

<표 1> 국립중앙도서관 시청각 자료 소장 현황

2021년 6월 30일 기준, 단위: 점

전체 수량	영상자료					음향자료		
	비디오 테이프	비디오 카세트	비디오 디스크	DVD	블루레이	LP	카세트 테이프	CD
843,440	136,776	1,456	5,320	396,678	4,802	9,248	130,854	158,306

3 국립중앙도서관. 장서개발 기본지침. [인용일 2021. 7. 12.]. 출처: <https://www.nl.go.kr/NL/contents/N50107010200.do>

## 2. 시청각 자료 보존처리 및 디지털화

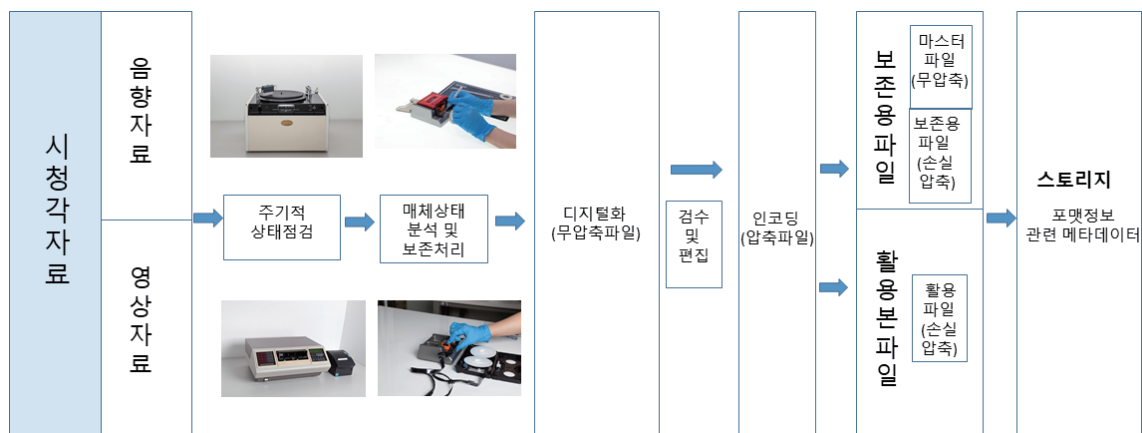
시청각 자료는 일반도서와 달리 저장매체가 온·습도에 민감하므로 보존환경에 더 엄격한 관리가 필요하다. 그렇기 때문에 서가에 있는 거의 모든 아날로그 시청각 자료는 장기간 보존하기 위해 필수적으로 디지털화 작업을 해야 한다. 시청각 자료의 효율적인 디지털화 작업은 다음과 같이 네 단계로 구성된다.

첫째, 디지털화 대상 우선순위를 정한다. 비디오테이프, 카세트테이프, LP는 영구 보존매체로 보기 어렵다. LP는 항온·항습 환경에서 잘 관리하고 재생 횟수가 적으면 가장 오래 보존할 수 있지만 재질의 특성상 백색잡음(white noise)이 많이 발생할 확률이 높다. 열화를 고려하여 디지털화 우선순위를 정한다면 비디오테이프, 카세트테이프, LP 순이라고 할 수 있다.

둘째, 원본자료의 보존상태를 점검해 매체의 물리적 변형을 확인한다. 매체 보존상태를 후각, 촉각 등으로 점검할 수도 있는데, 필름 종류의 경우 화학적 반응으로 초산(식초) 냄새가 날 수 있기 때문이다. 필름 재질이 산화하는 이러한 초산화 현상은 아세테이트 필름의 치명적 단점 중 하나이다. 초산화 현상을 지연시키는 약제는 있으나 예방할 수는 없으며 원래대로 복원하는 방법도 없다. 대체로 불규칙한 습도나 온도 때문에 나타나는 것으로 알려져 있으며, 특정 물질과 반응하여 산화 현상을 일으키기도 한다.

셋째, 디지털화 작업을 하기 전 최상의 상태를 유지하기 위한 세척(cleaning) 공정이다. 세척을 위해 자료를 케이스에서 꺼내 LP, VHS테이프 등의 매체를 자동세척기로 세척한다. 세척 후 반드시 처음으로 되감기를 해야 하는데, 이때 재생기기의 헤드는 에탄올과 면봉으로 닦아서 매체손상을 방지한다.

넷째, 디지털화(digitization) 과정이다. 재생장비와 최신 ADC(Analog-to-Digital Converter)를 사용하여 교정, 헤드 조정, 축음기 바늘 크기 등 최적의 재생 환경을 설정해야 한다. 이때 매체별 변환표준에 맞게(〈표2, 3〉 참조) 설정하되 원본 그대로 변환한다. 이렇게 디지털화한 파일을 서버에 저장한다.



〈그림 2〉 시청각 자료 보존처리 및 디지털화 프로세스

〈표 2〉 시청각 자료 디지털화 파일 포맷

자료유형	1차 디지털 변환			2차 디지털 변환	
	항 목	값	파일 포맷	코덱	파일 포맷
음향자료 (LP, CT)	샘플링 비율	48KHz	WAV	PCM	MP3
	비트심도	24Bit		MPEG-1 MPEG-2	
영상자료 (VHS, 방송용 테이프)	색공간	RGB	MXF	Video MPEG-2 Audio PCM	MP4
	서브 샘플링 비율	4:2:2			
	해상도	720*480		Video H.264 Audio AAC	
	비트심도	24Bit			

〈표 3〉 국립중앙도서관 시청각 디지털화 형식 기준

유형	구분	포맷	저장 방식	
음향	보존용	WAV	24bit	48kHz
	제공용	MP3	16bit	44.1kHz
		AAC	16bit	48kHz
영상	보존용	MXF	MPEG4(영상) / PCM(음성)	• SD: 4:3 • HD: 16:9
	제공용	MPEG-4	H.264(MPEG-4 / AVC)(영상) / AAC(음성)	• SD: 4:3 • HD: 16:9

### Ⅲ. 기술메타데이터 연구 개발 및 보존 전략

#### 1. 시청각 자료 기술메타데이터 선정

매년 시청각 자료를 보존하기 위하여 디지털화 변환업무를 추진하며, 올해 디지털화한 시청각 자료는 장기보존패키지<sup>4</sup>에 적용할 예정이다. 소장하고 있는 디지털화 시청각 자료를 장기보존하려면 디지털 자료의 구조적 특성을 분석한 정보를 보존해야 한다. 일반적으로 우리가 사용하는 디지털 자료는 디지털 포맷이라는 구조에 맞게 비트가 저장된 구조이다. 시청각 자료의 경우 MP4, AVI, WMV 등 포맷이 다양하며, 장기보존하려면 각 유형에 맞는 기술(technical) 메타데이터 추출 기능이 필요하다. 메타데이터의 다양한 종류 중 하나인 기술메타데이터는 디지털 객체에 기능적으로 요구되는 정보를 표현한 집합체로, 국립중앙도서관의 장기보존 패키지에 디지털 자료와 함께 저장되어 관리된다.

디지털 자료의 포맷에 대한 기술메타데이터는 디지털 포맷 관련 기술이 사장되어 더는 이용할 수 없을 때 포맷 관련 정보를 아는 데 필요하다. 이를 위해서 디지털 포맷에 대한 정보를 미리 추출하여 관리해야 한다. 따라서 디지털화 시청각 자료에서 요구되는 기술메타데이터 항목을 국가기록원, 한국영상자료원 등 유관기관의 메타데이터를 분석하여 <표 4>와 같이 도서관 자료에 맞게 선정하였다.

〈표 4〉 시청각 자료 기술메타데이터

구 분	기술메타데이터
음향자료 (7개)	오디오 비트레이트, 오디오 샘플링비트, 오디오 코덱 종류, 오디오 채널, 오디오 재생시간, 압축 유무, 디지털 변환날짜
영상자료 (15개)	오디오 포함 유무, 오디오 비트레이트, 오디오 샘플링비트, 오디오 채널, 오디오 코덱 종류, 오디오 압축 유무, 비디오 코덱 유형, 비디오 비트레이트, 파일 재생시간, 컬러 스페이스, 비디오 비트 깊이, 디지털 변환날짜, 해상도, 자막 유무, 자막 코덱 유형

#### 2. 장기보존패키지의 메타데이터 구성

장기보존패키지는 세부적으로 볼 때 크게 디지털화 원본 자료와 이들을 서술하는 메타데이터로 구성되어 있다. 각 장기보존패키지의 메타데이터는 장서자료를 표현하는 MODS(Metadata Object Description Schema)<sup>5</sup>와 디지털 포맷에 대한 정보를 표현하는 PREMIS(PREservation Metadata Implementation Strategies)<sup>6</sup>로 구성되어 있다.

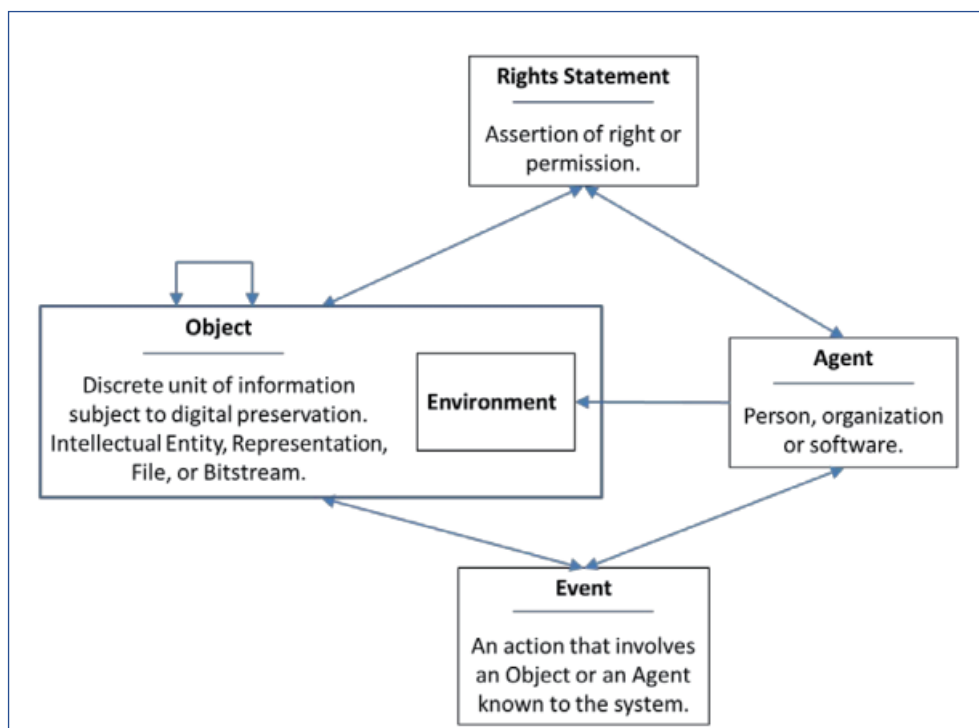
4 디지털 자료의 특성을 고려하여 보존에 요구되는 메타데이터를 디지털 자료와 함께 저장할 수 있도록 설계된 보존 포맷.

5 미의회도서관(MODS). <http://www.loc.gov/standards/mods>

6 미의회도서관(PREMIS). <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>



PREMIS 메타데이터의 목적은 디지털 자료에 해당하는 디지털 객체(digital object)를 장기간 보존한 후 다시 구동하는 데 필요한 정보를 관리하는 것이다. 기술발전에 따라 사장된 디지털 자원을 현재 환경에서 구현하기 위해서는 그 당시 사용하였던 운영체제 및 소프트웨어가 필요하다. 예를 들어 1990년대 사용하였던 아리랑 문서파일을 실행하기 위해서는 아리랑 문서편집 소프트웨어뿐만 아니라 소프트웨어가 실행될 수 있는 16비트 윈도우 운영체제 또한 동시에 요구된다. 즉, 디지털 자료는 저장만 하고 실행할 수 없다면 단지 0과 1로 표현된 단순한 비트 배열일 뿐이다. 따라서 PREMIS라는 정규화된 형태에 맞게 디지털 보존에 필요한 정보를 저장 관리해야 한다.



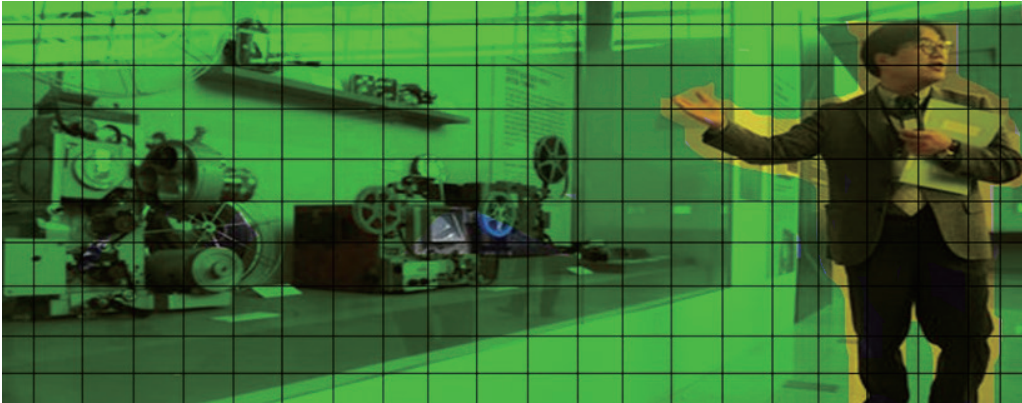
〈그림 3〉 PREMIS 데이터 모델

출처: PREMIS Data Dictionary for Preservation Metadata(<https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>)

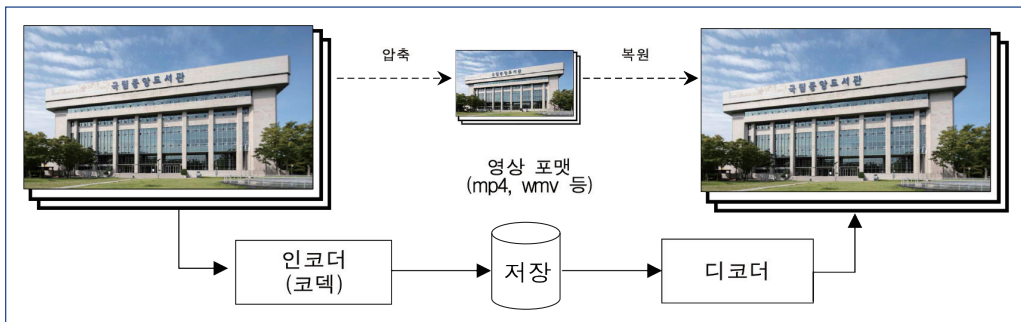
### 3. 시청각 기술메타데이터 구조 분석

시청각 디지털 자료 구조는 영상을 효율적으로 처리하기 위하여 코덱 압축 기술을 활용해 원본 영상의 용량을 경량화한다. 일반적으로 많이 사용하는 H.264 코덱의 경우 I(Intra coded frame), P(Predicted frame), B(Bi-predictive coded frame) 프레임 기준으로 영상에서 변화된 영역만 추출하여 영상을

압축 저장한다. 아래 그림과 같이 I 프레임에 해당하는 키 프레임 기준으로 영상에서 변화가 없는 녹색 영역을 제외한 움직이는 노란색 영역만 P, B 프레임에 저장하여 저장공간 및 전송 대역폭을 효율적으로 관리할 수 있게 한다.

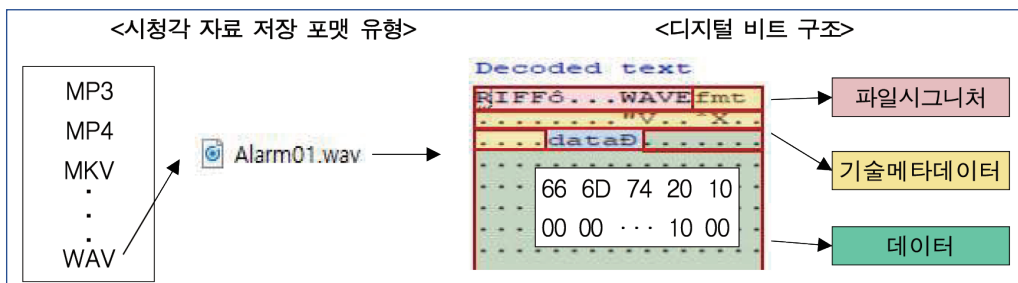


〈그림 4〉 시청각 자료 인코딩 원리



〈그림 5〉 인코더 및 디코더 관계

이렇게 코덱으로 인코딩되어 저장된 시청각 자료는 mkv, avi 등 컨테이너 디지털 포맷 내부에 저장되며, 각 저장 영역은 파일시그니처, 기술메타데이터, 데이터 영역으로 구분된다. 포맷 내부 세부 구조는 〈그림 6〉과 같다.



〈그림 6〉 디지털 포맷 내부 구조

컨테이너 디지털 포맷 유형에 따라 데이터와 함께 메타데이터가 〈그림 6〉과 같이 특정 패턴으로 저장된다. 파일시그니처 영역은 컨테이너 포맷에 대한 포맷 시작 값으로 〈그림 7〉, 〈그림 8〉과 같이 각 포맷 유형에 따라 시그니처값이 정해진다. 각 포맷 시그니처값은 포맷 유형을 분류하는 기준값으로 이용되며, 분류된 포맷 유형에 따라 두 번째 노란색 영역인 기술메타데이터를 분석하여 인코딩 유형, 비트레이트 정보를 추출한다. 또한 영상 재생 관련 정보가 세 번째 녹색 데이터와 함께 일정한 코덱 규칙에 따라 저장되기도 한다.

Offset(h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded text
00000000	00	00	00	18	66	74	79	70	69	73	6F	6D	00	00	00	00	....ftypisom....
00000010	69	73	6F	6D	33	67	70	34	06	B3	A1	DD	6D	64	61	74	isom3gp4.*;mdat
00000020	00	02	2A	33	65	B8	40	6C	A2	38	07	5F	85	C0	00	41	..*3e,@1c8._.À.A

〈그림 7〉 MP4 파일시그니처(MP4 ISO base media file)

Offset(h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded text
00000000	30	26	B2	75	8E	66	CF	11	A6	D9	00	AA	00	62	CE	6C	0&^užfĩ.;Û.*.bîl
00000010	F5	02	00	00	00	00	00	06	00	00	00	01	02	A1	DC		š.....;Û
00000020	AB	8C	47	A9	CF	11	8E	E4	00	C0	0C	20	53	65	68	00	«EG@İ.žä.À. Seh.

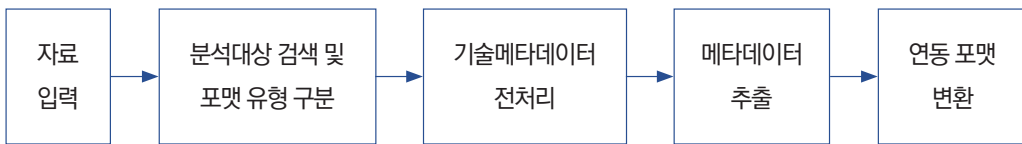
〈그림 8〉 WMV 파일시그니처

이와 같이 디지털 시청각 자료<sup>7</sup>의 경우 영상을 압축하는 코덱, 비트레이트, 컨테이너 포맷 등 영상 데이터를 처리하는 다양한 메타데이터를 저장하며, 시청각 자료를 재생하기 위한 정보를 포함하고 있다. 하지만 기술(technical) 메타데이터 정보는 소프트웨어 유형에 따라 디지털 포맷에 다양한 형태로 저장되어 있어 추출하기가 어렵다. 또한 대량의 시청각 자료를 자동으로 처리하려면 업무 프로세스 유형에 맞는 최적화가 요구된다.

7 디지털 시청각 자료의 경우 디지털 형태로 생성된 자료(born-digital) 및 디지털화한 시청각 자료 모두를 의미함.

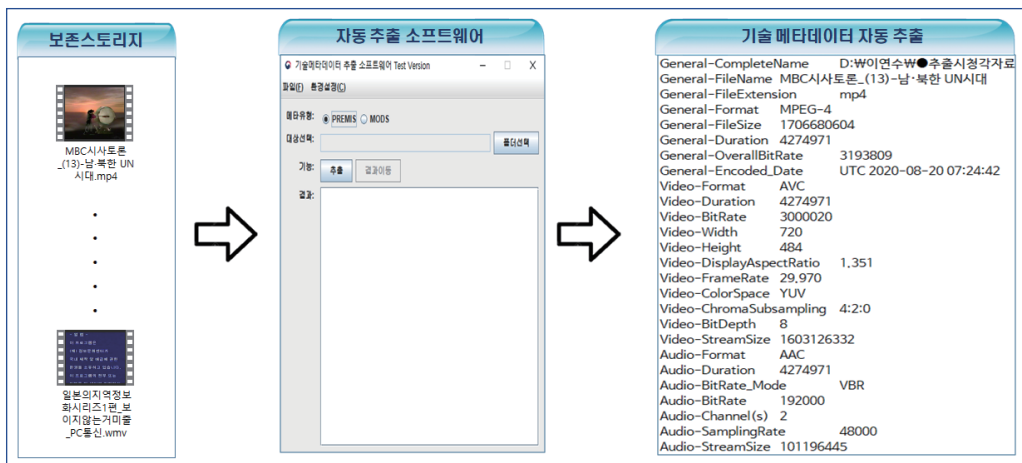
#### 4. 기술메타데이터 자동 추출 도구 개발

변환한 시청각 자료의 기술메타데이터를 추출하여 보존하려면 대량의 자료를 정확하게 처리하는 소프트웨어 또는 시스템이 필요하다. 국립중앙도서관은 올해 디지털화 시청각 자료 4천 점의 기술메타데이터<sup>8</sup>를 국립중앙도서관 통합자료관리시스템(KOrea Library Information System, KOLIS)<sup>9</sup>을 통해서 장기보존패키지로 변환하여 보존할 예정인데, 이는 시청각 디지털화 업무 프로세스에 적합하게 적용해야 한다. 이를 위해 디지털화 시청각 자료의 기술메타데이터를 자동으로 추출하여 장기보존패키지의 PREMIS 메타데이터와 연동해서 적용할 수 있는 도구를 개발, 업무에 적용하고 있다.



〈그림 9〉 기술메타데이터 분석 및 처리 순서

기술메타데이터 추출 도구는 C 기반의 미디어인포(MediaInfo) 라이브러리를 사용하여 개발하였으며, 스토리지 하위 폴더를 스캔하여 시청각 자료의 기술메타데이터를 추출·저장하는 형식으로 구현하였다.

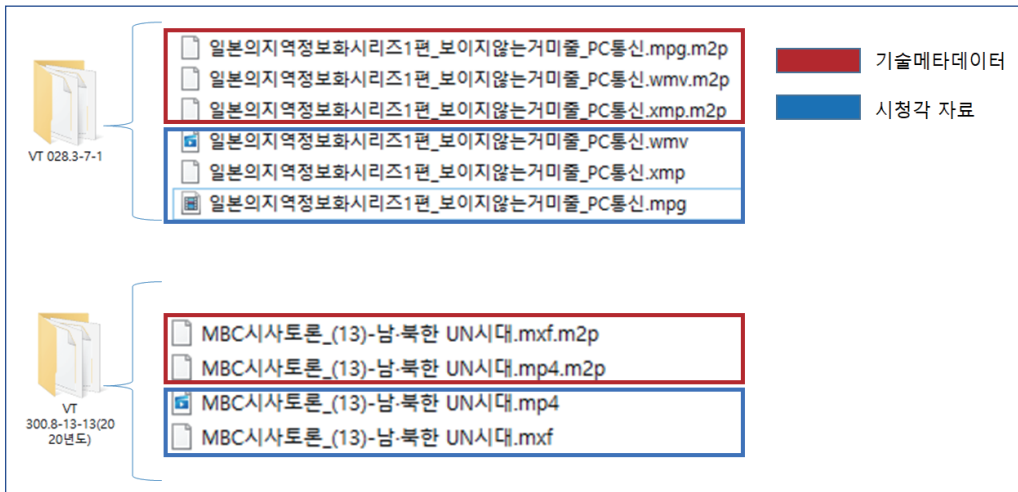


〈그림 10〉 시청각 기술메타데이터 추출 절차

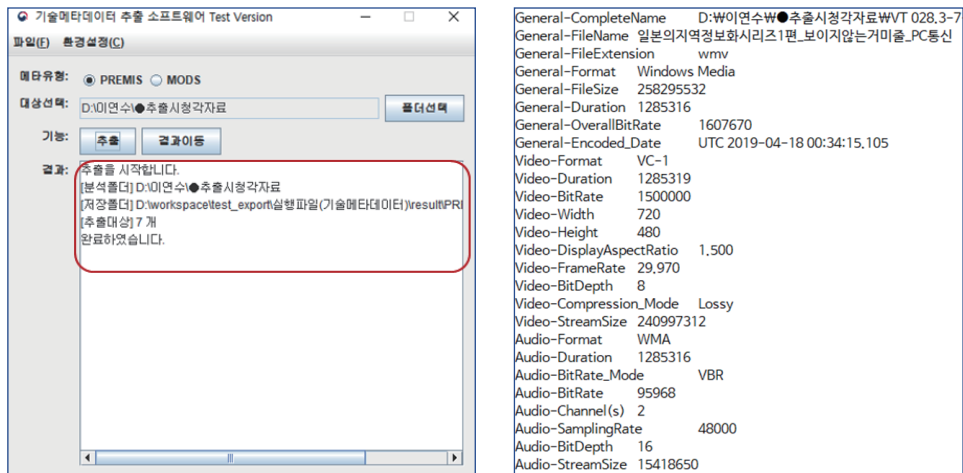
8 자동화 도구를 통해 추출된 시청각 기술메타데이터.

9 국립중앙도서관이 소장하고 있는 모든 자료를 체계적으로 관리하는 정보관리 시스템.

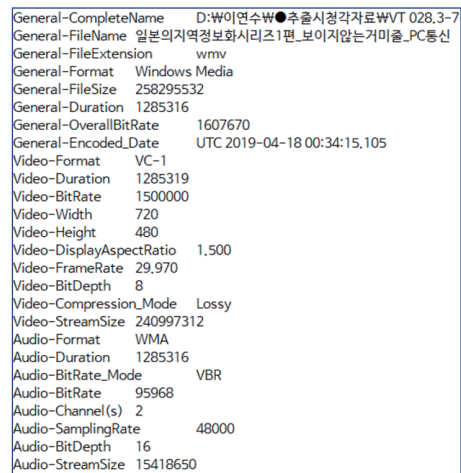
시청각 기술메타데이터는 <그림 11>과 같이 .m2p 확장자로 원본자료와 동일한 폴더에 저장되게 설계하였다. 각 디지털화 시청각 자료에 대한 기술메타데이터는 크게 일반, 비디오, 오디오 영역으로 구분할 수 있으며, <그림 13>과 같이 약 26개<sup>10</sup> 메타데이터가 추출된다.



<그림 11> 기술메타데이터 저장 형태



<그림 12> 소프트웨어 실행 화면



<그림 13> 기술메타데이터 자동추출 결과

10 시청각 디지털 자료 저장 포맷에 따라 자동으로 추출되는 정보가 상이할 수 있음.

## 5. 장기보존패키지 적용 및 보존 전략

추출된 기술메타데이터는 장기보존패키지의 PREMIS 메타데이터 내부에 포함되며, KOLIS와 연계된 정보로 만들어진 기존 필수 메타데이터와 함께 관리된다. 기술메타데이터의 세부 각 항목은 PREMIS의 디지털 보존 관련 요소인 Significant Properties 하위 값으로 추가 저장되어 관리된다.

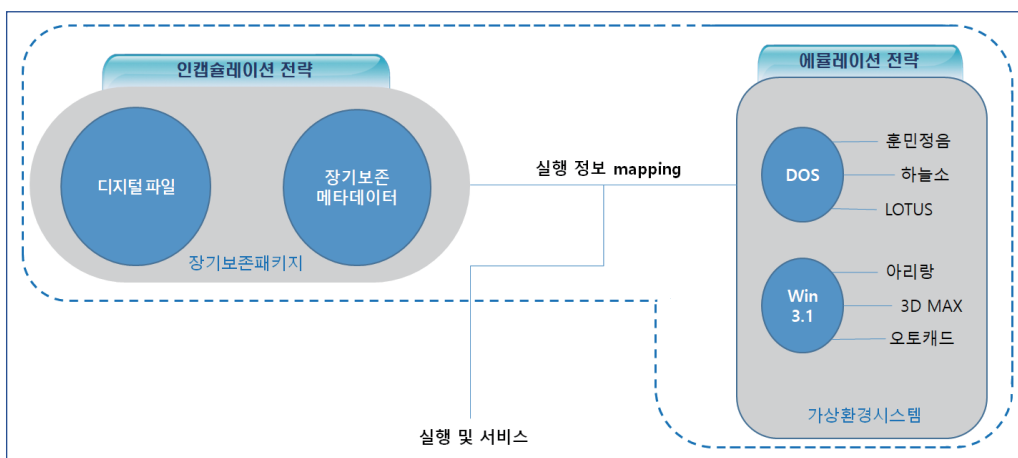
```
<?xml version="1.0" encoding="UTF-8" ?>
<mets:mdWrap MIMETYPE="text/xml" MDTYPE="PREMIS:OBJECT">
  <mets:xmlData>
    <premis:object xmlns:premis="info:lc/xmlns/premis-v2">
      <premis:objectIdentifier>
        <premis:objectIdentifierType>KOLIS_FILE_ID</premis:objectIdentifierType>
        <premis:objectIdentifierValue>FILE-00066143661</premis:objectIdentifierValue>
      </premis:objectIdentifier>
      <premis:objectCharacteristics>
        <premis:compositionLevel>0</premis:compositionLevel>
        <premis:fixity>
          <premis:messageDigestAlgorithm>SHA-1</premis:messageDigestAlgorithm>
          <premis:messageDigest>070bf0986ac150b2cbf6534aa171d2254d7fc92c</premis:messageDigest>
        </premis:fixity>
        <premis:size/>
      </premis:objectCharacteristics>
      <premis:format>
        <premis:formatDesignation>
          <premis:formatName>application/pdf</premis:formatName>
          <premis:formatVersion>1.0</premis:formatVersion>
        </premis:formatDesignation>
      </premis:format>
      <premis:storage>
        <premis:contentLocation>
          <premis:contentLocationType>URI</premis:contentLocationType>
          <premis:contentLocationValue>/jangseo7/R101/S1/202010/FILE-00066143661</premis:contentLocationValue>
        </premis:contentLocation>
      </premis:storage>
    </premis:object>
  </mets:xmlData>
</mets:mdWrap>
```

〈그림 14〉 장기보존패키지의 PREMIS 필수 메타데이터<sup>11</sup>

장기보존패키지에 저장된 PREMIS 메타데이터는 디지털 자료와 함께 인캡슐레이션 형태로 〈그림 15〉와 같이 동시에 관리된다. 이러한 패키지 형태 자료는 에뮬레이션 전략을 통해서 PREMIS 메타데이터에 해당하는 구동환경과 맵핑되어 가상환경에서 실행·서비스된다. 현재 DOS 기반 소프트웨어를 웹 환경에서 구동하는 테스트를 완료하였으며 SW, OS 등 다양한 실행환경이 동일한 플랫폼에서 구동되도록 연구를 계속할 예정이다.

11 PREMIS 필수 메타데이터: 장기보존패키지를 생성할 때 PREMIS 메타데이터 항목 중 KOLIS와 연동하여 자동으로 만들어지는 항목.





〈그림 15〉 기술메타데이터 기반 보존 전략

#### IV. 맺음말

지금까지 국립중앙도서관 시청각 자료의 매체별 소장 현황과 보존처리의 특징을 알아보았다. 또한 다양한 아날로그 매체를 장기보존하는 데 필요한 디지털화 작업과 고려사항을 심도 있게 논의하였다. 그리고 시청각 자료의 특성을 반영한 기술메타데이터를 분석함으로써 지속적인 접근 및 활용을 위한 보존메타데이터를 선정하였으며, 이를 기반으로 자동화된 기술메타데이터 추출 소프트웨어를 개발하여 업무에 적용하였다.

앞으로 다양한 디지털화 시청각 자료를 장기보존하려면 디지털 포맷 특성을 고려한 기술메타데이터 분석 기술을 선제적으로 연구·구축해야 한다. 과거 아날로그 시청각 자료는 매체를 읽을 수 있는 재생 장비만 필요하였지만, 현대의 디지털 자료는 구동 장비뿐만 아니라 저장된 디지털 포맷 특성을 고려한 운영체제와 소프트웨어 환경도 고려해야 한다.

이와 같이 유형이 다양한 디지털 자료의 특성을 분석하려면 자동 포맷 분석은 물론 이들을 분류하는 시스템이 필요한데, 본고에서 논의한 디지털화 시청각 자료 기술메타데이터 추출 소프트웨어 개발은 그 첫걸음이라고 볼 수 있다. 앞으로 디지털 형태로 생성된 자료(born-digital) 자료 수집 등 다양한 디지털 포맷을 분석하여 분류하는 장기보존 시스템 개발과 연구가 지속적으로 필요할 것으로 보인다.

### 〈참고문헌〉

국립중앙도서관. 장서개발 기본지침. [인용일 2021. 7. 12.]. 출처: <https://www.nl.go.kr/NL/contents/N50107010200.do>

기록관리 메타데이터 표준(v2.1). NAK 8:2021.

김상국 (2013). 디지털 시대 국가기록원 시청각 아카이빙 현황과 개선에 관하여. 기록물 보존복원, (6), 41-63.

서혜란 (2006). 문화원형콘텐츠의 장기보존에 관한 연구. 한국비블리아학회지, 17(2), 65-82.

정보통신정책연구원 (2021). 방송·미디어 산업 시장구조 변화에 대응한 정책방향 수립 연구, 한국방송통신전파진흥원.

최효진 (2018). 국내 공공영상아카이브 관리 체계 마련을 위한 과제. 기록학연구, (58), 95-145.

Information and Documentation—Records Management Processes—Metadata for Records—Part 2: Conceptual and Implementation Issues. ISO/TS 23081-2:2007.

International Association of Sound and Audiovisual Archives. Retrieved July 16, 2021, from <https://www.iasa-web.org>.

RIAA (n.d.). U.S. sales database. Retrieved July 14, 2021, from <https://www.riaa.com/u-s-sales-database>

Routhier Perry, S. (2014). Digitization and digital preservation. School of Information Student Research Journal, 4(1).

Space Data and Information Transfer Systems—Open Archival Information System(OAIS)—Reference Model. ISO 14721:2012.

The BagIt File Packaging Format(v1.0). IETF RFC 8493:2018.