디지털 보존시스템 현황 및 상호운용성 개념 이해

2019. 8. 28.

안대진 (아카이브랩)

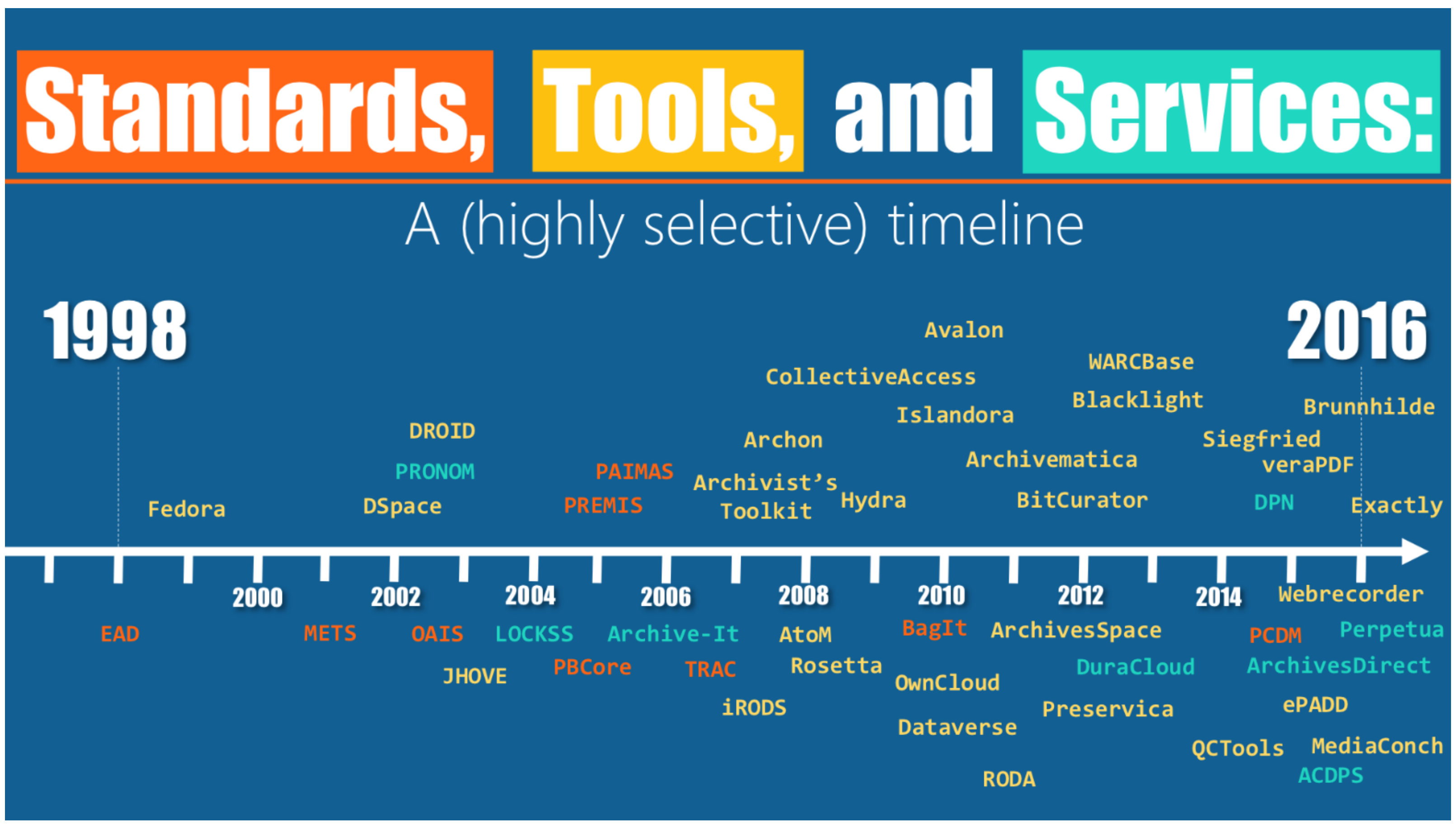
국가기록원 2019년 제4차 전자기록관리 연구포럼

목차

1. 표준, 툴, 서비스

디지털 보존 분야의 표준, 툴, 서비스

90년대 말부터 디지털 보존 커뮤니티는 표준을 공동 개발하고 이를 툴과 서비스로 발전시킴



주요 표준, 툴, 서비스

표준
- EAD, METS, PREMIS(메타데이터)
- OAIS, TRAC, PAIMAS (장기보존)
- BagIt(패키지)
- PCDM(데이터모델)

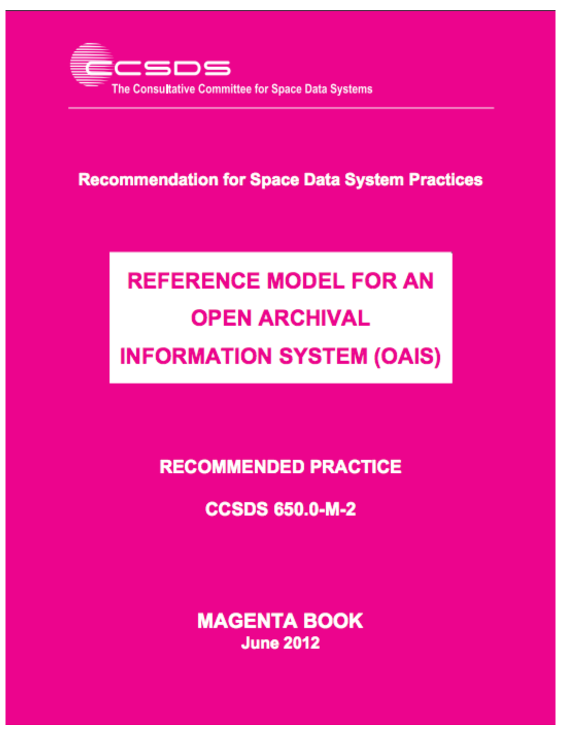
툴
- Fedora, Dspace 등 (리포지터리)
- Archon, AtoM 등 (카탈로그)
- DROID, JHOVE 등 (식별도구)
- veraPDF 등 (검증도구)
- Preservica 등 (아카이브시스템)

서비스
- PRONOM 등 (포맷레지스트리)
- Archive-It 등 (웹 아카이빙 서비스)
- ArchivesDIRECT 등 (보존 서비스)

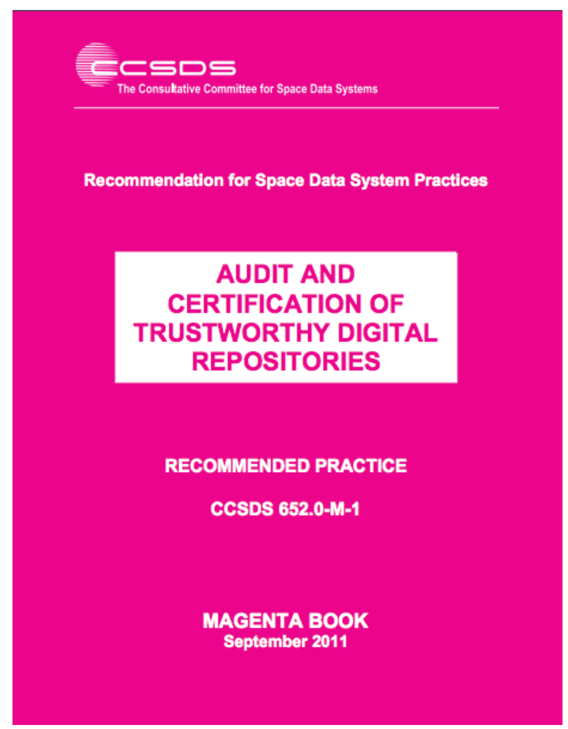
OAIS와 후속표준

디지털 보존 분야의 대표적 표준은 ISO 14721(OAIS 참조모델)과 ISO 16363(디지털저장소 신뢰성 감사 및 인증)임.
대부분의 디지털 아카이브시스템은 두 표준을 근간으로 설계됨.

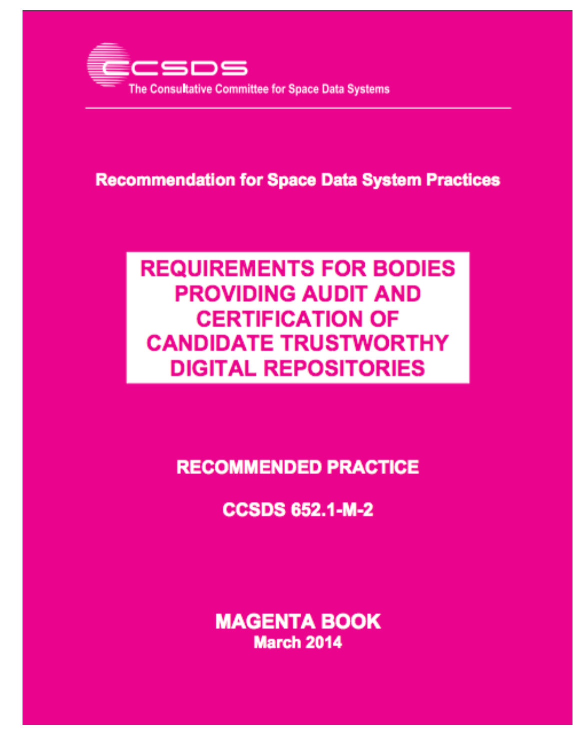
- ISO 14721은 전자기록의 장기보존과 지속적 접근을 위한 가장 포괄적인 표준(InterPARES 역시 OAIS를 기반으로 함)
- ISO 16363은 OAIS의 후속표준이며, TRAC(RLG & NARA)과 DRAMBORA 등의 평가도구를 근간으로 만들어짐.
- ISO 16919는 ISO 16363을 사용하여 디지털 저장소의 신뢰성을 평가/인증하는 조직의 요구사항을 정의함.



ISO 14721:2002 (OAIS)



ISO 16363:2012

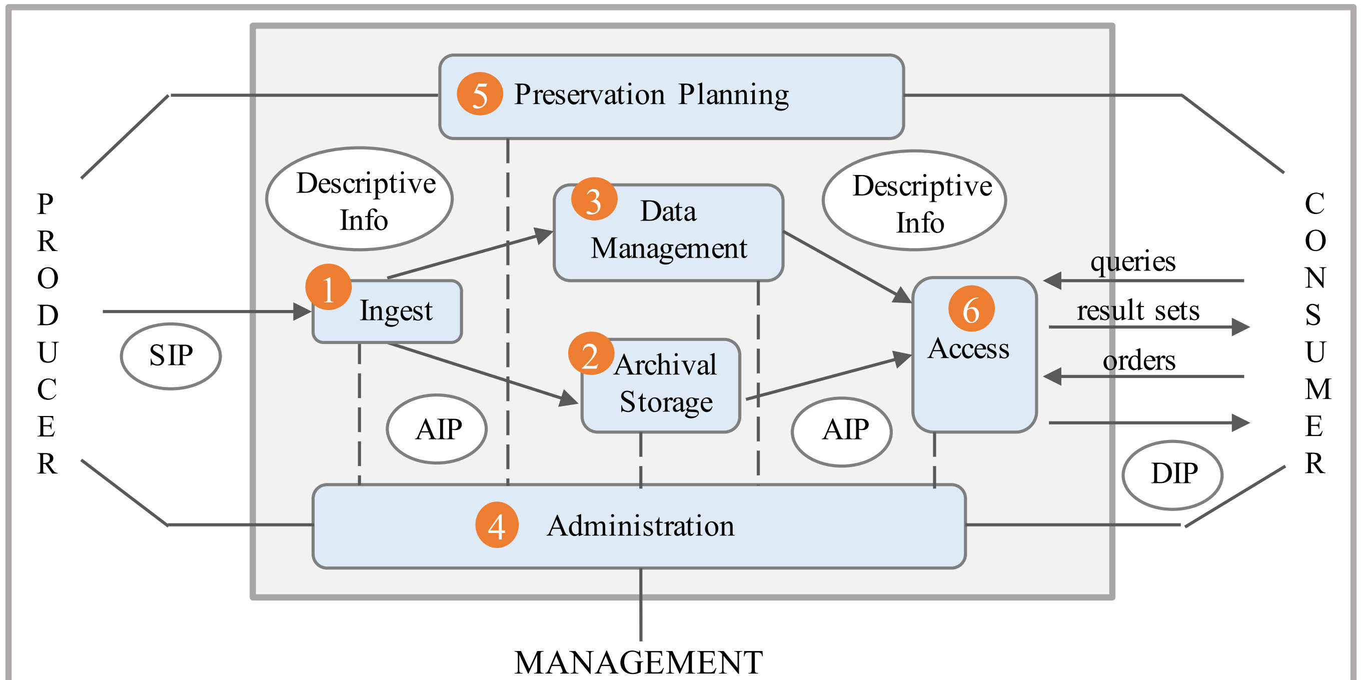


ISO 16919:2014

1. 표준, 툴, 서비스

OAIS 개요

ISO 14721(OAIS)은 전자기록을 장기보존하고 지속적 접근을 제공하기 위한 참조모델 제시



- SIP(Submission Information Package): 제출정보패키지
- AIP(Archival Information Package): 보존정보패키지
- DIP(Dissemination Information Package): 배부정보패키지

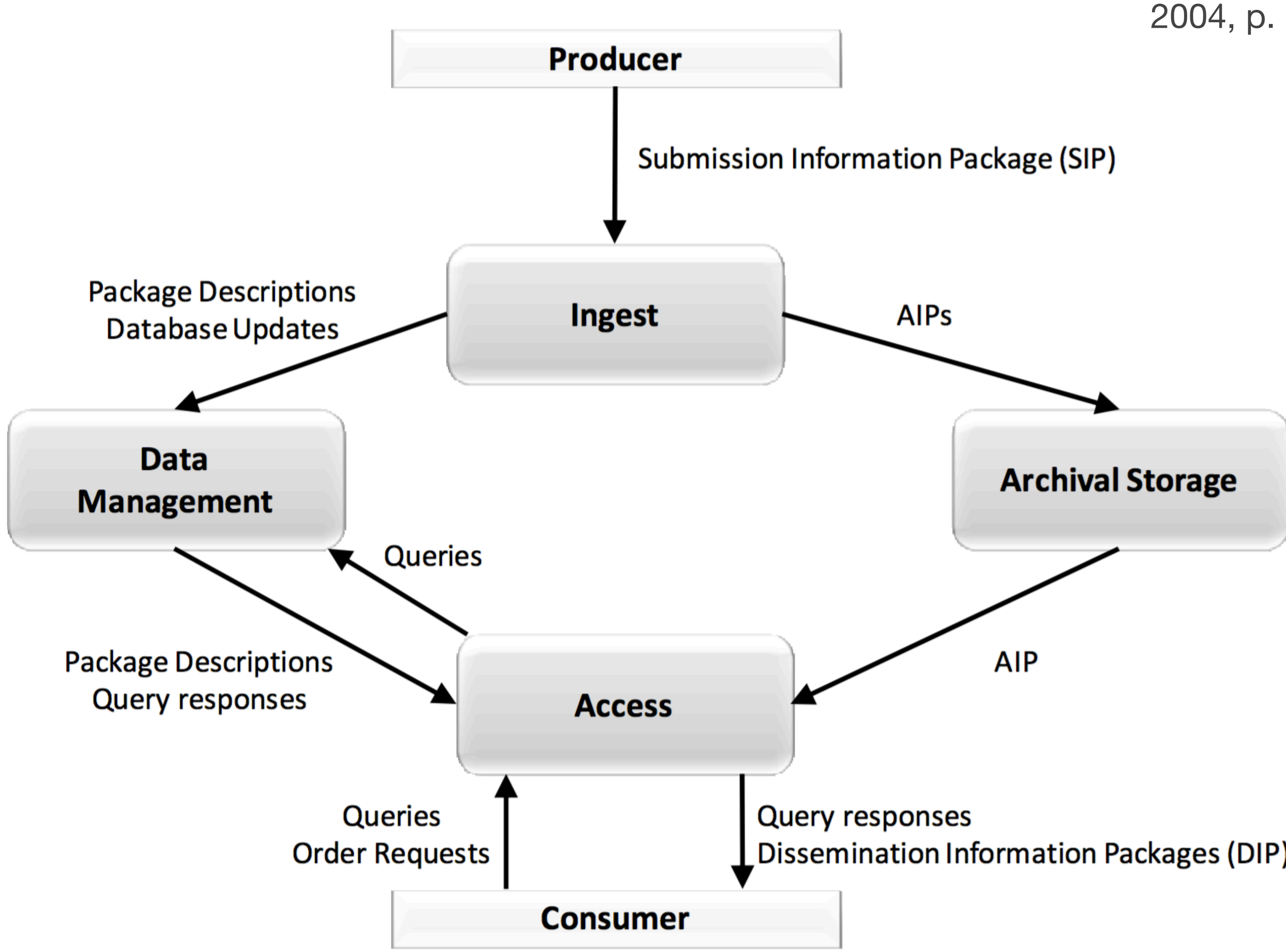
OAIS 기능별 역할

기능	역할
1 Ingest	Producer의 제출정보패키지(SIP)를 인수하는 기능과 서비스를 제공하고, 아카이브 내에서 저장 및 관리를 위한 내용을 작성
2 Archival Storage	AIP의 저장, 유지관리, 검색을 위한 기능과 서비스를 제공
3 Data Management	아카이브 소장물을 확인하고 기록하여 문서화하는 기술적인 정보와, 아카이브를 관리하는데에 사용되는 업무관리 데이터를 입력하고, 유지관리하고, 접근하는 기능과 서비스를 제공
4 Administration	아카이브 시스템의 전반적인 운영을 위한 서비스와 기능을 제공
5 Preservation Planning	OAIS 환경을 모니터링하기 위한 서비스와 기능을 제공
6 Access	Consumer가 OAIS에 저장된 정보의 존재여부, 기술내용, 저장장소, 이용가능성 등을 확인하도록 하고, 정보를 요청하고 받을 수 있도록 지원

OAIS의 문제

OAIS의 요구사항은 매우 추상적이어서 해석, 번역, 구체화 필요

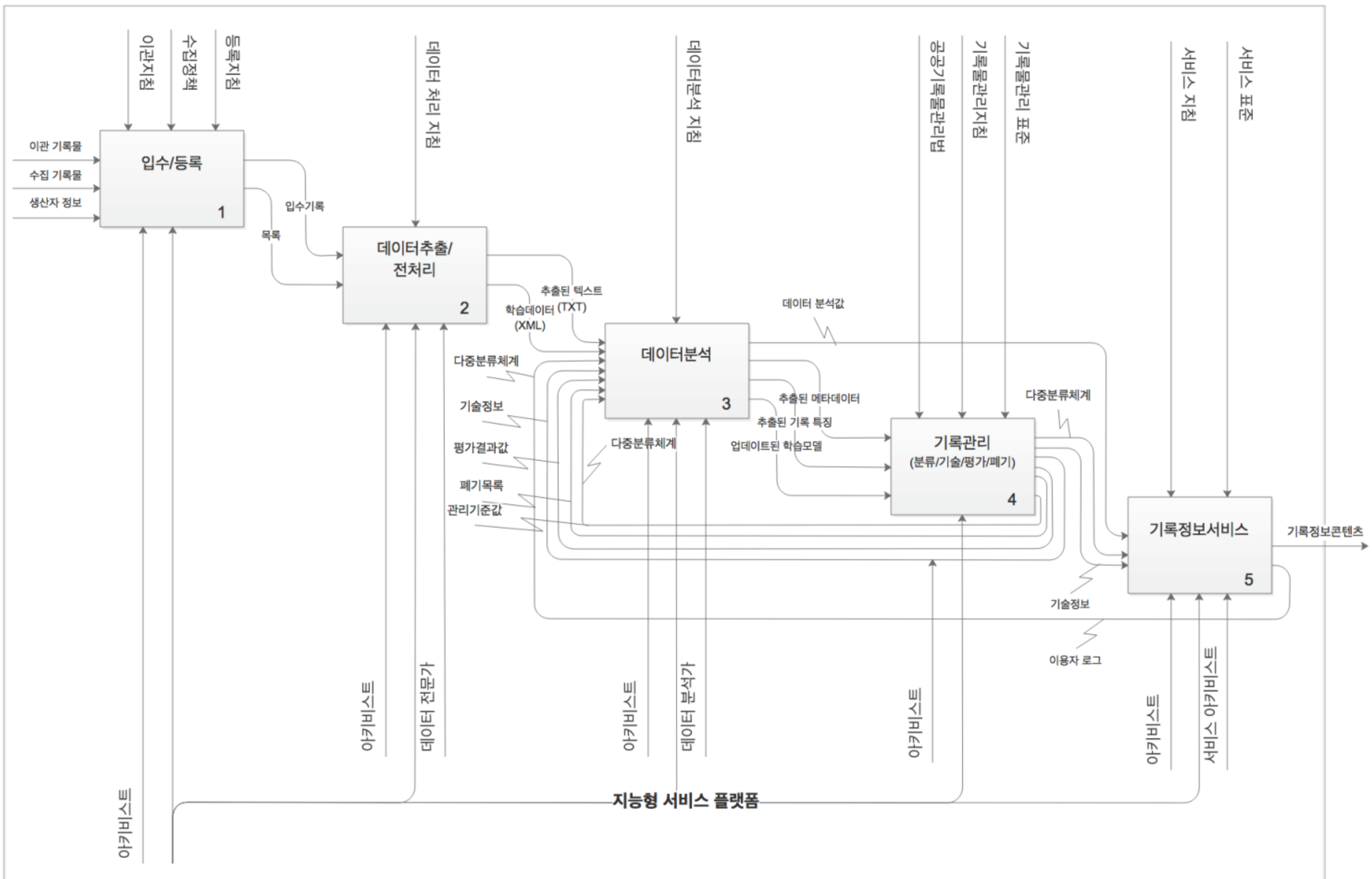
“ Because of the OAIS model’s abstract character, the notion of OAIS compliance is, as has been pointed out repeatedly, “necessarily vague” ” (Lavoie, 2004, p. 17).



[그림] High Level 데이터 흐름도

OAIS 해석 및 구현 : RM + AM

국가기록원 차세대 연구에서는 OAIS 참조모형에 따라 차세대 기록관리시스템의 서비스 모델 수립
기록물 입수, 데이터 분석, 기록물 분류, 평가폐기, 기록물 보존까지 RM과 AM의 기능을 수행하도록 해석 확장



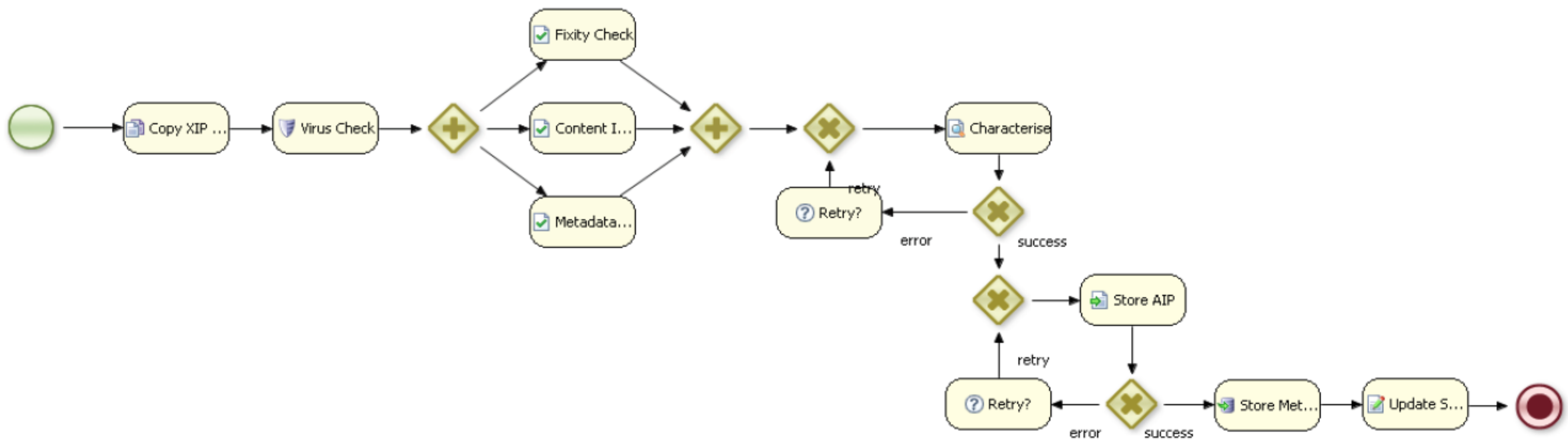
*출처: 국가기록원. 2017. 차세대기록관리모델 재설계.

OAIS 해석 및 구현 : 워크플로우

워크플로우는 특정한 비즈니스 요구사항을 충족하기 위해 시스템 상의 여러 가지 처리 작업(job)을 구체적으로 정렬한 것

워크플로우 맥락(contexts)은 워크플로우를 정의하는 것과 이것을 실행하는데 필요한 구성 정보의 조합을 설명시 사용

워크플로우 인스턴스(instance)는 워크플로우 맥락의 특정한 단일 행위에 따라 실행



[그림] 일반적인 워크플로우 다이어그램

OAIS 해석 및 구현 : 정보 패키지

OAIS 참조모델에서는 정보패키지의 유형을 SIP, AIP, DIP로 구분하고 있으나 CAMS 등에는 이의 구분이 모호함

구분	SIP 제출정보패키지(Submission)	AIP 보존정보패키지(Preservation)	DIP 배포정보패키지(Dissemination)
정의	• 생산자가 OAIS로 보내는 정보패키지 (형태와 내용은 OAIS와 생산자 간의 협약에따름)	• SIP를 기반으로 OAIS에 보존하기 위해 변형된 정보패키지	• 이용자의 주문에 대해 AIP의 전체 또는 일부를 제공하는 정보패키지
정보 요소	• 내용정보와 보존기술정보로 구성 • 표현정보와 연결	• 내용정보와 보존기술정보의 묶음, 혹은 AIP들의 집합 • PREMIS/METS 메타데이터 적용	• AIU(AIP들의 집합)이거나 PDI(보존기술정보) 없는 AIP 등 • AtoM 등 서비스 시스템의 계정정보 및 보안정보(DRM 등) • METS 메타데이터 적용
비고	• 다수 SIP가 동일한 표현정보 를 이용하면 표현정보는 한 번만 제공함 • 다수 SIP에 동일한 보존기술 정보가 적용되면 내용정보 없는 SIP 형태로 제공함 • 기술정보는 SIP 제출시점에 관계없이 제공됨	• 패키징 정보는 내부의 표준을 따름 • AIP 기술정보를 이용해 이용자가 원하는 내용정보를 찾고주문 가능하도록 함	• 배포매체와 이용자 요구에 따라 패키징정보가 다양함 • DIP 기술정보는 DIP 전송 이전, 이후, 전송시 제공가능

시사점

메타데이터 구분 어려움

- OAIS에서는 메타데이터 모형을 제공하지 않아 SIP, AIP, DIP별로 메타데이터를 나누기 어려움

규격 확정 필요

- 시스템 구축과 병행하여 세 규격 확정 필요(특히 SIP는 생산자와 AMS 간의 협약된 규격에 따르며, 유연하게 적용 가능)
- 시청각기록, 데이터세트 등이 포함될 수 있는 패키지 규격 필요(현재는 문서보존포맷 중심으로 설계됨)

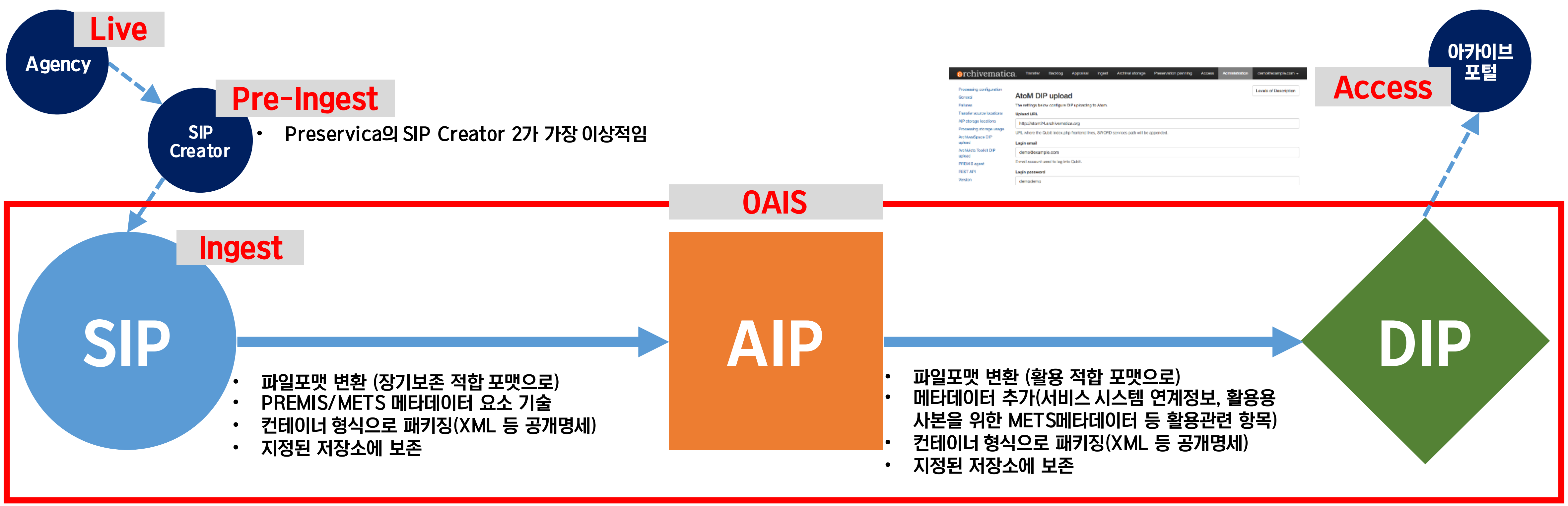
적용사례(Implementation) 벤치마킹 필요

- OAIS는 추상적 수준 요건이므로 실제 구현 사례 검토 필요
- 각국 아카이브즈 및 연구 (E-Ark), rm, AMS 소프트웨어(Archivematica 등)의 SIP, AIP, DIP 규격 벤치마킹 필요

*출처: 대통령기록관. 2019. 대통령기록 정보시스템 고도화 ISP

OAIS 해석 및 구현 : 정보 패키지

E-Ark 프로젝트에서는 Preservica, Archivematica 등 현행 7종의 AIP와 3종의 컨테이너 포맷(.ace, .rar, .tar)를 분석한 후 범 유럽 공통의 SIP, AIP, DIP 명세서를 개발 및 공개(장기보존, 벤더독립, 사본호환, 재난복구 등 4대 목표)



시사점	E-Ark의 SIP, AIP, DIP 명세 최우선 참고 요망 • 공개된 명세 최신버전 확인 필요 (https://www.eark-project.com/resources/project-deliverables) E-Ark의 AIP 비교연구 참고 요망 • Project 620998: European Archival Records and Knowledge Preservation – E-ARK (https://www.eark-project.com/resources/project-deliverables/93-d34-1/file)

*출처: 대통령기록관. 2019. 대통령기록 정보시스템 고도화 ISP

OAIS 해석 및 구현 : 정보 패키지

각국의 기록소장기관들은 OAIS 참조모형의 정보패키지를 기반으로 SIP, AIP, DIP 설계
XML로 묶는 형태와 폴더 형태로 구분

- **보존패키지 유형**

METS	PREMIS	XIP	AAP	NEO / VEO	EAD / EAC	SIARD	XML	PDF/A	Windows folder	BagIt

- **장기보존포맷 요건 및 특징 (OAIS 참조모형 기준)**

개요	• 대부분 아카이브는 OAIS를 따르며, 정보 패키지는 SIP, AIP, DIP로 구성 • 국가별 정보 패키지 규격을 설계하여 XML로 표현
AIP 구성요소	• 콘텐츠정보(Content Information), 보존기술정보(PDI)
PDI 구성요소*	• 참조정보, 출처정보, 맥락정보, Fixity 정보 (기관별 재량)
Fixity check	• MD5, SHA-1, SHA-256, SHA-512 등 (기관별 재량)
인캡슐레이션	• ZIP, TAR, Base64, 또는 무압축(Unzipped Bag 등) (기관별 재량)

* PDI 내용구성은 개별 아카이브의 재량이며, 전자서명(PKI) 인증정보를 패키지 내에 포함하기도 함(NEO, VEO 등). 전자서명 대신 해시로 위변조 여부 파악하기도 함

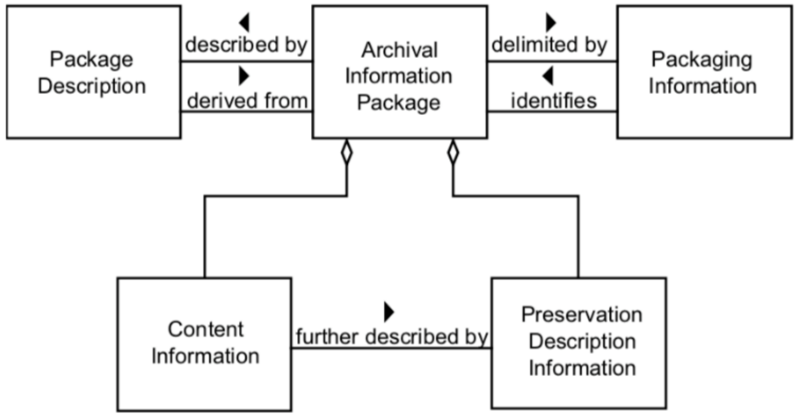


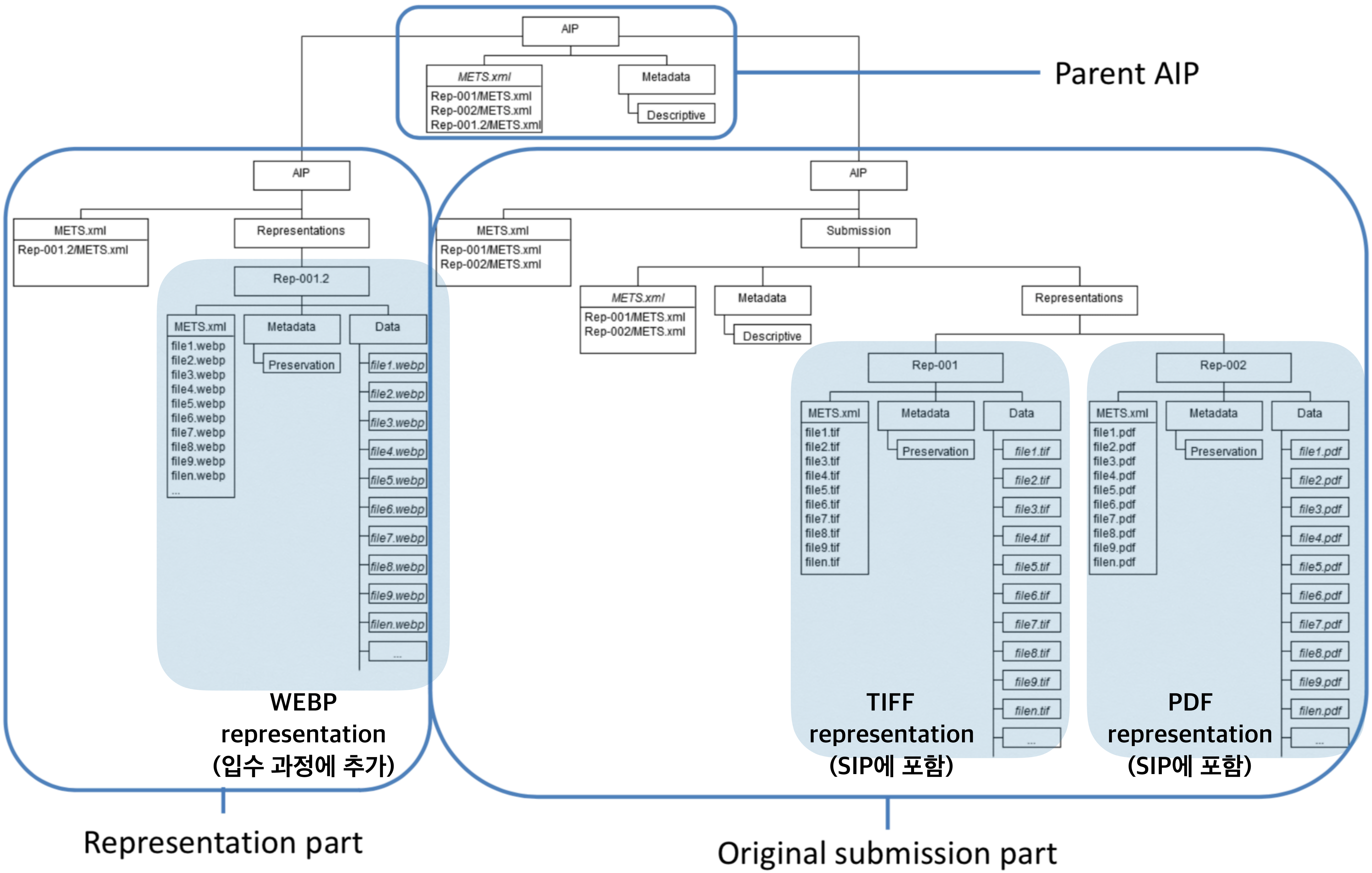
Figure 4-15: Archival Information Package (AIP)

<그림> OAIS의 보존정보패키지 개념도

*출처: 대통령기록관. 2019. 대통령기록 정보시스템 고도화 ISP

OAIS 해석 및 구현 : 정보 패키지 (E-Ark)

E-Ark는 SIP, AIP, DIP 명세서와 변환 툴 제공(변환 워크플로우 커스터마이징 가능)



OAIS 해석 및 구현 : 정보 패키지 (VEO3)

VEO3 생성과정은 8단계를 거침

VEO3는 콘텐츠와 메타데이터 디렉토리 전체를 ZIP으로 인캡슐레이션

Step 1.
VEO 디렉토리 생성

- .veo 디렉토리 생성
- 예 : xyzzy.veo

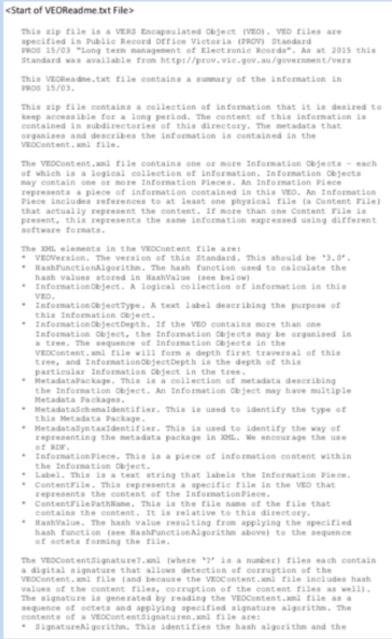
Step 2.
디렉토리에 콘텐츠 복사

- 하위 디렉토리 생성
- 원하는 대로 콘텐츠 파일 정렬
- 디렉토리 이름 지정

- → 정보가 처음 생성된 구조로 복사한 후 추가되는 정보는 별도의 추가 하위 디렉토리로 구성

Step 3.
VEOReadme.txt 파일 생성

- 사람이 읽고 이해할 수 있는 텍스트 형식으로 VEO 파일의 구조 설명



<Start of VEOReadme.txt File>

This zip file is a VERS Encapsulated Object (VEO). VEO files are specified in Public Record Office Victoria (PROV) Standard PROS 15/03 "Long term management of Electronic Rcords". As at 2015 this Standard was available from http://prov.vic.gov.au/government/vers

This VEOReadme.txt file contains a summary of the information in PROS 15/03.

This zip file contains a collection of information that it is desired to keep accessible for a long period. The content of this information is contained in subdirectories of this directory. The metadata that organises and describes the information is contained in the VEOContent.xml file.

The VEOContent.xml file contains one or more Information Objects - each of which is a logical collection of information. Information Objects may contain one or more Information Pieces. An Information Piece represents a piece of information contained in this VEO. An Information Piece includes references to at least one physical file (a Content File) that actually represent the content. If more than one Content File is present, this represents the same information expressed using different software formats.

The XML elements in the VEOContent file are:
* VEOVersion. The version of this Standard. This should be '3.0'.
* HashFunctionAlgorithm. The hash function used to calculate the hash values stored in HashValue (see below)
* InformationObject. A logical collection of information in this VEO.
* InformationObjectType. A text label describing the purpose of this Information Object.
* InformationObjectDepth. If the VEO contains more than one Information Object, the Information Objects may be organised in a tree. The sequence of Information Objects in the VEOContent.xml file will form a depth first traversal of this tree, and InformationObjectDepth is the depth of this particular Information Object in the tree.
* MetadataPackage. This is a collection of metadata describing the Information Object. An Information Object may have multiple Metadata Packages.
* MetadataSchemaIdentifier. This is used to identify the type of this Metadata Package.
* MetadataSyntaxIdentifier. This is used to identify the way of representing the metadata package in XML. We encourage the use of RDF.
* InformationPiece. This is a piece of information content within the Information Object.
* Label. This is a text string that labels the Information Piece.
* ContentFile. This represents a specific file in the VEO that represents the content of the InformationPiece.
* ContentFilePathName. This is the file name of the file that contains the content. It is relative to this directory.
* HashValue. The hash value resulting from applying the specified hash function (see HashFunctionAlgorithm above) to the sequence of octets forming the file.

The VEOContentSignature?.xml (where '?' is a number) files each contain a digital signature that allows detection of corruption of the VEOContent.xml file (and because the VEOContent.xml file includes hash values of the content files, corruption of the content files as well). The signature is generated by reading the VEOContent.xml file as a sequence of octets and applying specified signature algorithm. The contents of a VEOContentSignaturen.xml file are:
* SignatureAlgorithm. This identifies the hash algorithm and the

Step 4.
VEOContent.xml 파일 생성

- 최상위폴더에 생성
- 서브디렉토리 내 포함된 정보목록 제공
- 모든 콘텐츠 파일의 암호화 해시값 포함
- 재현포맷(PDF 등) 연결
- 정보/메타 논리적 구성

Step 5.
VEOContent.xml 파일 서명

- VEOContent.xml 생성 후 최소 1번 디지털 서명
- 해당 서명의 유효성 확인을 위한 서명과 인증서는 VEOContentSignature?.xml 파일에 저장
- 서명파일에 순번 부여

Step 6.
VEOHistory.xml 파일 생성

- 생성 이후 VEO 파일의 이벤트 히스토리 저장

히스토리 포함정보
- 1)이벤트 날짜/시각
- 2)이벤트 책임자/조직
- 3)이벤트 설명
- 4)이벤트 관련 오류

Step 7.
VEOHistory.xml 파일 서명

- VEOHistory.xml 생성 후 최소 1번 디지털 서명
- 해당 서명의 유효성 확인을 위한 서명과 인증서는 VEOHistorySignature?.xml 파일에 저장
- 서명파일에 순번 부여

Step 8.
ZIP 압축

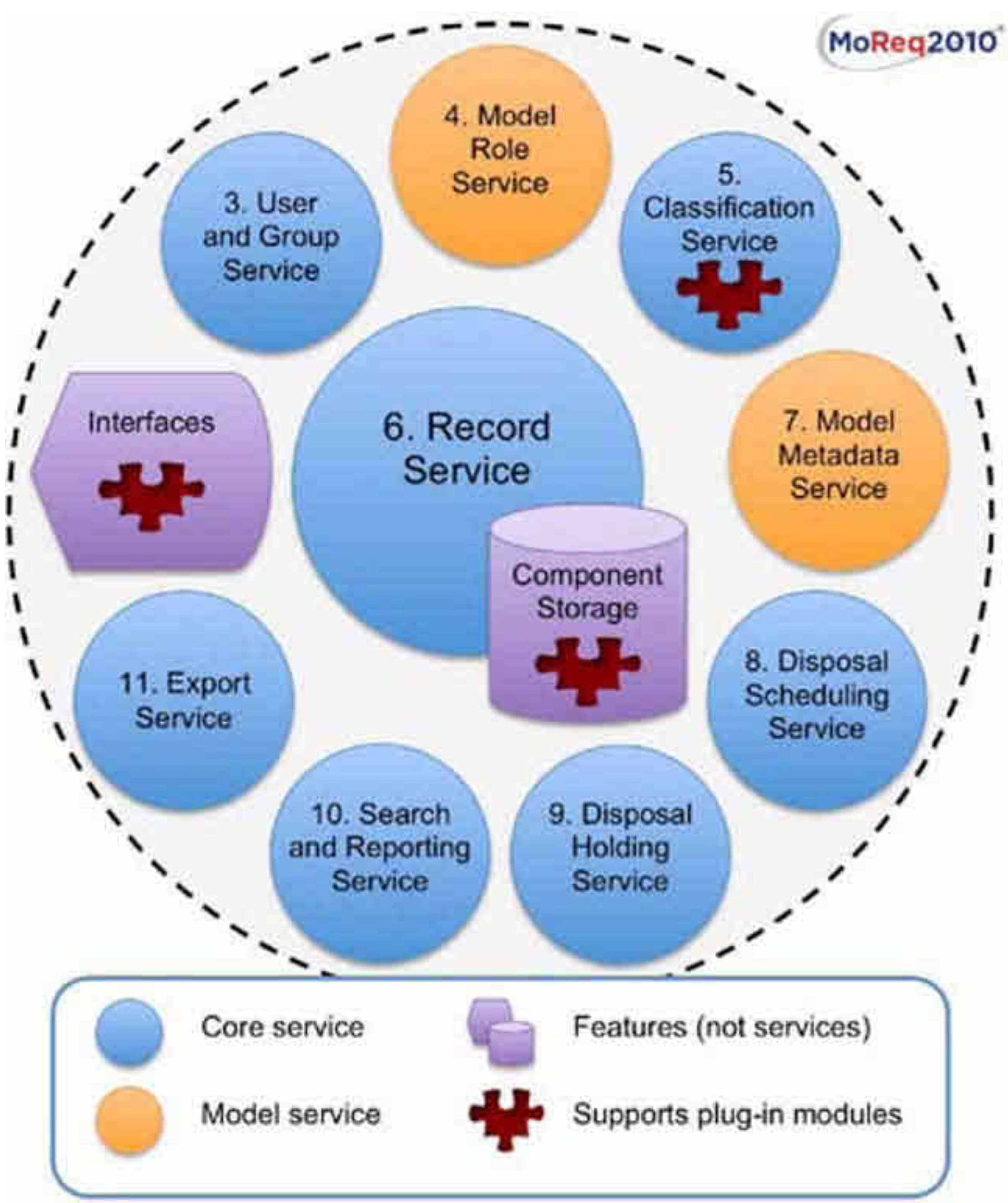
- VEO 디렉토리 압축
- 예 : TestVEO.veo.zip
- PKWARE APPNOTE.TXT 준수

ZIP 제한사항
- ZIP 항목명은 VEO 디렉토리명으로 시작
- 표준 'deflate' 메커니즘 사용
- 암호화 금지
- 디지털 서명 금지
- 패치 데이터 금지

1. 표준, 툴, 서비스

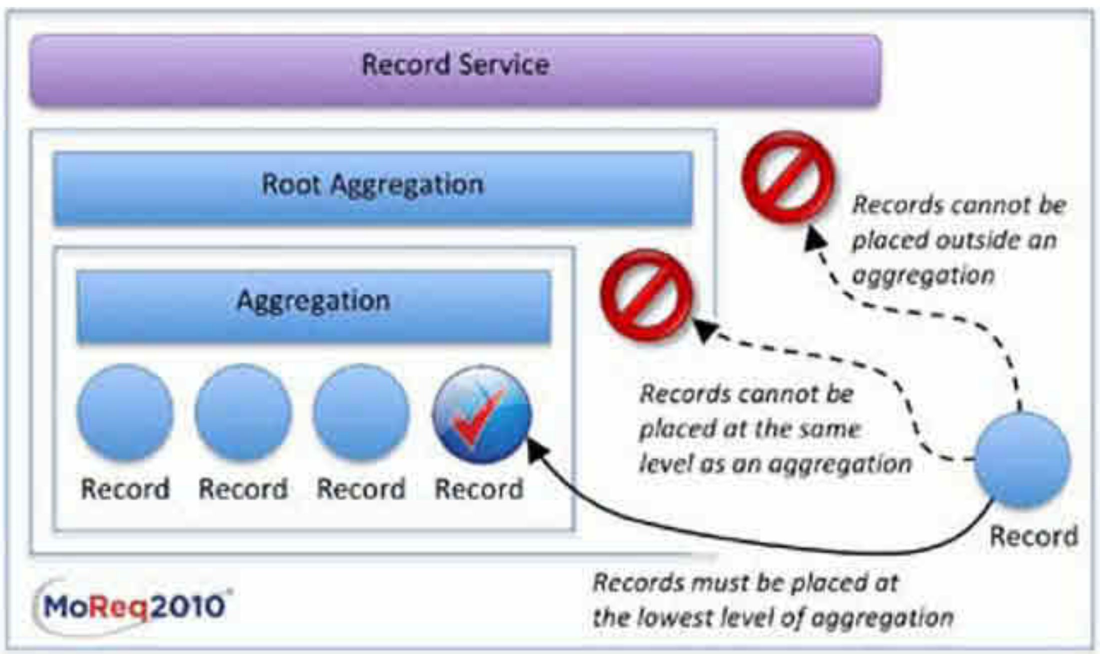
MoReq2010

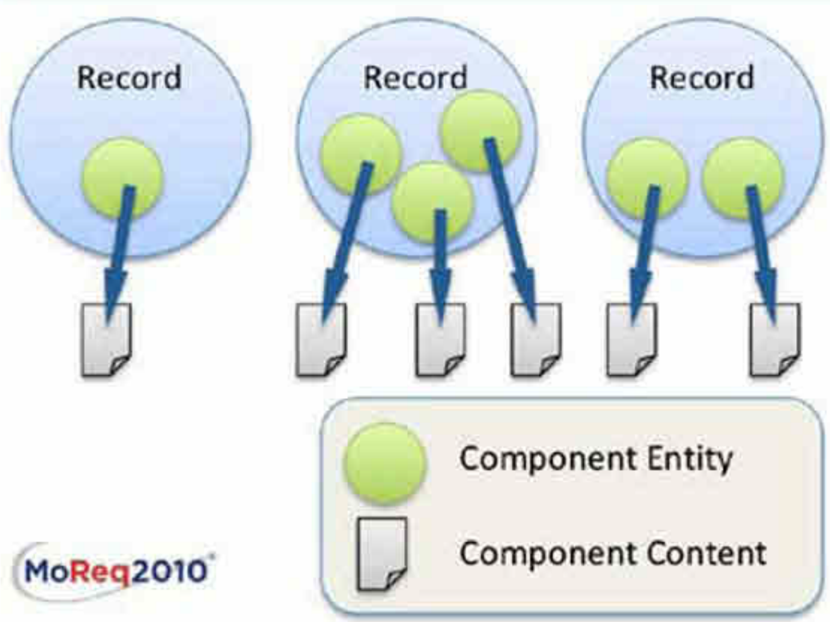
DLM Forum 등이 주도적으로 개발한 기록시스템 기능요건(장기보존과 구분)
MoReq2010은 모듈화, 집합체(Aggregation), 메타데이터와 콘텐츠 구분, 콘텐츠 단위 식별과 통제, 진본성 확인체계 등 기록시스템의 이상적인 개념 제시

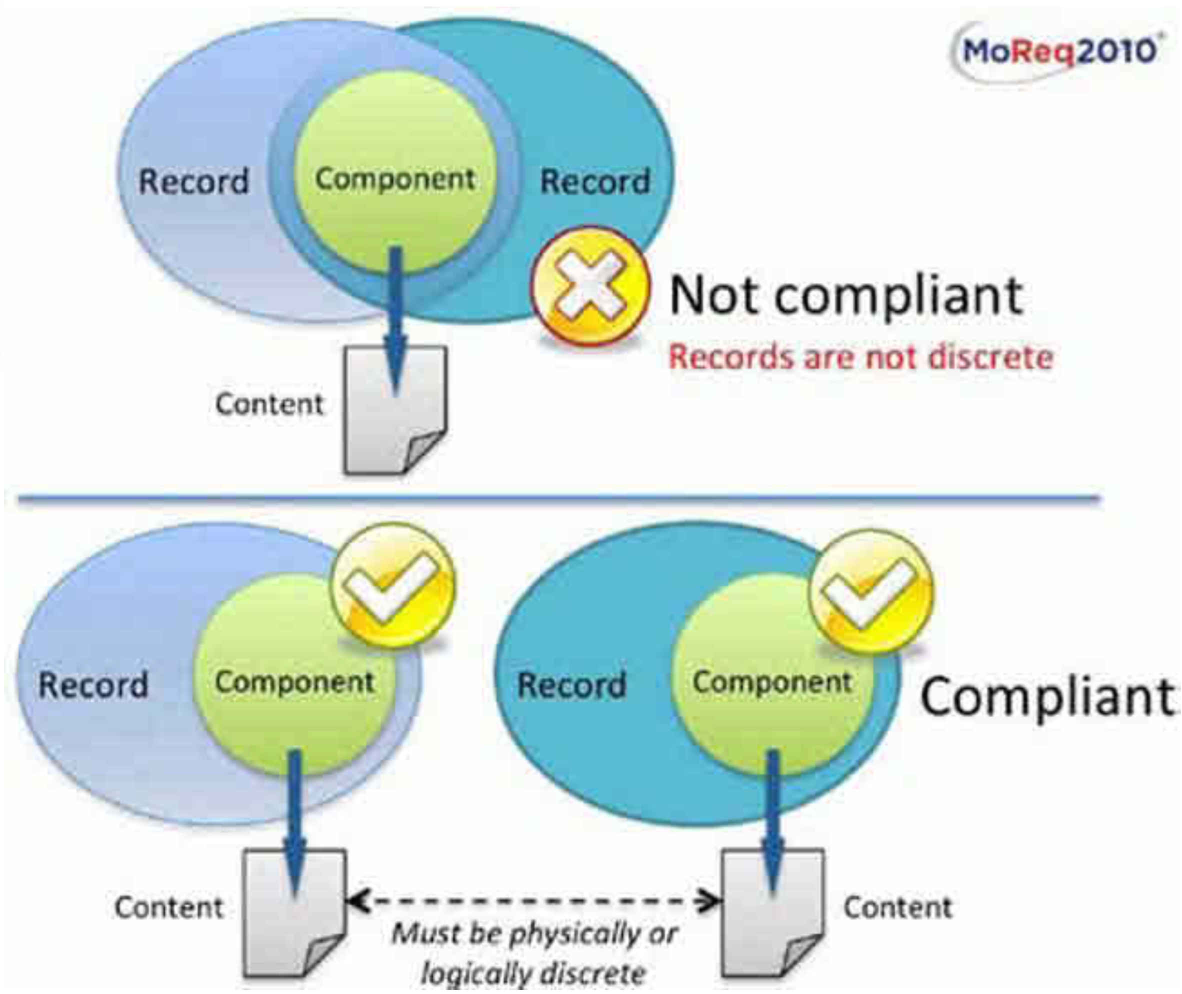


MoReq2010

집합체(Aggregation)와 레코드, 컴포넌트, 콘텐츠 개념은 RM과 AM 단계에 공통 적용 가능







1. 표준, 툴, 서비스

OAIS와 ISO 16363 컴플라이언스 측정 : Langley, 2018

최근 ISO 14721과 ISO 16363 표준을 기반으로 디지털 보존 수준을 평가/인증하는 다양한 지표들이 개발됨
2018년에는 이 지표들의 항목을 비교한 연구가 수행됨(Langley, 2018)

- Langley는 신뢰저장소(TDR), 성숙도, 위험관리, 워크플로우, DPOC 등 5개 분야 12개 평가지표의 123개 요건을 크로스워크 분석
- 보존계획 수립 및 시스템 설계 시 디지털 보존 정책, 전략, 계획 등을 총체적으로 고려하여 시스템 요구사항 개발 필요

Categories	Trusted Digital Repository (TDR) Certification			Digital Preservation Maturity Models					Risk Management Model		Digital Preservation Workflow Models	DPOC Requirements
Models	CoreTrustSeal (CTS) *formerly Data Seal of Approval (DSA) harmonised with World Data System (WDS)*	Trusted Repositories Audit and Certification (TRAC) *(basis for ISO 16363)*	Nestor Seal *(based on the DIN 31644 standard)*	Digital Preservation Capability Maturity Model (DPCMM) (Dollar & Ashley)	National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation	Three-Legged Stool for Digital Preservation (Survey of Institutional Readiness)	Assessing Organisational Readiness (AOR) *formerly Assessing Institutional Digital Assets (AIDA)*	Five Organisational Stages of Digital Preservation (Kenney & McGovern)	Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)	Controlled Lots Of Copies Keeps Stuff Safe (CLOCKSS) Threats Model and Mitigation Strategy	Digital Stewardship End-to-End Workflow Model (Langley)	Digital Preservation at Oxford and Cambridge (DPOC)
						Background Information						
Context	Context – R0											
								Organisational Stages				
Acknowledge								Acknowledge: Understanding that digital preservation is a local concern				
Act								Act: Initiating digital preservation projects				
Consolidate								Consolidate: Seguing from projects to programs				
Institutionalise								Institutionalise: Incorporating the larger environment				
Externalise								Externalise: Embracing inter-institutional collaboration and dependency				
									Risks Stages			
Identify Organisational Context									Stage 1 - Identify Organisational Context			
Document Policy and Regulatory Framework									Stage 2 - Document Policy and Regulatory Framework			
Identify Activities, Assets and Owners									Stage 3 - Identify activities, assets and their owners			
Identify Risks									Stage 4 - Identify Risks			
Assess Risks									Stage 5 - Assess the Identified Risks			
Manage Risks									Stage 6 - Manage Risks			
	Organisational Infrastructure	Organizational Infrastructure				Organisational Infrastructure						
Mission	Mission/Scope – R1					Mission	ORG 01: Mission Statement					
Mandate									T1: What is the mandate of your repository or the organisation in which it is embedded?			
Goals									T2: List goals and objectives of your repository			
Policy		Procedural accountability & policy framework		Policy		Policies and Procedures	ORG 02: Policies and Procedures					
Strategy				Strategy					T3: List your repository's strategic planning documents			
Governance		Governance & organizational viability		Governance								
Collaboration				Collaboration								
Authority						Authority	ORG 03: Sharing of Digital Content					
Implementation						Implementation						
Licences	Licences – R2	Contracts, Licenses and Liabilities	C6 Legal and contractual basis				ORG 10: Intellectual Property Rights and Rights Management					Access Rights

디지털 보존 평가지표 크로스워크

출처: Langley, S. (2018). Digital Preservation Maturity Model, Risks, Trusted Digital Repository Certification and Workflow Crosswalk [Dataset]. https://doi.org/10.17863/CAM.34266

1.신뢰저장소
- CTS
- TRAC
- Nester Seal

2.성숙도
- DPCMM
- NDSA레벨
- Three-Legged Stool for DP
- AOR
- Five Organisational Stages of DP

3.위험관리
- DRAMBORA
- CLOCKSS

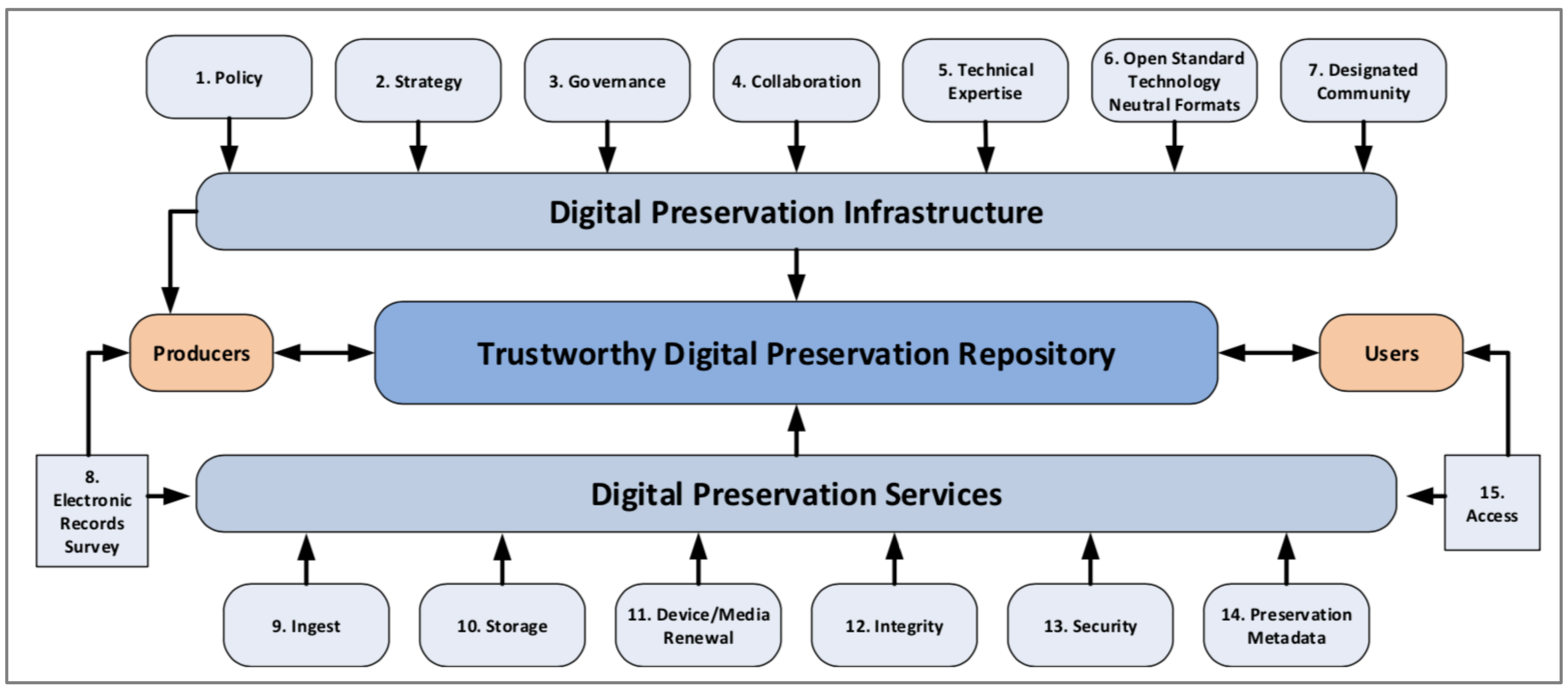
4.워크플로우
- Digital Stewardship End-to-End Workflow Model

5.DPOC 요건
- DPOC
-

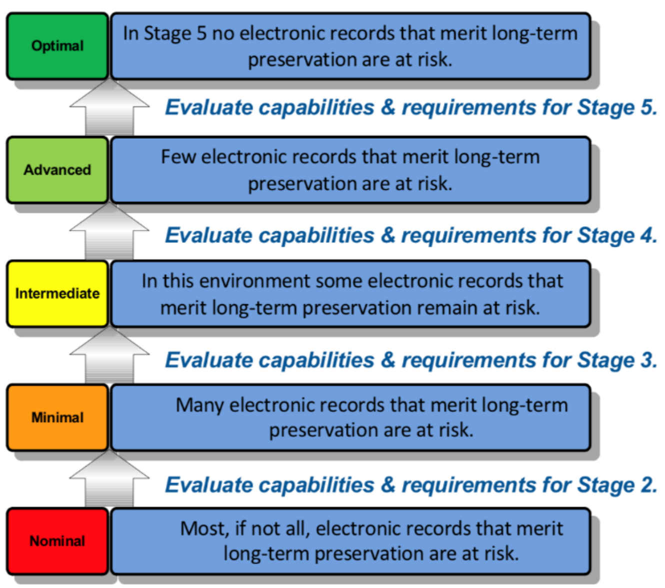
OAIS와 ISO 16363 컴플라이언스 측정 : DPCMM

DPCMM(Digital Preservation Capability Maturity Model)은 OAIS와 ISO 16363 두 표준의 추상성을 15개 구성요소로 상세화하여 실무담당자 스스로 평가 및 지속 모니터링이 가능하도록 함

- OAIS(ISO 24721)와 신뢰저장소인증기준(ISO 16363)의 보존 요구사항을 측정하는 메트릭(xlsx 서식) 제공



DPCMM 개념도
출처: http://www.securelyrooted.com/dpcmm

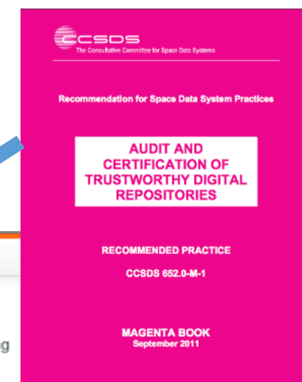


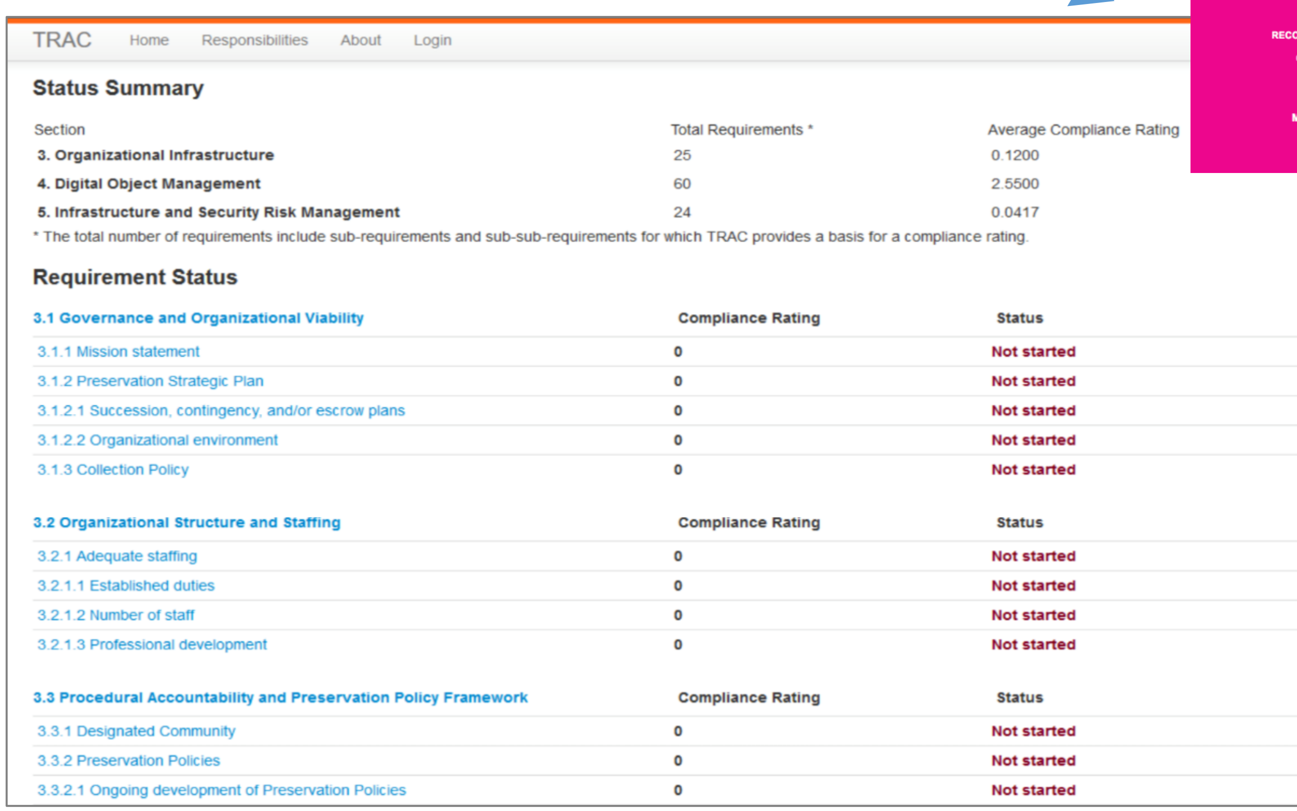
DPCMM 5단계
출처: http://www.securelyrooted.com/dpcmm

OAIS와 ISO 16363 컴플라이언스 측정 : Drupal TRAC Review Tool

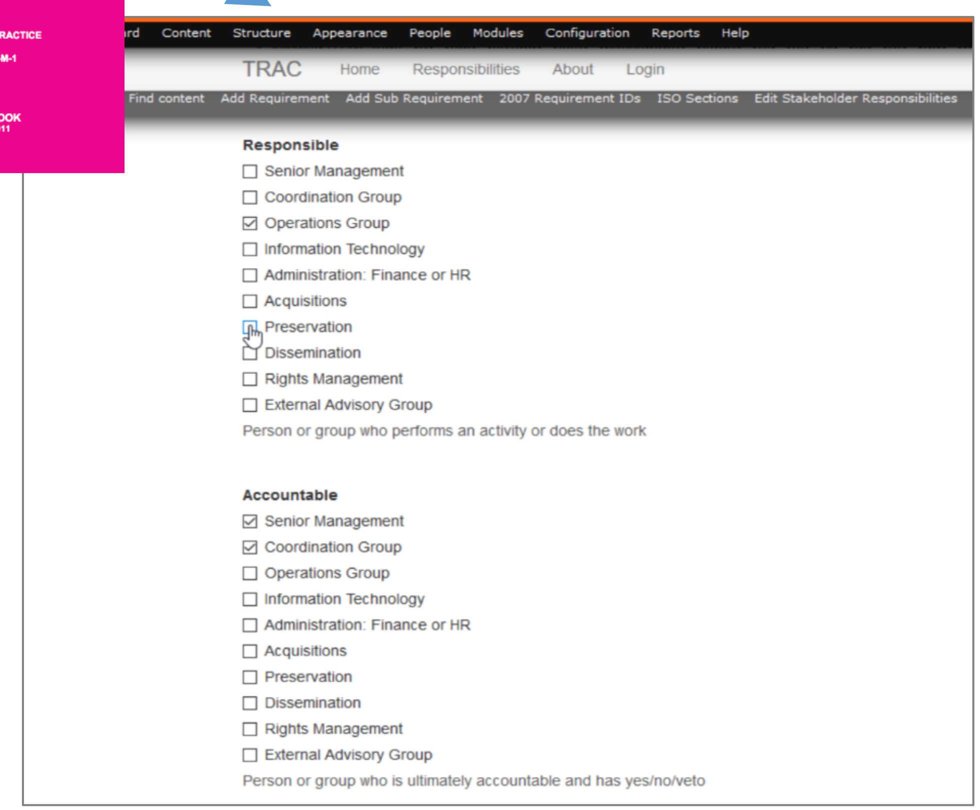
최근엔 TRAC 등 평가모델을 실행할 수 있는 웹 기반 소프트웨나 호스팅 서비스가 배포됨

- MIT 도서관에서 개발 및 배포, Artefactual에서 호스팅 서비스 제공 (Drupal 7.5 버전 기반)
- ISO 16363 의 체크리스트를 온라인 구현하여 각 요구사항별 조직의 역할을 통해 TRAC 준수 여부를 5단계 점수(0~4점)로 평가





TRAC 리뷰 툴 –개요 페이지
출처: Dan Gillean(2017). Digital Preservation with Archivematica: An Introduction.



TRAC 리뷰 툴 –실행 페이지
출처: Dan Gillean(2017). Digital Preservation with Archivematica: An Introduction.

1. 표준, 툴, 서비스

소프트웨어 레지스트리 : COPTR, POWRR

COPTR : 52개 기능별 600여 개 툴/서비스 목록 제공
POWRR : 24단계 보존 프로세스별 66개 툴 매핑

	Access, Use and Reuse	Create or Receive (Acquire)	Cross-Lifecycle Functions	Dispose	Ingest	Preservation Action	Preservation Planning	Store
Audio	2	4	3		11	10		
Binary Data			4					
Container						5		
Database		1	3		3	13		
Disk Image		6	3		2	1		
Document	3	1	4		33	14		
EBook					5	2		
Email			5		2	4		1
Geospatial					1			
Image	2	2	3		23	23		
Project Management Data	1					1		
Research Data	2	8	13		4		16	17
Software					1	2		
Spreadsheet					6	3		
Video	1	3	1		8	8		
Web	3	21	3		7	3	1	1
~Not Content Type Specific~	18	37	82	9	65	58	29	49

COPTR, http://www.digipres.org/tools/#dcc-ref

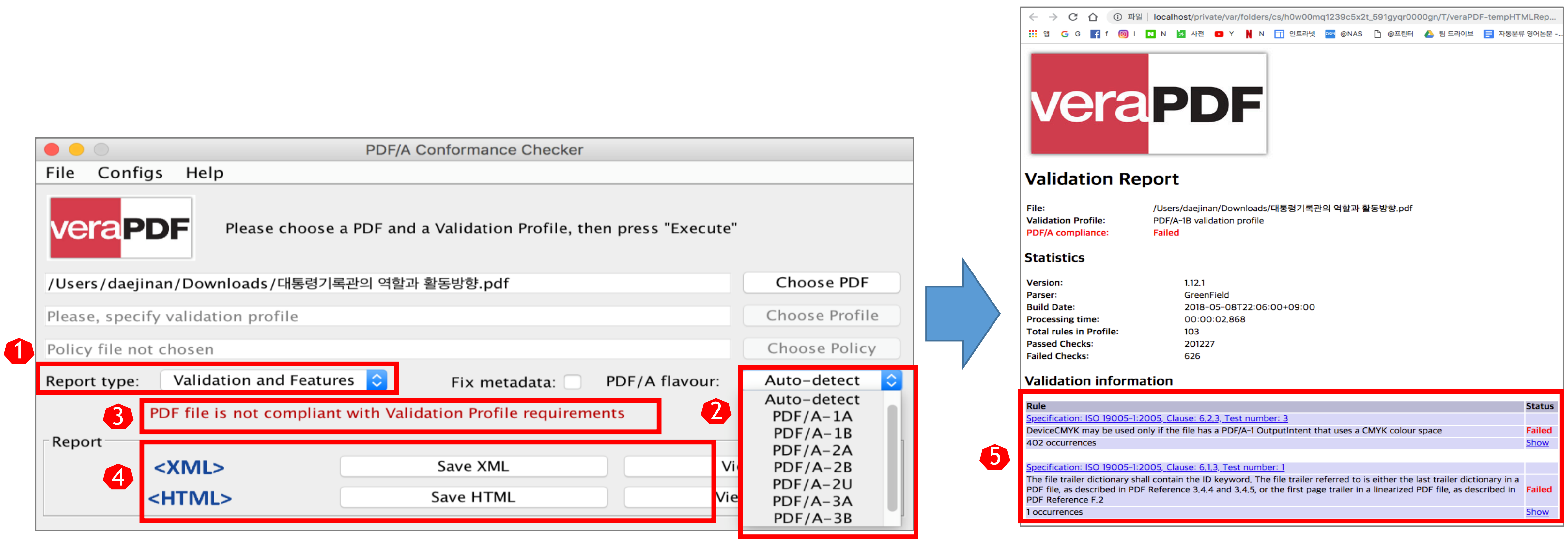
Digital POWRR Tool Evaluation Grid	Ingest					Processing						Access		Storage					Maintenance			Other		
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost
ACE (Audit Control Environment)		x																		x		x		Free
AFF Open Source Computer Forensics Software							x															x		Free
Amazon S3	x	x			x	x									x	x	x	x		x	x		x	Varies
Archive-It	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x		x		x	x	Varies
**Archivematica	x	x	x		x	x	x	x	x	x	x		x	x					x			x	x	Free
BagIt Library		x			x	x				x				x				x						
BagIt Transfer Utilities																								Free
BitCurator	x	x				x	x	x		x		x												free
BWF MetaEdit		x						x	x													x		free
Carbonite	x	x										x			x	x		x		x	x			$149 – $599
Chronopolis	x	x		x	x	x	x	x	x	x	x			x	x	x	x	x	x	x	x		x	$1,500 – $2,200
Cinch		x	x	x	x		x																x	Free
ContextMiner																						x	x	Free
**Curator's Workbench	x	x			x	x	x	x		x	x											x		Free
DAITSS	x	x	x		x	x				x	x		x	x			x	x						
DCape (ingest only)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				x		x		
DataVerse		x			x			x	x	x		x										x		Free

POWRR, http://digitalpowrr.niu.edu/tool-grid/

포맷 검증 : VeraPDF

유럽연합의 PREFORMA 프로젝트에서는 PDF 인코더별 변환 품질의 문제를 해결하고자 PDF 검증도구인 veraPDF를 개발하여 오픈소스로 배포

1. 리포트 유형 : Validation, Features, Validation and Features, Policy 중 선택
2. PDF/A 유형 선택 : ISO 19005 (PDF/A)의 세부항목 선택
3. 실행 결과 메시지 (2번에서 선택한 요건의 적합여부)
4. 리포트 저장 및 열람 (XML 또는 HTML)
5. ISO 19005 요건별 적합여부 및 Failed의 원인 설명

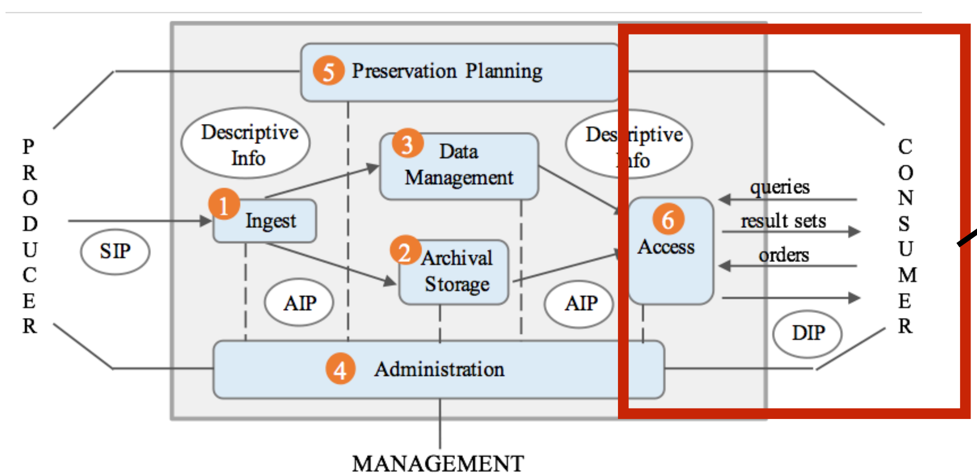


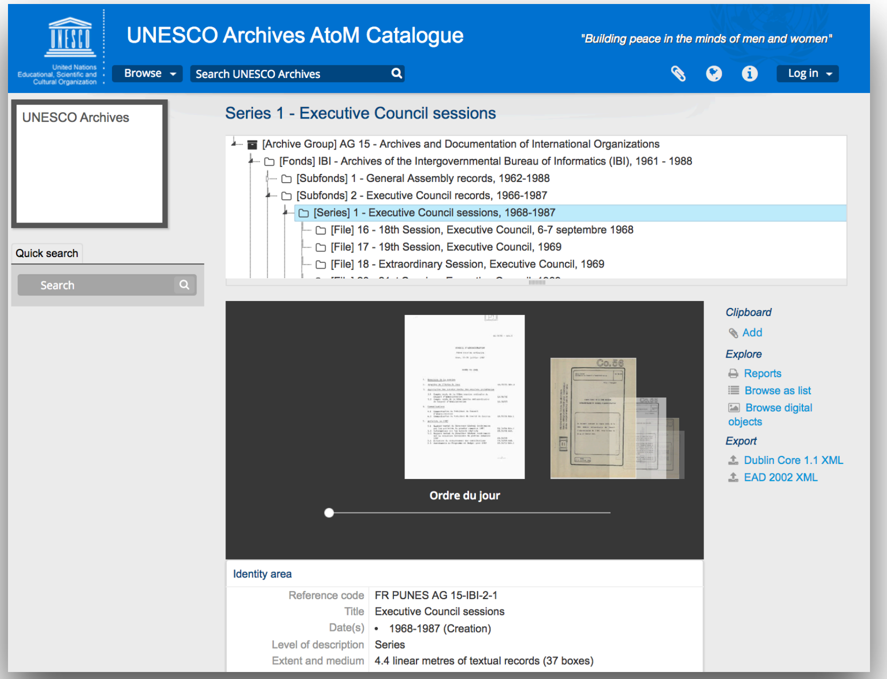
1. 표준, 툴, 서비스

카탈로깅 : AtoM, Rosetta, ArchivesSpace, CollectiveAccess 등

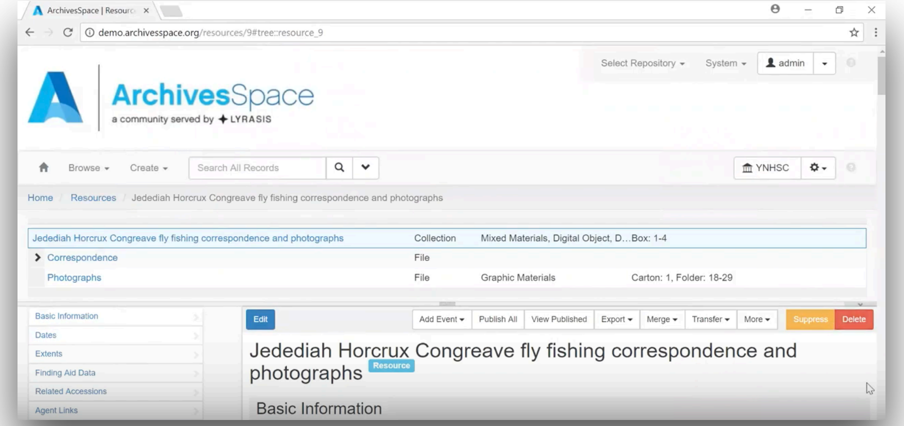
디지털 보존시스템과 카탈로깅시스템 연계 활용(ex. Archviematica to AtoM, Preservica to ArchivesSpace)

카탈로깅 시스템 : 국제/국가기록물기술표준, 더블린 코어 등을 지원하는 온라인 카탈로깅 소프트웨어





AtoM



ArchivesSpace

가상화 : Docker, Kubernetes 등

도커 등 컨테이너 기술은 OAIS를 지원하는 툴을 가상화하여 서비스 형태로 제공

- Docker : 컨테이너 기술

- Kubernetes : 컨테이너 관리(오케스트레이션) 툴

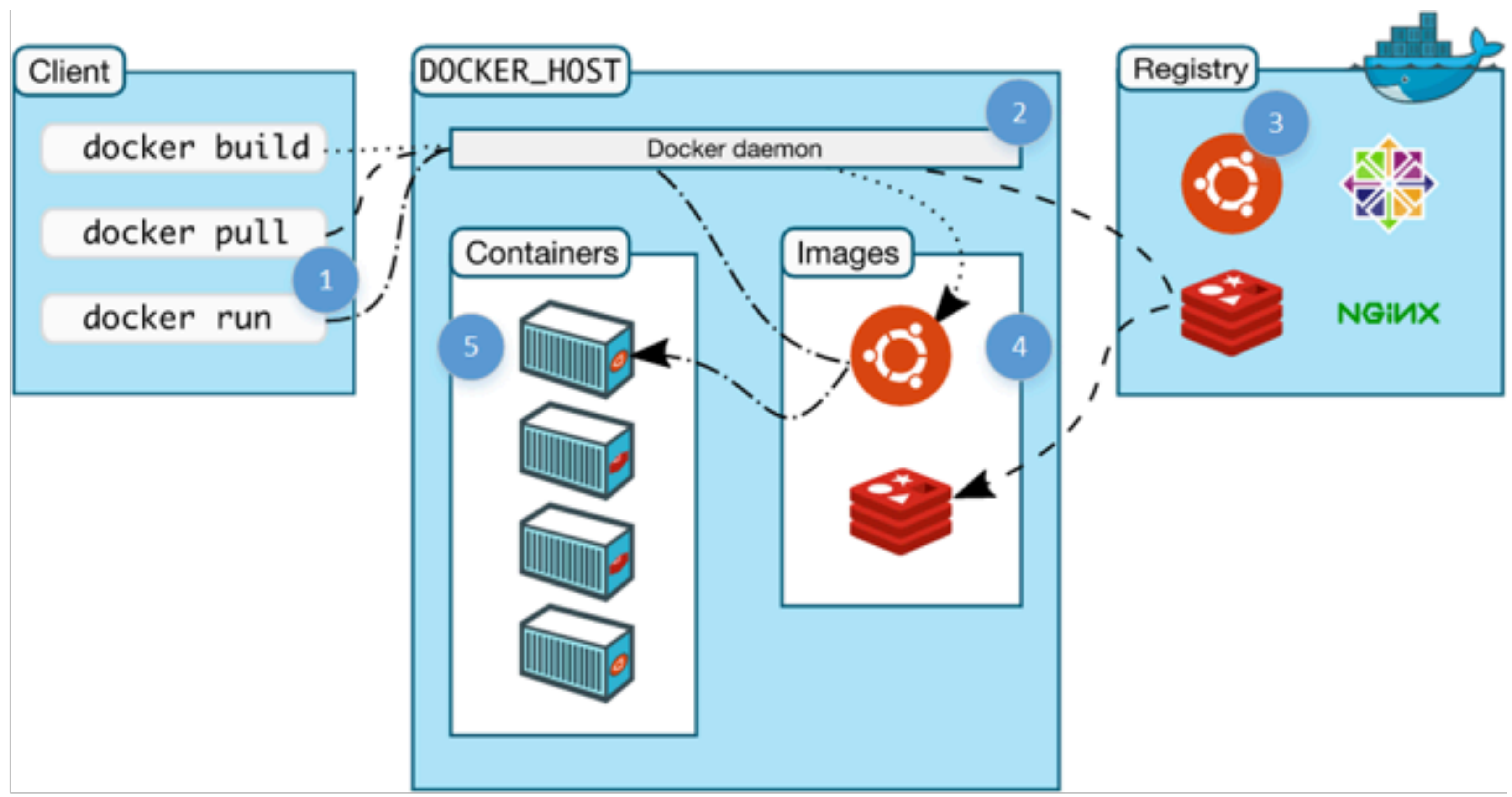
Docker Client
- 도커 조작 인터페이스(CLI)

Dockerfile
- 도커 이미지 조합을 위한 명령어 (텍스트 파일)

Images
- 도커파일로부터 정의된 도커 이미지(OS, App, DBMS 등)

Containers
- 도커 이미지 실행 인스턴스

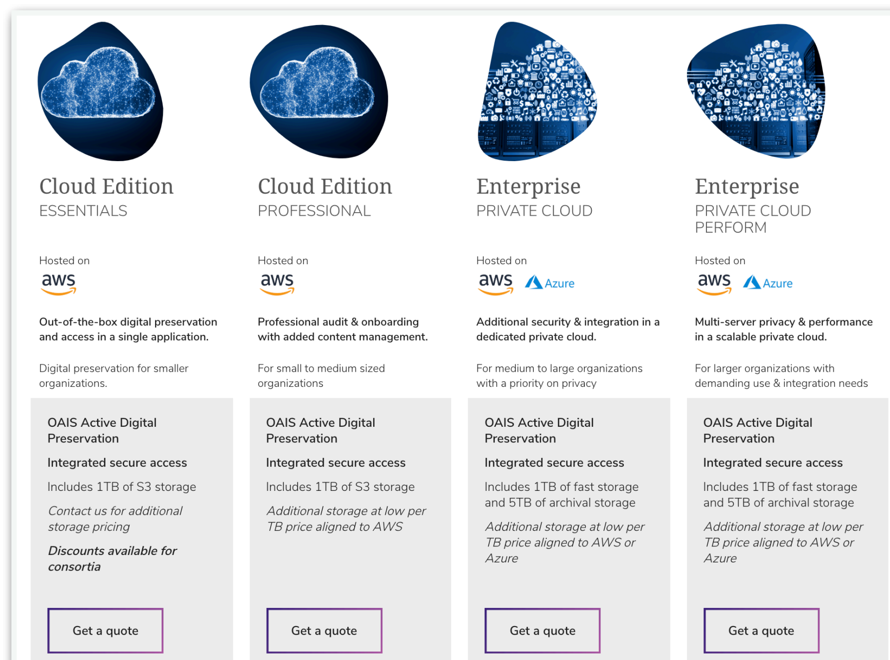
Registry
- 도커 이미지 저장소



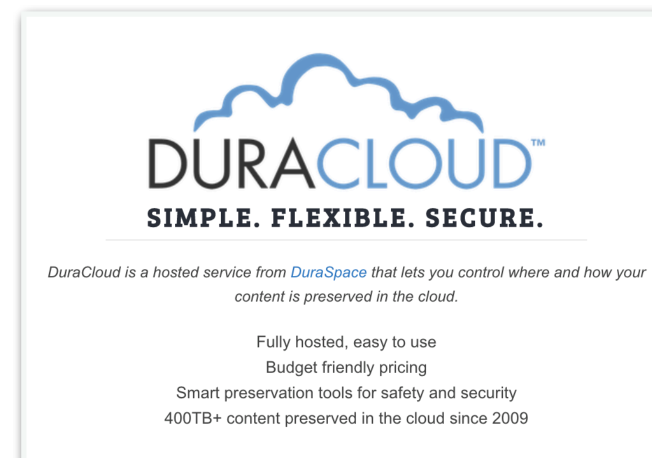
디지털 보존시스템

유럽, 북미의 기록소장기관을 중심으로 OAIS 요건을 충족하는 디지털 보존시스템, 저장소 등을 클라우드 형태의 구독 서비스로 활용

ArchivesDIRECT
(Archivematica 디지털 보존 서비스)



Preservica
(디지털 보존 서비스))



DuraCloud
(클라우드 스토리지 서비스)

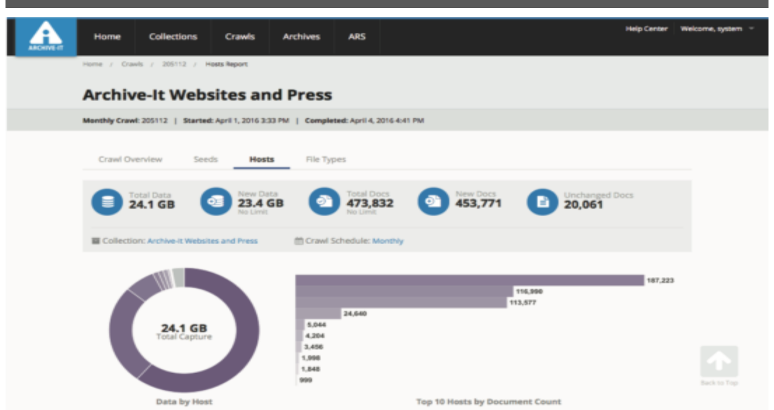
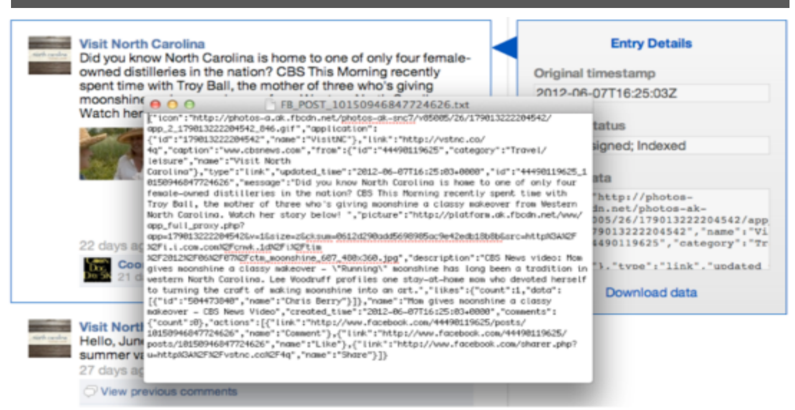
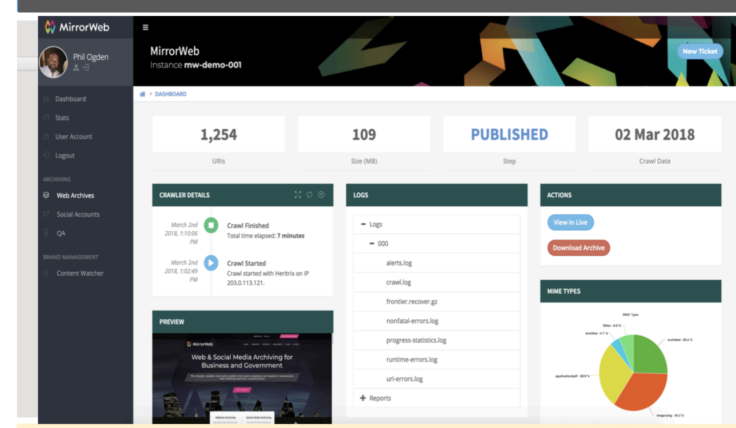
*출처: 국가기록원. 2017. 차세대기록관리모델 재설계.

1. 표준, 툴, 서비스

웹 아카이빙

2010년 이후 웹 아카이빙 기관들은 웹 크롤링, 보존, 재현 등 웹 아카이빙 업무를 전문 서비스 구독 형태로 전환

- Archive-It : 미국 샌프란시스코(웹/소셜미디어)
- ArchiveSocial : 미국 노스캐롤라이나(소셜미디어)
- Erado : 미국 댈러스(웹/소셜미디어)
- MirrorWeb : 영국 맨체스터(웹/소셜미디어)

Archive-It	ArchivesSocial	MirrorWeb
		
• 인터넷 아카이브(IA, Internet Archive)에서 제공하는 클라우드 웹 아카이빙 서비스 • 이용 기관들은 사용자 친화적인 웹 기반 인터페이스와 자동화된 품질평가도구 등을 이용하여 쉽게 웹 자원을 수집, 보존, 관리할 수 있음 • 수집 과정에서의 인덱싱을 통해 보존된 자원의 풀 텍스트 검색 제공 • Archive-It을 통해 수집된 웹 자원은 인터넷 아카이브에 축적되며 Wayback-machine 툴을 통해 일반인들에게 공개됨 • 2016년 5월 현재 전 세계 16개국의 도서관, 박물관, 아카이브, 정부기관, 자치단체, 비영리단체 등 459 기관이 파트너로 참여하여 3,487개 컬렉션을 수집함 • Internet Archive와의 협약을 통해 서비스 이용이 가능하며 클라우드 저장 용량에 따라 1년 단위 과금	• 미국 노스캐롤라이나 주 더럼(Durham) 소재의 웹 아카이빙 서비스 제공자로 정부기관, 법률기관, 금융기관, 교육기관 대상 웹 아카이빙 서비스 제공 • Facebook, Twitter, Instagram, Youtube 등 소셜 미디어 아카이빙 기능에 특화되어 있으며 오바마 대통령의 소셜 미디어 아카이브(Obama White-House Social Media Archive) 구축 사례로 널리 알려짐 • 월별 웹 기록 수집 건 수에 따라 절약형, 표준형, 고급형의 요금제를 제공하며 경제형의 경우 10개의 사이트만을 아카이빙할 수 있고 모든 요금제에 무제한 저장 용량 제공	• 영국 맨체스터 소재의 웹 아카이빙 서비스 제공자로 정부기관, 금융기관, 기업 웹페이지 및 소셜 미디어 아카이빙 서비스 제공 • **eDiscovery 지원 – MirrorWeb을 통해 아카이빙된 웹 기록을 eDiscovery 전문가, 변호사, 기타 조사원들이 검색/활용하는 기능 제공** • **영국국립기록관리청(TNA), 영국은행(Bank of England) 등 유럽의 다수 정부기관 활용** • **Facebook, Twitter, Youtube, LinkedIn, Instagram, Vimeo 등 소셜미디어 6종 지원**

*출처: 대통령기록관. 2019. 대통령기록 정보시스템 고도화 ISP

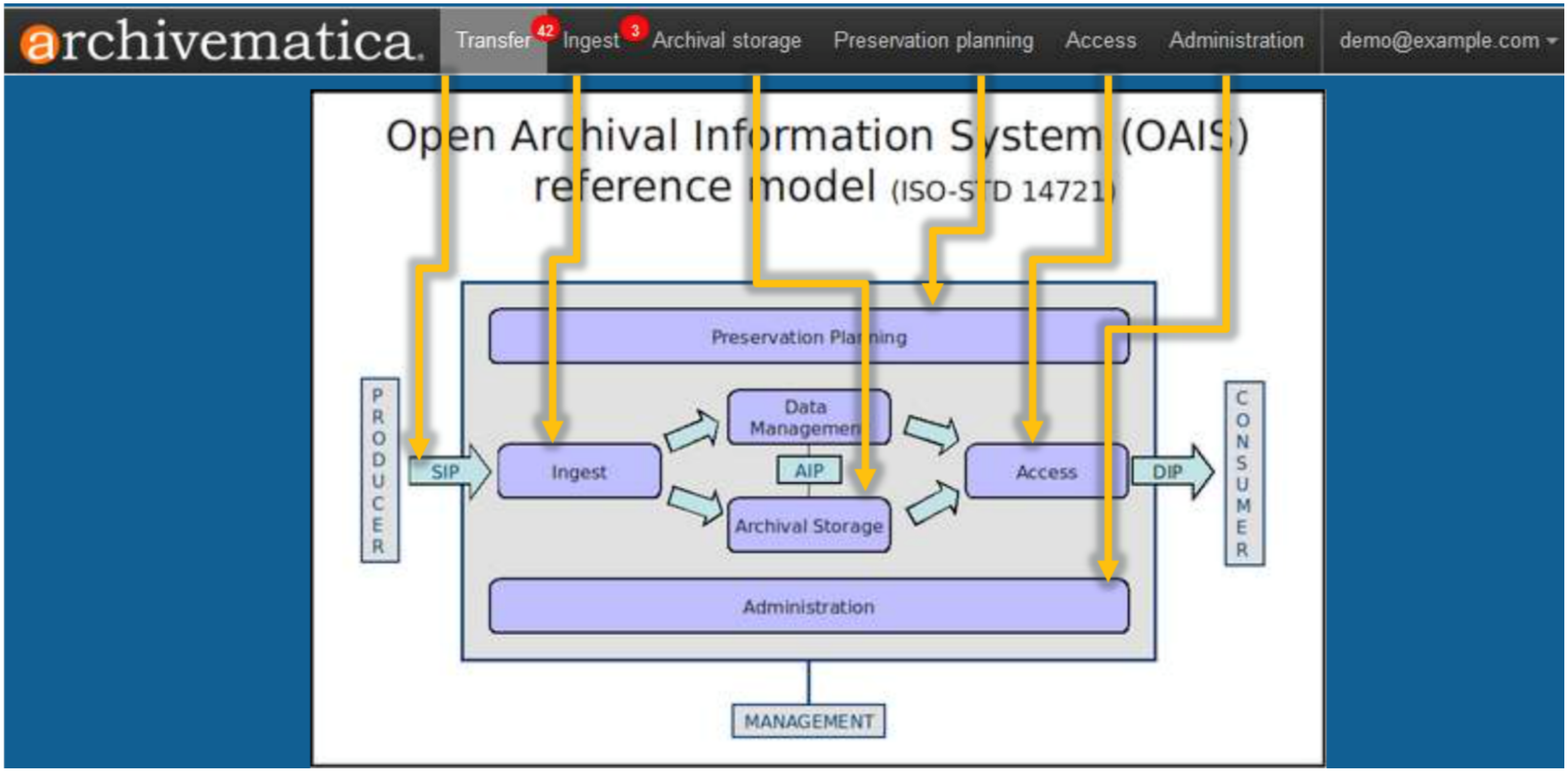
2. 시스템

개요

UNESCO 등의 펀딩으로 2012년 캐나다의 Artefactual 사에서 개발

OAIS의 주요 기능 엔티티를 시스템의 메뉴로 구성

(Ingest, Archival Storage, Preservation Planning, Access, Administration)



사용자

주로 북미지역 대학도서관, 국제기구, 미술관/박물관/아카이브에서 사용

artefactual systems inc.

CITY OF VANCOUVER ARCHIVES

UBC

ROCKEFELLER ARCHIVE CENTER

UNESCO
United Nations Educational, Scientific and Cultural Organization

SFU

UNIVERSITY OF ALBERTA

UNIVERSITY OF SASKATCHEWAN University Library

MoMA The Museum of Modern Art

Yale UNIVERSITY LIBRARY

ZIB

HARVARD BUSINESS SCHOOL Knowledge & Library Services BAKER 3.0

COPPUL

BENTLEY HISTORICAL LIBRARY UNIVERSITY OF MICHIGAN

DURASPACE™

MIT Libraries

OCUL Ontario Council of University Libraries

UNIVERSITY of York

UNIVERSITY OF Hull

LLGC NLW LLYFRGELL GENEDLAETHOL CYMRU THE NATIONAL LIBRARY OF WALES

Canadian Council of Archives
Conseil canadien des archives

international institute of social history

2. 시스템

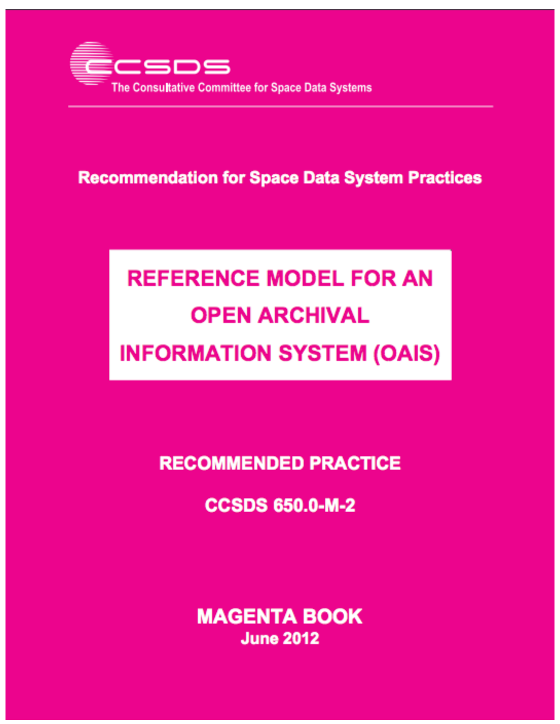
특징

표준 : OAIS, PREMIS, METS, BagIT 등 국제기록관리 커뮤니티 표준 적용

오픈소스 : 기 개발된 30여 종의 오픈소스 활용

FPR 서비스 : 수백 종의 파일포맷 식별 및 처리

마이크로서비스 : 각 기능을 세분하여 독립적 처리



CCSDS
The Consultative Committee for Space Data Systems

Recommendation for Space Data System Practices

REFERENCE MODEL FOR AN OPEN ARCHIVAL INFORMATION SYSTEM (OAIS)

RECOMMENDED PRACTICE

CCSDS 650.0-M-2

MAGENTA BOOK
June 2012

ISO 14721 (OAIS)

PREMIS Data Dictionary for Preservation Metadata

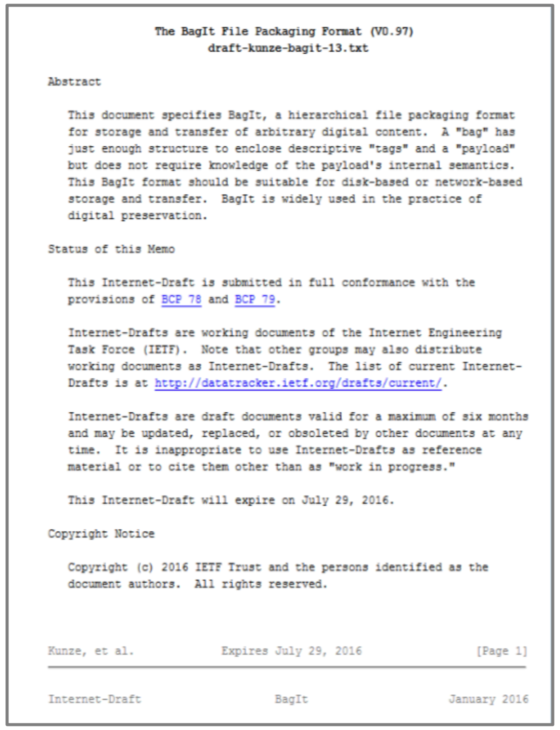
Contents:

Acknowledgments
Introduction
Background
The PREMIS Data Model
General Topics on Structure & Use
Implementation Considerations
The PREMIS Data Dictionary Version 3.0
Special Topics
Glossary

PREMIS

<METS>
METADATA ENCODING AND TRANSMISSION STANDARD:
PRIMER AND REFERENCE MANUAL

METS



The BagIt File Packaging Format (V0.97)
draft-kunze-bagit-13.txt

Abstract

This document specifies BagIt, a hierarchical file packaging format for storage and transfer of arbitrary digital content. A "bag" has just enough structure to enclose descriptive "tags" and a "payload" but does not require knowledge of the payload's internal semantics. This BagIt format should be suitable for disk-based or network-based storage and transfer. BagIt is widely used in the practice of digital preservation.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on July 29, 2016.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

Kunze, et al. Expires July 29, 2016 [Page 1]

Internet-Draft BagIt January 2016

BagIT

마이크로서비스 아키텍처

전자기록 입수부터 제공까지의 각 기능을 마이크로서비스 수준까지 잘게 나누어 처리
마이크로서비스는 각 기능의 독립적 실행,배포,운영,개선이 가능하여 MoReq2010의 핵심개념을 실현

Submission Information Package	UUID	Ingest start time	
Test_transfer	b7d5f60d-3f40-4ac5-b49b-b8079c52be55	2016-10-06 14:50	
Micro-service: Store AIP			
Micro-service: Upload DIP			
Micro-service: Prepare AIP			
Micro-service: Prepare DIP			
Micro-service: Generate AIP METS			
Micro-service: Verify checksums			
Micro-service: Process metadata directory			
Micro-service: Process submission documentation			
Micro-service: Transcribe SIP contents			
Micro-service: Add final metadata			
Micro-service: Process manually normalized files			
Micro-service: Normalize			
Micro-service: Clean up names			
Micro-service: Rename SIP directory with SIP UUID			
Micro-service: Verify SIP compliance			
Micro-service: Remove cache files			
Micro-service: Include default SIP processingMCP.xml			
Micro-service: Verify transfer compliance			

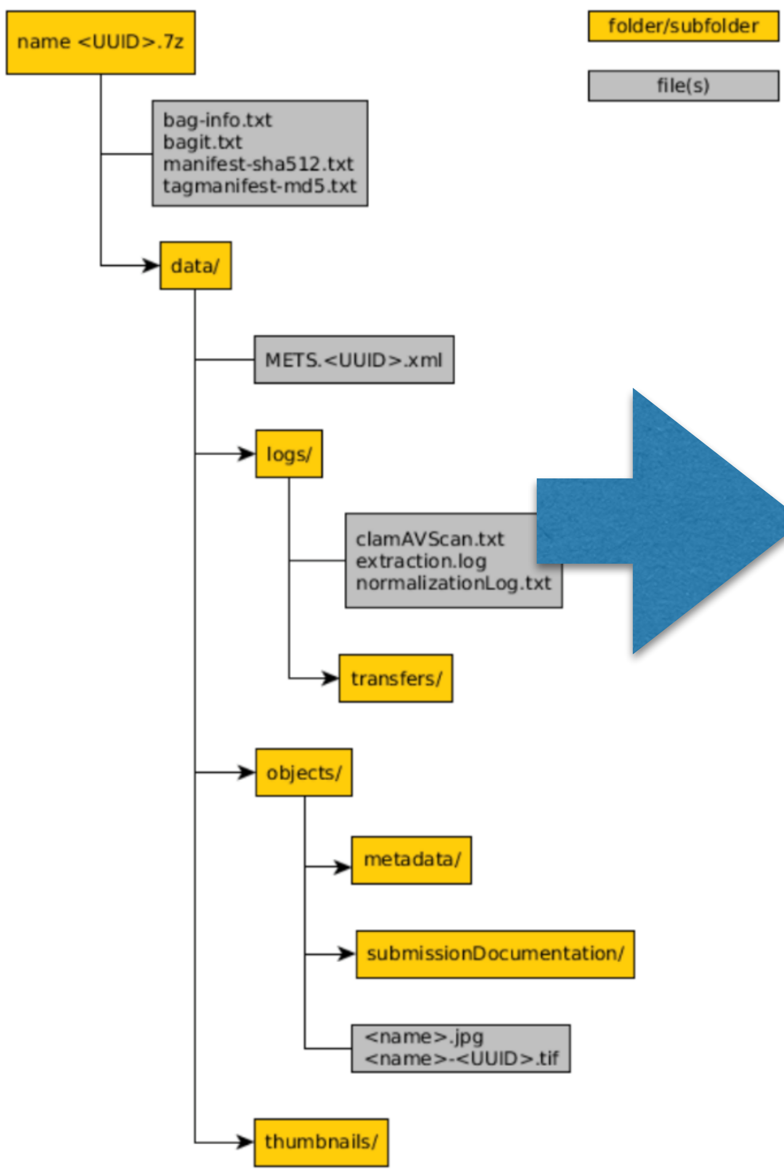
마이크로서비스 아키텍처

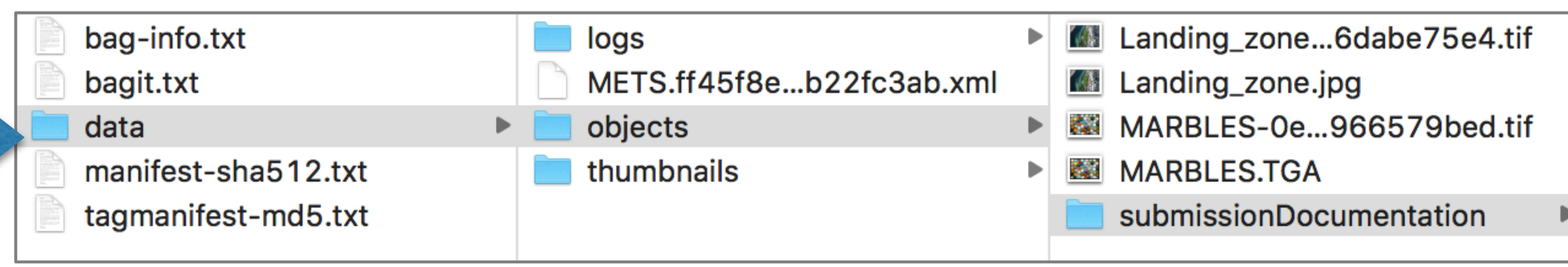
전자기록 입수부터 제공까지의 각 기능을 마이크로서비스 수준까지 잘게 나누어 처리
마이크로서비스는 각 기능의 독립적 실행,배포,운영,개선이 가능하여 MoReq2010의 핵심개념을 실현

Micro-service: Normalize		
Set file permissions	Completed successfully	
Move to processing directory	Completed successfully	
Approve normalization [?]	Completed successfully	
Move to approve normalization directory	Completed successfully	
Remove files without linking information (failed normalization artifacts etc.)	Completed successfully	
Set file permissions	Completed successfully	
Normalize for preservation	Completed successfully	
Normalize for access	Completed successfully	
Normalize for thumbnails	Completed successfully	
Create thumbnails directory	Completed successfully	
Create DIP directory	Completed successfully	
Move to processing directory	Completed successfully	
Normalize [?]	Completed successfully	
Resume after normalization file identification tool selected.	Completed successfully	
Identify file format	Completed successfully	
Select pre-normalize file format identification command	Completed successfully	
Move to select file ID tool	Completed successfully	
Grant normalization options for no pre-existing DIP	Completed successfully	
Set remove preservation and access normalized files to renormalize link.	Completed successfully	
Check for Access directory	Completed successfully	
Check for Service directory	Completed successfully	
Identify manually normalized files	Completed successfully	

SIP, AIP, DIP (정보 패키지)

Archivematica는 BagIt, METS, PREMIS 기반으로 SIP, AIP, DIP를 구현



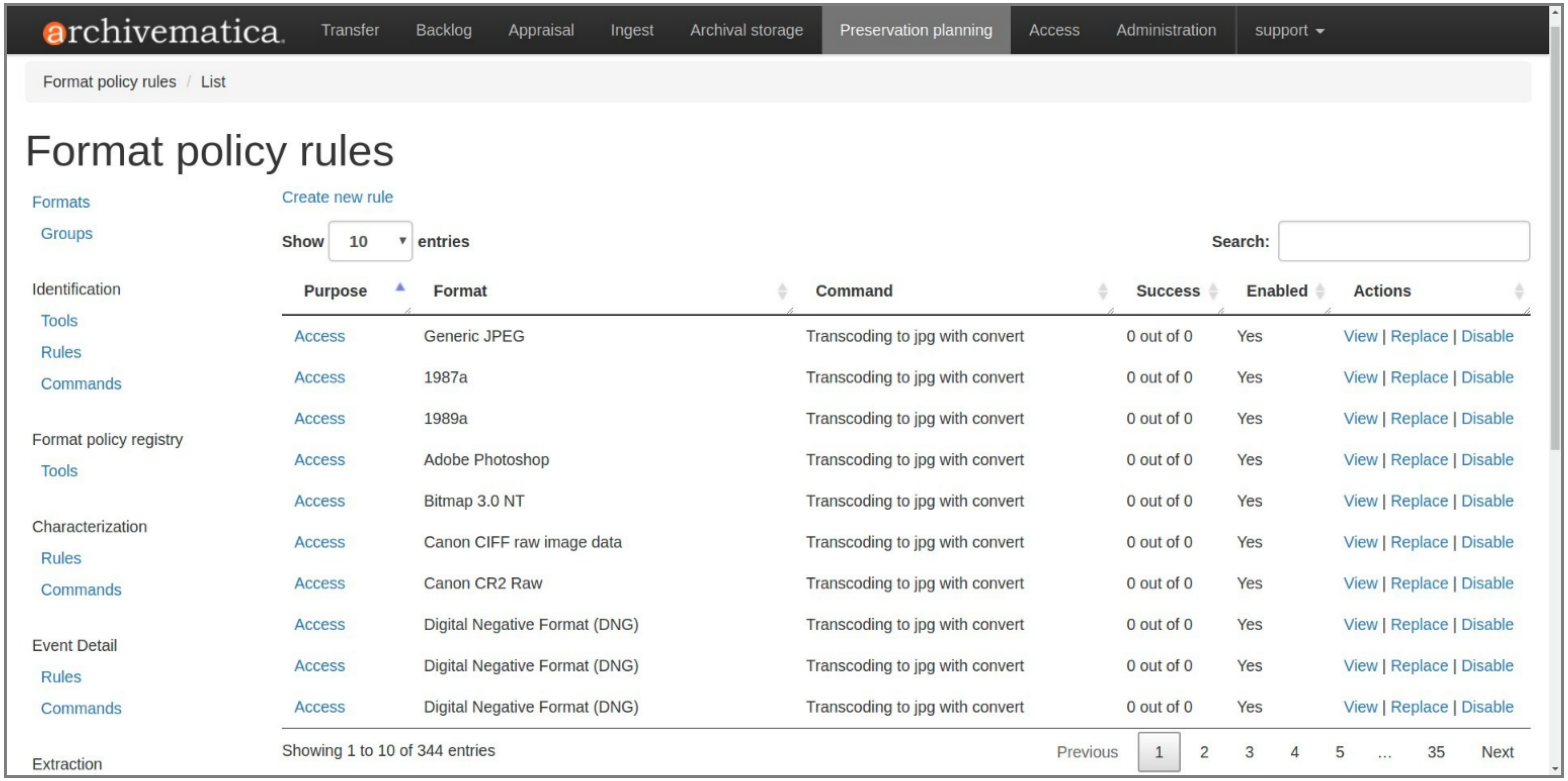


AIP 적용 사례

AIP 개념도

FPR(Format Policy Registry, 포맷정책 레지스트리)

FPR서버를 통해 파일 포맷의 식별, 특성 추출, 불변성체크(fixity), 포맷변환 수행
FPR은 디지털 보존 계획(Preservation Planning)의 핵심도구로 쓰임

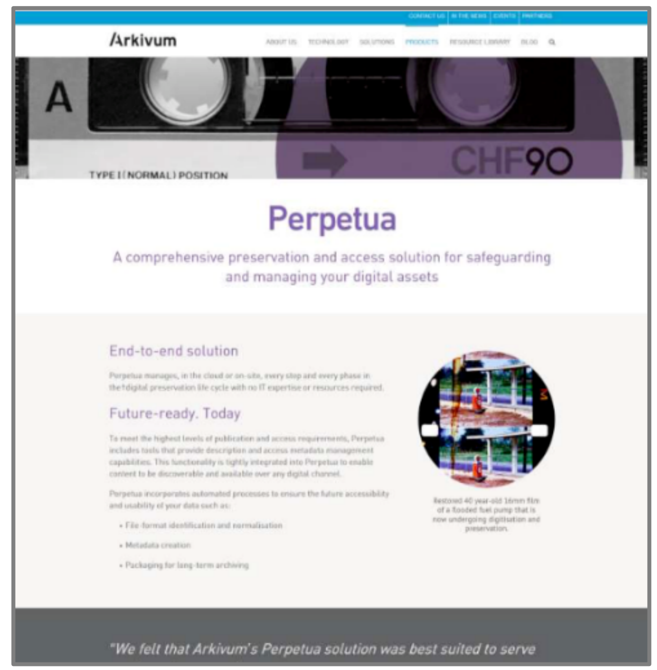


Archivematica 의 포맷정책규칙(FPR)
출처: https://www.archivematica.org/en/docs/archivematica-1.9/_images/normalization-rules.png

클라우드 서비스

Artefactual사는 2014년 미국(ArchivesDIRECT), 2016년 영국(Perpetua)과 캐나다(ArchivesCANADA)의 로컬 클라우드 서비스 공급자와 제휴하여 클라우드 기반 디지털 보존 호스팅 서비스 제공

ArchivesDIRECT
(미국)



Perpetua
(U.K.)

ArchivesCANADA
(캐나다)

구독모델 (ArchivesDIRECT)

- **디지털보존 평가 플랜 :** 3개월 컨설팅, -500GB 제공 ($4,500)
- **표준 플랜 :** Archivematica 전기능 활용, -1TB 제공 ($9,999)
- **프로페셔널 :** 대규모 또는 복잡한 구현

2. 시스템

개요

TNA는 영국의 아카이브시스템 개발업체인 Tessela와 함께 Preservica를 개발
이후 Preservica는 이를 엔터프라이즈 버전으로 확장하여 유럽 다수의 내셔널 아카이브에 클라우드 서비스 제공

National & Pan-National

UK National Archives
Malaysian Archives
European Commission
Dutch National Archives
Swiss Federal Archives
Austrian Archives
National Archives of Hungary
Finnish National Archives
Latvian National Archives
Estonian National Archives

Libraries, Museums & Education

DC public library
wellcome library
National Library of Australia
Grand Valley State University
Colby-Sawyer College
Hagley
Emerson College
Bates College
UCD Dublin
Aberystwyth University
University of Glasgow
University of Dundee
northumbria University
University of Liverpool
University College London

State & Government

Archives of Michigan
Kentucky Department for Libraries and Archives
Wisconsin Historical Society
State of Vermont Archives
UK Met Office
Dorset
UK Parliament
Rotterdam City Archive
Budapest City Archive

Business & Corporate

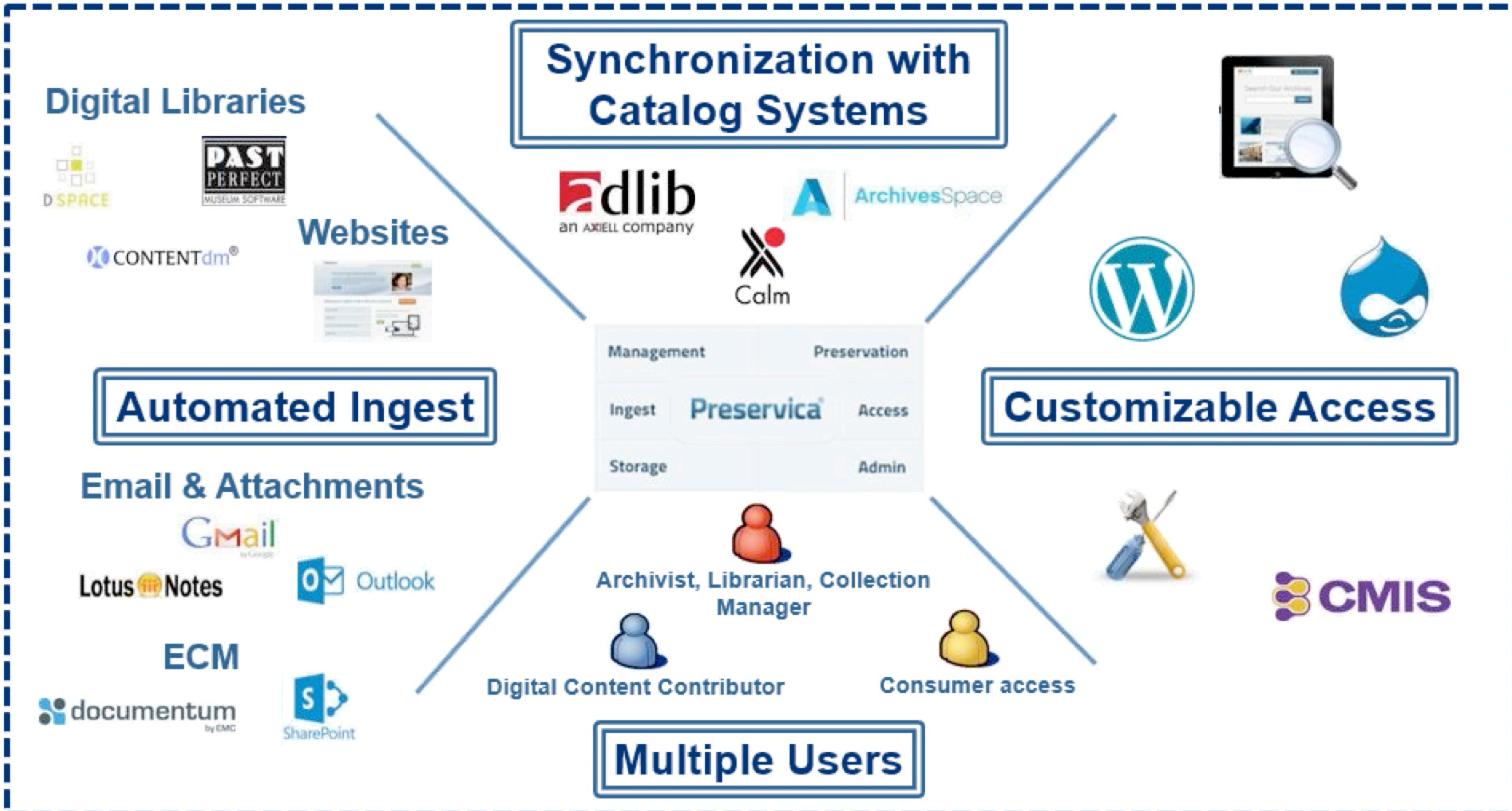
Unilever
HSBC Corporate Archives
Science & Technology Facilities Council
FamilySearch

OAIS 확장

Preservica는 OAIS의 모든 요건을 만족시킬 뿐만 아니라, 대부분의 업무를 워크플로우로 자동화하고, Access 등 아카이브 소장물 검색 및 이용자 편의기능을 대폭 향상시킴

CMIS

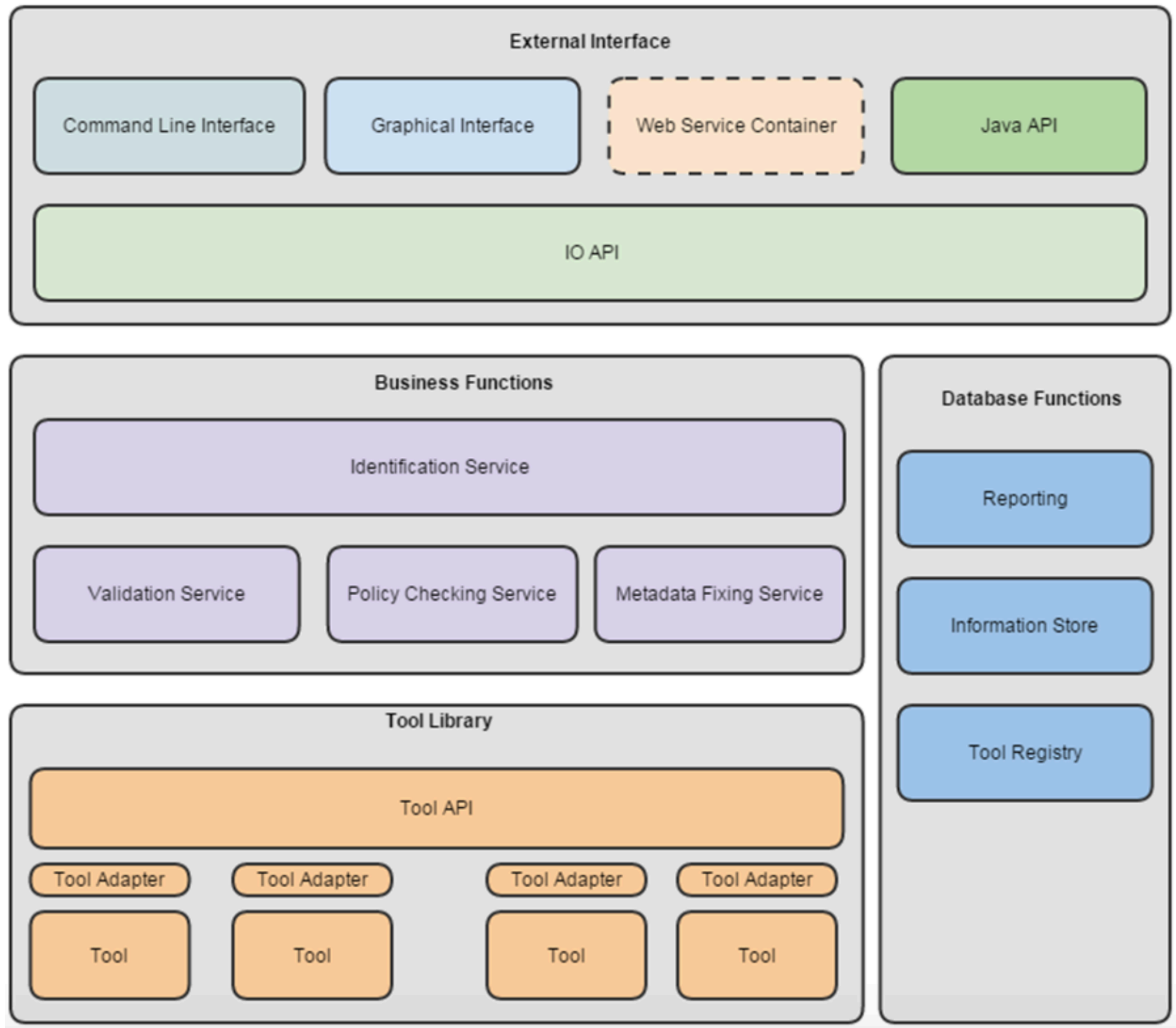
주요 기록 생산시스템, 서비스시스템 등과 연결 인터페이스 제공



2. 시스템

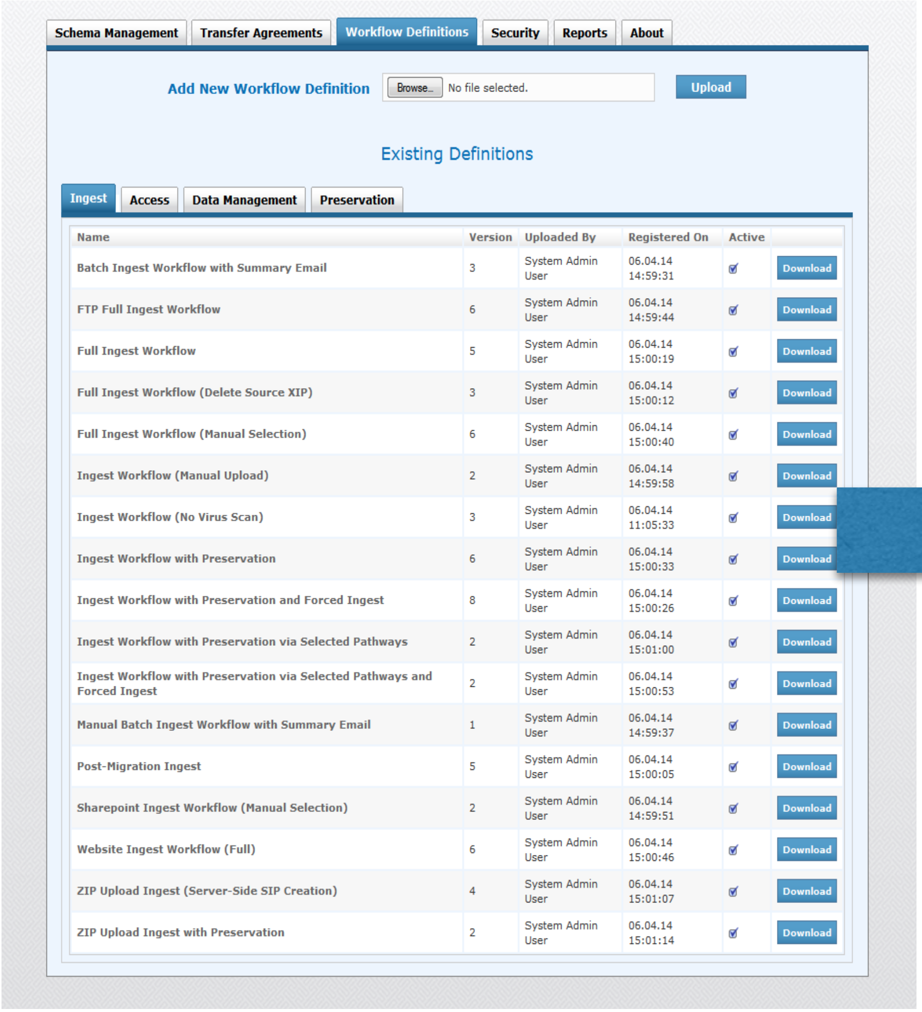
아키텍처

Tool Library, Business Function, Database Function, External Interface로 구성

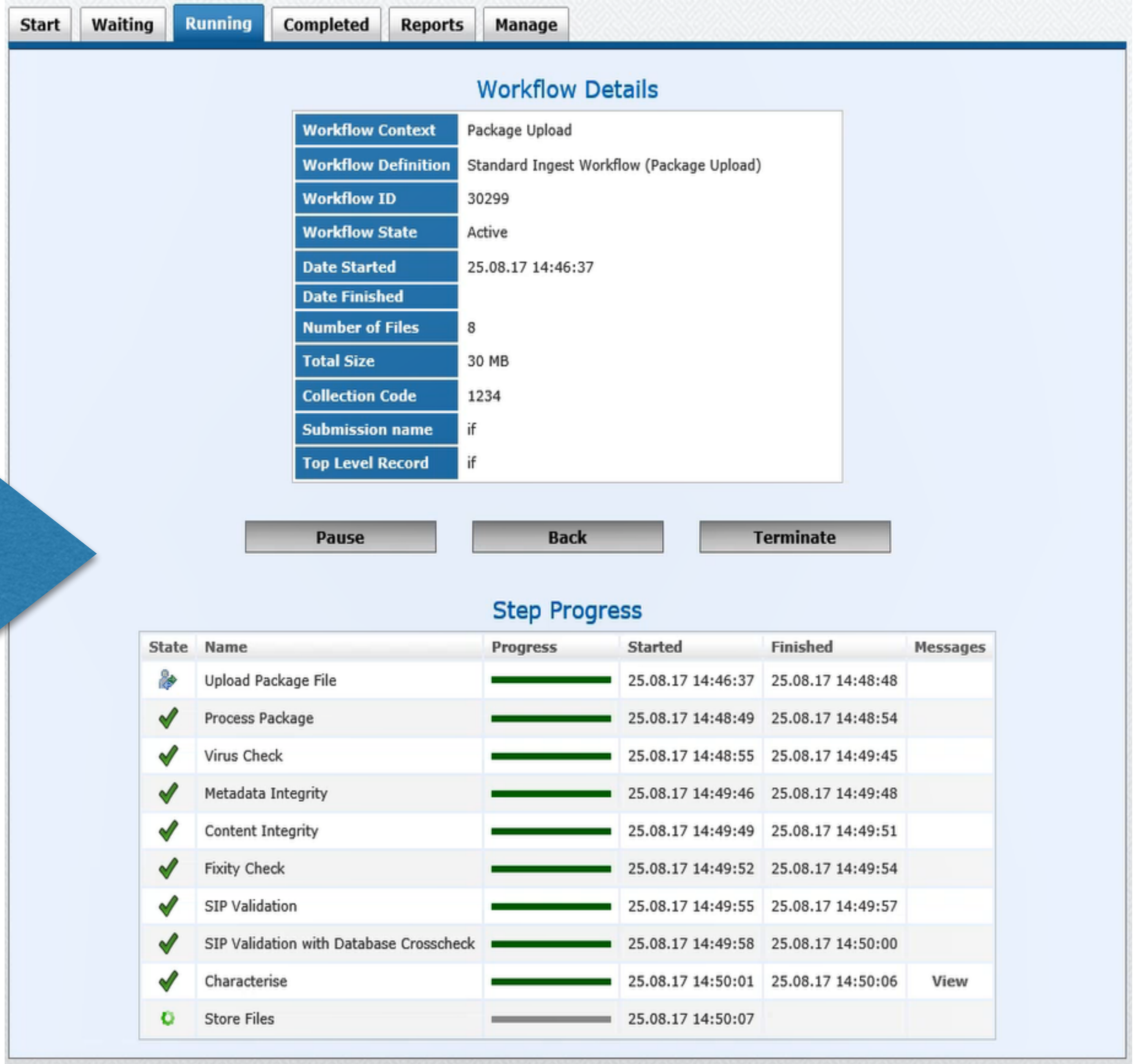


워크플로우

아키비스트는 개별 업무를 워크플로우로 정의하고, Preservica시스템이 실행

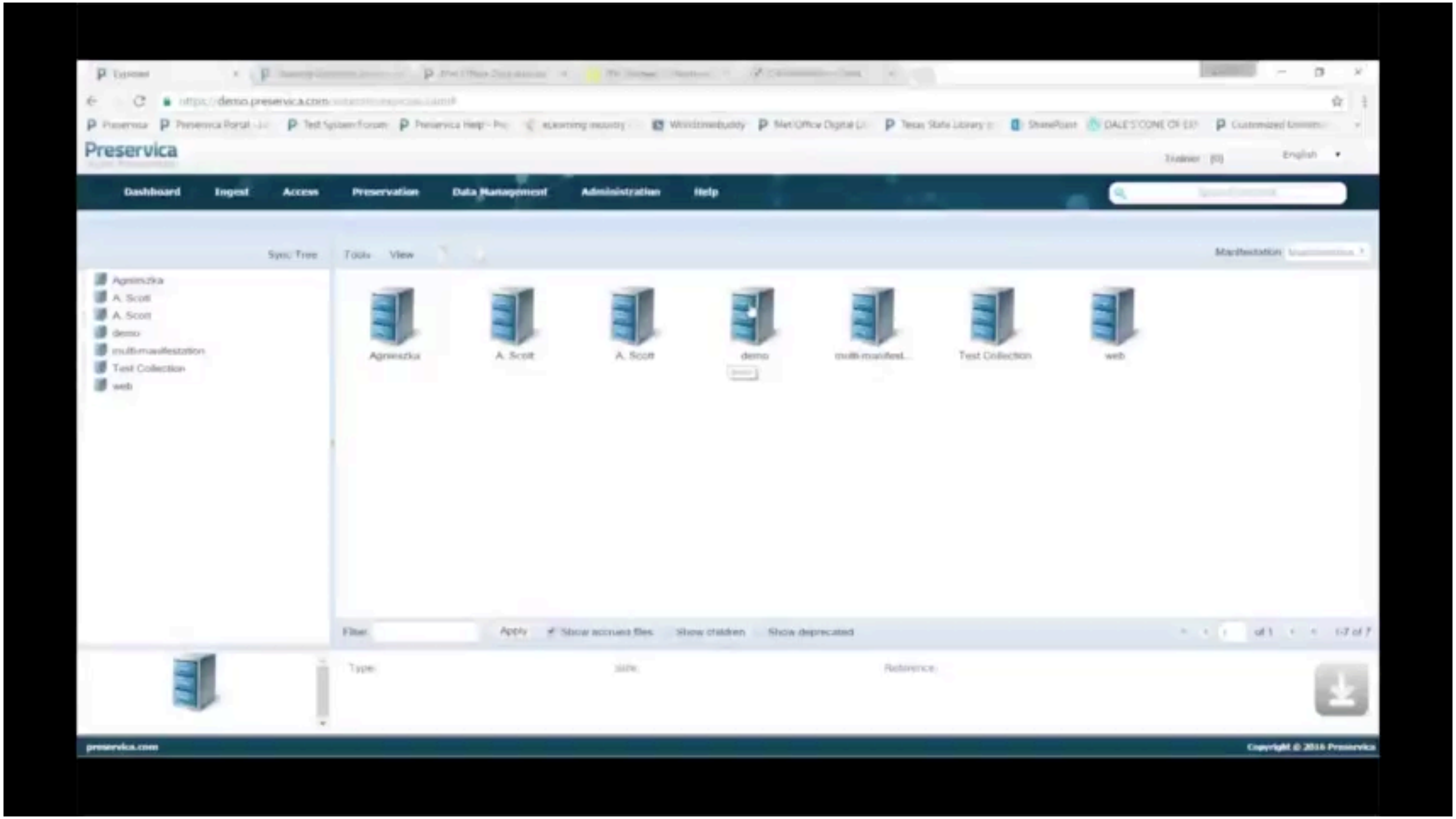


워크플로우 정의



워크플로우 실행

시연 (입수 워크플로우)



Preservica
Dashboard
Ingest
Access
Preservation
Data Management
Administration
Help
Agnieszka
A. Scott
A. Scott
demo
Test Collection
web

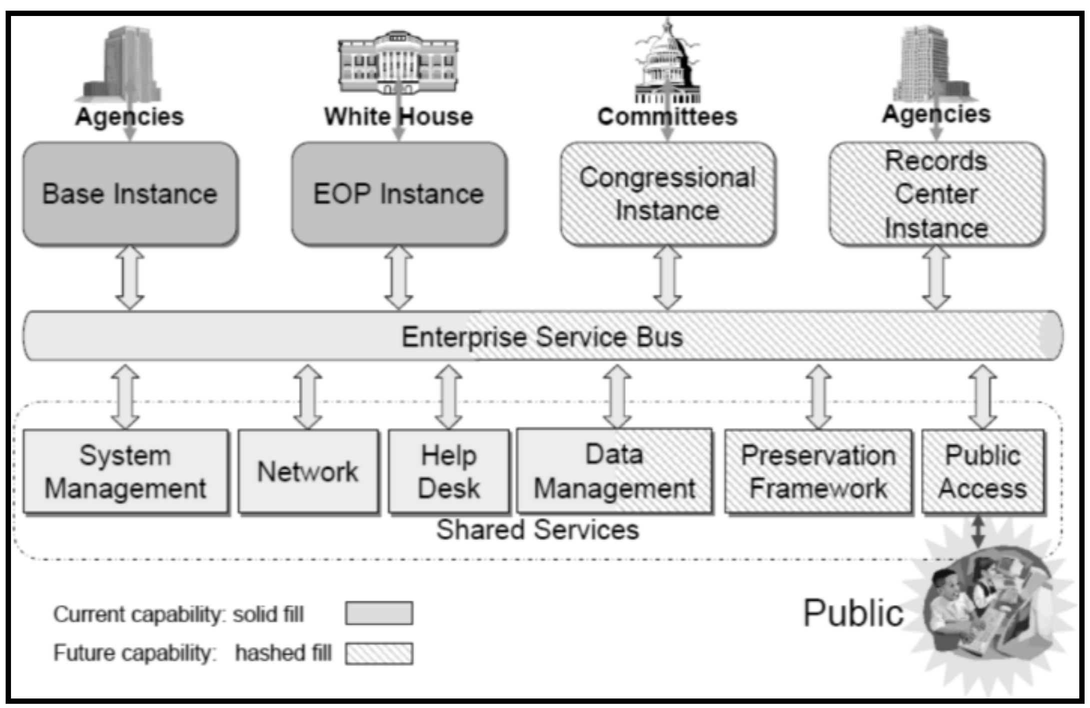
2. 시스템

개요

NARA는 2017년부터 클라우드 기반의 ERA 2.0으로 고도화 추진 중
(소장기록 : 백악관, 의회 및 연방기관 전자기록(160억건/8억철, 400TB) 및 비전자기록)

ERA 1.0

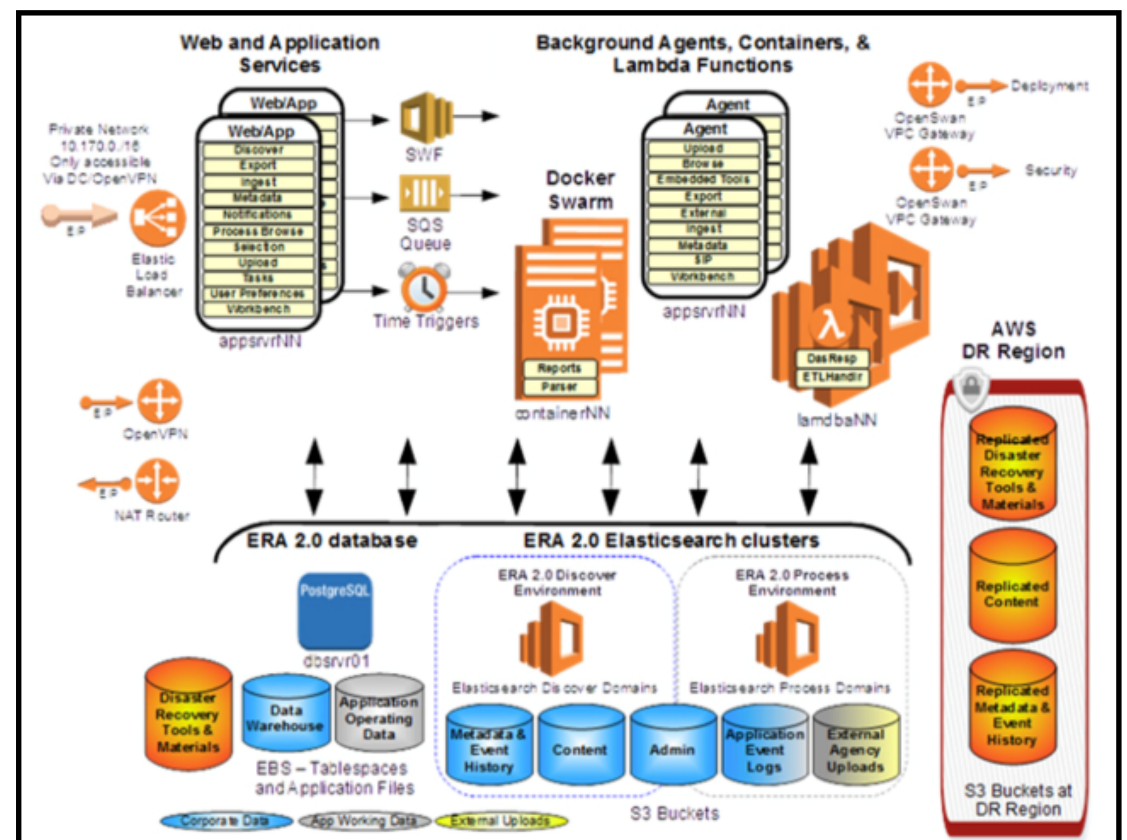
개발시기 : 2005 ~ 2011 (예산 4천 억원, 록히드 마틴)
주요 이용자 : 백악관, 의회, 연방기관
개발배경 : 90년대 말 전자정부 도래로 2000년 클린턴 행정부 종료 시 방대한 전자기록 이관을 대비하여 ERA 프로그램 준비
기능 범위 : RM(전자/비전자) + AM(자동화된 영구기록관리)
아키텍처 : ESB 기반의 서비스 지향 아키텍처(SOA)
ERA와 각 기관 RMS의 관계 : ERA는 GRS와 패키징/이관 수행



<그림> ERA 1.0 개요
출처: Thibodeau, Kenneth(2009) "NARA's Electronic Records Archives Status and Plans"

ERA 2.0

개발시기 : 2017 ~ 2011
주요 이용자 : 백악관, 의회, 연방기관
개발배경 : 클라우드 환경의 기민한 개발 프로세스로 고도화
기능 범위 : RM(전자/비전자) + AM(자동화된 영구기록관리)
아키텍처 : 애자일 프로세스 기반의 마이크로서비스(MSA)
기능 아키텍처 : DPE(처리), DOR(보존), BOM(업무지원) 구분



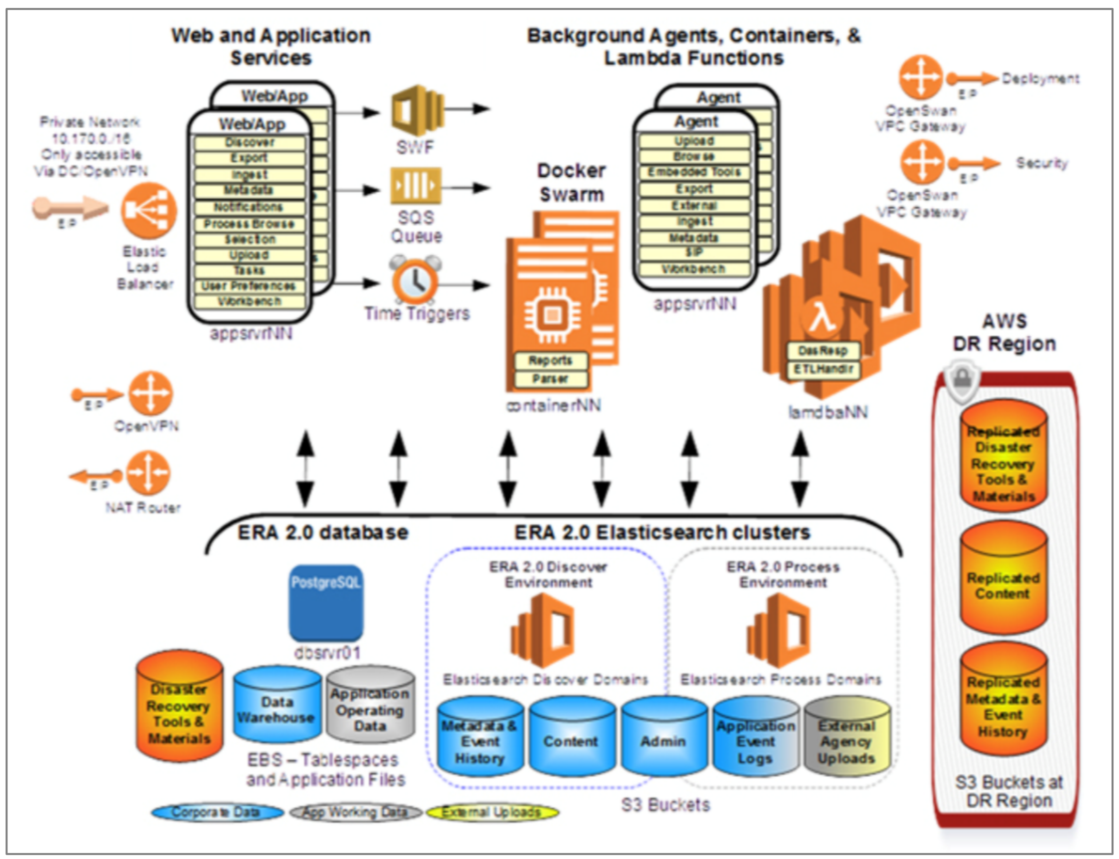
<그림> ERA 2.0 개요
출처: Leslie Johnston (2018) "ERA 2.0: A Cloud-Based Preservation Repository at the NARA"

2. 시스템

특징

1. 모듈 구성 : 입수 단계를 별도 모듈로 분리

2. 클라우드 : 클라우드 컴퓨팅 기반의 인프라, MicroService 소프트웨어 아키텍처, 배포(Docker 컨테이너 기반) 등 차세대 기술로 전면 전환



<그림> ERA 2.0 아키텍처

출처: Leslie Johnston (2018) "ERA 2.0: A Cloud-Based Preservation Repository at the NARA"

DPE (입수)

- DPE(Digital Processing Environment)는 입수 단계에 필요한 소프트웨어 툴과 메타데이터 처리 등의 기능을 수행

DOR (보존)

- DOR(Digital Object Repository)은 불변성 검사, 평가, 리포팅 등 소장기록의 보존과 관리를 위한 모든 기능을 수행

BOM (업무지원)

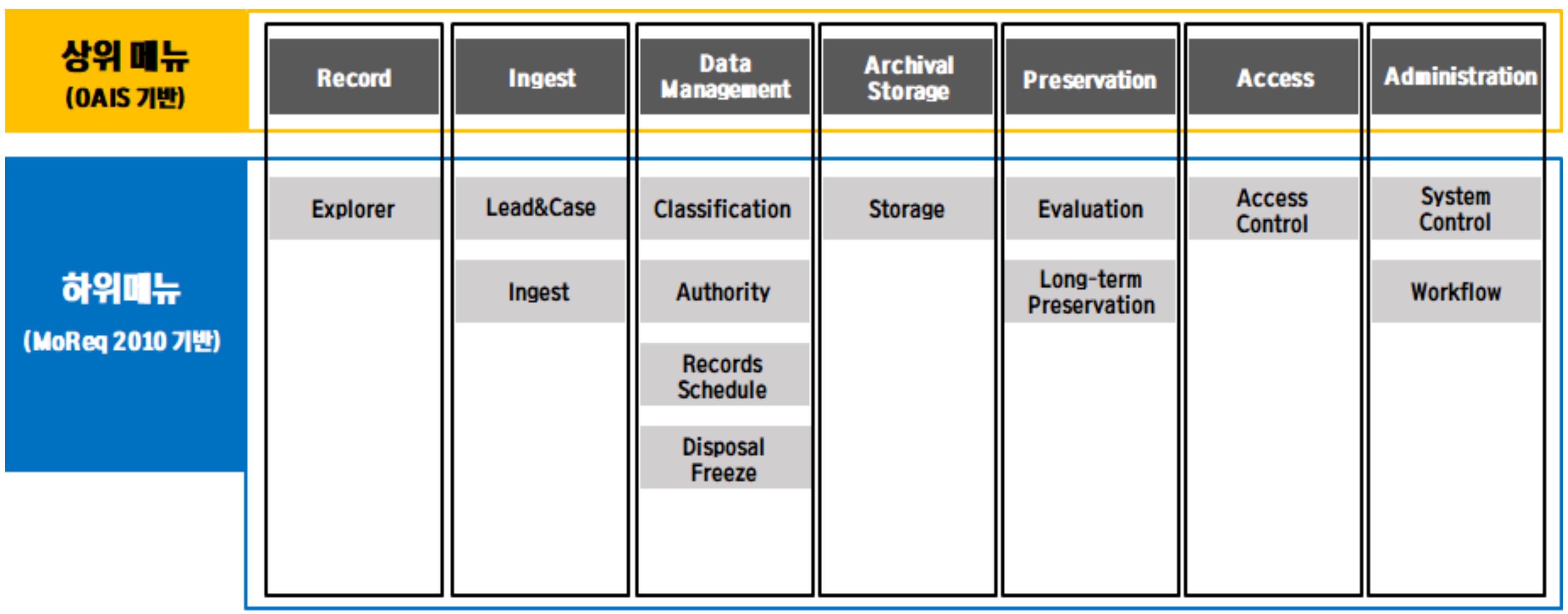
- BOM(Business Object Management)는 아카이브의 업무를 지원하는 기능을 수행(현재 단일(Federal) 워크플로에서 멀티(Federal, Presidential, Legislative, Judicial)로 전환

<그림> ERA 2.0 모듈 구성

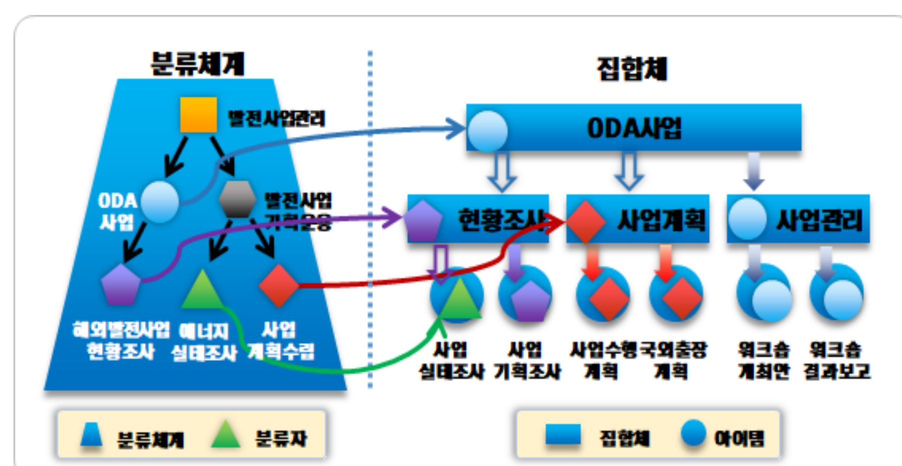
4) SDAS(서울기록원)

개요 : 2019년 아카이브시스템 개발 3년차 수행 중

- OAIS 참조모형을 기반으로 시스템 상위 메뉴 구성
- Moreq2010에서 제시하는 모듈화된 서비스 기반하여 하위 메뉴 구성
- Preservica의 Explorer, 워크플로우 기능 등을 응용하여 기록관리 기능 구현
- 기록물 집합체(Aggregation) 개념 도입하여 기록조직화 및 기록분류



SDA시스템 기록조직화



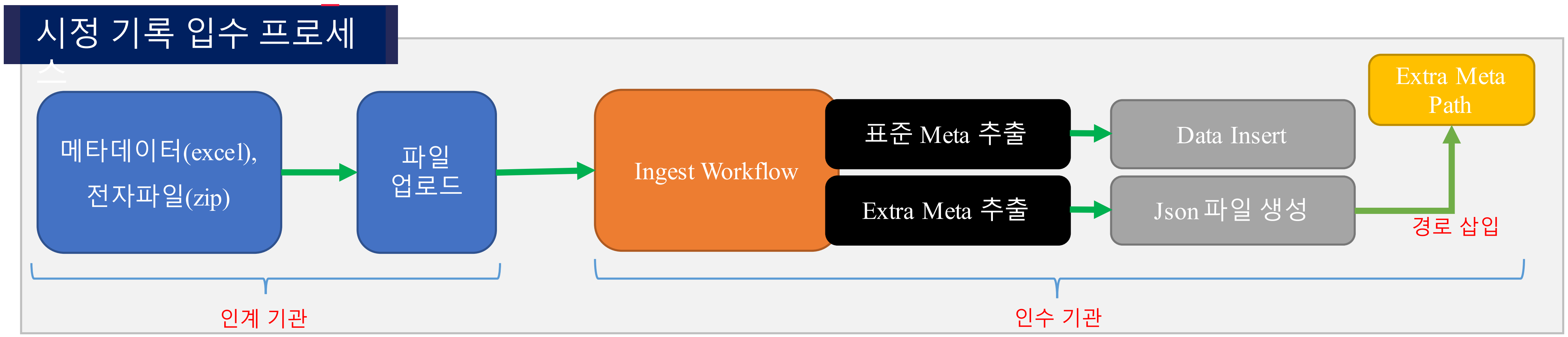
SDA시스템 기록분류

2. 시스템

4) SDAS(서울기록원)

워크플로우

시정기록 입수 워크플로우 사례



집합체(Aggregation) 테이블

UUID	Ref Code	Title	Record Schedule	Creation Date	Creation End Date	Description	…	Extra Meta Path

자동생성

UUID_ + ExtMeta.json

```
{
 UUID : 8234-9239-FF02-FFE2
 부제목 : 임야 세부측량원도
 기록유형 : 도면류
 전자기록물야부 :  비전자
 생산부서 : 서울시 도시계획국
 소장처 : 기록관
 서고관리번호 : 2-6111365-1961
 …….
}
```

철ID	철제목	부제목	기록유형	전자기록물 여부	생산부서	소장처	서고 관리번호	…

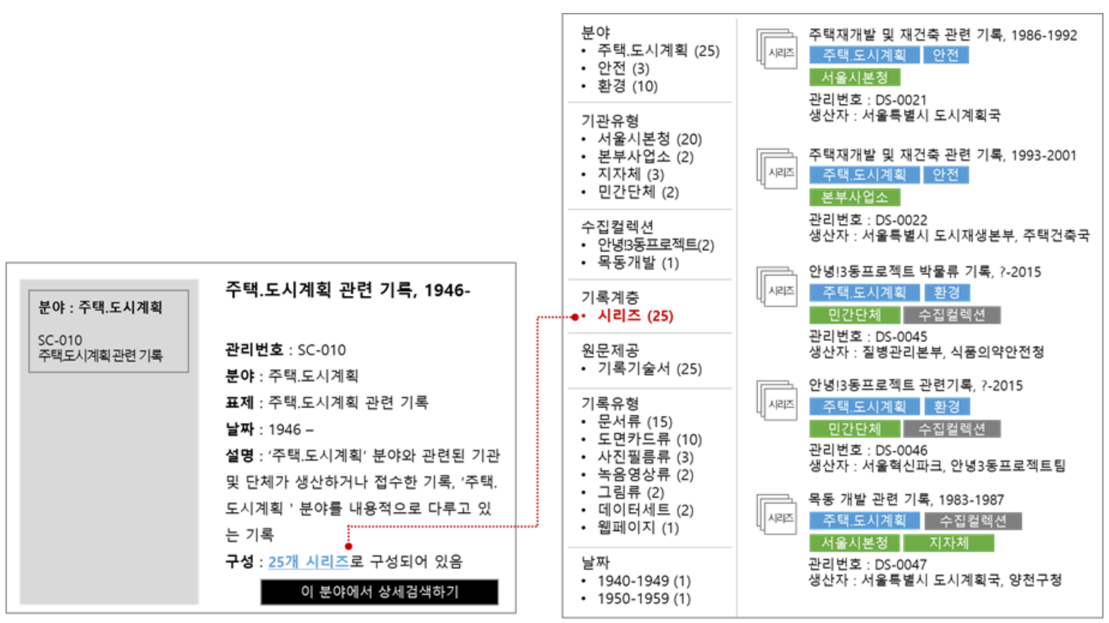
행정기록 정보(기본,생산,보존,이력 등) - 엑셀파일

* 기본적으로 철,건은 상위관계코드 설정을 제외하곤 동일한 컨셉

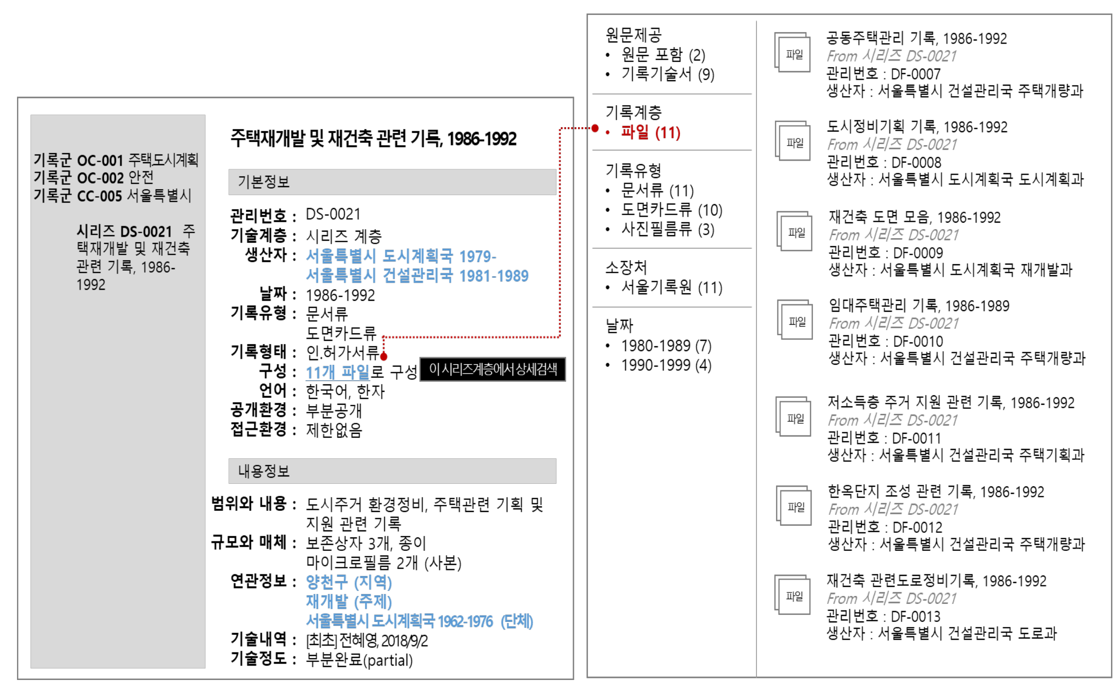
디지털 카탈로그

2018년부터 디지털 카탈로그 모형을 수립하고 시범구축 진행
기록활용의 핵샘도구로 내외부 이용자에게 서비스 예정

- 외부이용자 : 검색도구, 기록자원 네비게이션, Primary 기록콘텐츠 제공
- 내부이용자 : 기록조직·통제도구, 기록맥락 확보, 기록활용의 기본도구로서 활용
- 전거(기관, 지역, Topic, 인물 업무기능) Authority와 기록기술을 통해 기록의 통합적인 활용 가능



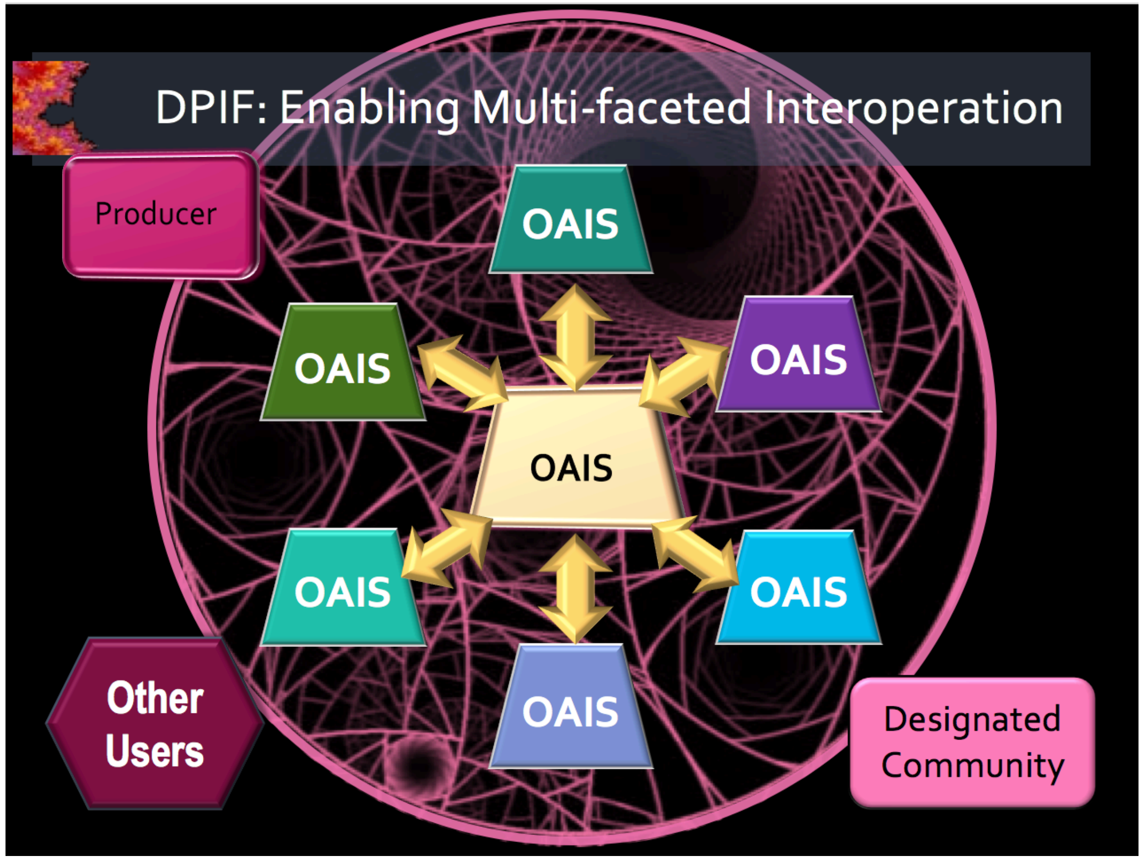
Classfication Description

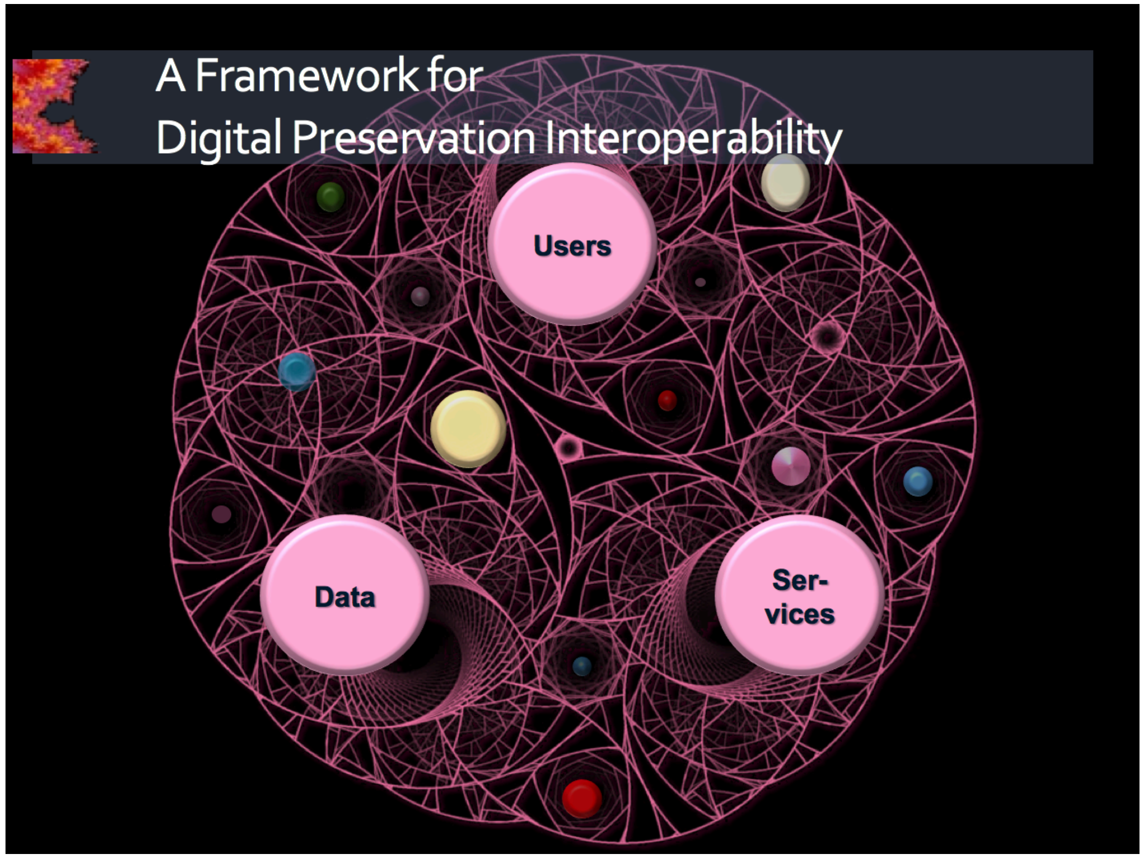


Archival Description

DPIF

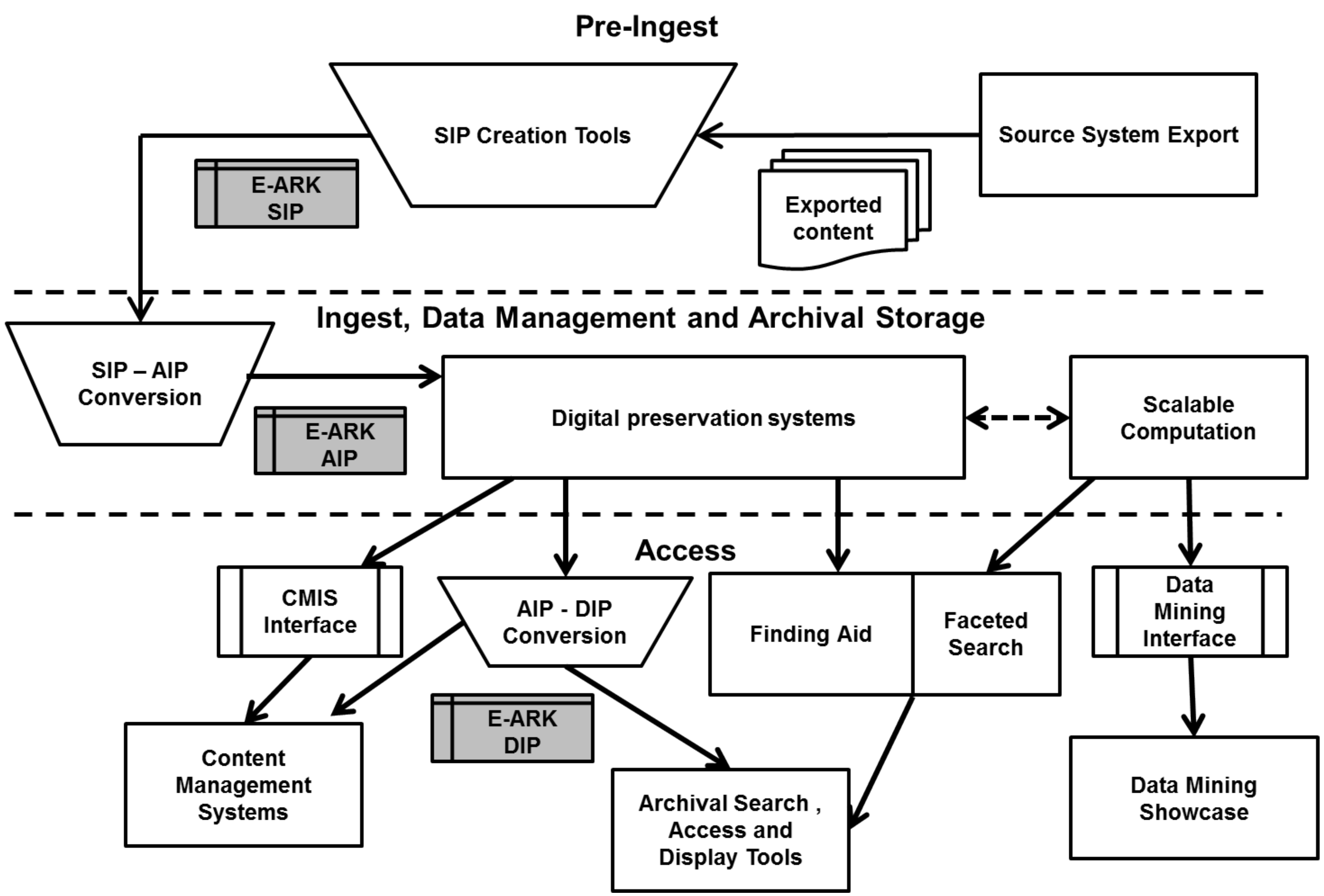
ISO/IEC Digital Preservation Interoperability Framework(DPIF)
OAIS를 참조하는 복수 시스템 간 이용자의 요구에 따른 데이터 보존처리 방법과 도구, 서비스를 상호 운용 가능하도록 하기 위한 프레임워크 (2009~2010)





E-Ark : OAIS의 실행을 위한 툴

E-Ark 프로젝트는 OAIS를 확장하고 공유하기 위한 EU의 공동연구로, 디지털 보존 실무를 보다 상세하게 정의하고 실무에 쓰일 수 있는 도구들을 식별 또는 직접 개발하여 오픈소스로 배포함



[그림] E-Ark 아키텍처

3. 상호운용성

E-Ark

E-Ark : 오픈소스 툴

OAIS의 주요 엔티티별 34종의 오픈소스 툴과 포맷 명세서 제공

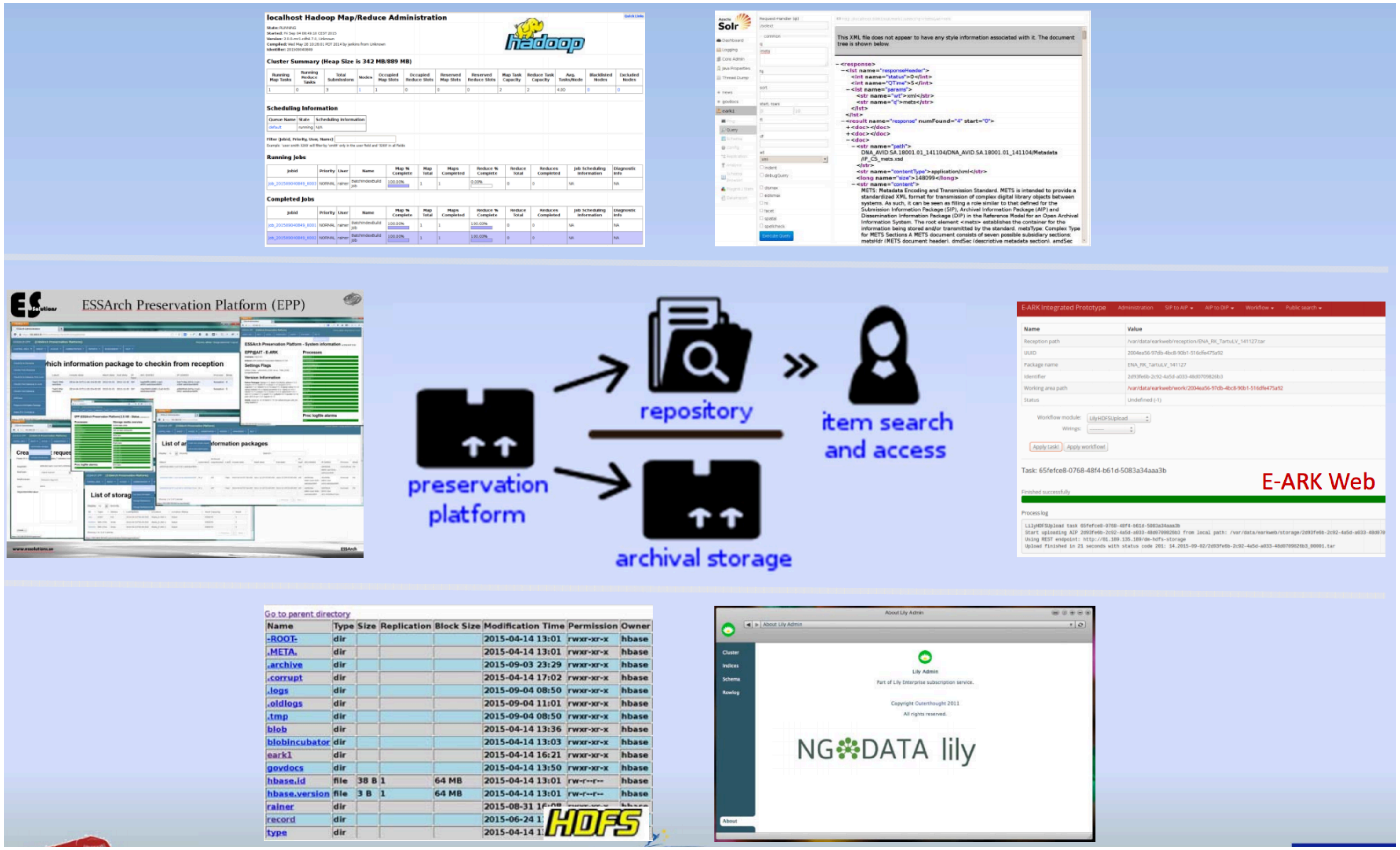
E-ARK Tools and Integrated Prototype Elements

Next

	E-ARK Tools	Integrated Prototype	Third party tools used by E-ARK
Pre-Ingest	Database Preservation Toolkit (DBPTK) ERMS Export Module Universal Archiving Module (UAM) RODA-In ESSArch Tools for Producer	SIP Creator (E-ARK web)	
Ingest	ESSArch Tools for Archive	SIP to AIP Converter (E-ARK web)	
Storage	ESSArch Preservation Platform RODA Repository	HDFS Storage (E-ARK web) SOLR index (E-ARK web)	
Access	Search and Display GUI Order Management Tool	Lily-Ingest E-ARK Web Search AIP2DIP (E-ARK Web)	QGIS Geoserver
Viewer	IP Viewer (also ERMS and single file viewer) Database Visualisation Toolkit (DBVTK) CMIS portal/viewer		Oracle OLAP Viewer Peripleo

E-Ark : 통합 프로토타입

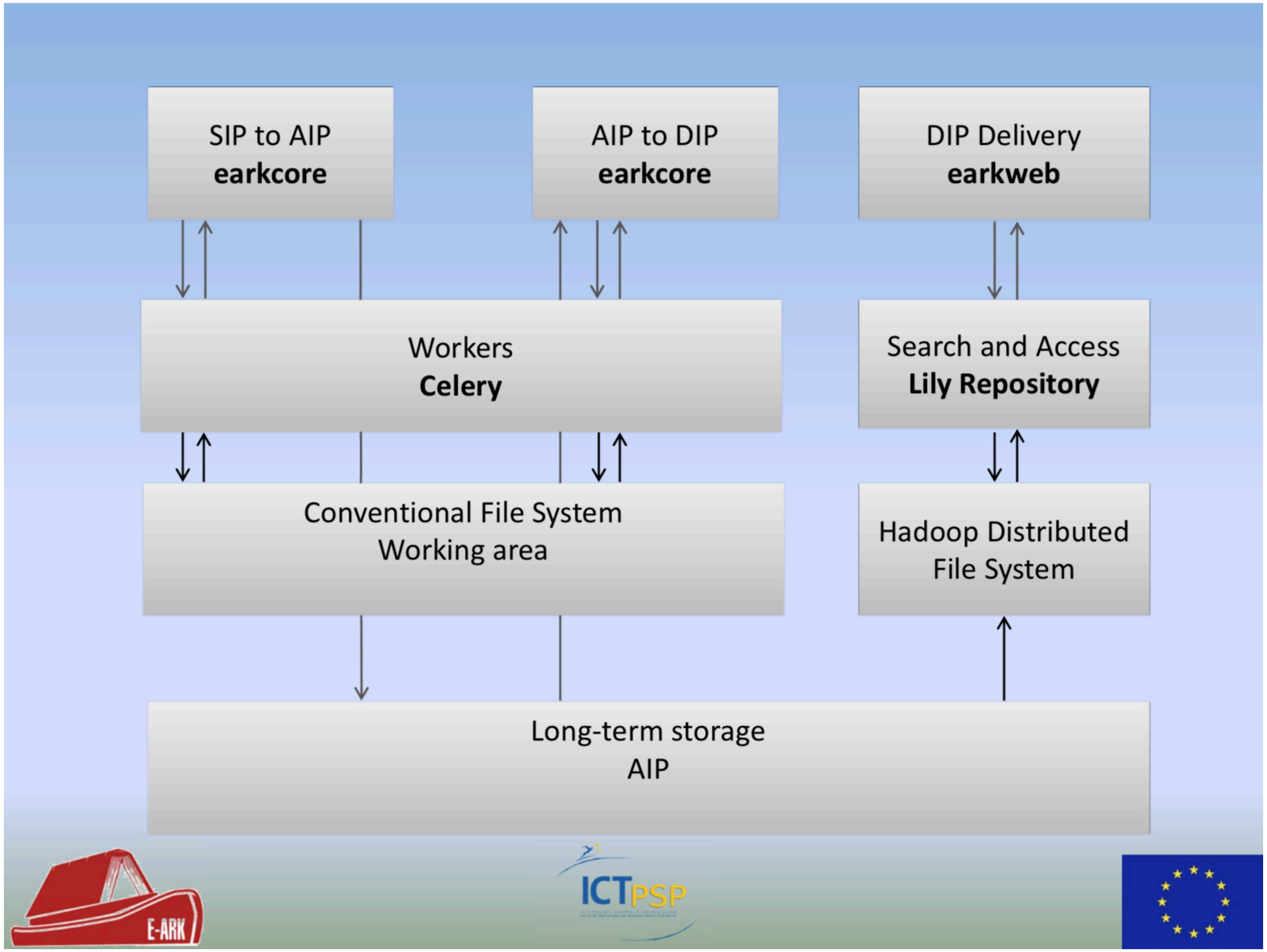
OAIS 전반을 수행하는 통합된 프로토타입 제시



출처 : https://www.eark-project.com/resources/conference-presentations/dlm-oct15/43-earkips-1/file

E-Ark : 통합 프로토타입

OAIS 전반을 수행하는 통합된 프로토타입 제시



E-Ark : 워크플로우

E-Ark 프로젝트에서는 입수 단계의 베스트 프랙티스를 아래 6단계로 정의하고, 각 단계에서 쓰일 수 있는 아카이브 커뮤니티 내의 식별된 도구들을 배치함

Workflow step	Name of the tool
Transfer to SIP	Elev SIP Creator; Preservica, MetsCreation; UAM; DRI; rsync*.
Ingest	DRI; Archaeology Data Service (ADS) Data Seal of Approval (after ADS DSOA); kleio; SFTP*, maior memorix.
Identification	DROID; maior memorix.
Normalisation	Preservica; METS, DRI, ADS DSOA; maior memorix; AdobePhotoshop; AdobePremiereCS5.5.2; Matrox MAX H264 Capture.
Characterisation	JHOVE (via kleio); Preservica; MODS; ADS DSOA; PRONOM; DROID.
Additional metadata	Preservica, DRI, UAM, MARC to MODS; maior.memorix; PIT+AIS; Adobe Bridge; EZID, gencat (Catalonian metadata); FTK Imager.

* file transfer protocol not specific to archiving

E-Ark 프로젝트에서 정의한 이관 프로세스 및 도구

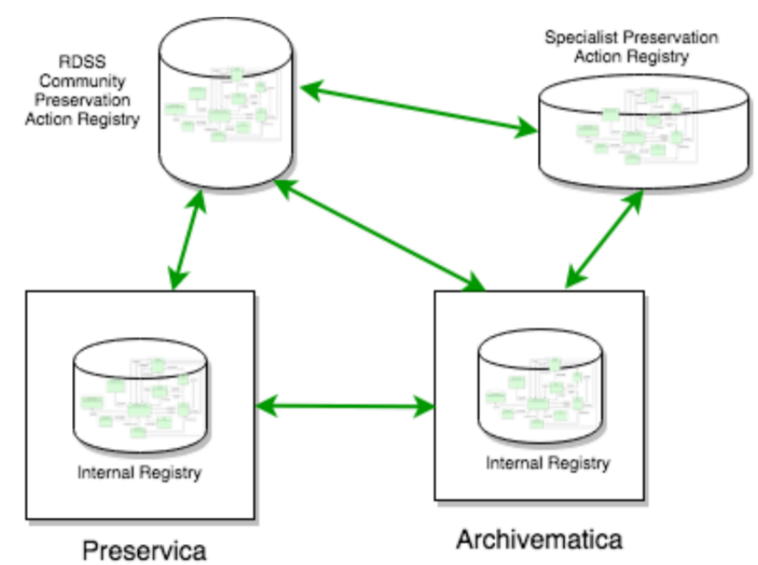
출처: https://www.eark-project.com/

개요

- 영국 JISC의 최신 프로젝트 (http://parcore.org/)
- 디지털 보존의 실무/기술을 주도하는 5개 기관/업체가 디지털 객체 보존기술과 모범사례를 포착하고 공유하기 위한 보존조치 레지스트리(Preservation Action Registry, PAR)를 연구

개발 배경

- 배경 : Archivematica의 FPR (Format Policy Register)과 Preservica의 Linked Data Registry (LDR)는 디지털 보존을 수행할 때 사용할 도구와 규칙을 정의하는데, 이 두 가지 방법은 유사한 문제를 해결하기 위한 것임에도 상호운용 안됨
- Artefactual, Arkivum, Preservica, Jisc 및 Open Preservation Foundation은 디지털 객체를 보존하기위한 기술적 모범 사례를 포착하고 공유할 수있는 PAR(Preservation Action Registries)을 연구하기로 함



데이터 모델

- PAR 레지스트리의 데이터 모델은 다양한 디지털 객체의 1)처리도구, 2)속성, 3)보존기능을 정의
- 이 정보에는 버전이 지정되며 다양한 보존 조치를 실행하는 데 필요한 도구 매개 변수 및 소프트웨어 환경이 포함되어 있어 보존 시스템에서 직접 실행 가능

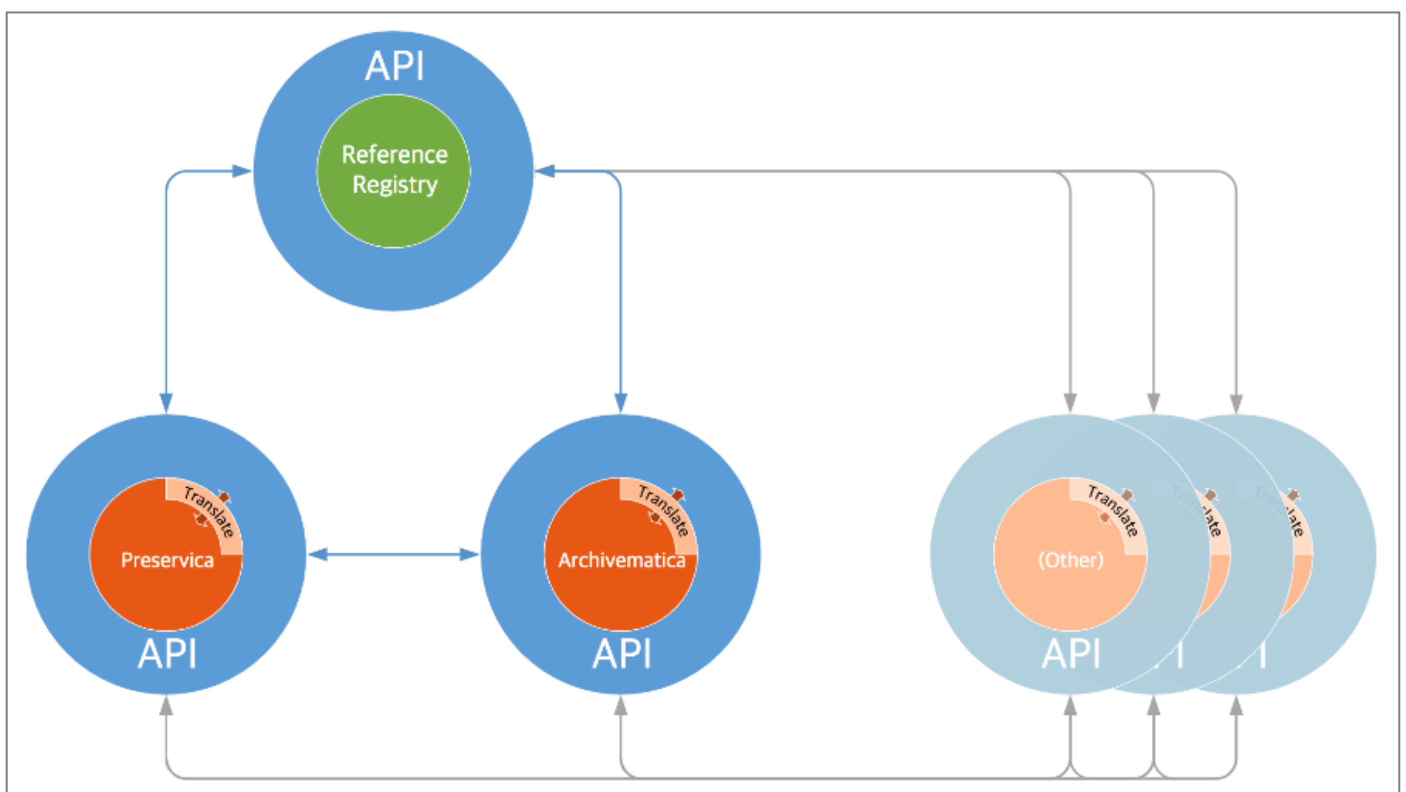
산출물

산출물	내용
개념모델	프레임워크, 논의이슈 (다국어 제공)
JSON 스킴	API 페이로드에 적용될 PAR 모델 정의
API	보존시스템 간 교환 가능한 범용인터페이스
수행 가능한 DP 액션	플랫폼 간 배포/실행 가능한 툴 (벤더독립)
PoC	개념증명 결과물

개념도

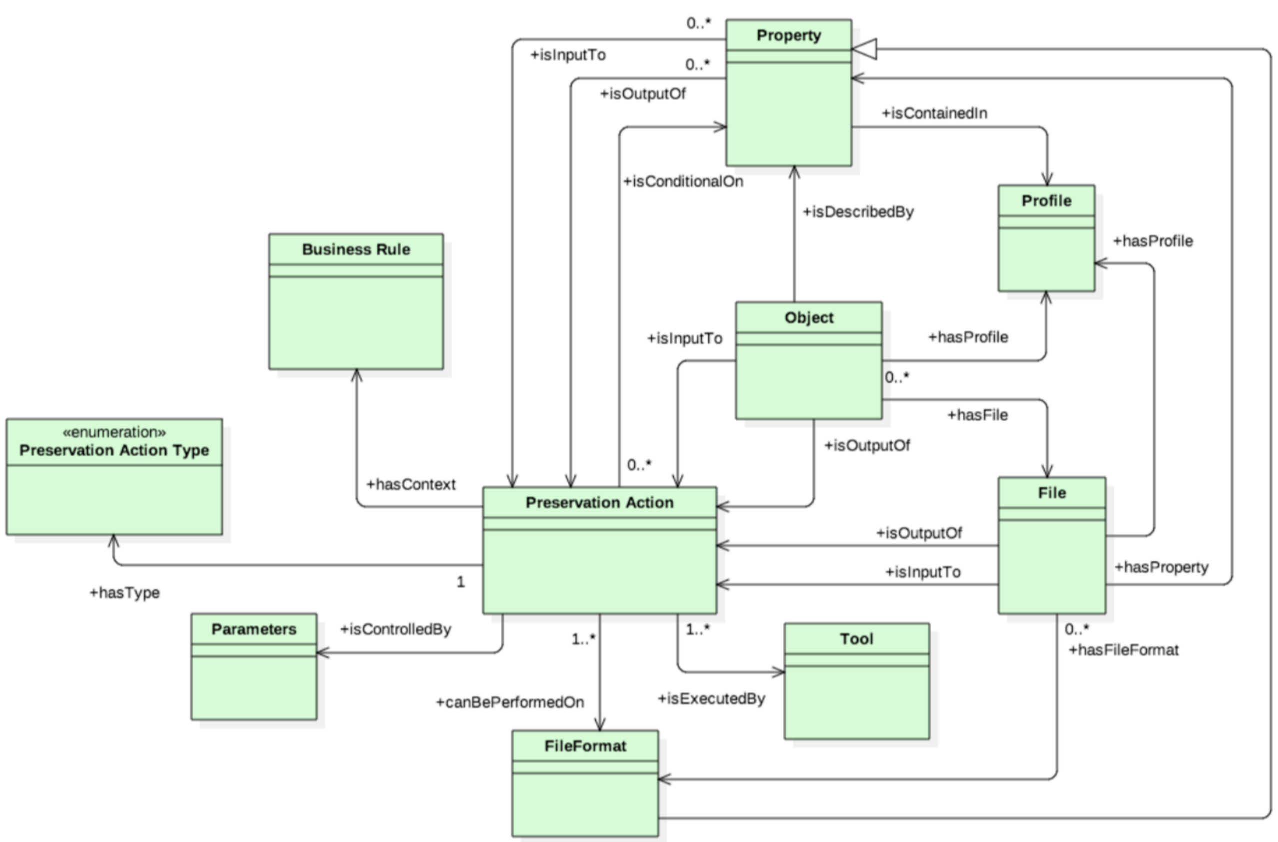
▶ Archivematica, Preservica 등 보존시스템들이 보존정책 및 조치방법을 상호 참조

▶ 향후 다양한 보존시스템 및 기관들이 선진실무를 상호 공유할 수 있도록 API 연계 확대 예정



개념모델

▶ 보존조치는 md5sum, DROID, JHOVE, ffmpeg 등 웹서비스, 또는 컨테이너 형태의 툴을 호출하여 Object 또는 File에 디지털 보존조치 실행(명령줄 실행 또는 그 이상의 조치)



<그림> PAR 개념모델

Preservation Action	par:GenerateMD5checksum	Preservation Action Type	premis:fixity
		Input Type	pcdm:File
		Output Type	par:Property
Tool	par:toolRegistryName	COPTR	
	par:toolRegistryKey	Md5sum Unix Command	
	par:toolVersion	md5sum (GNU coreutils) 8.25	
	par:toolDocumentation	https://man.cx/md5sum	
Property (Output)	par:propertyType	premis:messageDigest	
	premis:messageDigestAlgorithm	MD5	
	xsd:anyURI	http://id.loc.gov/vocabulary/preservation/cryptographicHashFunctions/md5.html	

Simplified example of a Preservation Action (generating a MD5 checksum on a File)

<그림> 파일의 MD5 체크섬을 생성하는 보존조치 사례

4. 결론

1. OAIS 해석 및 기능요건 구체화

2. 표준/툴/서비스 (METS, PREMIS, DROID, JHOVE, Archive-It 등)

3. 신기술 (클라우드, 마이크로서비스 아키텍처, 도커/쿠버네티스, 오픈소스)

4. 모니터링과 협력

감사합니다

안대진

daejin@archivelab.co.kr