

2024年度京都大学図書館機構講演会
「研究データの持続可能なライフサイ
クル構築へ向けて」
オプションセミナー
長期保存の概念的枠組みであるOAIS参
照モデルの概要

2025/1/16

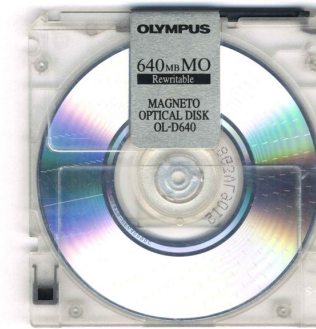
京都大学大学文書館

橋本 陽

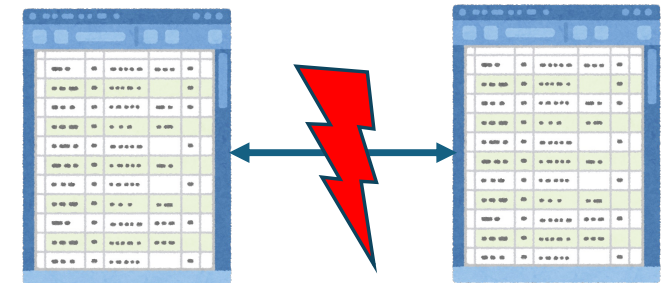
hashimoto.yo.4a@kyoto-u.ac.jp

悲しき実話

1. MOディスクが読み取れない
→記憶媒体の陳腐化 (media obsolescence)
2. パソコンに日本酒をかけた
→記憶媒体の劣化 (media degradation)
3. 昔のマイクロソフト・アクセスの電子ファイルが開けない
→フォーマットの陳腐化 (format obsolescence)
4. 2をフリーソフトで無理やり開いた
 - 各シート間のリンクが切れる
 - 作成時間も開いた時間になった→出所 (来歴) の消失 (loss of provenance)



Old



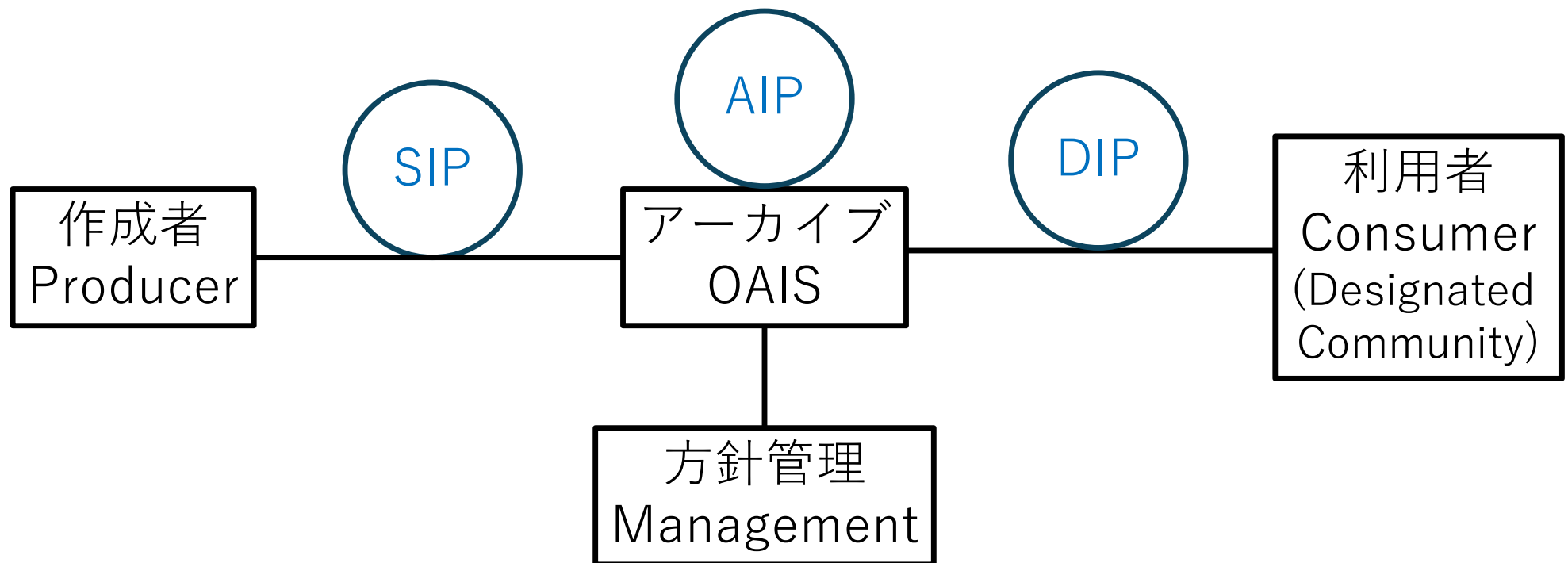
New!

New!

The Open Archival Information System Reference Model (OAIS参照モデル)

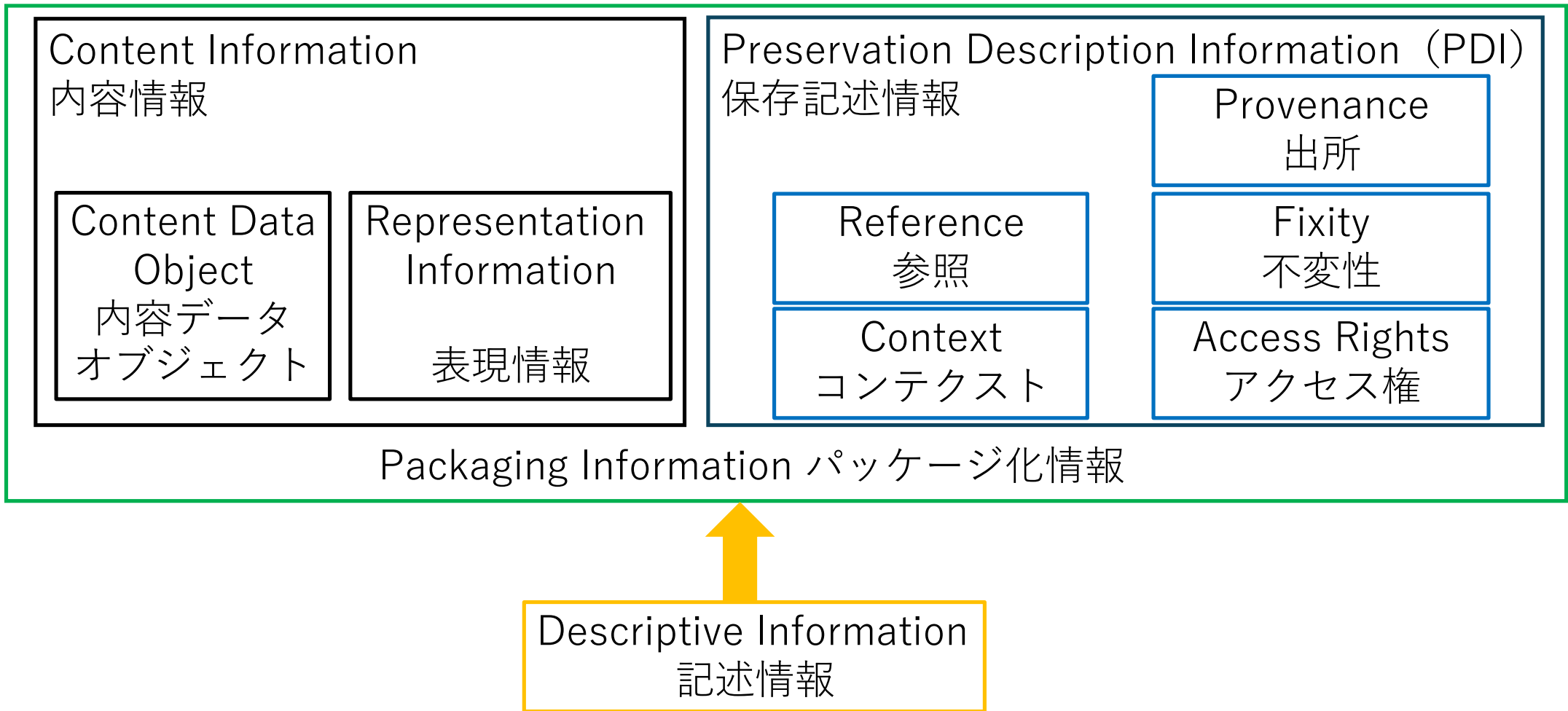
- Open Archival Information System
 - 「公開の」「アーカイブ情報システム」
 - データの長期保存・利用（長期＝変わり続ける技術環境と利用者に関連）
- 宇宙データシステム諮問委員会（The Consultative Committee for Space Data Systems (CCSDS)）が決定
- 背景
 - 1950～70年代、コンピューター依存のデータが蓄積
 - MLA関係者やコンピューター関連分野（科学など）が保存に関心
 - 抽象的なモデル開発の要望
 - 1990年代にNASAが主導
 - アメリカ国立公文書館（NARA）も作成に関与
 - 2003年に国際標準ISO 14721として採択

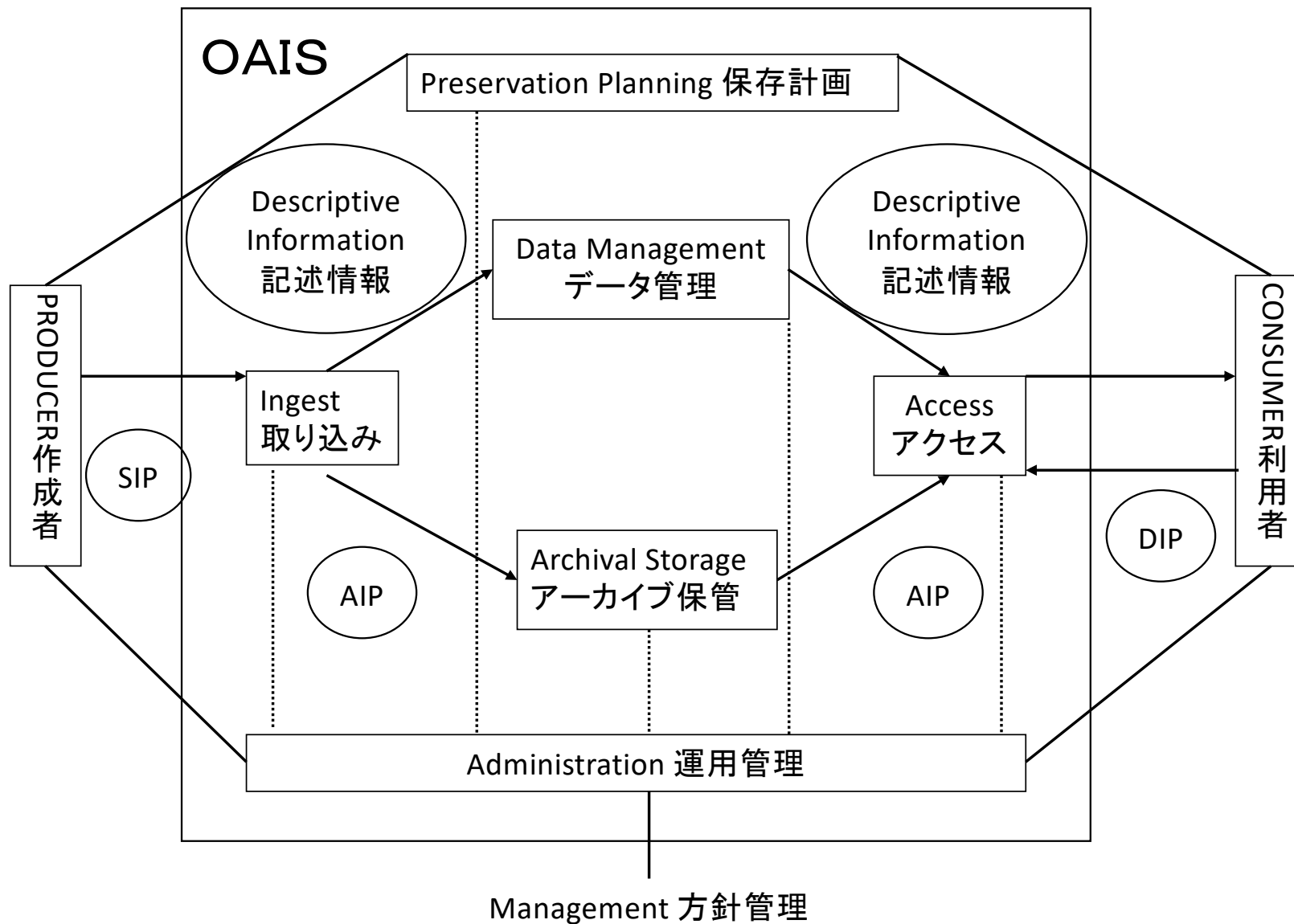
OAIS環境 [Lavoie, p. 9]



- SIP (Submission Information Package) : 提供情報パッケージ
- AIP (Archival Information Package) : アーカイブ情報パッケージ
- DIP (Dissemination Information Package) : 配布情報パッケージ

情報パッケージ [Lavoie, p. 16]





メタデータ (Metadata)

- 別の情報資源の特徴を示す情報
- 記述 (Descriptive) メタデータ
 - 検索 (次のスライド)
- 管理運営 (Administrative) メタデータ
 - 技術 (Technical) メタデータ
 - 権利 (Rights) メタデータ
 - 保存 (Preservation) メタデータ
- 構造 (Structural) メタデータ
 - 構造内にある複数のデータの関係性



藤子・F・不二雄、「念録マイク」、『ドラえもん』、てんとう虫コミックス・短編第14巻第6話、53頁

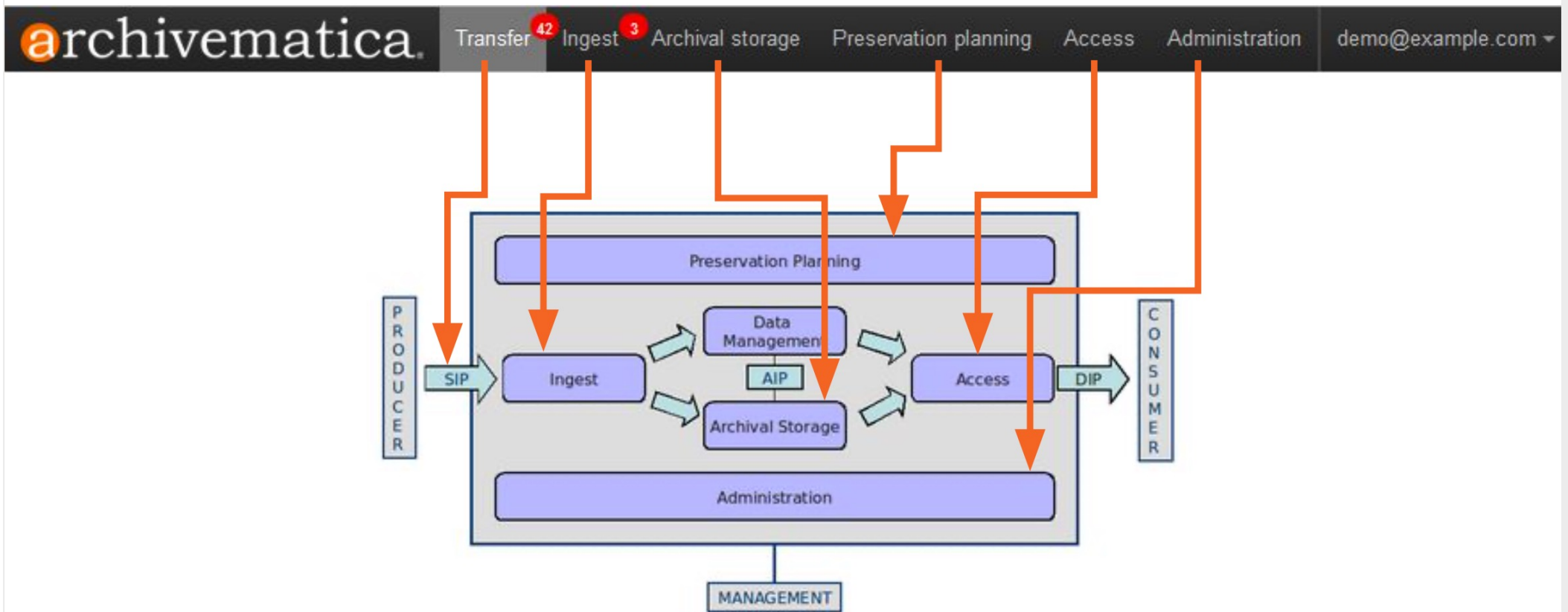
Archivematica (アーカイブマティカ)

- Artefactual Systems社が開発
 - ブリティッシュ・コロンビア大学出身のアーキビスト
 - オープンソース；ウェブベース；長期保存のためのソフトウェア



Archivematicaのシステム概念図

OAIS model



メタデータの標準

- ArchivematicaではMETS.xmlをメタデータのデジタルファイルとして使用
- METS : Metadata Encoding and Transmission Standard
 - セクション : 記述 ; 管理運営 (技術 ; 出所 ; 権利) ; ファイル ; 構造
 - PDI (保存記述情報) と Descriptive Information (記述情報) を記載
 - 内部にPREMISとDublin Coreを含める
- PREMIS : Preservation Metadata: Implementation Strategies
 - デジタル・データ保存のためのメタデータの基準
 - 辞書
 - 出所 ; 技術 ; 権利
- Dublin Core (DC、ダブリン・コア)
 - 記述メタデータの標準
 - 主に検索用

内容情報 (Content Information)

- 内容データオブジェクト (Content Data Object)
 - 保存対象となる情報
 - PDFファイル；HTMLファイル；JPEGファイル
- 表現情報 (Representation Information)
 - 内容データオブジェクトを構成するビット列を表示し、理解するのに必要となる情報
 - 見読性（可読性）→再生可能性・フォーマットの陳腐化
 - フォーマット情報をメタデータとしてMETS.xmlに入力

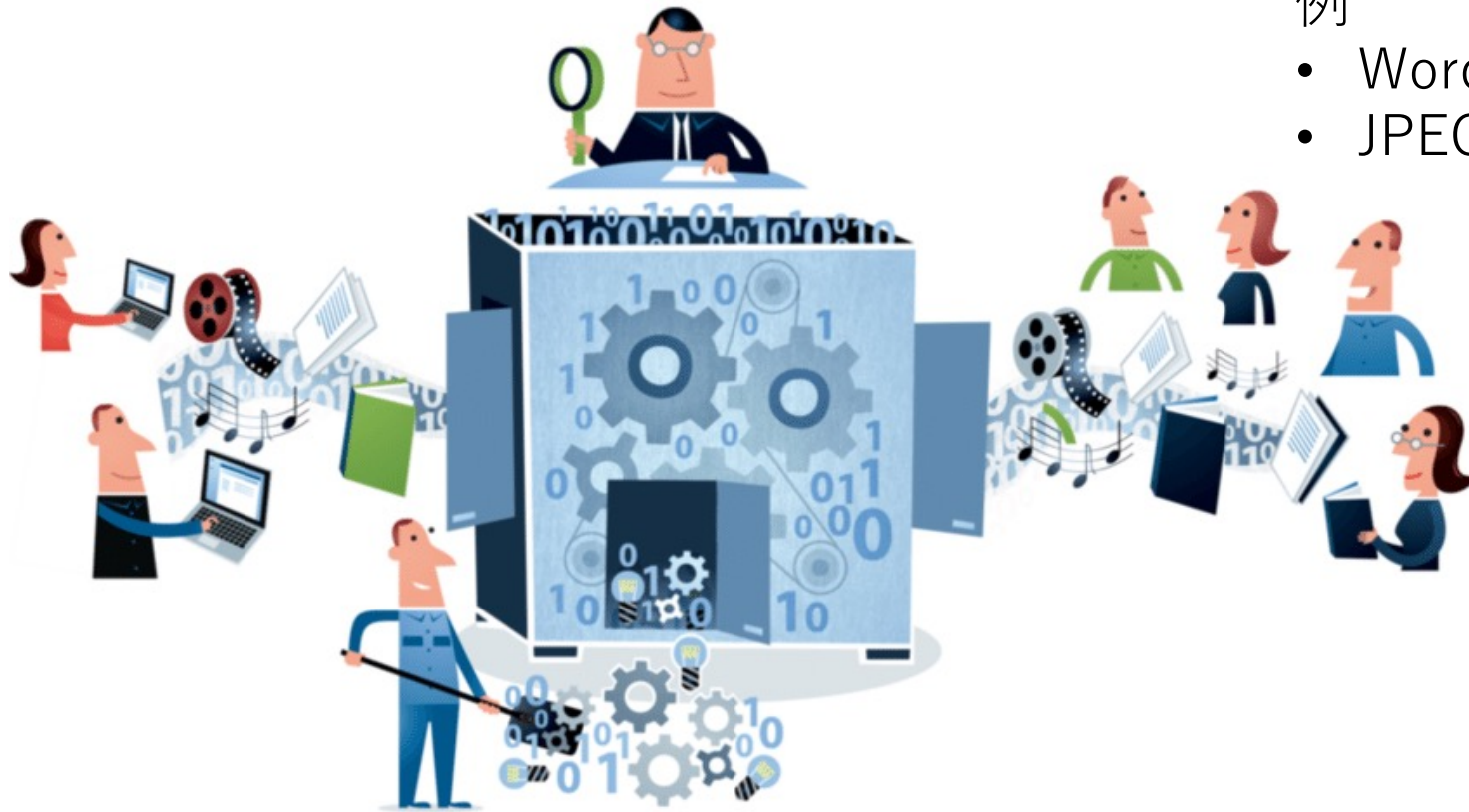
保存記述情報（PDI）→METS.xmlに入力

- 参照情報（Reference Information）：
 - 識別番号（Unique Identifier）→例えば、UUIDを使用
- コンテキスト情報（Context Information）：
 - 内容情報間の関係性；複製やオリジナルの所在
- 出所（来歴）情報（Provenance Information）：
 - 内容情報の歴史（作成や変更）；保管の連鎖；保存活動
 - フォーマット転換（MigrationまたはConversion）、Archivematicaでは正規化（Normalization）（次のスライド）をメタデータとして記載
- 不変性情報（Fixity Information）：
 - メタデータ未記載の変更をしていない
 - ハッシュ値（次々のスライド）；デジタル署名
- アクセス権情報（Access Rights Information）
 - 内容情報の利用条件と制限

フォーマット転換・正規化

例

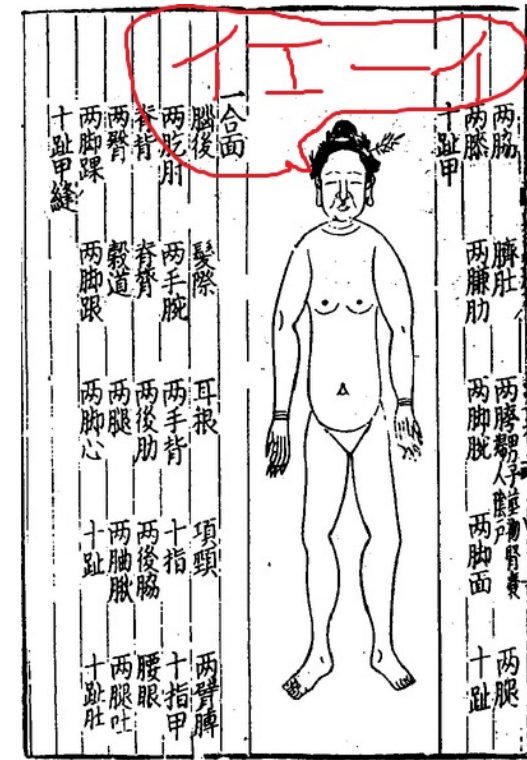
- Word→PDF/A
- JPEG; PNG; GIF→TIFF



Migration and Normalization of File Formats for Long Term Preservation. Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

SHA-1で計算

2da6a8e8f255237d697a06d71d5d7b5ffc29e453



『大元聖政国朝典章』より

パッケージ化情報 (Packaging Information)

- 内容情報 (Content Information) と保存記述情報 (PDI) をつなげて、一つの論理的なパッケージ (まとまり) にする
- Archivematicaでは、BagItでパッケージ化→METS.xmlに入力

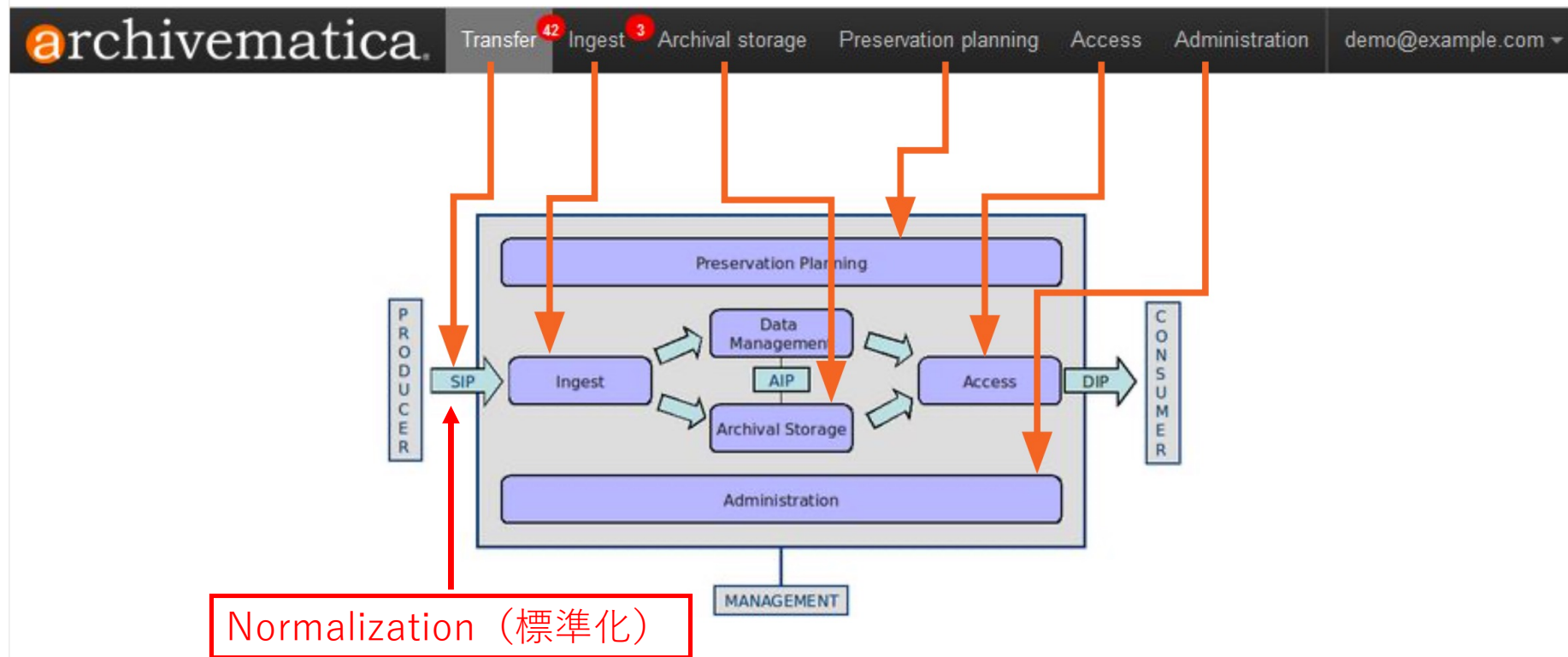
記述情報 (Descriptive Information)

- 内容情報 (Content Information) の検索を補助。
- ダブリン・コアを使用した記述メタデータとしてMETS.xmlに
入力

システム概念図

OAIS model

メタデータ付与の過程 (METS.xml)
PDI、Descriptive Information



この画像を入れると仮定



Archivematica: transfer

Archivematica Dashboard - Transfer - Mozilla Firefox

Archivematica Storage ... x Archivematica Dash... x

localhost/transfer/#

Search 検索

archivematica

Transfer Ingest Archival storage Preservation planning Access Administration wonnami

Connected

Standard koala 0012 /home Browse Start transfer

Type Transfer name Accession no.

/home/ippt_tokyo/koala

Transfer	UUID	Transfer start time
20150415-6	5677b0ce-0cd1-4546-a0b0-86ab77d2ea7a	2015-04-15 21:25
Micro-service: Create SIP from Transfer		
Job: Create SIP(s) [?]	Awaiting decision	Actions
Job: Load options to create SIPs	Completed successfully	
Job: Check transfer directory for objects	Completed successfully	
Job: Load options to create SIPs	Completed successfully	
Job: Check transfer directory for objects	Completed successfully	
Job: Load options to create SIPs	Completed successfully	
Job: Check transfer directory for objects	Completed successfully	
Micro-service: Complete transfer		
Micro-service: Examine contents		
Micro-service: Validation		
Micro-service: Characterize and extract metadata		
Micro-service: Extract packages		
Micro-service: Identify file format		
Micro-service: Clean up names		
Micro-service: Generate transfer structure report		
Micro-service: Scan for viruses		
Micro-service: Quarantine		
Micro-service: Generate METS xml document		

Archivematica: 取り込み (SIP生成)

Archivematica Dashboard - Transfer - Mozilla Firefox

Archivematica Storage ... x Archivematica Dash... x Archivematica Dash... x Re: [archivematica] J... x

localhost/transfer/#

Search 検索

archivematica Transfer Ingest 1 Archival storage Preservation planning Access Administration wonnami Connected

Type Standard Transfer name Accession no. /home Browse Start transfer

Transfer	UUID	Transfer start time
koala	92cc8510-6f04-4452-8829-56ad24cf1126	2015-04-19 14:52
Micro-service: Create SIP from Transfer		
Job: Create SIP(s) [?]	Awaiting decision	Actions
Job: Load options to create SIPs	Completed successfully	
Job: Check transfer directory for objects	Completed successfully	
Micro-service: Complete transfer		
Micro-service: Examine contents		
Micro-service: Validation		
Micro-service: Characterize and extract metadata		
Micro-service: Extract packages		
Micro-service: Identify file format		
Job: Identify file format	Completed successfully	
Job: Determine which files to identify	Completed successfully	
Job: Select file format identification command	Completed successfully	
Job: Move to select file ID tool	Completed successfully	
Micro-service: Clean up names		
Micro-service: Generate transfer structure report		
Micro-service: Scan for viruses		
Micro-service: Quarantine		
Micro-service: Generate METS.xml document		
Micro-service: Verify transfer checksums		

Archivematica : 標準化

Archivematica Dashboard - Ingest - Mozilla Firefox

localhost/ingest/

archivematica Transfer ¹ Ingest ² Archival storage Preservation planning Access Administration wonnaml Connected

Any Keyword Search transfer backlog

Add New

originals arrange

Hide View File Delete Create SIP Add Directory

Submission Information Package	UUID	Ingest start time	
koala	3ed45959-a557-4bd1-853a-b8654e0c15c3	2015-04-19 14:54	
▶ Micro-service: Normalize			
Job: Normalize [?]		Awaiting decision	⚙ Actions
Job: Resume after normalization file identification tool selected.		Completed successfully	⚙
Job: Identify file format		Completed successfully	⚙
Job: Select pre-normalize file format identification command		Completed successfully	⚙
Job: Move to select file ID tool		Completed successfully	⚙
Job: Set resume link after tool selected.		Completed successfully	⚙
Job: Find options to normalize as		Completed successfully	⚙
Job: Move to workflowDecisions-createDip directory		Completed successfully	⚙
Job: Grant normalization options for no pre-existing DIP		Completed successfully	⚙
Job: Set remove preservation and access normalized files to renormalize link.		Completed successfully	⚙
Job: Check for Access directory		Completed successfully	⚙
Job: Check for Service directory		Completed successfully	⚙
Job: Identify manually normalized files		Completed successfully	⚙
▶ Micro-service: Clean up names			
▶ Micro-service: Remove cache files			
▶ Micro-service: Include default SIP processingMCP.xml			

Archivematica：標準化と保存計画

文書ファイル→PDF/A、画像ファイル→TIFF

Format policy rules

[Formats](#)

[Groups](#)

[Create new rule](#)

Show entries

Search:

Identification

[Tools](#)

[Rules](#)

[Commands](#)

Format policy registry

[Tools](#)

Characterization

[Rules](#)

[Commands](#)

Event Detail

[Rules](#)

[Commands](#)

Extraction

Purpose	Format	Command	Success	Enabled	Actions
Preservation	Acrobat PDF/X Exchange 3:2002	Transcoding to pdfa with Ghostscript	0 out of 0	Yes	View Replace Disable
Preservation	Acrobat PDF/X Exchange 3:2003	Transcoding to pdfa with Ghostscript	0 out of 0	Yes	View Replace Disable
Preservation	Acrobat PDF/X Exchange PDF/X-4	Transcoding to pdfa with Ghostscript	0 out of 0	Yes	View Replace Disable
Preservation	Acrobat PDF/X Exchange PDF/X-4p	Transcoding to pdfa with Ghostscript	0 out of 0	Yes	View Replace Disable
Preservation	Acrobat PDF/X Exchange PDF/X-5g	Transcoding to pdfa with Ghostscript	0 out of 0	Yes	View Replace Disable
Preservation	Acrobat PDF/X Exchange PDF/X-5n	Transcoding to pdfa with Ghostscript	0 out of 0	Yes	View Replace Disable
Preservation	Acrobat PDF/X Exchange PDF/X-5pg	Transcoding to pdfa with Ghostscript	0 out of 0	Yes	View Replace Disable
Preservation	Generic PDF	Transcoding to pdfa with Ghostscript	0 out of 0	Yes	View Replace Disable
Preservation	Nikon Digital SLR Camera Raw Image File	Transcoding to tif with convert	0 out of 0	Yes	View Replace Disable
Preservation	Kodak Digital Camera Raw Image File	Transcoding to tif with convert	0 out of 0	Yes	View Replace Disable

Showing 161 to 170 of 344 entries

Previous 1 ... 16 **17** 18 ... 35 Next

Archivematica: 記述情報 (データ管理)

Archivematica Dashboard -- Mozilla Firefox

localhost/ingest/3ed45959-a557-4bd1-853a-b8654e0c15c3/metadata/add/

archivematica. Transfer ¹ Ingest ² Archival storage Preservation planning Access Administration wonnami

Ingest / koala / Metadata / Add

Metadata

koala

Title

コアラ

Part of AIC

あ 般

Optional: leave blank if unsure

Creator

お父さんとお母さん

Subject

動物

Description

Publisher

Contributor

A name given to the resource.
(ISO15836)

Archivematica : AIPとDIPの生成

Archivematica Dashboard - Ingest - Mozilla Firefox

Archivematica Storage ... x Archivematica Dash... x Archivematica Dash... x Re: [archivematica] J... x

localhost/ingest/ 検索

archivematica Transfer ¹ Ingest ² Archival storage Preservation planning Access Administration wonnaml Connected

Any Keyword Search transfer backlog

Add New

originals arrange

Hide View File Delete Create SIP Add Directory

Submission Information Package	UUID	Ingest start time	
koala	3ed45959-a557-4bd1-853a-b8654e0c15c3	2015-04-19 14:54	
▶ Micro-service: Store AIP			
Job: Store AIP (review) [?]		Awaiting decision	⚙ Actions
Job: Move to the store AIP approval directory		Completed successfully	⚙
▶ Micro-service: Prepare AIP			
▶ Micro-service: Upload DIP			
Job: Upload DIP		Awaiting decision	⚙ Actions
▶ Micro-service: Prepare DIP			
▶ Micro-service: Process metadata directory			
▶ Micro-service: Process submission documentation			
Job: Verify checksums generated on ingest		Completed successfully	⚙
Job: Remove files without linking information (failed normalization artifacts etc.)		Completed successfully	⚙
Job: Characterize and extract metadata on submission documentation		Completed successfully	⚙
Job: Identify file format		Completed successfully	⚙
Job: Select file format identification command		Completed successfully	⚙
Job: Scan for viruses in submission documentation		Completed successfully	⚙
Job: Sanitize file and directory names in submission documentation		Completed successfully	⚙

アーカイブ保管

archivematica

TransferBacklogAppraisalIngestArchival storagePreservation planningAccessAdministrationtest

Archival storage / Search

Any

Keyword

Show files? ☐

Search archival storage

[Add new](#)

Browse archival storage

Total size: 233.0 MB | Files indexed: 80

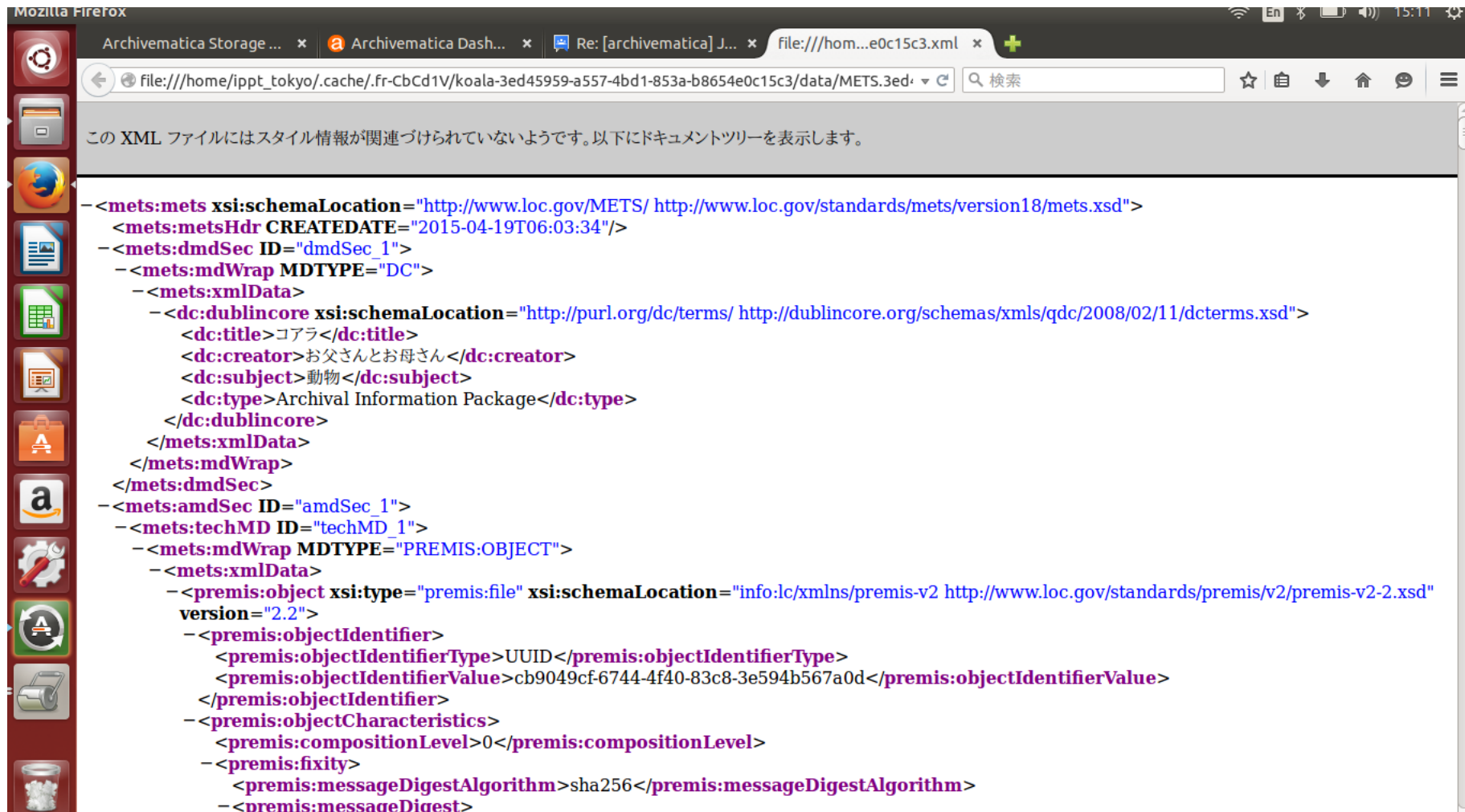
Create an AICDownload CSV

Name	UUID	Size	Created	Status	Encrypted	Actions
transfer1	30350521-8ede-4313-a263-45359b96bb6c	41.6 MB	2020-08-19 16:08	Stored	False	View
transfer10	fb8711a7-c82a-4e30-86c6-1ff9f3fc1071	830.4 KB	2020-08-19 16:47	Stored	False	View
transfer11	e083b0ce-56e8-4318-85bc-f8109cdd3f8f	10.9 MB	2020-08-19 16:48	Stored	False	View
transfer12	266ea405-e21b-42d9-9728-f82d4ae9f4cd	3.1 MB	2020-08-19 16:49	Stored	False	View
transfer3	a443f1cb-fc5e-406b-967e-9c828360f1ae	23.0 MB	2020-08-19 16:12	Stored	False	View
transfer4	e8d75607-fbd5-4c52-b53a-ce4291ff53fe	75.3 MB	2020-08-19 16:12	Stored	False	View
transfer5	e8f9c097-11e3-466b-9e6b-845193d1b391	10.9 MB	2020-08-19 16:24	Stored	False	View
transfer6	d0874403-4f95-49ea-81f4-c1a1201000cc	4.8 MB	2020-08-19 16:45	Stored	False	View
transfer6	803a0041-aecb-4f2d-97bd-d73d919dbc84	627.3 KB	2020-08-19 16:47	Stored	False	View
transfer7	33992409-5673-4424-98b9-e7d440cf2183	1.0 MB	2020-08-19 16:47	Stored	False	View

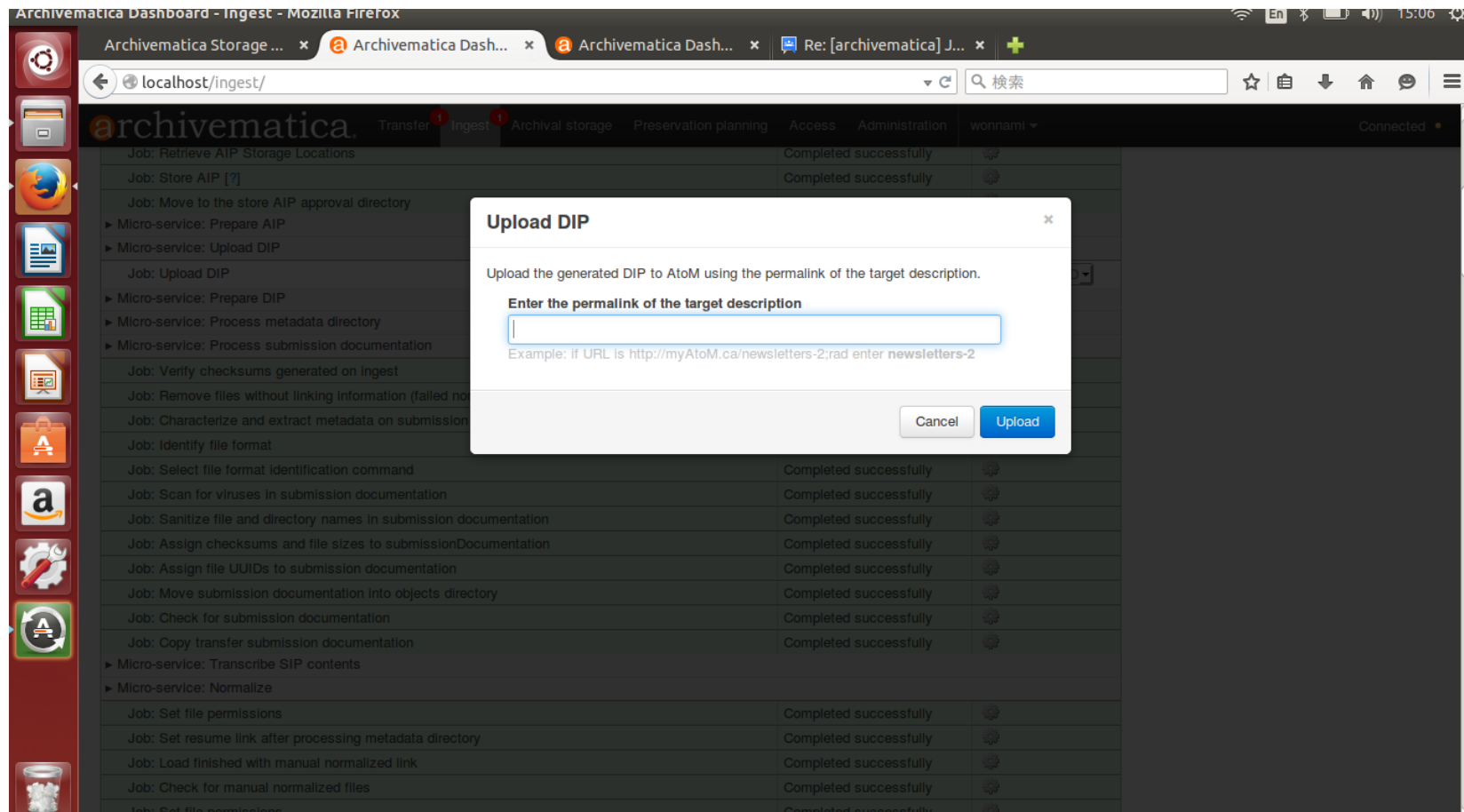
Showing 1 to 10 of 12 entriesSelect columnsFirstPrevious12NextLast

https://www.archivematica.org/en/docs/archivematica-1.13/_images/archival-storage-tab.png

メタデータ：METS.xmlファイル



アクセス：DIPをAtoMにアップロード



アクセス：DIPのアップロード

Holdings Quick search

▼ Fonds [F1 - Atomic Bomb Casualty Com...](#)

▼ Series [03 - ABCC Program Components ...](#)

▼ File [Tumor Registry: 1956-1963 \(Draft\)](#)

image 01

Atomic Bomb Casualty Commission » ABCC Program Components » Tumor Registry: 1956-1963 » image 01



Reports

Import

XML

CSV

Export

Dublin Core 1.1 XML

EAD 2002 XML

Identity elements »

Name and location of repository	National Academy of Sciences
Title	image 01
Name of creator	Atomic Bomb Casualty Commission

Digital object metadata »

Filename	Koala.jpg
Media type	Image
Mime-type	image/jpeg
Filesize	762.5 KiB
Uploaded	October 20, 2014 2:03 AM

運用管理

archivematica

Transfer 1BacklogAppraisalIngest 25Archival storagePreservation planningAccessAdministrationtest ▾

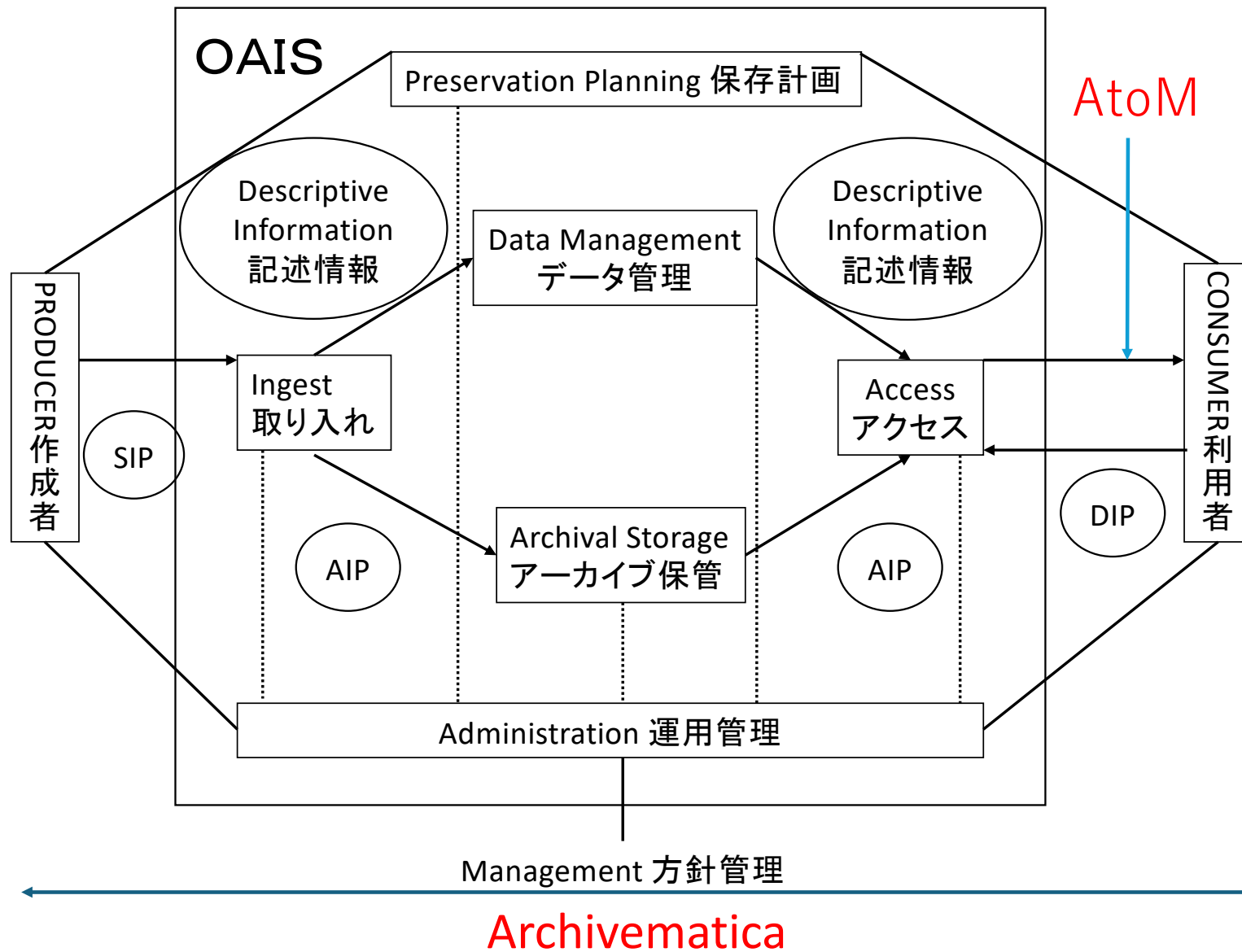
Processing configuration

GeneralFailuresTransfer source locationsAIP storage locationsProcessing storage usageDIP uploadAtoMArchivesSpaceArchivists' ToolkitPREMIS agentREST APIUsersHandle server configLanguageVersion

Processing configuration

Name:	default
Assign UUIDs to directories:	No ▾
Send transfer to quarantine:	No ▾
Remove from quarantine after (days):	days
Generate transfer structure report:	No ▾
Select file format identification command (Transfer):	Identify using Fido ▾
Extract packages:	None ▾
Delete packages after extraction:	None ▾
Perform policy checks on originals:	No ▾
Examine contents:	None ▾
Create SIP(s):	Create single SIP and continue processing ▾
Select file format identification command (Ingest):	Identify using Siegfried ▾
Normalize:	None ▾
Approve normalization:	Yes ▾

https://www.archivematica.org/en/docs/archivematica-1.13/_images/ProcessingConfig.png



OAIS環境→Borealis

