



—議論の前段として—
長期保存のための国際標準の概要

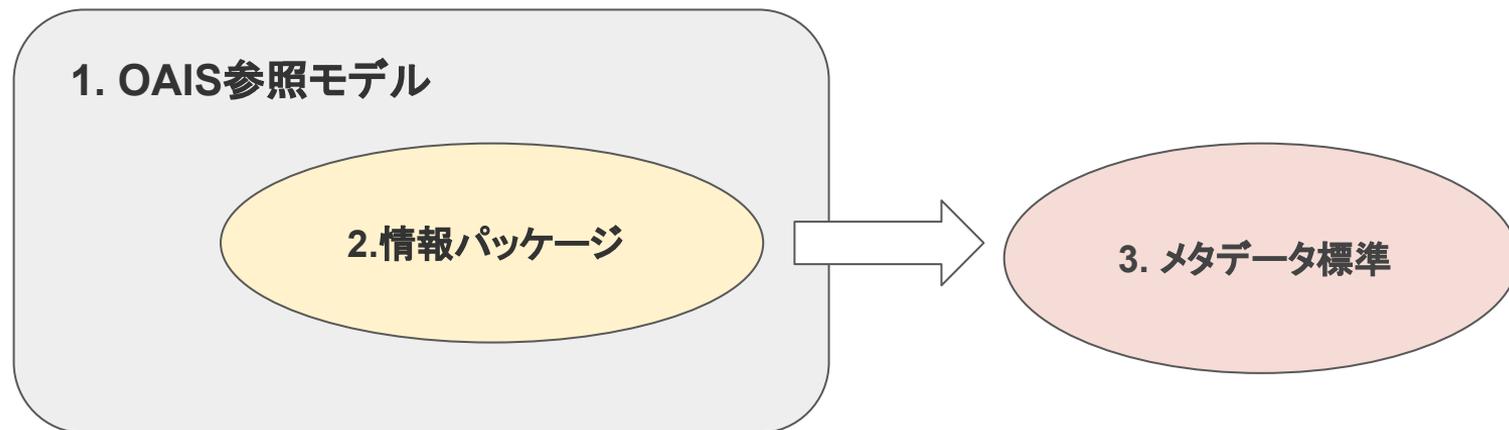
【D メタデータとプラットフォーム】D4 長期保存のための情報パッケージ再考

2024/6/21(金) 14:00~15:30 ハイブリッド

発表者: 公益財団法人渋沢栄一記念財団 / 東京大学 金甫榮(キムボヨン)

発表概要

議論のための、前段となる理論や背景について説明





1. OAIS参照モデル

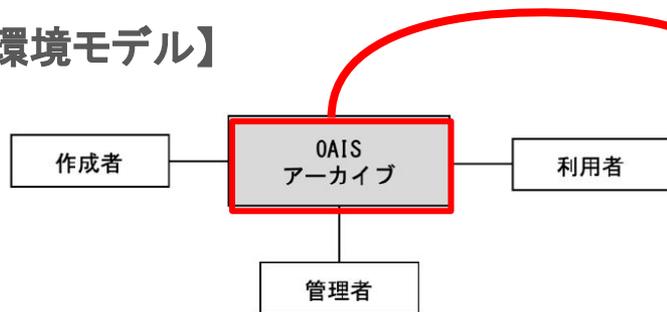
OAIS reference model

OAIS参照モデルの概要

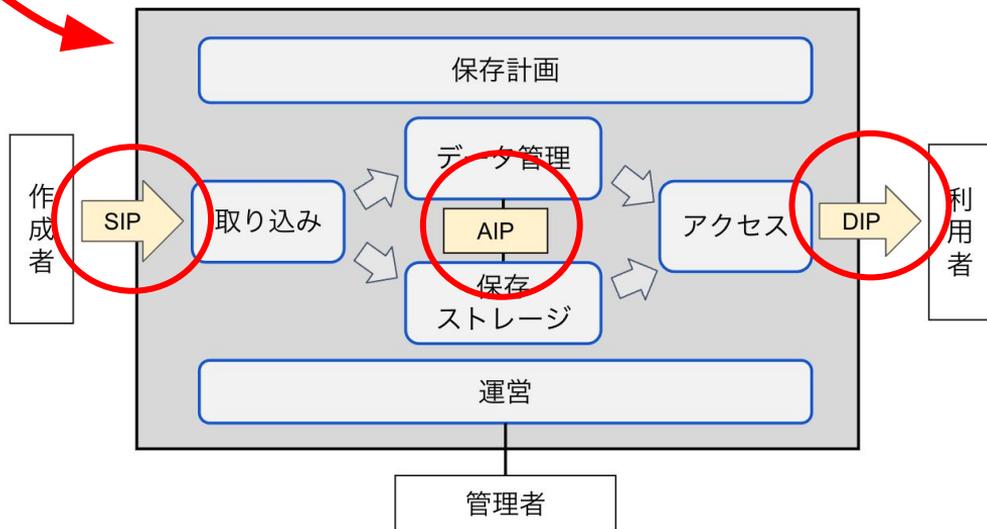
- Open Archival Information Systemの略
- デジタル情報を保存するアーカイブズ機関やシステムに求められる環境、機能、情報に関する概念モデルを提供
- 宇宙データシステム諮問委員会 (Consultative Committee for Space Data Systems : CCSDS) により、宇宙関連データ保存用として開発
- 2002年1月、ISO14721として認定 → 改訂版 ISO14721:2012
- OAIS準拠レポジトリの主な2つの機能
 - 長期保存
 - アクセスを提供

概念モデルと関係

【環境モデル】



【機能モデル】



SIP: Submission Information Package
提出用情報パッケージ

AIP: Archival Information Package
保存用情報パッケージ

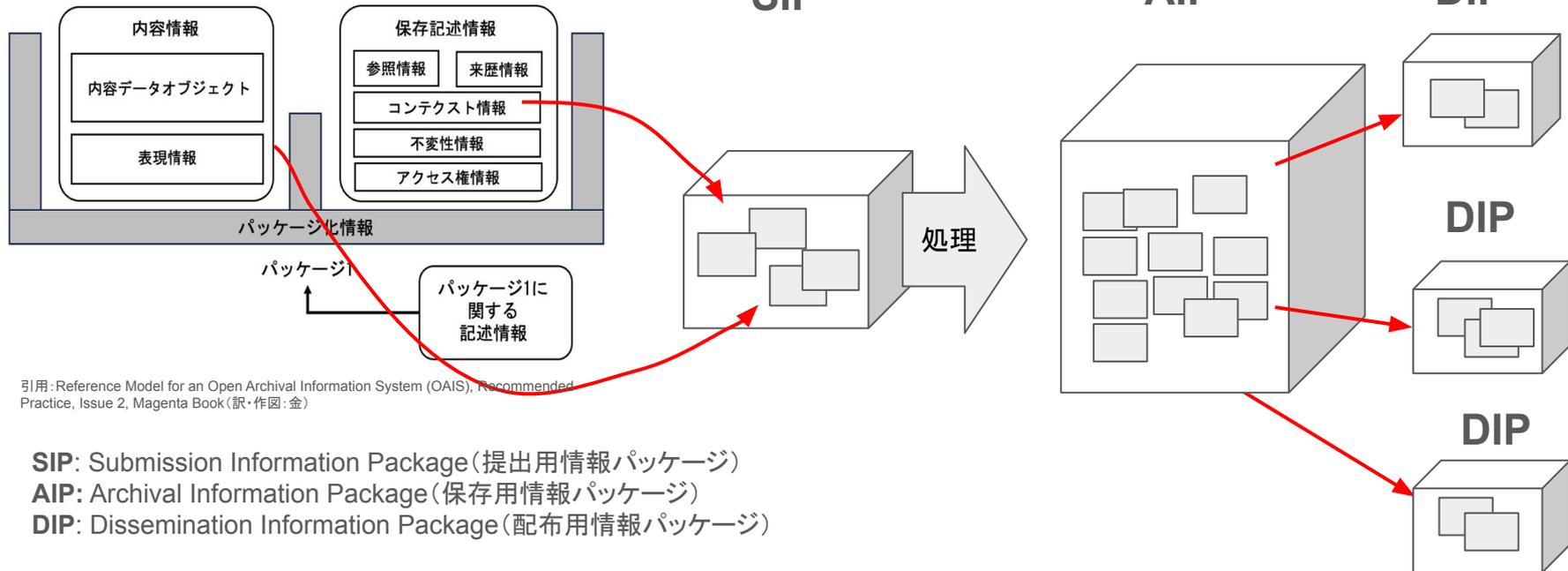
DIP: Dissemination Information Package
配布用情報パッケージ

引用: Reference Model for an Open Archival Information System (OAIS), Recommended Practice, Issue 2, Magenta Book, p.4-1. (訳・作図: 金)

情報パッケージ: AIPの重要性

【情報モデル】

※SIP、AIP、DIPを作成する流れの一例



引用: Reference Model for an Open Archival Information System (OAIS), Recommended Practice, Issue 2, Magenta Book (訳・作図: 金)

SIP: Submission Information Package (提出用情報パッケージ)

AIP: Archival Information Package (保存用情報パッケージ)

DIP: Dissemination Information Package (配布用情報パッケージ)

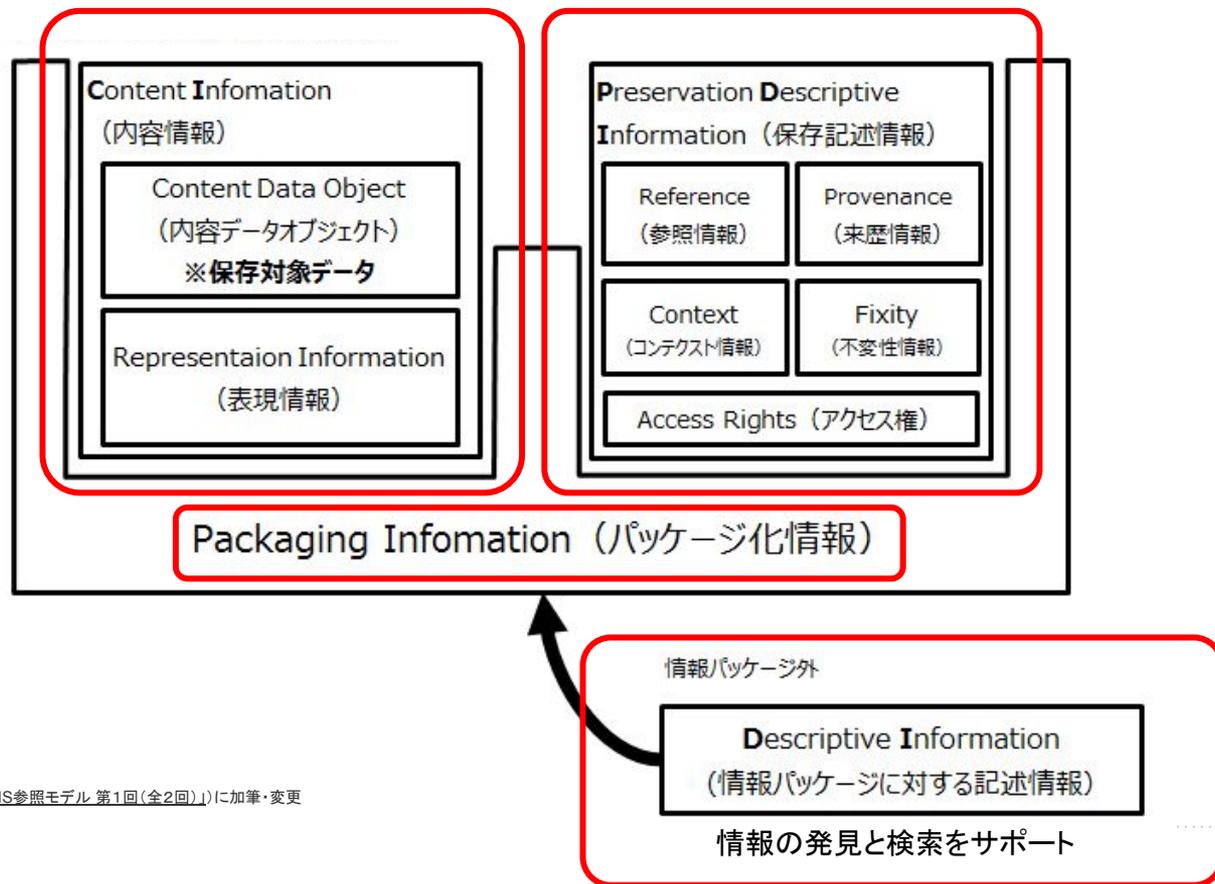
→ 特にAIPを作成することは、デジタル情報の長期保存の要となる



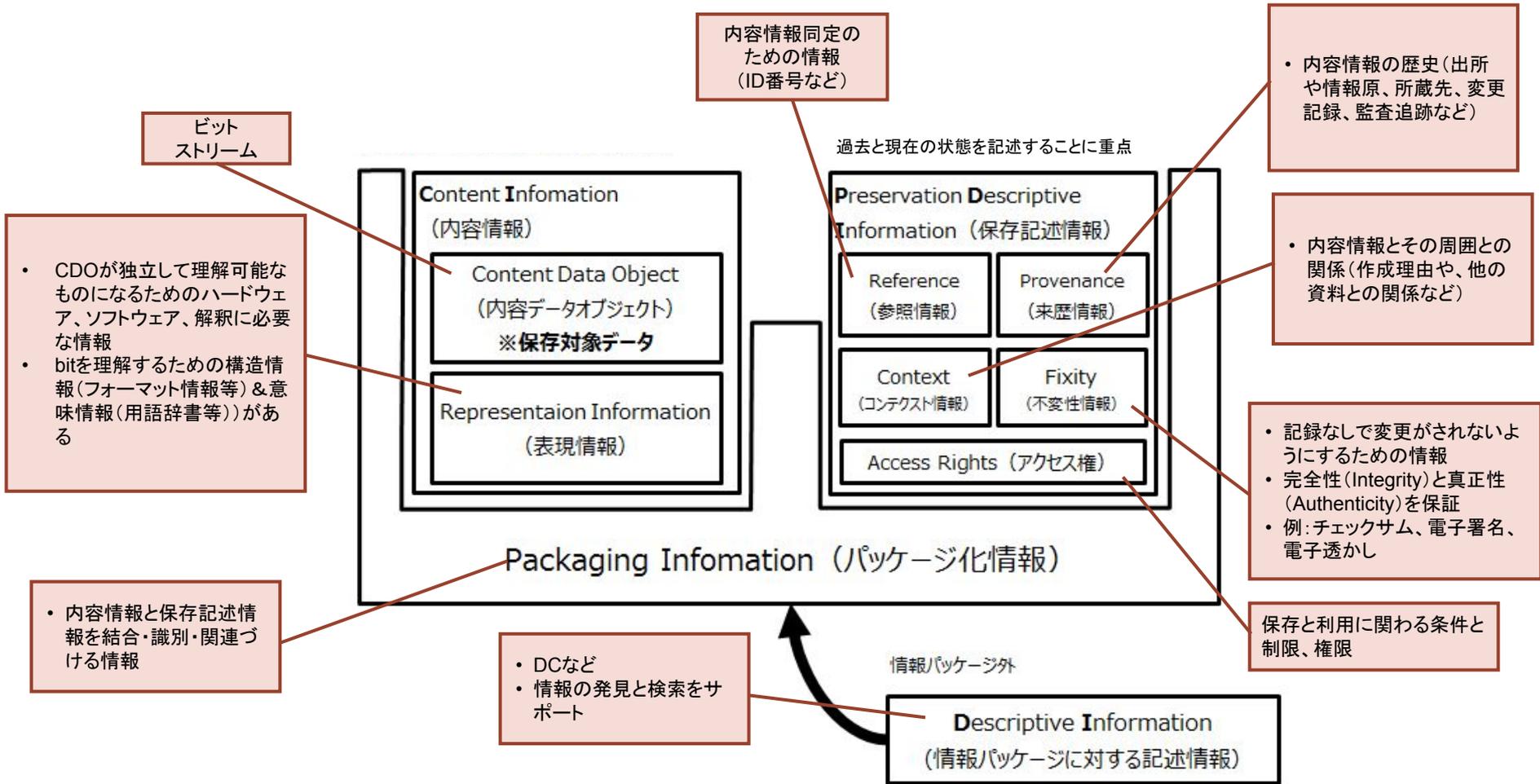
2. 情報パッケージ

Information Packages

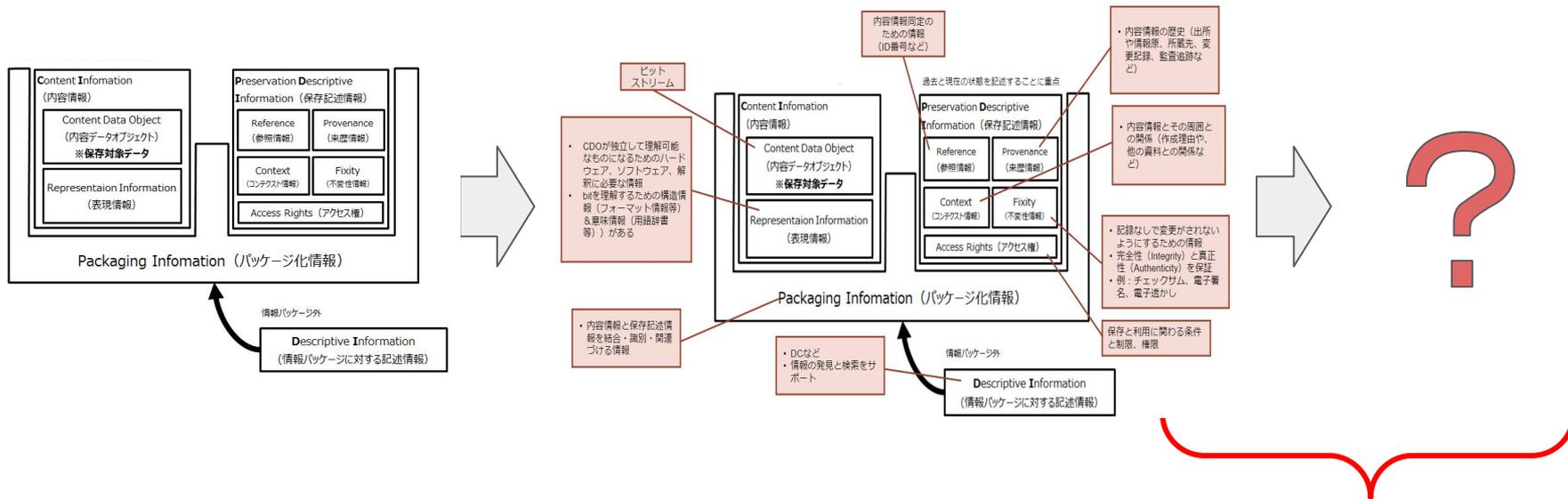
情報パッケージの構成要素



引用図(BDCブログ「情報の長期保存に有力なOAIS参照モデル 第1回(全2回)」)に加筆・変更



概念の具体化が必要

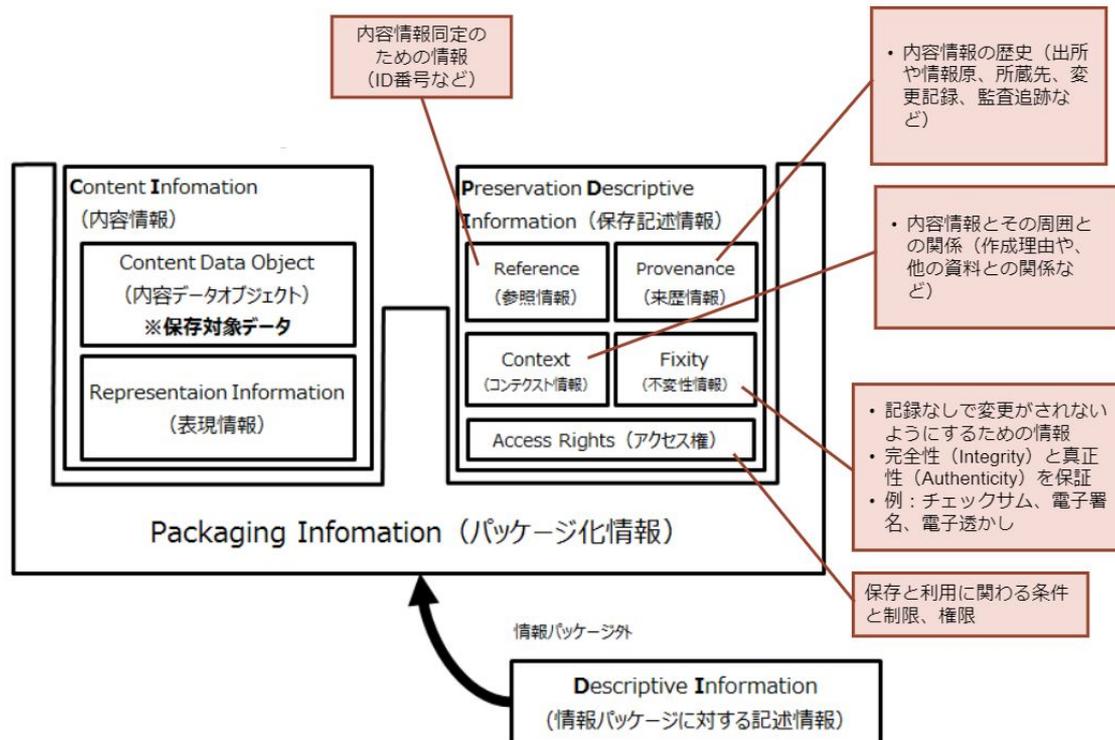


実装のためには具体化が必要
→ 該当するメタデータの決定

特に、具体化が難しいメタデータ

→ 保存記述情報

Preservation Description Information (PDI)



保存記述情報(PDI)の例

【宇宙科学データの例】

※仮訳であることを、ご理解の上ご参考ください。

PDI要素	例	
参照情報	オブジェクト識別子、参照論文、ミッション、観測装置 タイトル、属性セット	
来歴情報	観測装置の記述 研究代表者 処理履歴 保管・取り扱い履歴 センサーの説明	観測装置 観測装置モード デコピュータマップ ソフトウェア・インターフェース仕様 情報プロパティ記述
コンテキスト情報	(装置の)調整履歴、関連データセット ミッション、資金調達履歴	
不変性情報	CRC、チェックサム、リード・ソロモン符号	
アクセス権情報	指定コミュニティの識別(アクセス・コントロール) 保存と流通のための権限付与 不変性と来歴情報へのポイント(デジタル署名や権利者など)	

保存記述情報(PDI)の例

【デジタルライブラリーコレクションの例】

※仮訳であることを、ご理解の上ご参考ください。

PDI要素	例	
参照情報	書誌記述、永続的識別子	
来歴情報	スキャンされたコレクションの場合： <ul style="list-style-type: none">デジタル化に関するメタデータマスターバージョンへのプロセスポイント ボーンデジタル出版物の場合： <ul style="list-style-type: none">デジタル原本へのポイント	保存プロセスに関するメタデータ： <ul style="list-style-type: none">コレクションアイテムの以前のバージョンへのポイント変更履歴情報プロパティ説明
コンテキスト情報	公開時のオリジナル環境における関連文書へのポイント	
不変性情報	電子署名、チェックサム、真正性指標	
アクセス権情報	法的枠組み、ライセンス条項、配布時に適用される権利行使手段の仕様 保存と流通のための権限付与、提出時および保存時に適用される電子透かしに関する情報、不変性と来歴情報(電子署名、権利者など)へのポイント	

保存記述情報(PDI)の例

【ソフトウェアパッケージの例】

※仮訳であることを、ご理解の上ご参考ください。

PDI要素	例	
参照情報	名称 作者/発案者	バージョン番号 シリアル番号
来歴情報	改訂履歴 登録	著作権 情報プロパティ説明
コンテキスト情報	ヘルプファイル ユーザーガイド	関連ソフトウェア 使用言語
不変性情報	証明書、チェックサム、暗号化、CRC	
アクセス権情報	指定コミュニティ、配布時に適用される権利行使手段の仕様 法的枠組み、ライセンス条項 不変性と来歴情報(電子署名、権利者など)へのポインタ	

【宇宙科学データの例】

来歴情報	観測装置の記述 研究代表者 処理履歴 保管・取り扱い履歴 センサーの説明	観測装置 観測装置モード デコミュータマップ ソフトウェア・インターフェース仕様 情報プロパティ記述
コンテキスト情報	(装置の) 調整履歴、関連データセット ミッション、資金調達履歴	

【デジタルライブラリーコレクションの例】

来歴情報	スキャンされたコレクションの場合： <ul style="list-style-type: none">デジタル化に関するメタデータマスターバージョンへのプロセスポイント ボーンデジタル出版物の場合： <ul style="list-style-type: none">デジタル原本へのポイント	保存プロセスに関するメタデータ： <ul style="list-style-type: none">コレクションアイテムの以前のバージョンへのポイント変更履歴情報プロパティ説明
コンテキスト情報	公開時のオリジナル環境における関連文書へのポイント	

【ソフトウェアパッケージの例】

来歴情報	改訂履歴 登録	著作権 情報プロパティ説明
コンテキスト情報	ヘルプファイル ユーザーガイド	関連ソフトウェア 使用言語

OAIS参照モデルの影響

- 保存メタデータ: **PREMIS**
- メタデータのエンコードまたはパッケージ化をサポート: **METS**
- OAIS準拠レポジトリ構築のためのアプリケーション: **Archivematica**
- OAISタイプのアーカイブと作成者 (producer) 間の相互作用に関する記述標準
:ISO20652:2006: Producer-Archive Interface - Methodology Abstract Standard (PAIMAS)
- 信頼できるレポジトリの監査と認証ツール (Trustworthy Repositories Audit & Certification: Criteria & Checklist, TRAC → Trusted Digital Repository (TDR):ISO16363:2012)



3. メタデータ標準

METS, PREMISの概要

Preservation Metadata: Implementation Strategies (PREMIS)

- デジタル保存のためのメタデータ標準
(<https://www.loc.gov/standards/premis/>)
- 2005年Ver.1.0発表、最新Ver.3.0
- データモデル
 - **オブジェクト** (またはデジタルオブジェクト) : デジタル保存の対象となる情報の個別単位
 - **環境** : オブジェクトをサポートする技術 (ソフトウェアまたはハードウェア)
 - **エージェント** : イベント、権利ステートメント、環境に関連する人、組織、またはソフトウェアプログラムシステム
 - **イベント** : オブジェクトまたはエージェントに関与または影響するアクション
 - **権利ステートメント** : オブジェクトやエージェントに関連する権利または許諾

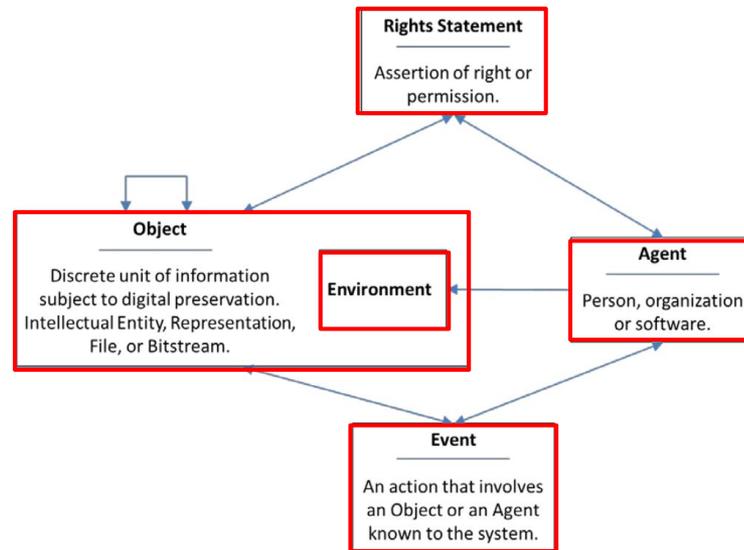


Figure 1: The PREMIS Data Model

引用: PREMIS Editorial Committee, PREMIS Data Dictionary for Preservation Metadata version 3.0, 2015, p.6

PREMIS記述例

```
<premis:object xsi:type="premis:file">
```

```
<premis:objectIdentifier>
```

```
<premis:objectIdentifierType>local</premis:objectIdentifierType>
```

```
<premis:objectIdentifierValue>001</premis:objectIdentifierValue>
```

```
</premis:objectIdentifier>
```

```
<premis:event>
```

```
<premis:eventIdentifier>
```

```
<premis:eventIdentifierType>UUID</premis:eventIdentifierType>
```

```
<premis:eventIdentifierValue>1d1f888-fbf7-47e5-987e-5e5bc72ac2c5</premis:eventIdentifierValue>
```

```
</premis:eventIdentifier>
```

```
<premis:eventType>format identification</premis:eventType>
```

```
<premis:eventDateTime>2014-04-16T22:22:22</premis:eventDateTime>
```

```
<premis:eventDetailInformation>
```

```
<premis:eventDetail>program="fido"</premis:eventDetail>
```

```
</premis:eventDetailInformation>
```

PREMIS記述例

<premis:agent>

```
<premis:agentIdentifier>
  <premis:agentIdentifierType>local</premis:agentIdentifierType>
  <premis:agentIdentifierValue>NRI</premis:agentIdentifierValue>
</premis:agentIdentifier>
<premis:agentName>Not a Real Institution</premis:agentName>
<premis:agentType authority="agentType"
authorityURI="http://id.loc.gov/vocabulary/preservation/agentType"
valueURI="http://id.loc.gov/vocabulary/preservation/agentType/org">organization</premis:agentType>
</premis:agent>
```

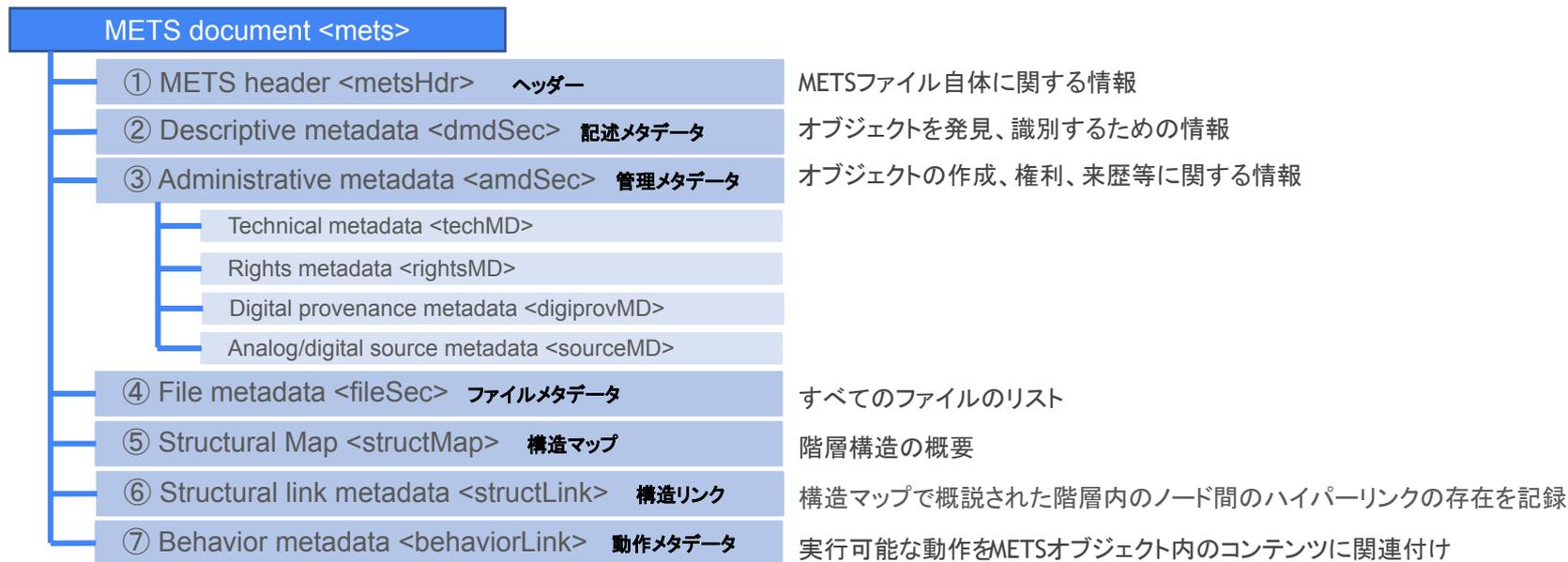
<premis:rights>

```
<premis:rightsStatement>
  <premis:rightsStatementIdentifier>
    <premis:rightsStatementIdentifierType>local</premis:rightsStatementIdentifierType>
    <premis:rightsStatementIdentifierValue>1234</premis:rightsStatementIdentifierValue>
  </premis:rightsStatementIdentifier>
  <premis:rightsBasis authority="rightsBasis"
authorityURI="http://id.loc.gov/vocabulary/preservation/rightsBasis"
valueURI="http://id.loc.gov/vocabulary/preservation/rightsBasis/cop">Copyright</premis:rightsBas
is>
```

Metadata Encoding and Transmission Standard (METS)

- <https://www.loc.gov/standards/mets/>
- 異なるメタデータを共通のメタデータとして一元的に扱い、リポジトリ間での交換を容易にするための符号化と送信のフレームワーク

※METSファイルの構成



METS記述例

```
<metsHdr CREATEDATE="2003-07-04T15:00:00" RECORDSTATUS="Complete">
  <agent ROLE="CREATOR" TYPE="INDIVIDUAL">
    <name>Jerome McDonough</name>
  </agent>
  <agent ROLE="ARCHIVIST" TYPE="INDIVIDUAL">
    <name>Ann Butler</name>
  </agent>
</metsHdr>

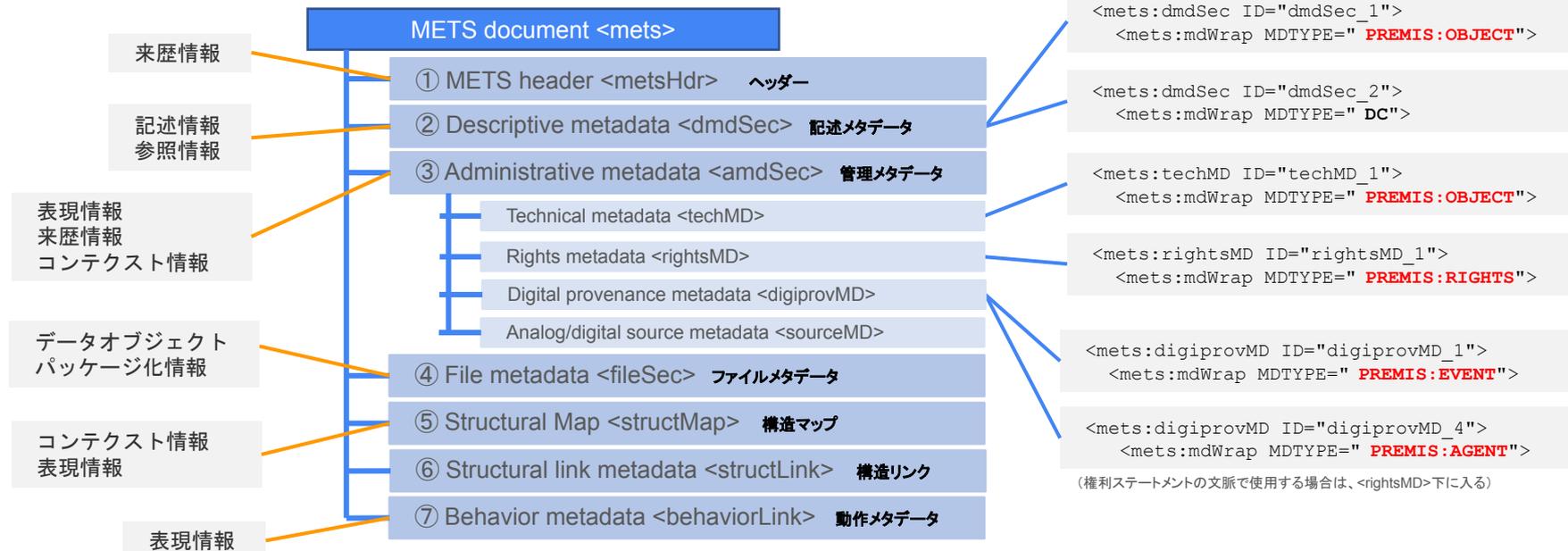
<dmdSec ID="dmd002">
  <mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Dublin Core Metadata">
    <xmlData>
      <dc:title>Alice's Adventures in Wonderland</dc:title>
      <dc:creator>Lewis Carroll</dc:creator>
      <dc:date>between 1872 and 1890</dc:date>
      <dc:publisher>McCloughlin Brothers</dc:publisher>
      <dc:type>text</dc:type>
    </xmlData>
  </mdWrap>
</dmdSec>

<techMD ID="AMD001">
  <mdWrap MIMETYPE="text/xml" MDTYPE="NISOIMG" LABEL="NISO Img. Data">
    <xmlData>
      <niso:MIMETYPE>image/tiff</niso:MIMETYPE>
      <niso:Compression>LZW</niso:Compression>
      <niso:PhotometricInterpretation>8</niso:PhotometricInterpretation>
      <niso:Orientation>1</niso:Orientation>
      <niso:ScanningAgency>NYU Press</niso:ScanningAgency>
    </xmlData>
  </mdWrap>
</techMD>
```

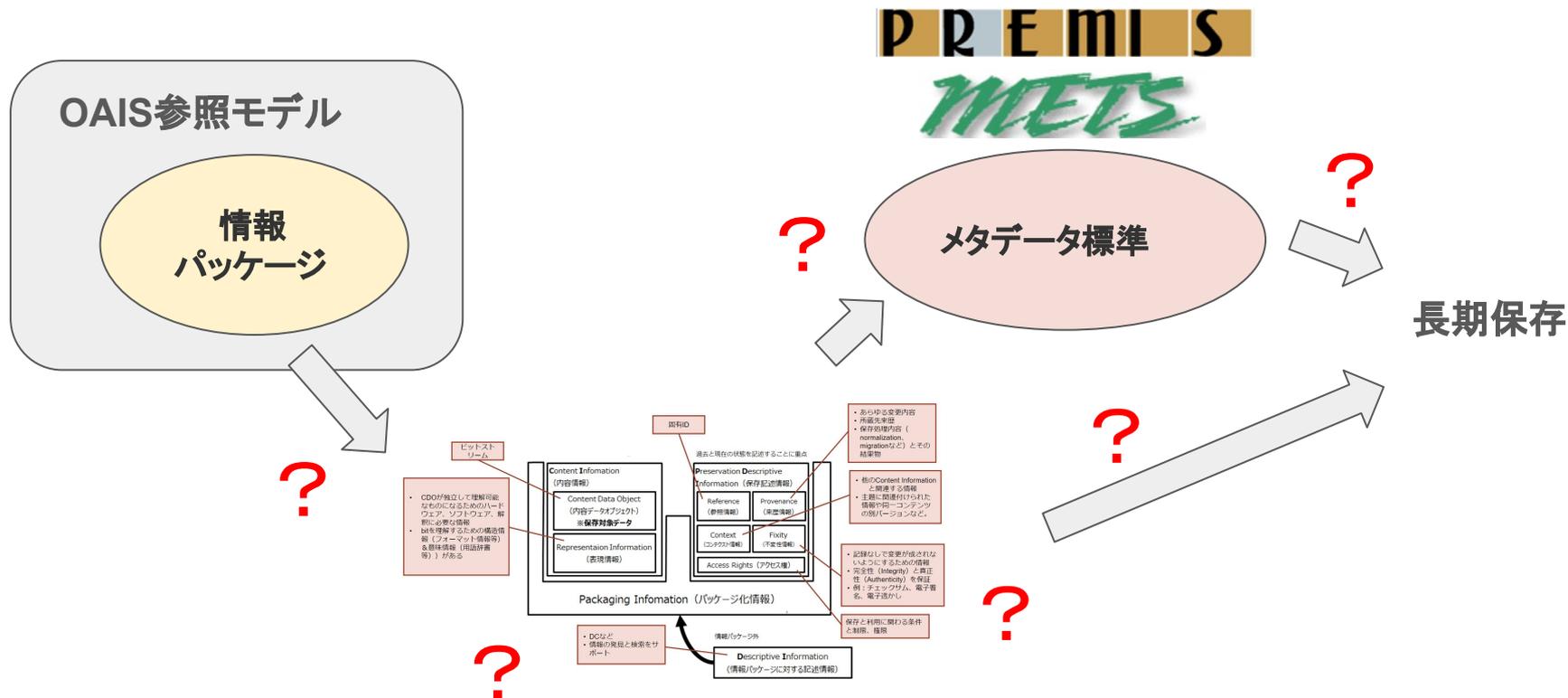
METSとPREMISの活用例

※対応するOAIS情報要素の例 ※METSファイルの構成

※PREMISメタデータ



まとめ:本日の議論について



ご清聴ありがとうございます。