



# **PREMIS implementation in pre-ingestion workflow**

**Angela Di Iorio**



**PREMIS Implementation Fair 2013**

Lisbon Technical University Congress center

*September 5° 2013, Lisbon*



## SapienzaDigitalLibrary

Collezioni ■ Servizi ■ Dona una risorsa ■ News ■ Community

### RICERCA

Digital Library  Web site

Tutti



RICERCA AVANZATA

### NAVIGA



TESTI



IMMAGINI



VIDEO



AUDIO



MAPPE



3D



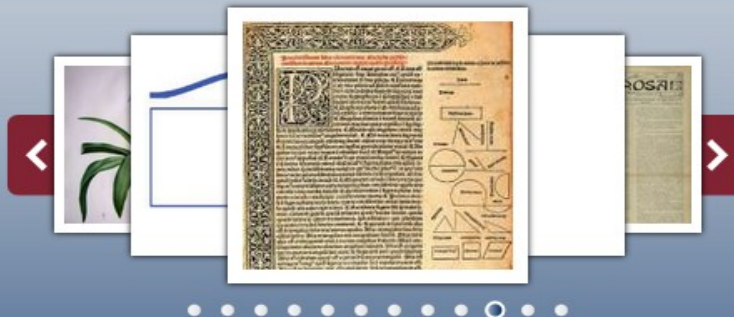
SOFTWARE

### ESPLORA

#### Il fondo librario antico della Biblioteca del dipartimento di matematica Guido Castelnuovo

La biblioteca possiede un pregiato fondo antico: circa 2300 opere datate tra il 1482 ed il 1830. Le opere provengono, per la maggior parte, dalla Biblioteca della Reale Scuola di Applicazione degli In...

Composta da 34 risorse digitali  
Copertura temporale: 1482 - 1687



### NEWS

- Energie per la ricerca
- Università, trasferimento della conoscenza e diritto d'autore dell'era digitale
- Google Research Awards alla Sapienza

[Leggi tutte le news](#)

### ULTIMI ARRIVI

L'ultima collezione aggiunta in Sapienza Digital Library è



"Architettura brasiliana vista da Carlo Severati", composta da 96 risorse digitali.

### DONA UNA RISORSA

Condividi in rete le tue creazioni. Valorizza il tuo patrimonio culturale sulla Sapienza Digital Library.



### ESPLORA I TEMI DELLE COLLEZIONI

architettura astrofisica astronomia bahia barocco chimica **cinema** danza documentari fisica fisica sperimentale fisica teorica lazio **lirica** necropoli oggetti di uso tecnico e scientifico preistoria **prosa** strumenti di misurazione **videoarte**

Screenshot of the portal prototype



**Rome**

geographically dispersed  
replicated storage space

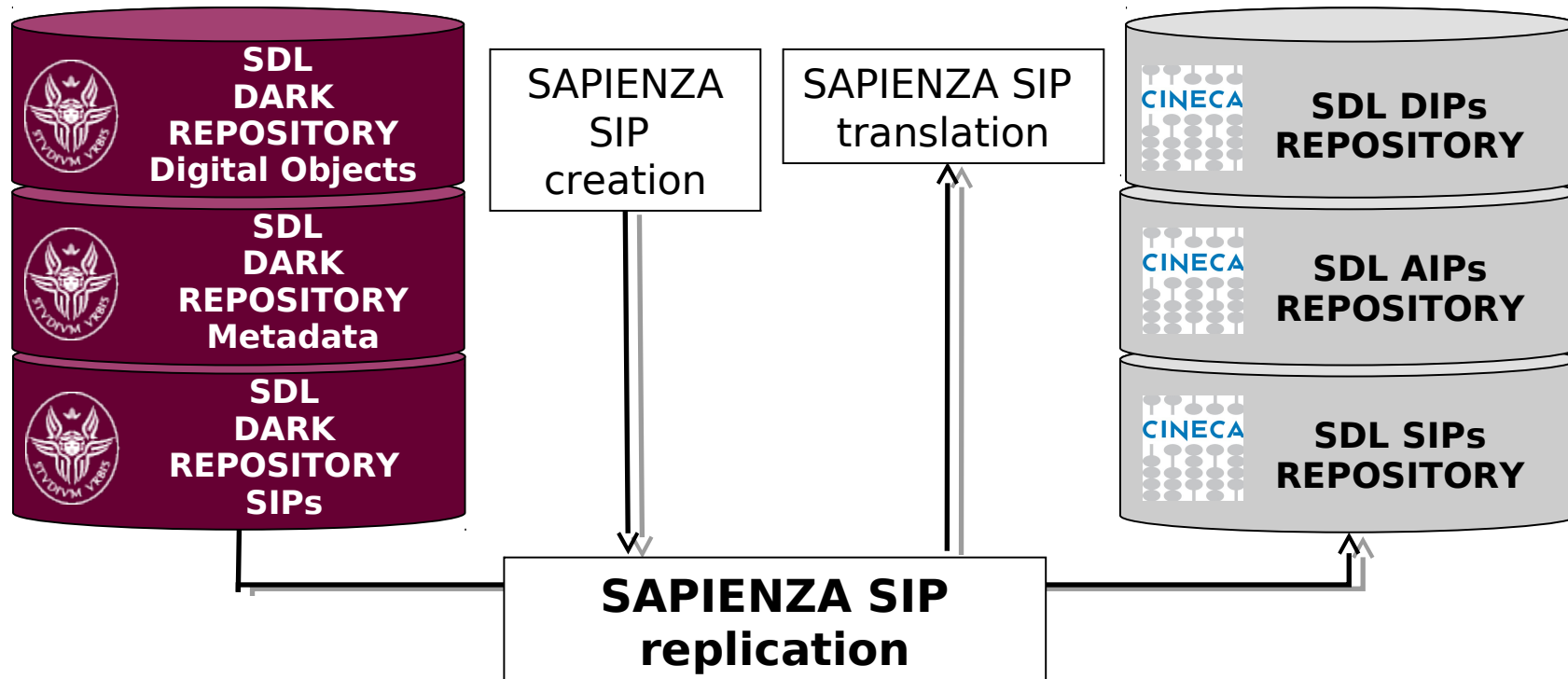
**Bologna**



SAPIENZA  
UNIVERSITÀ DI ROMA

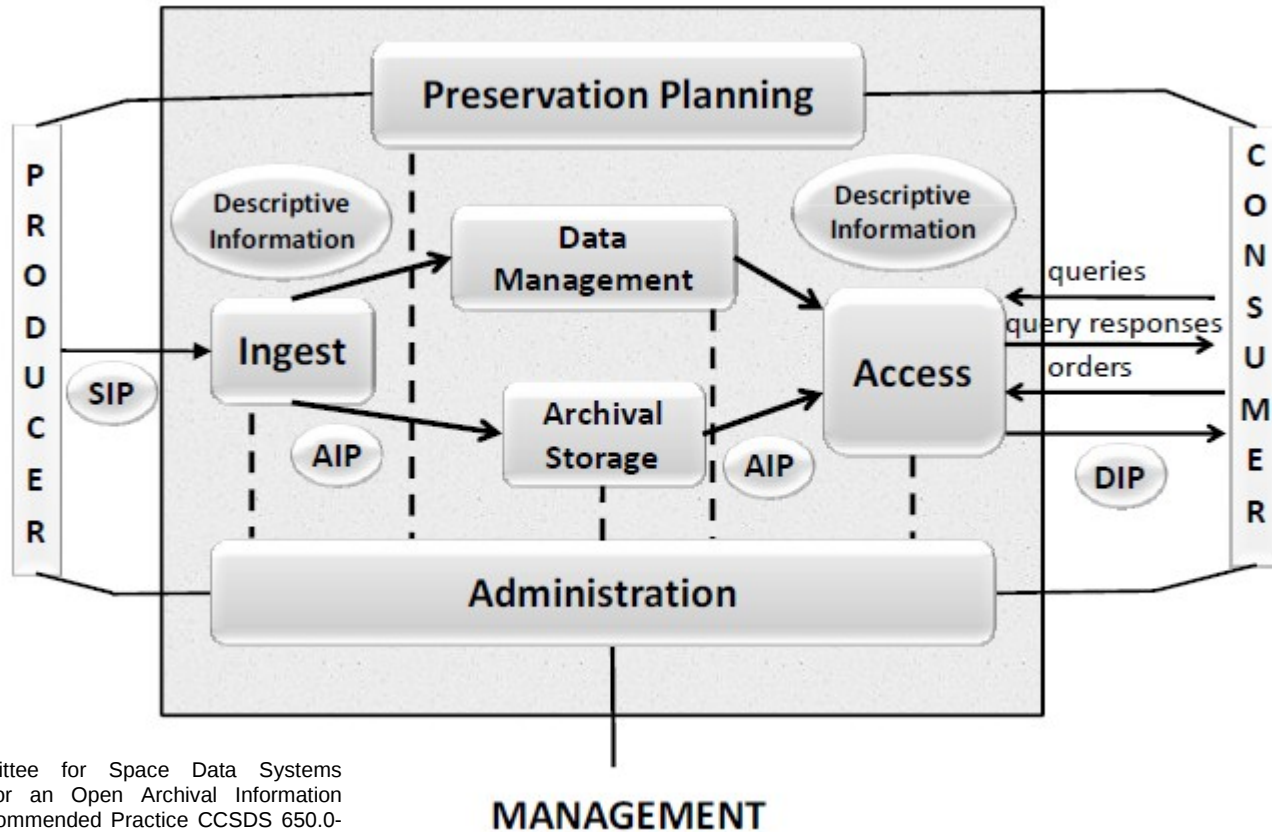


**Inter  
University  
Consortium**

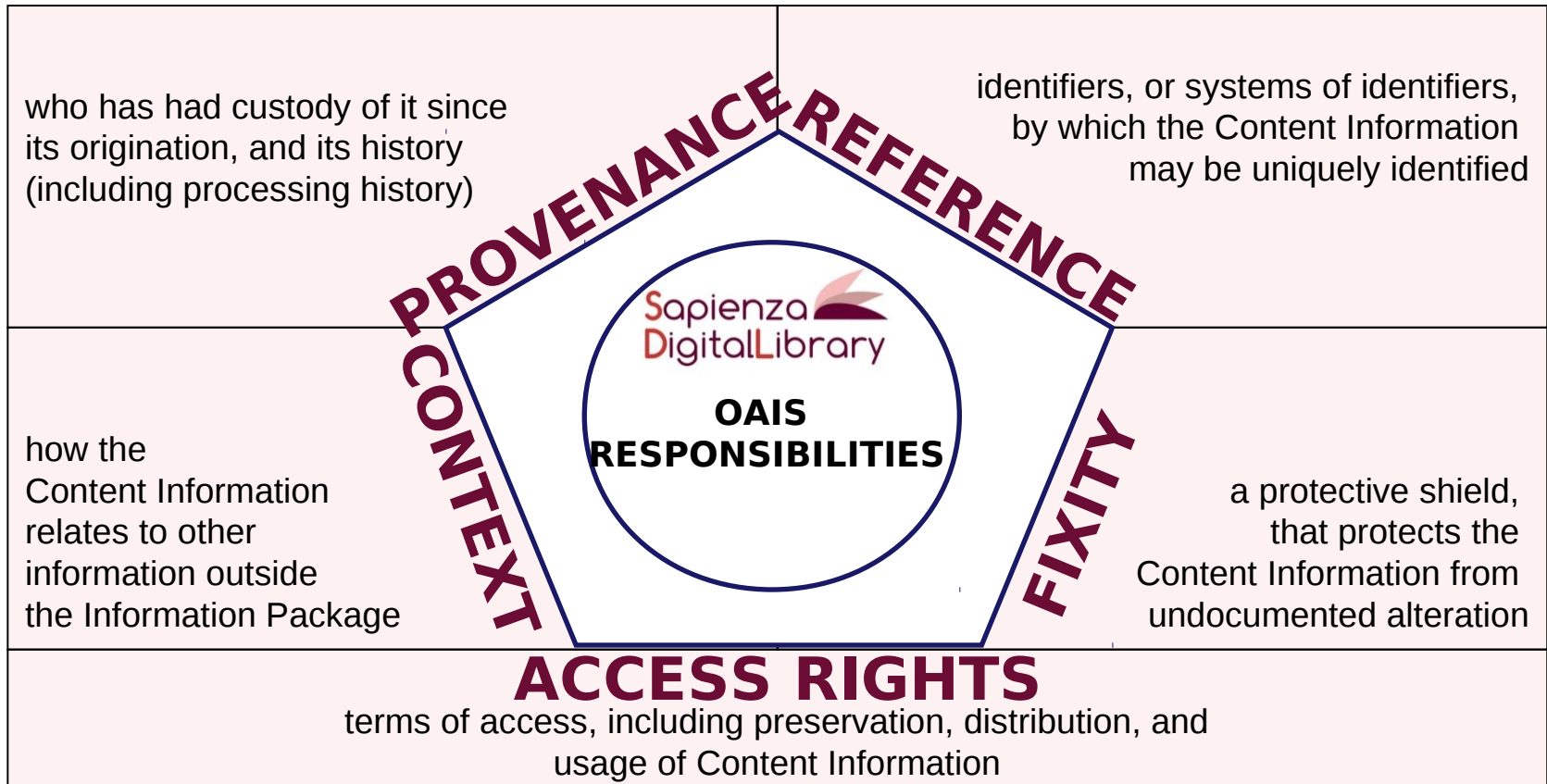




## Open Archival Information System



Consultative Committee for Space Data Systems  
Reference Model for an Open Archival Information  
System (OAIS), Recommended Practice CCSDS 650.0-  
M-2 Magenta Book, June 2012,  
<http://public.ccsds.org/publications/archive/652x0m1.pdf>



a. Di Iorio, M. Schaerf, M. Guercio, et al., „A digital infrastructure for trustworthiness The Sapienza Digital Library experience“, will be published in Springer LNCS Volume Number: 385: Bridging between Cultural Heritage Institutions -- Proceedings of the 9th Italian Research Conference on Digital Libraries ([IRCDL 2013](#))



## 1.1 objectIdentifier (M, R)

### 1.1.1 objectIdentifierType (M, NR)

### 1.1.2 objectIdentifierValue (M, NR)

## 1.2 objectCategory (M, NR)

### 1.3 preservationLevel (O, R) [representation, file]

#### 1.3.1 preservationLevelValue (M, NR) [representation, file]

#### 1.3.2 preservationLevelRole (O, NR) [representation, file]

#### 1.3.3 preservationLevelRationale (O, R) [representation, file]

#### 1.3.4 preservationLevelDateAssigned (O, NR) [representation, file]

### 1.4 significantProperties (O, R)

#### 1.4.1 significantPropertiesType (O, NR)

#### 1.4.2 significantPropertiesValue (O, NR)

#### 1.4.3 significantPropertiesExtension (O, R)

## 1.5 objectCharacteristics (M, R) [file, bitstream]

### 1.5.1 compositionLevel (M, NR) [file, bitstream]

### 1.5.2 fixity (O, R) [file, bitstream]

#### 1.5.2.1 messageDigestAlgorithm (M, NR) [file, bitstream]

#### 1.5.2.2 messageDigest (M, NR) [file, bitstream]

#### 1.5.2.3 messageDigestOriginator (O, NR) [file, bitstream]

### 1.5.3 size (O, NR) [file, bitstream]

### 1.5.4 format (M, R) [file, bitstream]

#### 1.5.4.1 formatDesignation (O, NR) [file, bitstream]

##### 1.5.4.1.1 formatName (M, NR) [file, bitstream]

##### 1.5.4.1.2 formatVersion (O, NR) [file, bitstream]

#### 1.5.4.2 formatRegistry (O, NR) [file, bitstream]

##### 1.5.4.2.1 formatRegistryName (M, NR) [file, bitstream]

##### 1.5.4.2.2 formatRegistryKey (M, NR) [file, bitstream]

##### 1.5.4.2.3 formatRegistryRole (O, NR) [file, bitstream]

#### 1.5.4.3 formatNote (O, R) [file, bitstream]

### 1.10 relationship (O, R)

#### 1.10.1 relationshipType (M, NR)

#### 1.10.2 relationshipSubType (M, NR)

#### 1.10.3 relatedObjectIdentification (M, R)

##### 1.10.3.1 relatedObjectIdentifierType (M, NR)

##### 1.10.3.2 relatedObjectIdentifierValue (M, NR)

##### 1.10.3.3 relatedObjectSequence (O, NR)

#### 1.10.4 relatedEventIdentification (O, R)

##### 1.10.4.1 relatedEventIdentifierType (M, NR)

##### 1.10.4.2 relatedEventIdentifierValue (M, NR)

##### 1.10.4.3 relatedEventSequence (O, NR)

### 1.5.5 creatingApplication (O, NR) [file, bitstream]

#### 1.5.5.1 creatingApplicationName (O, NR) [file, bitstream]

#### 1.5.5.2 creatingApplicationVersion (O, NR) [file, bitstream]

#### 1.5.5.3 dateCreatedByApplication (O, NR) [file, bitstream]

#### 1.5.5.4 creatingApplicationExtension (O, R) [file, bitstream]

### 1.5.6 inhibitors (O, R) [file, bitstream]

#### 1.5.6.1 inhibitorType (M, NR) [file, bitstream]

#### 1.5.6.2 inhibitorTarget (O, R) [file, bitstream]

#### 1.5.6.3 inhibitorKey (O, NR) [file, bitstream]

### 1.5.7 objectCharacteristicsExtension (O, R) [file, bitstream]

## 1.6 originalName (O, NR) [representation, file]

## 1.7 storage (O, R) [file, bitstream]

### 1.7.1 contentLocation (O, NR) [file, bitstream]

#### 1.7.1.1 contentLocationType (M, NR) [file, bitstream]

#### 1.7.1.2 contentLocationValue (M, NR) [file, bitstream]

### 1.7.2 storageMedium (O, NR) [file, bitstream]

## 1.8 environment (O, R)

### 1.8.1 environmentCharacteristic (O, NR)

### 1.8.2 environmentPurpose (O, R)

### 1.8.3 environmentNote (O, R)

### 1.8.4 dependency (O, R)

#### 1.8.4.1 dependencyName (O, R)

#### 1.8.4.2 dependencyIdentifier (O, R)

##### 1.8.4.2.1 dependencyIdentifierType (M, NR)

##### 1.8.4.2.2 dependencyIdentifierValue (M, NR)

### 1.8.5 software (O, R)

#### 1.8.5.1 swName (M, NR)

#### 1.8.5.2 swVersion (O, NR)

#### 1.8.5.3 swType (M, NR)

#### 1.8.5.4 swOtherInformation (O, R)

#### 1.8.5.5 swDependency (O, R)

### 1.8.6 hardware (O, R)

#### 1.8.6.1 hwName (M, NR)

#### 1.8.6.2 hwType (M, NR)

#### 1.8.6.3 hwOtherInformation (O, R)

### 1.8.7 environmentExtension (O, R)

## 1.9 signatureInformation (O, R) [file, bitstream]

### 1.9.1 signature (O, R)

#### 1.9.1.1 signatureEncoding (M, NR) [file, bitstream]

#### 1.9.1.2 signer (O, NR) [file, bitstream]

#### 1.9.1.3 signatureMethod (M, NR) [file, bitstream]

#### 1.9.1.4 signatureValue (M, NR) [file, bitstream]

#### 1.9.1.5 signatureValidationRules (M, NR) [file, bitstream]

#### 1.9.1.6 signatureProperties (O, R) [file, bitstream]

#### 1.9.1.7 keyInformation (O, NR) [file, bitstream]

### 1.9.2 signatureInformationExtension (O, R) [file, bitstream]

## 1.11 linkingEventIdentifier (O, R)

### 1.11.1 linkingEventIdentifierType (M, NR)

### 1.11.2 linkingEventIdentifierValue (M, NR)

## 1.12 linkingIntellectualEntityIdentifier (O, R)

### 1.12.1 linkingIntellectualEntityIdentifierType (M, NR)

### 1.12.2 linkingIntellectualEntityIdentifierValue (M, NR)

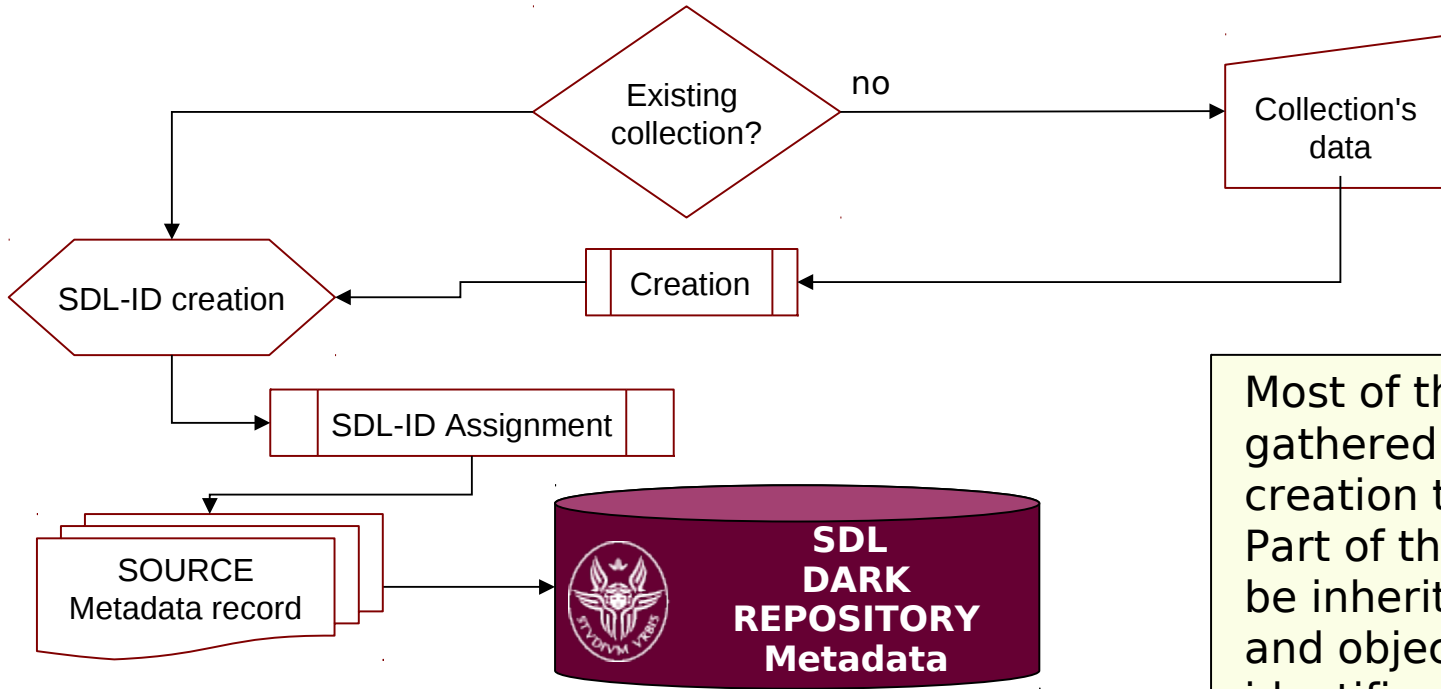
## 1.13 linkingRightsStatementIdentifier (O, R)

### 1.13.1 linkingRightsStatementIdentifierType (M, NR)

### 1.13.2 linkingRightsStatementIdentifierValue (M, NR)



# Collection aware workflow



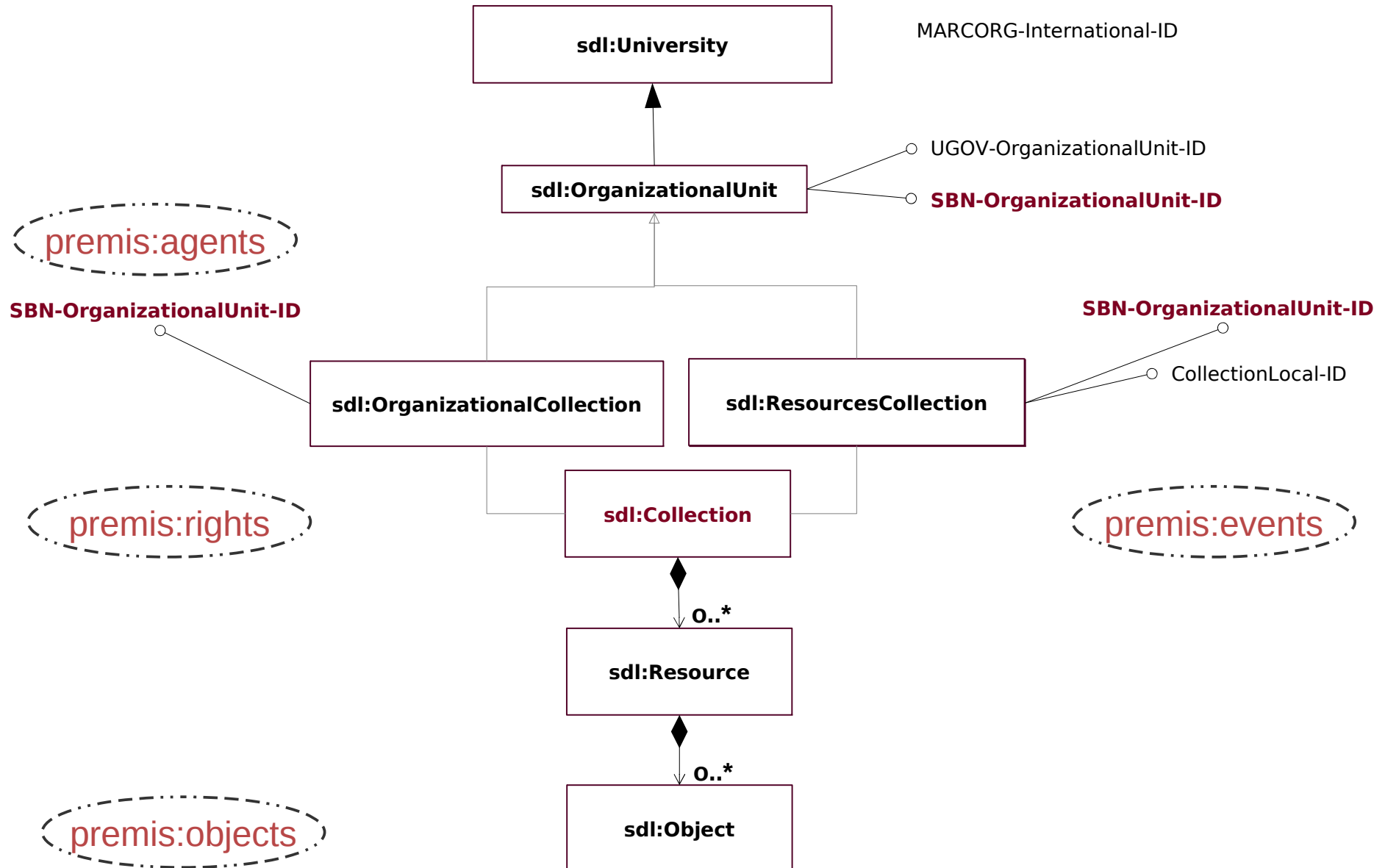
Most of the information is gathered at the collection creation time. Part of the information will be inherited by resources and objects: ENTITIES' identifiers family, common RIGHTS references, information about Provenance

licenseInformation-nr  
licenseDocumentationIdentifierType-r  
licenseDocumentationIdentifierValue-r  
licenseDocumentationRole-r[nr]  
licenseTerms-nr  
licenseNote-nr  
licenseApplicableStartDate-nr  
licenseApplicableEndDate-nr

statuteJurisdiction\*nr  
statuteCitation\*nr  
statuteInformationDeterminationDate-nr  
statuteNote-r  
statuteDocumentationIdentifierType-r  
statuteDocumentationIdentifierValue-r  
statuteDocumentationRole-r[nr]  
statuteApplicableStartDate-nr  
statuteApplicableEndDate-nr

copyrightStatus-nr  
copyrightJurisdiction\*nr  
copyrightStatusDeterminationDate-nr  
copyrightNote-r  
copyrightDocumentationIdentifierType-r  
copyrightDocumentationIdentifierValue-r  
copyrightDocumentationRole-r[nr]  
copyrightApplicableStartDate-nr  
copyrightApplicableEndDate-nr

otherRightsDocumentationIdentifierType\*nr  
otherRightsDocumentationIdentifierValue\*nr  
otherRightsDocumentationRole-nr  
otherRightsBasis\*nr  
otherRightsApplicableStartDate-nr  
otherRightsApplicableEndDate-nr  
otherRightsNote-r

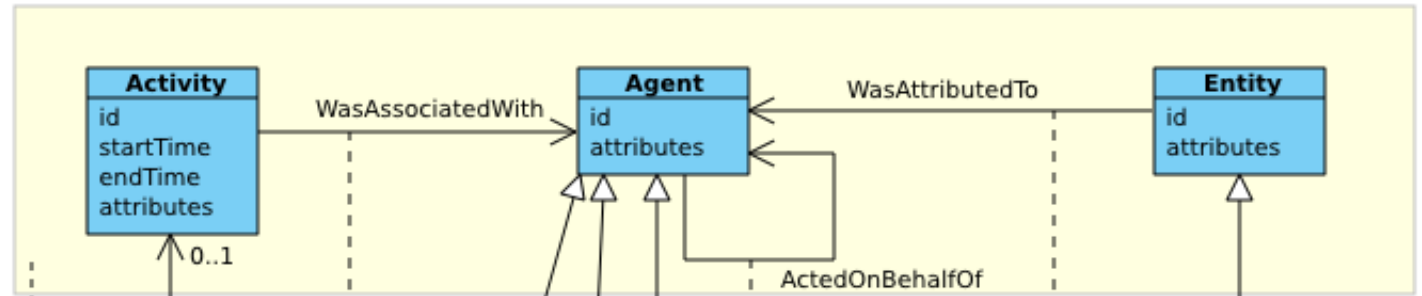






sdl:University

MARCORG-International-ID

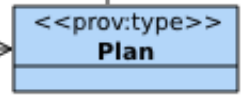
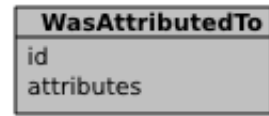
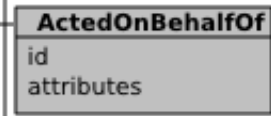
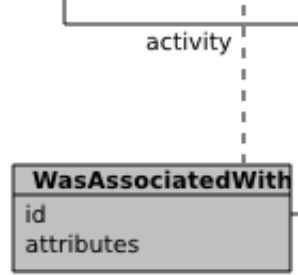


SBN-OI

Unit-ID

al-ID

Classes and associations within this area belong to PROV core



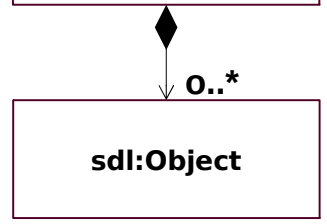
premis:rights

premis:events



premis:agents

premis:objects

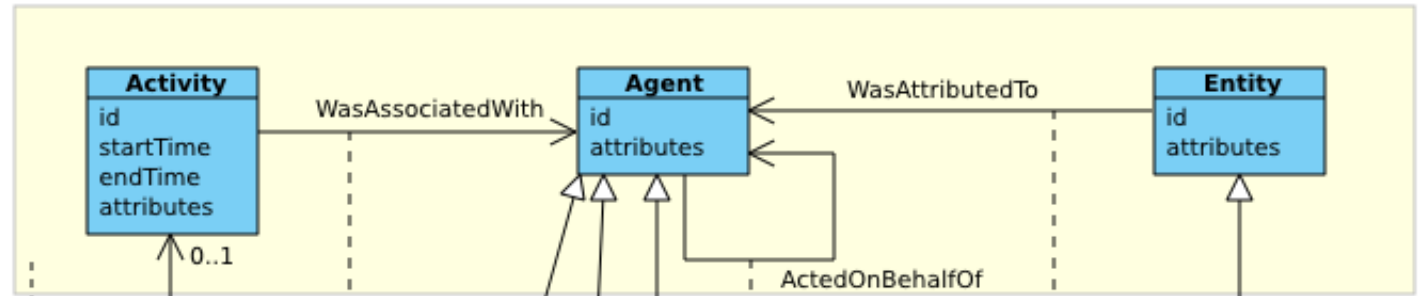


premis-prov ontology alignment



sdl:University

MARCORG-International-ID

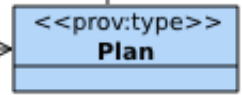
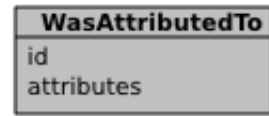
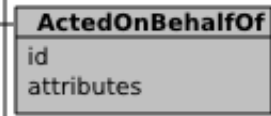
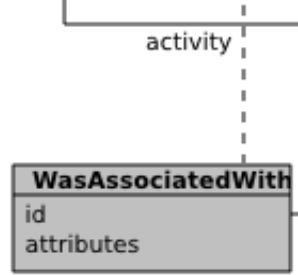


SBN-OI

Unit-ID

al-ID

Classes and associations within this area belong to PROV core



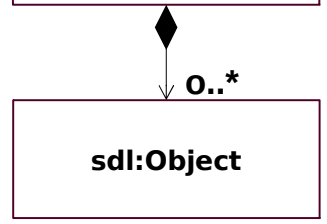
premis:rights

premis:events



premis:objects

premis:agents



premis-prov ontology alignment



## INTELLECTUAL ENTITY

### OBJECTS

oldentifierType\*nr  
 oldentifierValue\*nr  
 oCategory\*nr  
 compositionLevel\*nr  
 messageDigestAlgorithm\*nr  
 messageDigest\*nr  
 messageDigestOriginator\*nr  
 size-nr  
 formatName\*nr  
 formatVersion-nr  
 formatRegistryName\*nr  
 formatRegistryKey\*nr  
 formatRegistryRole-nr  
 formatNote-r  
 relationshipType-r  
 relationshipSubType-r  
 relatedObjectIdentifierType[\*r]  
 relatedObjectIdentifierValue[\*r]  
 relatedObjectSequence[-r]  
 relatedEventIdentifierType[-r]  
 relatedEventIdentifierValue[-r]  
 relatedEventSequence[-r]  
 linkEventIdentifierType-r  
 linkEventIdentifierValue-r  
 linkIntellectulaEntityIdentifierType-r  
 linkIntellectualIdentifierValue-r  
 linkRightsStatementIdentifierType-r  
 linkRightsStatementIdentifierValue-r

### RIGHTS

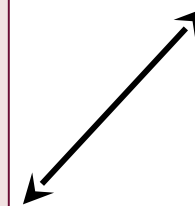
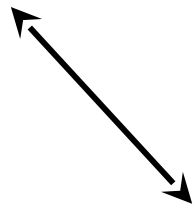
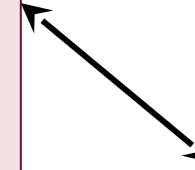
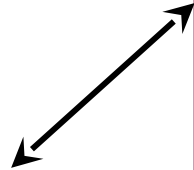
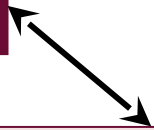
rIdentifierType\*nr  
 rIdentifierValue\*nr  
 rightsBasis\*nr  
 act\*nr  
 restriction-r  
 termOfGrantStartDate\*nr  
 TermOfGrantEndDate-nr  
 termOfRestrictionStartDate-nr  
 termOfRestrictionEndDate-nr  
 rGrantedNote-r  
 linkObjectIdentifierType-r  
 linkObjectIdentifierValue-r  
 linkObjectIdentifierRole-r  
 linkAgentIdentifierType-r  
 linkAgentIdentifierValue-r  
 linkAgentIdentifierRole-r[nr]

### EVENTS

eIdentifierType\*nr  
 eIdentifierValue\*nr  
 eType\*nr  
 eDateTime\*nr  
 eDetail-nr  
 eOutcome-nr  
 eOutcomeDetailNote-nr  
 eOutcomeDetailExtension-r  
 linkAgentIdentifierType  
 linkAgentIdentifierValue  
 linkAgentIdentifierRole  
 linkObjectIdentifierType  
 linkObjectIdentifierValue  
 linkObjectIdentifierRole

### AGENTS

aIdentifierType\*nr  
 aIdentifierValue\*nr  
 aName  
 aType  
 aNote  
 aExtension  
 linkEventIdentifierType-r  
 linkEventIdentifierValue-r  
 linkRightsStatementIdentifier  
 Type-r  
 linkRightsStatementIdentifier  
 Value-r





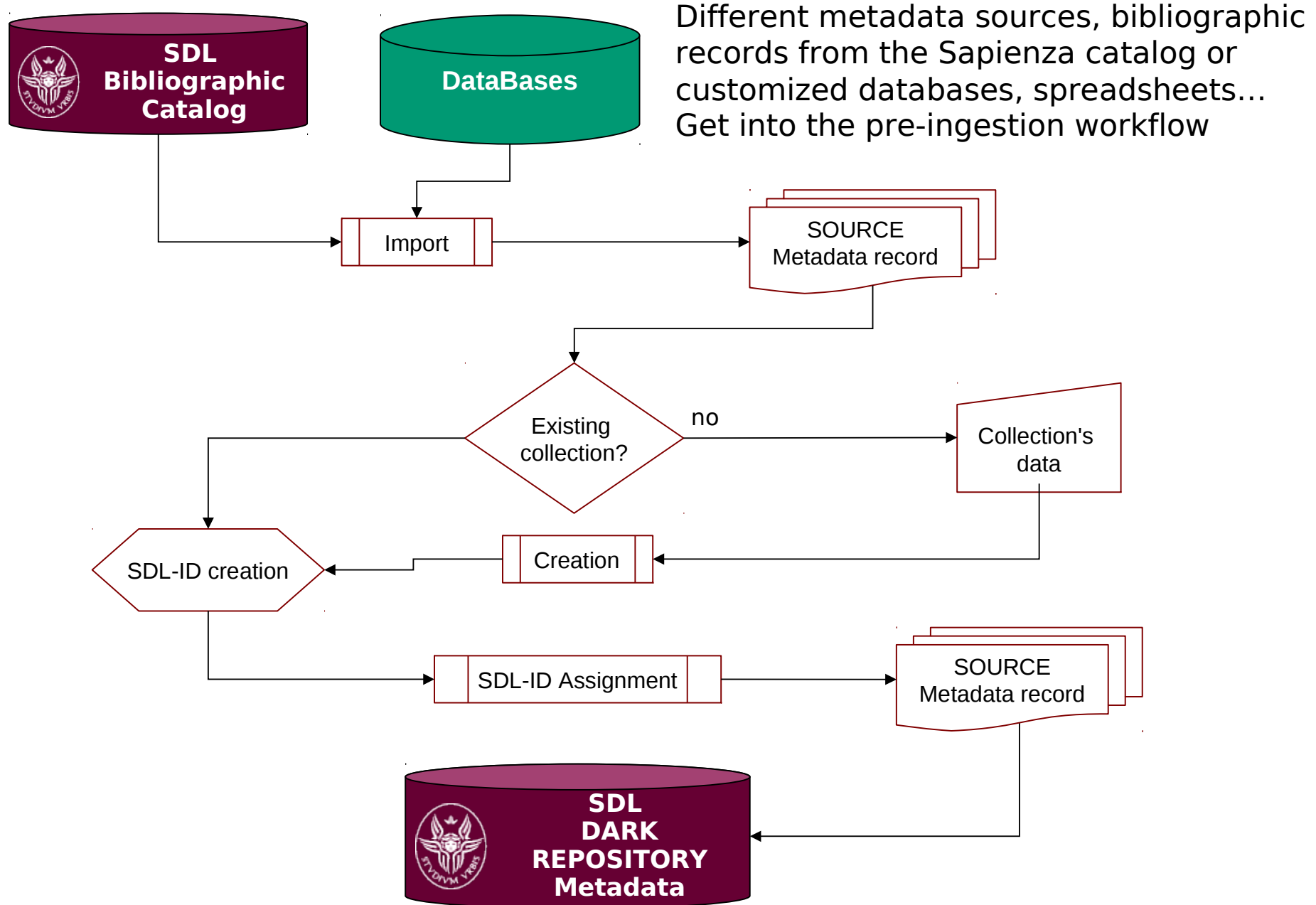
*... which means, by the system point of view*

## «conditio sine qua non»

- The submission of resources depends on the collection existence strictly connected to its Provenance

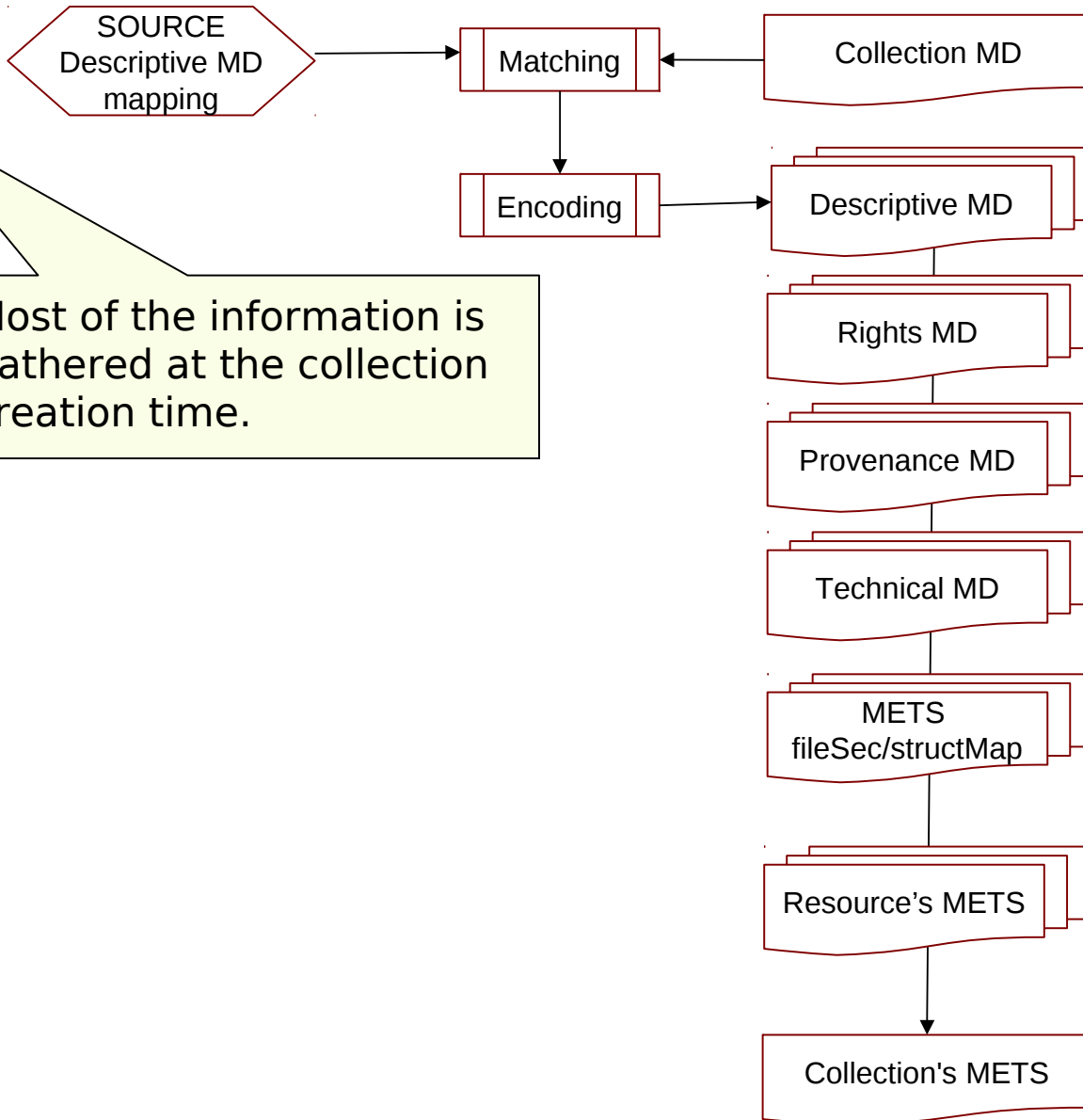
## «conditio sine qua non»

- A digital object must belong to a well identified resource described by metadata (IE, objects, agents, events, rights)



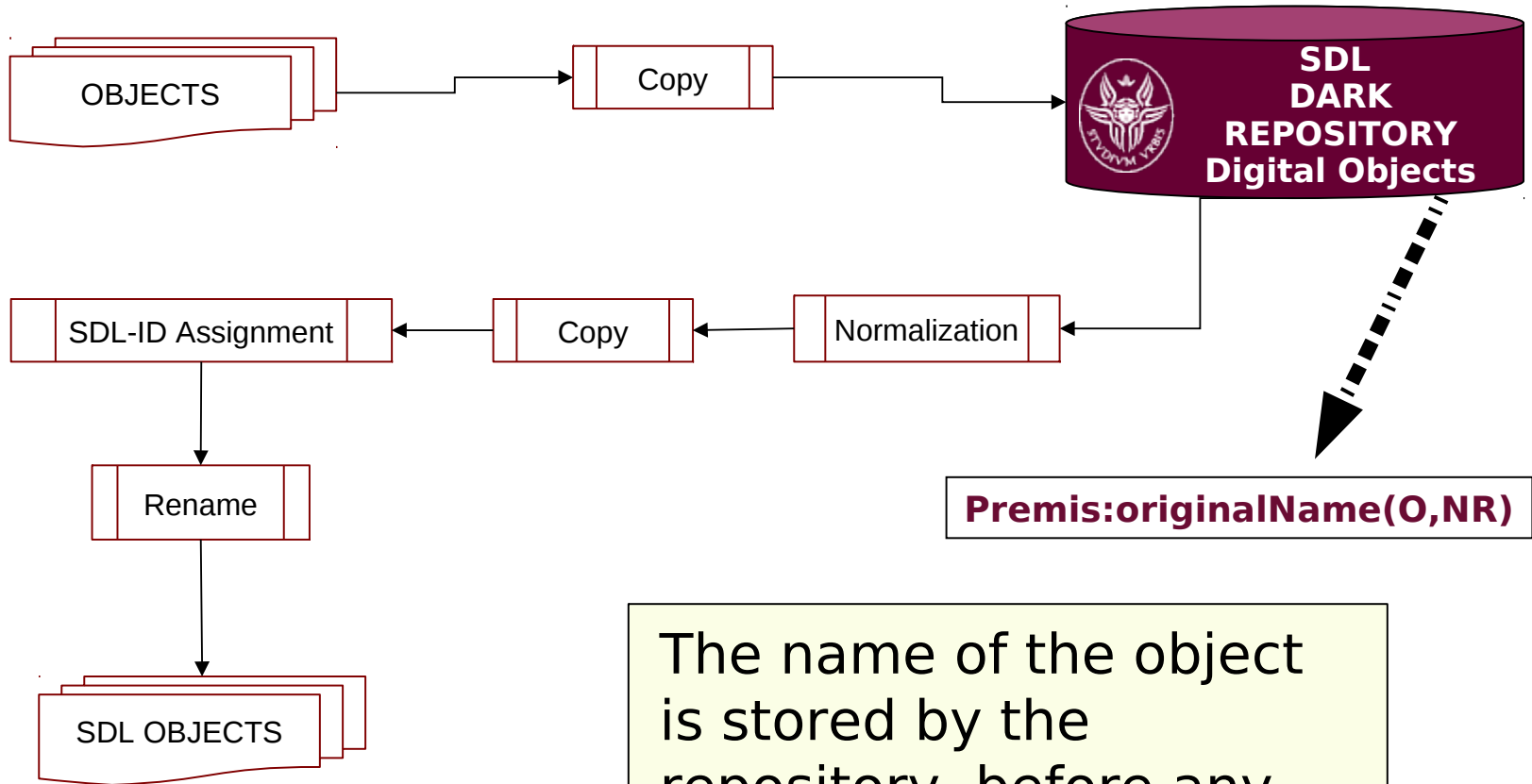


```
▼<premis:ObjectIdentifier>
  <premis:objectIdentifierType>Sapienza Digital Library accession system</premis:objectIdentifierType>
  <premis:objectIdentifierValue>RMS45/RMS45_00000001/JPEG/RMS45_00000001_0001.jpg</premis:objectIdentifierValue>
</premis:ObjectIdentifier>
<premis:fileSize>2260142</premis:fileSize>
▼<premis:FormatDesignation>
  <premis:formatName>jpg</premis:formatName>
  <premis:formatVersion>-</premis:formatVersion>
</premis:FormatDesignation>
▼<premis:FormatRegistry>
  <premis:formatRegistryName>PRONOM</premis:formatRegistryName>
  <premis:formatRegistryKey>fmt/42</premis:formatRegistryKey>
</premis:FormatRegistry>
▼<premis:Fixity>
  <premis:messageDigestAlgorithm>MD5</premis:messageDigestAlgorithm>
  <premis:messageDigest>6c78cb2a59533331d60f912a2601fa03d</premis:messageDigest>
  <premis:messageDigestOriginator>PHP 5.2.10 function</premis:messageDigestOriginator>
</premis:Fixity>
▼<premis:Fixity>
  <premis:messageDigestAlgorithm>SHA-1</premis:messageDigestAlgorithm>
  <premis:messageDigest>b41f4678628d470dcd3ab33a71f27b4c461e4aa8</premis:messageDigest>
  <premis:messageDigestOriginator>PHP 5.2.10 function</premis:messageDigestOriginator>
</premis:Fixity>
▼<premis:Fixity>
  <premis:messageDigestAlgorithm>CRC32</premis:messageDigestAlgorithm>
  <premis:messageDigest>-559115941</premis:messageDigest>
  <premis:messageDigestOriginator>PHP 5.2.10 function</premis:messageDigestOriginator>
</premis:Fixity>
</premis:BasicDigitalObjectInformation>
```



Most of the information is gathered at the collection creation time.

.....It results also in automatic building of the relations between the objects->resources->collections



The name of the object is stored by the repository, before any action of renaming





METSs

premis:objectCategory = **representation**  
premis:objectCategory = **file**

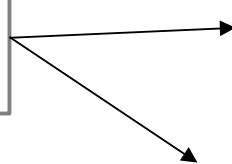
Objects

premis:objectCategory = **file**  
premis:compositionLevel = **0**

ZIPs

premis:compositionLevel = **1**

METSs



premis:eventType = **validation** (the process of comparing an object with a standard and noting compliance or exceptions)

premis:eventType = **creation** (the act of creating a new object)

premis:eventType = **replication** (the process of creating a copy of an object that is, bitwise, identical to the original)

premis:eventType = **message digest calculation** (the process by which a “hash” is created)

Objects



premis:linkingAgentRole = **authorizer, implementer, validator, executing program**  
premis:linkingObjectRole = **source, outcome**



premis:messageDigestAlgorithm = MD5  
premis:messageDigestAlgorithm = SHA-1  
premis:messageDigestAlgorithm = CRC32

premis:formatNote = format not validated

premis:agentType = organization  
premis:agentType = person  
premis:agentType = software

premis:rightsBasis = copyright  
premis:copyrightStatus = copyrighted (under copyright)  
premis:copyrightStatus = public domain (in the public domain)  
premis:copyrightStatus = unknown (copyright status of the resource is unknown)  
premis:copyrightDocumentationIdentifier = link alla legge diritto d'autore

premis:rightsBasis = license  
premis:licenseDocumentationIdentifier = agreements' IDs: sapienza-cineca, digilab-siae, sapienza-rightsHolder

premis:licenseDocumentationRole = preservation repository-licensee-licensor,  
premis:rightsBasis = statute                      licensor-national authority, licensor-rightsHolder  
premis:statuteDocumentationIdentifier = link institutional policies (delibera senato accademico)

premis:linkingAgentRole = contact, creator, publisher, rightsholder, grantor  
premis:linkingObjectRole = owned, created, held, preserved???



***But hopefully...  
never Ending Preservation Care...***

***MANY THANKS, for you kind attention***

Question time ?