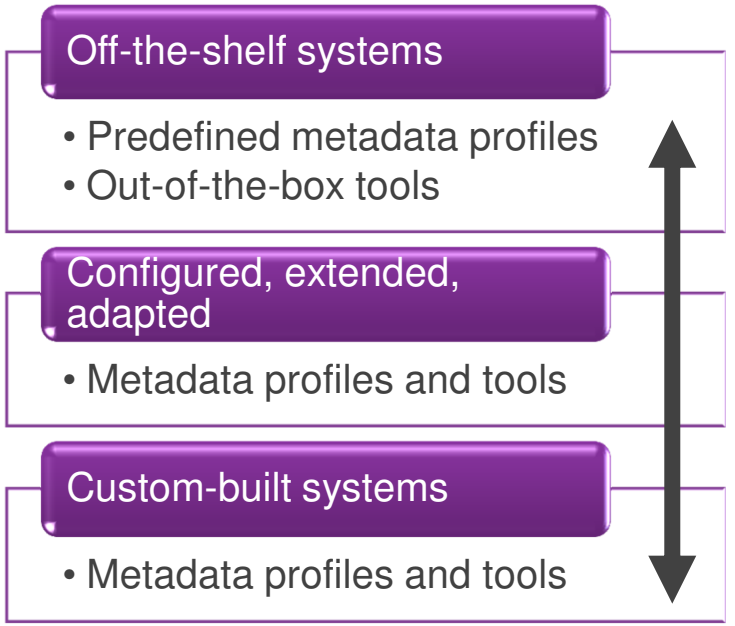
Emerging Nuances of Digital Preservation Metadata – PREMIS Version 3.0

Angela Dappert
The British Library
3 December 2015

Tayloring PREMIS to needs

- Evolving metadata
  - Increasing experience ensuring the longevity of digital objects
  - Changing future technical possibilities
  - Changing future legal framework
  - Always user-driven

- Tayloring solutions
  - Varying needs
    - Content-types
    - Institutional policies
    - Intended use
  - Off-the-shelf (OS / commercial) or custom-built



PREMIS: From V2 to V3 based on user needs

- Add preservationLevelType semantic unit
- Add agentVersion semantic unit
- Add “unknown” values
- Add eventDetailInformation semantic unit

minor

- Add authority for controlled vocabulary
- Make Intellectual Entity an Object category
- Make Environments independent Objects
- Add physical Objects
- Update conformance statement

Add preservationLevelType semantic unit

- 1.3 preservationLevel
  -
  - 1.3.1 preservationLevelValue
  - 1.3.2 preservationLevelRole
  - 1.3.3 preservationLevelRationale
  - 1.3.4 preservationLevelDateAssigned

Add preservationLevelType semantic unit

- 1.3 preservationLevel
  - 1.3.1 preservationLevelType
  - 1.3.2 preservationLevelValue
  - 1.3.3 preservationLevelRole
  - 1.3.4 preservationLevelRationale
  - 1.3.5 preservationLevelDateAssigned

- Associate type of preservation function with preservation level.

- objectIdentifier
  - objectIdentifierType: ARK
  - objectIdentifierValue: ark:/9999/c1
- objectCategory: file
- preservationLevel
  - preservationLevelType: Bit preservation
  - preservationLevelValue: medium
- preservationLevel
  - preservationLevelType: Functional preservation
  - preservationLevelValue: migration
- objectCharacteristics
  - compositionLevel: 0
  - size: 726970368
  - format
    - formatDesignation
      - format name: application/vnd.ms-excel

Add agentVersion semantic unit

- If agentType is software,
  - agentVersion can be used to refine agentName.

- 3.1 agentIdentifier
- 3.2 agentName
- 3.3 agentType

- 3.4 agentNote
- 3.5 agentExtension
- 3.6 linkingEventIdentifier
- 3.7 linkingRightsStatementIdentifier
-

Add agentVersion semantic unit

- If agentType is software,
  - agentVersion can be used to refine agentName.

- 3.1 agentIdentifier
- 3.2 agentName
- 3.3 agentType
- 3.4 agentVersion
- 3.5 agentNote
- 3.6 agentExtension
- 3.7 linkingEventIdentifier
- 3.8 linkingRightsStatementIdentifier
- 3.9 linkingEnvironmentIdentifier

Unknown compositionLevel and format

compositionLevel and format:

- A value of *unknown* added
  if the information is not available.

Add eventDetailInformation semantic unit .

- 2.1 eventIdentifier
- 2.2 eventType
- 2.3 eventDateTime
- 2.4 eventDetail

- 2.5 eventOutcomeInformation
- 2.6 linkingAgentIdentifier
- 2.7 linkingObjectIdentifier

Add eventDetailInformation semantic unit .

- 2.1 eventIdentifier
- 2.2 eventType
- 2.3 eventDateTime
- 2.4 eventDetailInformation
  - 2.4.1 eventDetail
  - 2.4.2 eventDetailExtension
- 2.5 eventOutcomeInformation
- 2.6 linkingAgentIdentifier
- 2.7 linkingObjectIdentifier

PREMIS: From V2 to V3 based on user needs

- Add preservationLevelType semantic unit
- Add agentVersion semantic unit
- Add “unknown” values
- Add eventDetailInformation semantic unit

minor

- Add authority for controlled vocabulary

bonus

- Make Intellectual Entity an Object category
- Make Environments independent Objects
- Add physical Objects
- Update conformance statement

Implementation specific change: Add authority for controlled vocabulary

- eventIdentifier:
  - eventIdentifierType: UUID
  - eventIdentifierValue: 908985d3-9600-4da4-a
- eventType: validation

authority="premisEventType"
authorityURI=
"http://id.loc.gov/vocabulary/preservation/eventType.html"
valueURI= "http://id.loc.gov/vocabulary/preservation/eventType/val.html

- eventDateTime: 2014-07-03T23:18:19
- eventDetailInformation:
  - eventDetail: program="Jhove"; version="1.5"
- eventOutcomeInformation:
  - eventOutcome: fail
  - eventOutcomeDetail:
    - eventOutcomeDetailNote:
      format="JPEG"; version="1.02"; result="Not well formed"

- capture
- compression
- creation
- fixity check
- ingestion
- message digest calculation
- migration
- normalization
- replication
- validation
- virus check

PREMIS: From V2 to V3 based on user needs

- Add preservationLevelType semantic unit
- Add agentVersion semantic unit
- Add “unknown” values
- Add eventDetailInformation semantic unit

minor

- Add authority for controlled vocabulary

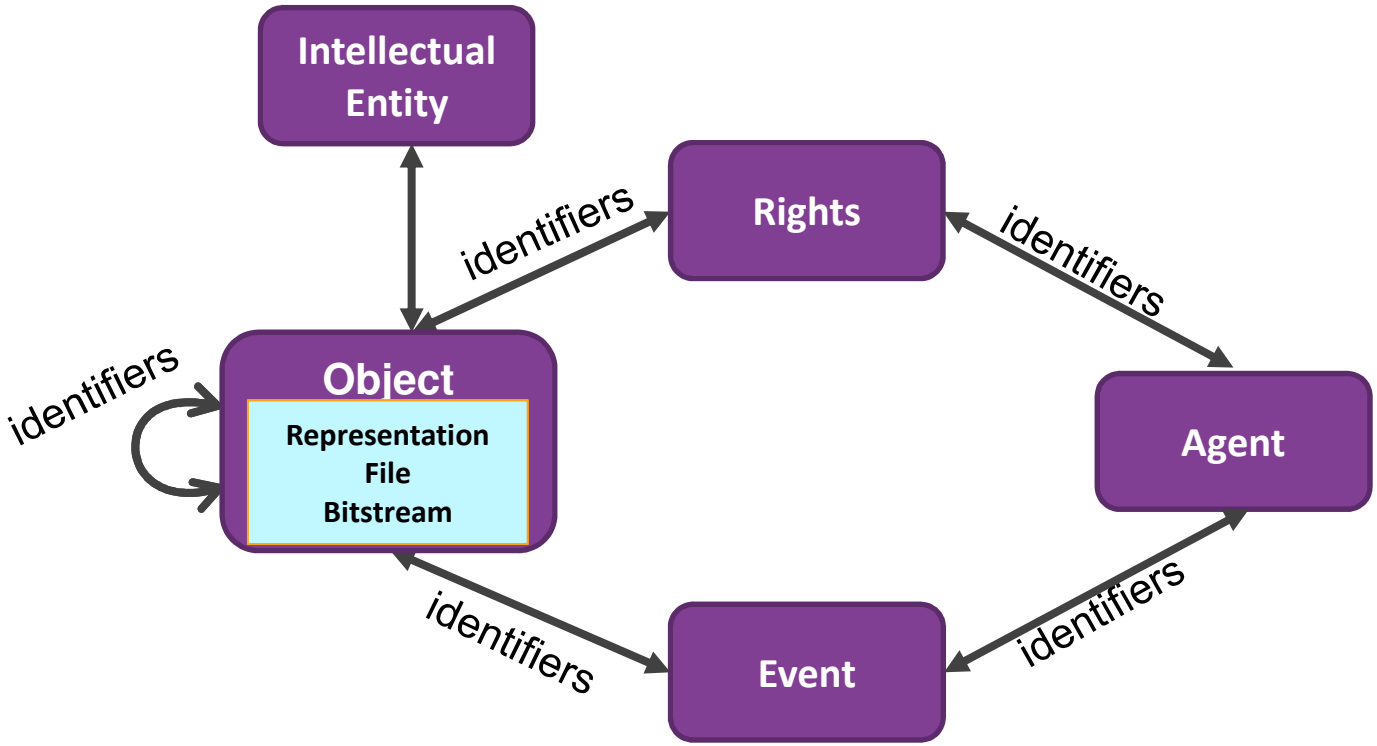
bonus

- Make Intellectual Entity an Object category
- Make Environments independent Objects
- Add physical Objects

major

- Update conformance statement

Make Intellectual Entity an Object category

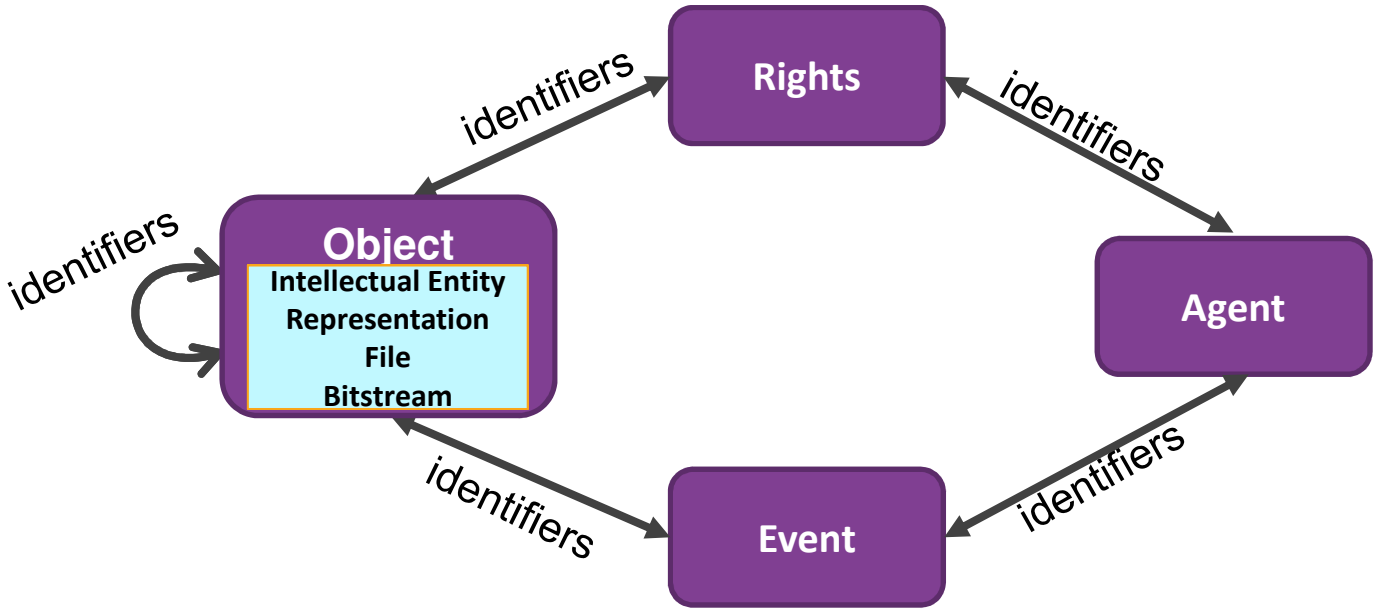


V2:

- Assumed to be held in a container metadata schema
- No Intellectual Entity semantic units
- Exception: identifier to enable linking to a description
- PREMIS Objects link to it.

- A set of content that is considered a single intellectual unit for purposes of management and description
- For example, a particular book, map, photograph, or database.

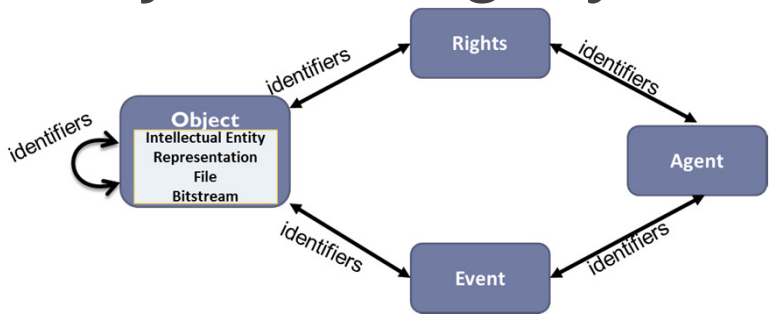
Make Intellectual Entity an Object category



V3:

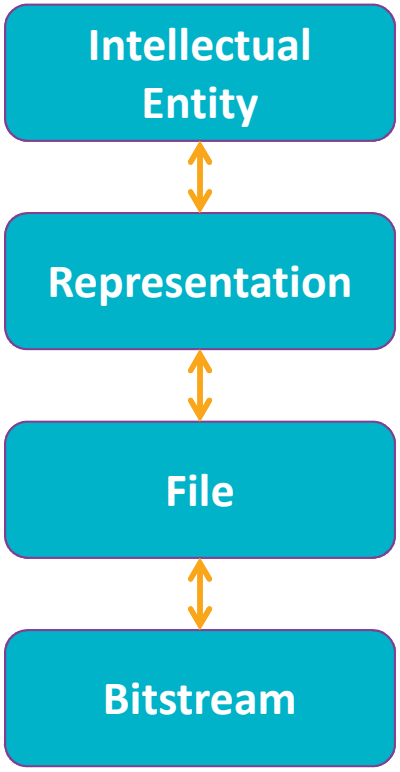
- Possibility to describe preservation aspects of intellectual entities
- Same semantic units as Representations

Make Intellectual Entity an Object category



- Relate to PREMIS Events and RightsStatements.
- Support structural and derivative relationships with Objects.
- Represent an aggregate, such as a collection, FRBR work, FRBR expression, fonds or series.
- Capture versioning information and metadata update events at the Intellectual Entity level
- Associate business requirements with them.
  - Significant characteristics, risk definitions, guidelines for preservation actions, etc..

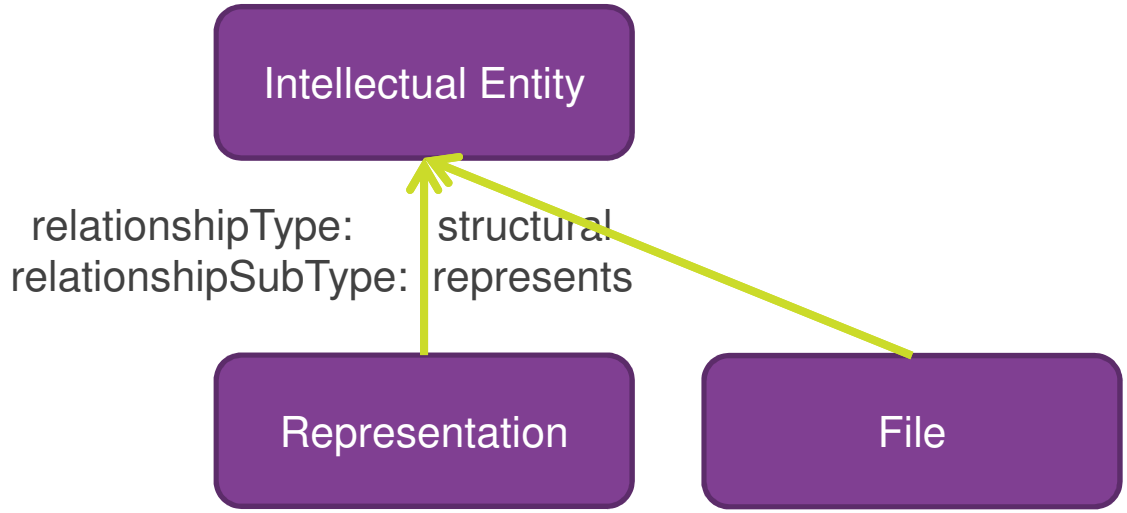
Make Intellectual Entity an Object category



Like before, an intellectual

An intellectu
entity Obje
relates to
Objects thro
relationsh

An Intellectual Entity may have one or more digital representations.



Make Environments independent Objects

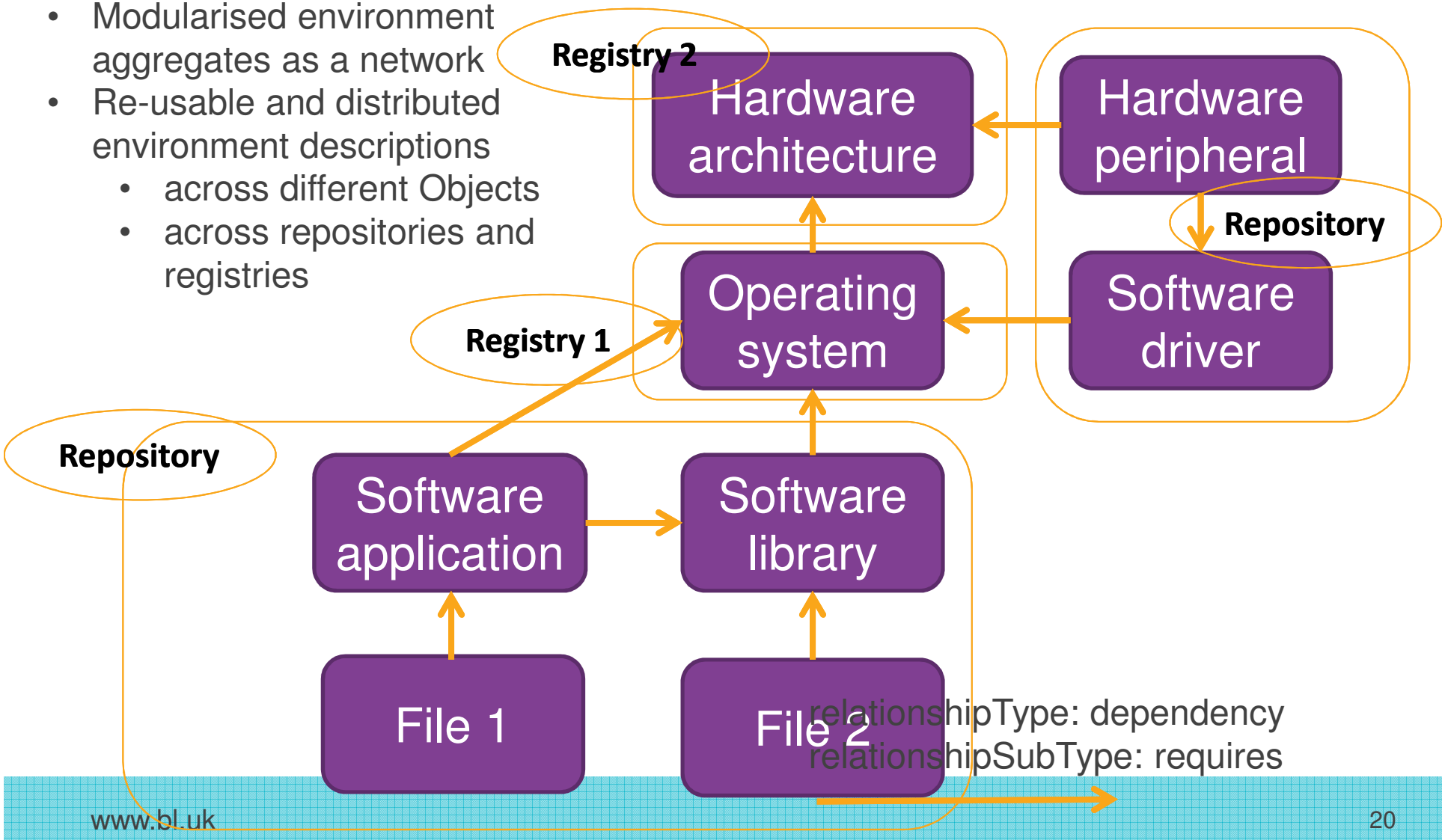
- What is needed to render or use an object
  - Operating system
  - Application software
  - Hardware
  - Computing resources

- A high-level data model

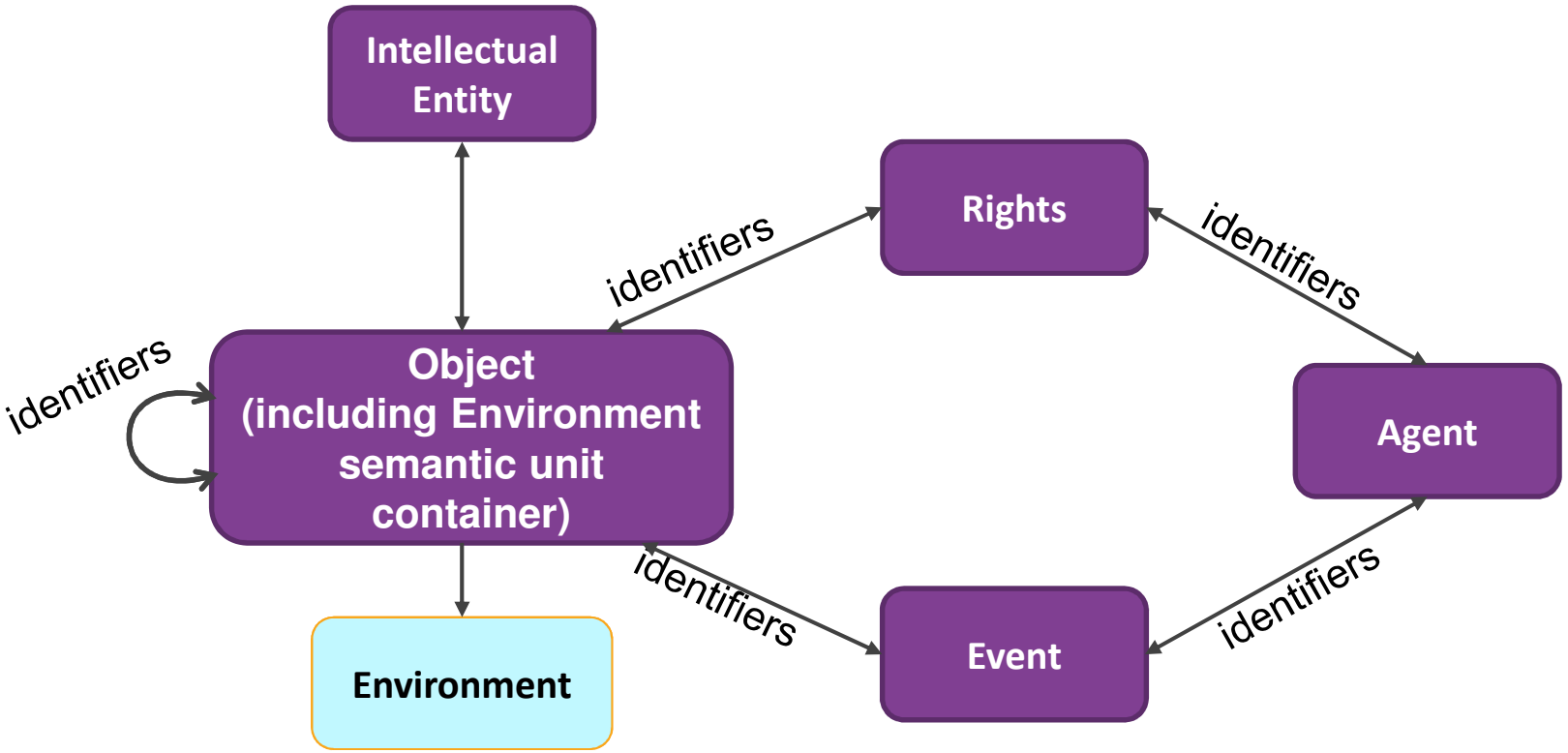
- **No** detailed characteristics specific to an environment type

Example: Environment stack and dependency relationships

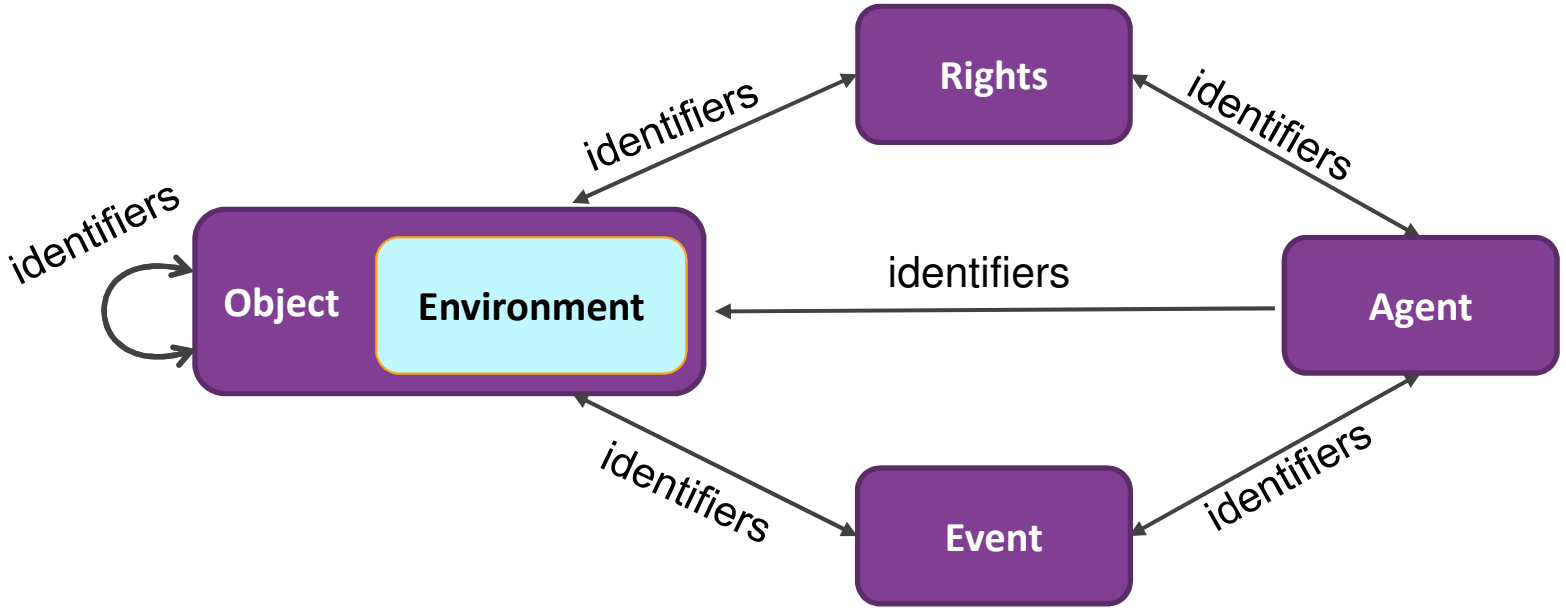
- Modularised environment aggregates as a network
- Re-usable and distributed environment descriptions
  - across different Objects
  - across repositories and registries



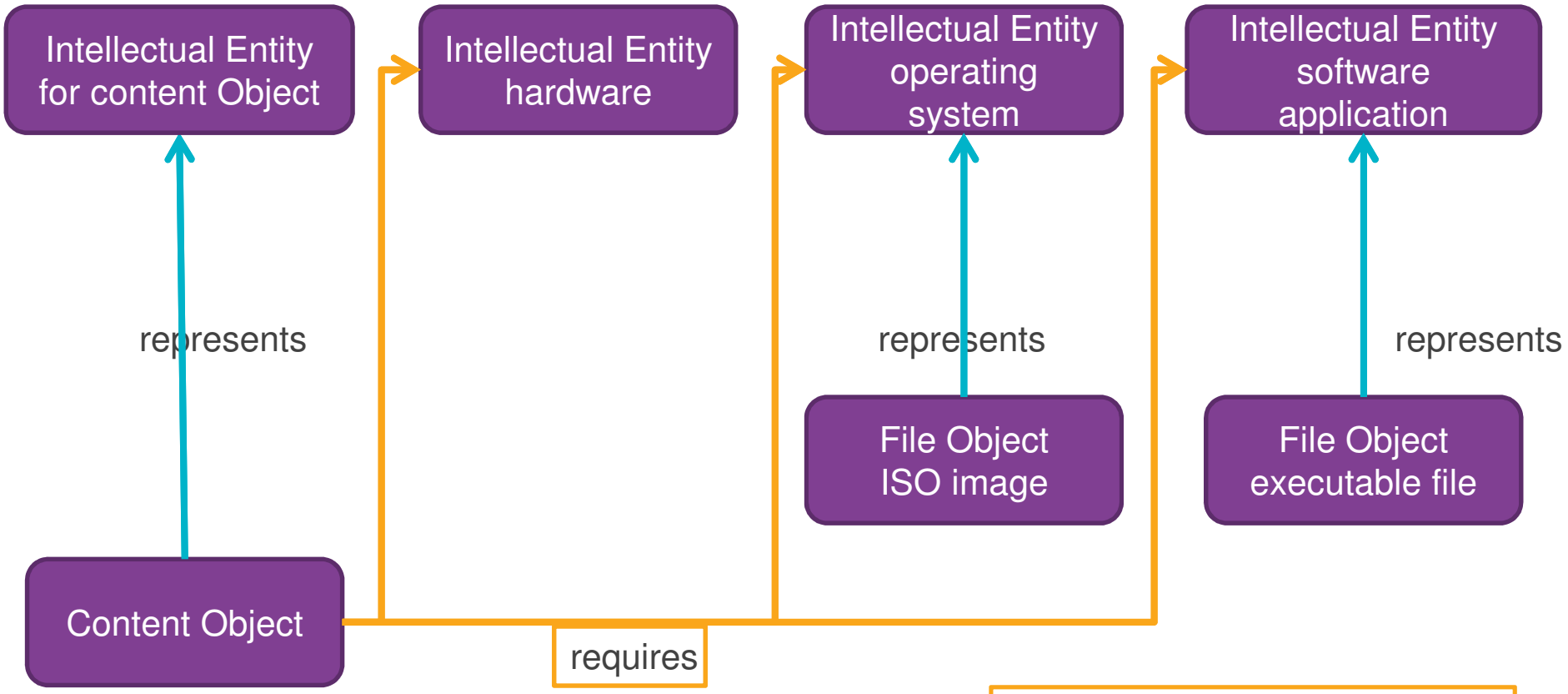
Data Model in PREMIS V2



Data Model in PREMIS V3

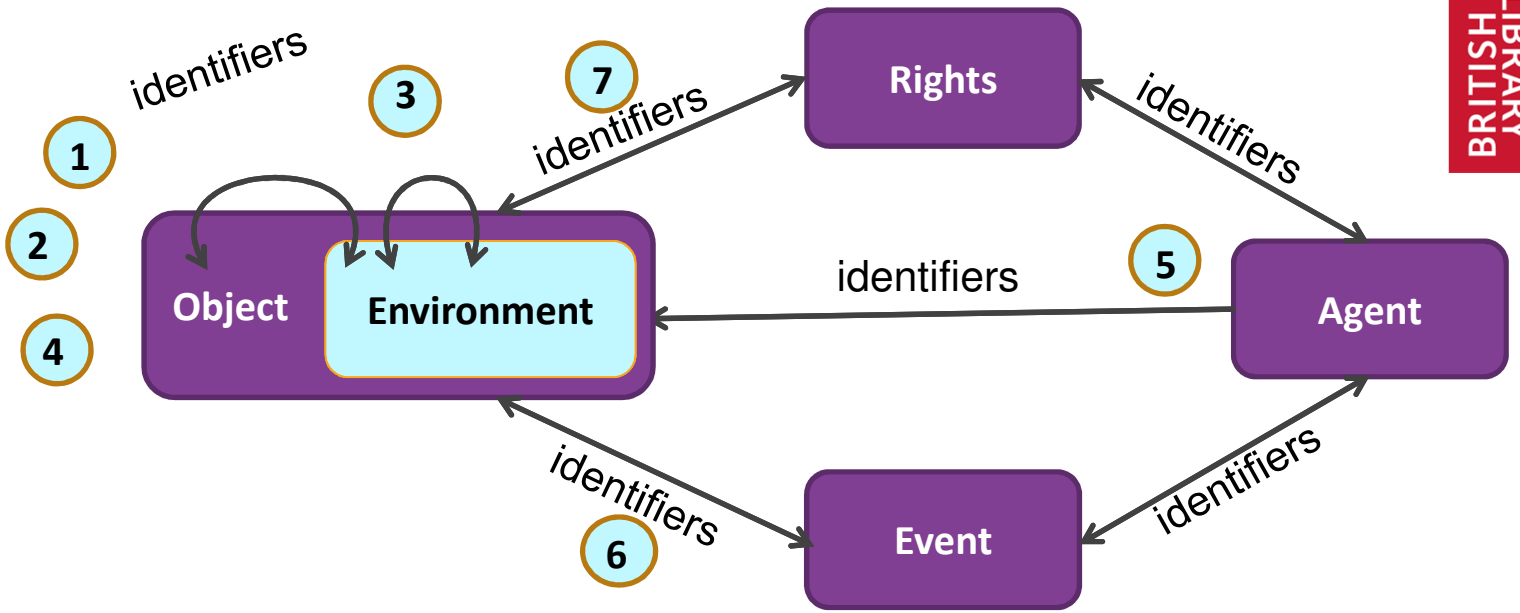


Example:
An object and its rendering environment



represents =
relationshipType: structural
relationshipSubType: represents

requires =
relationshipType: dependency
relationshipSubType: requires



1. Object to environment - specify computational context
2. environment to Object - documentation, specifications, surrogates
3. environment to environment - inclusion, dependency, derivation,other
4. environment is an Object – preserved software source code
5. Agent to Environment - role of an Agent
6. environment to Event - environment specific Events (provenance)
7. environment to RightsStatement - software license, policy

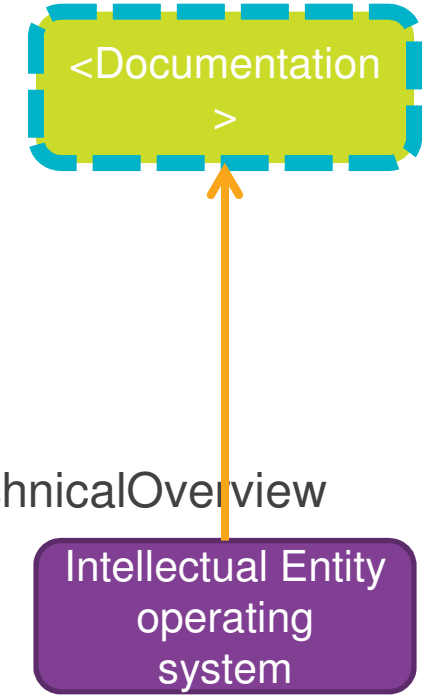
"Object": here a traditional content Object

Expanded relationship types for environment Objects

- Dependency
  - Requires, is required by
  - Is deployed on
- Derivation
  - Is source of, has source
- Logical
  - generalises,
    is generalised by
- Reference
  - Documents,
    is documented in
- Replacements
  - Supercedes,
    is superceded by
- Structural
  - Includes, is included in
  - Represents,
    is represented as

Expanded relationship types for environment Objects

relationshipType: reference
relationshipSubType: is documented in
relatedObjectIdentifier
    relatedObjectIdentifierType: URL
    relatedObjectIdentifierValue:
https://wiki.ubuntu.com/QuantalQuetzal/TechnicalOverview



Semantic units only applicable to environment Intellectual Entities

- 1.9 environmentFunction
  - environmentFunctionType
  - environmentFunctionLevel

objectIdentifier
    objectIdentifierType: ARK
    objectIdentifierValue: ark:/9999/b1
objectCategory: intellectual entity
environmentFunction
    environmentFunctionType: software
    environmentFunctionLevel: 1
environmentFunction
    environmentFunctionType: operating system
    environmentFunctionLevel: 2

XP Professional, Service Pack 3

Semantic units only applicable to environment Intellectual Entities

- 1.9 environmentFunction
  - environmentFunctionType
  - environmentFunctionLevel
- 1.10 environmentDesignation
  - environmentName
  - environmentVersion
  - environmentOrigin
  - environmentDesignationNote
  - environmentDesignationExtensio

objectCategory: intellectual entity
environmentFunction
environmentFunctionType: software
environmentFunctionLevel: 1
environmentFunction
environmentFunctionType: operating system
environmentFunctionLevel: 2
environmentDesignation
environmentName: Windows XP Professional
environmentVersion: Service Pack 3
environmentDesignationNote:
maintenance deadline: 2014-04

Semantic units only applicable to environment Intellectual Entities

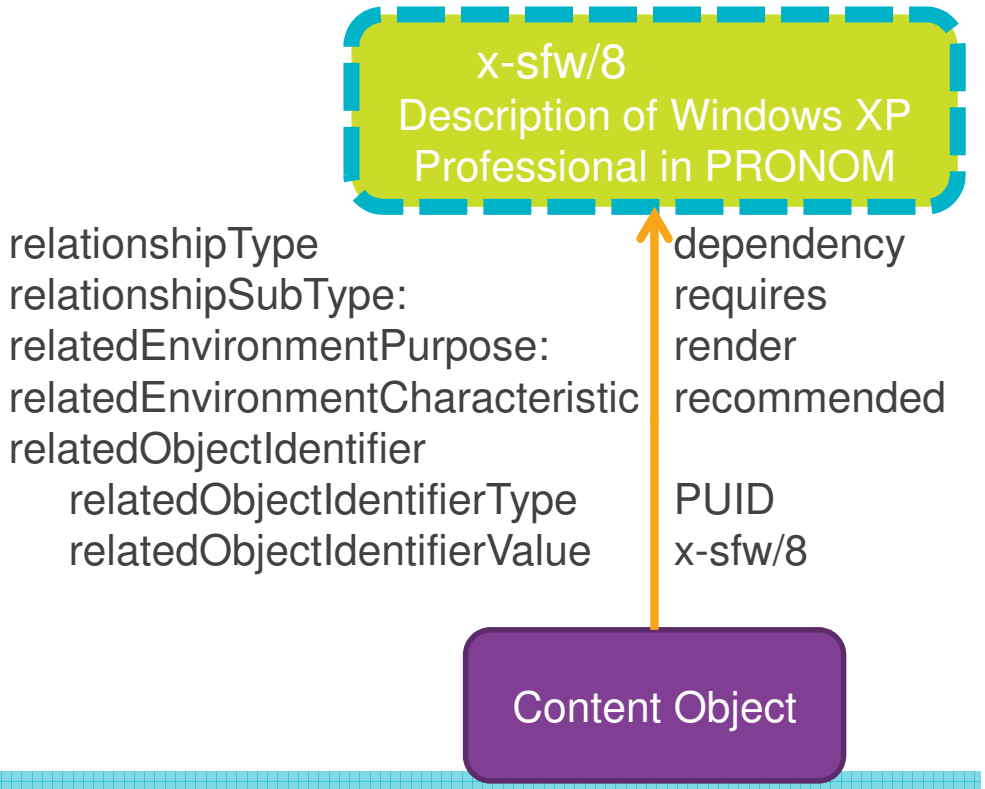
- 1.9 environmentFunction
  - environmentFunctionType
  - environmentFunctionLevel
- 1.10 environmentDesignation
  - environmentName
  - environmentVersion
  - environmentOrigin
  - environmentDesignationNot
  - environmentDesignationExte
- 1.11 environmentRegistry
  - environmentRegistryName
  - environmentRegistryKey
  - environmentRegistryRole

objectCategory: intellectual entity
environmentFunction
environmentFunctionType: software
environmentFunctionLevel: 1
environmentFunction
environmentFunctionType: operating system
environmentFunctionLevel: 2
environmentDesignation
environmentName: Windows XP Professional
environmentVersion: Service Pack 3
environmentRegistry
environmentRegistryName: PRONOM
environmentRegistryKey: x-sfw/8
environmenttRegistryRole: identity

Semantic units only applicable to environment Intellectual Entities

- 1.9 environmentFunction
  - environmentFunctionType
  - environmentFunctionLevel
- 1.10 environmentDesignation
  - environmentName
  - environmentVersion
  - environmentOrigin
  - environmentDesignationNote
  - environmentDesignationExtension
- 1.11 environmentRegistry
  - environmentRegistryName
  - environmentRegistryKey
  - environmentRegistryRole

Alternative:
Link to an external registry

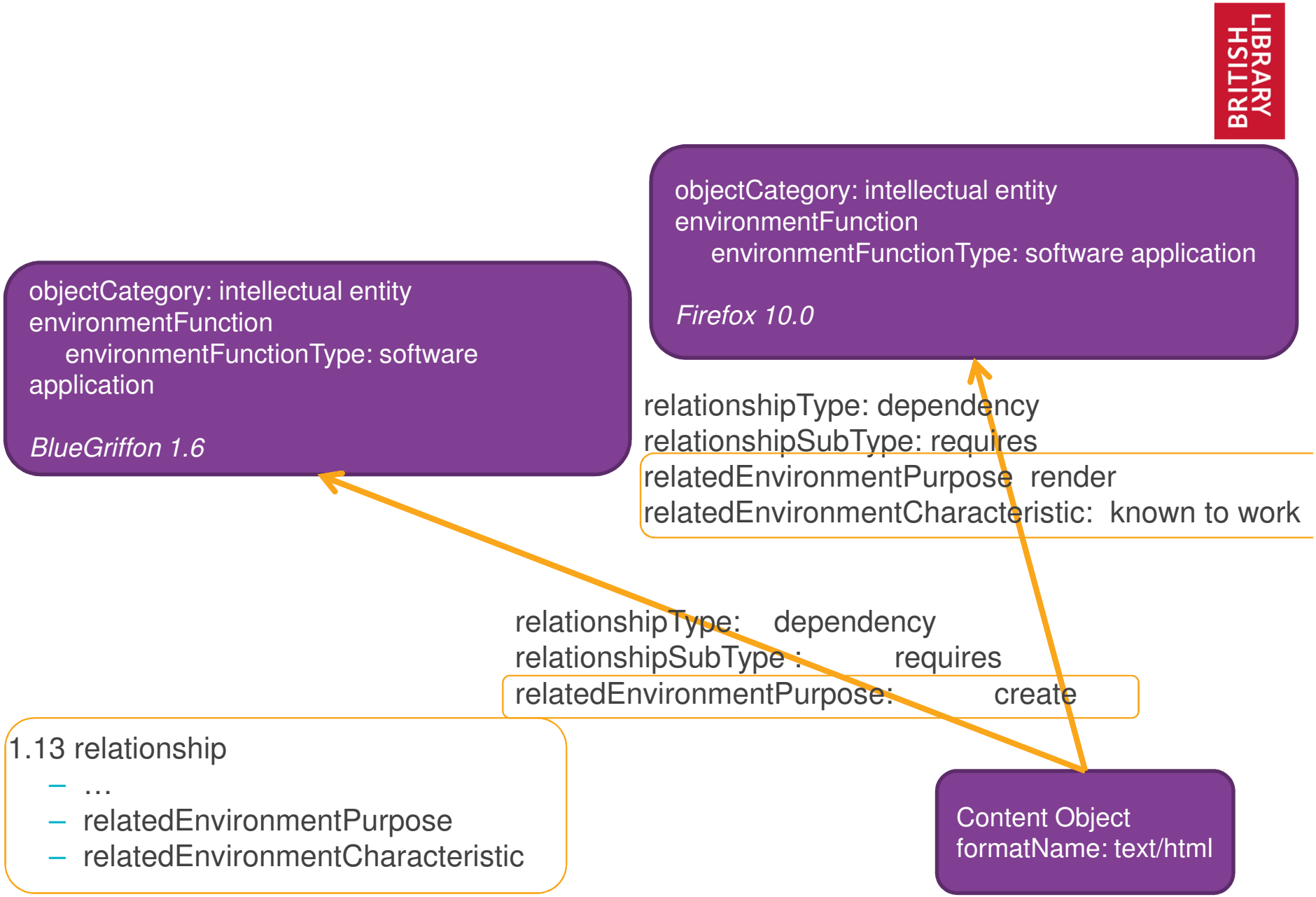


Semantic units only applicable to environment Intellectual Entities

- 1.9 environmentFunction
  - environmentFunctionType
  - environmentFunctionLevel
- 1.10 environmentDesignation
  - environmentName
  - environmentVersion
  - environmentOrigin
  - environmentDesignationNote
  - environmentDesignationExtension
- 1.11 environmentRegistry
  - environmentRegistryName
  - environmentRegistryKey
  - environmentRegistryRole
- 1.12 environmentExtension
- 1.13 relationship

…

  - relatedEnvironmentPurpose
  - relatedEnvironmentCharacteristic

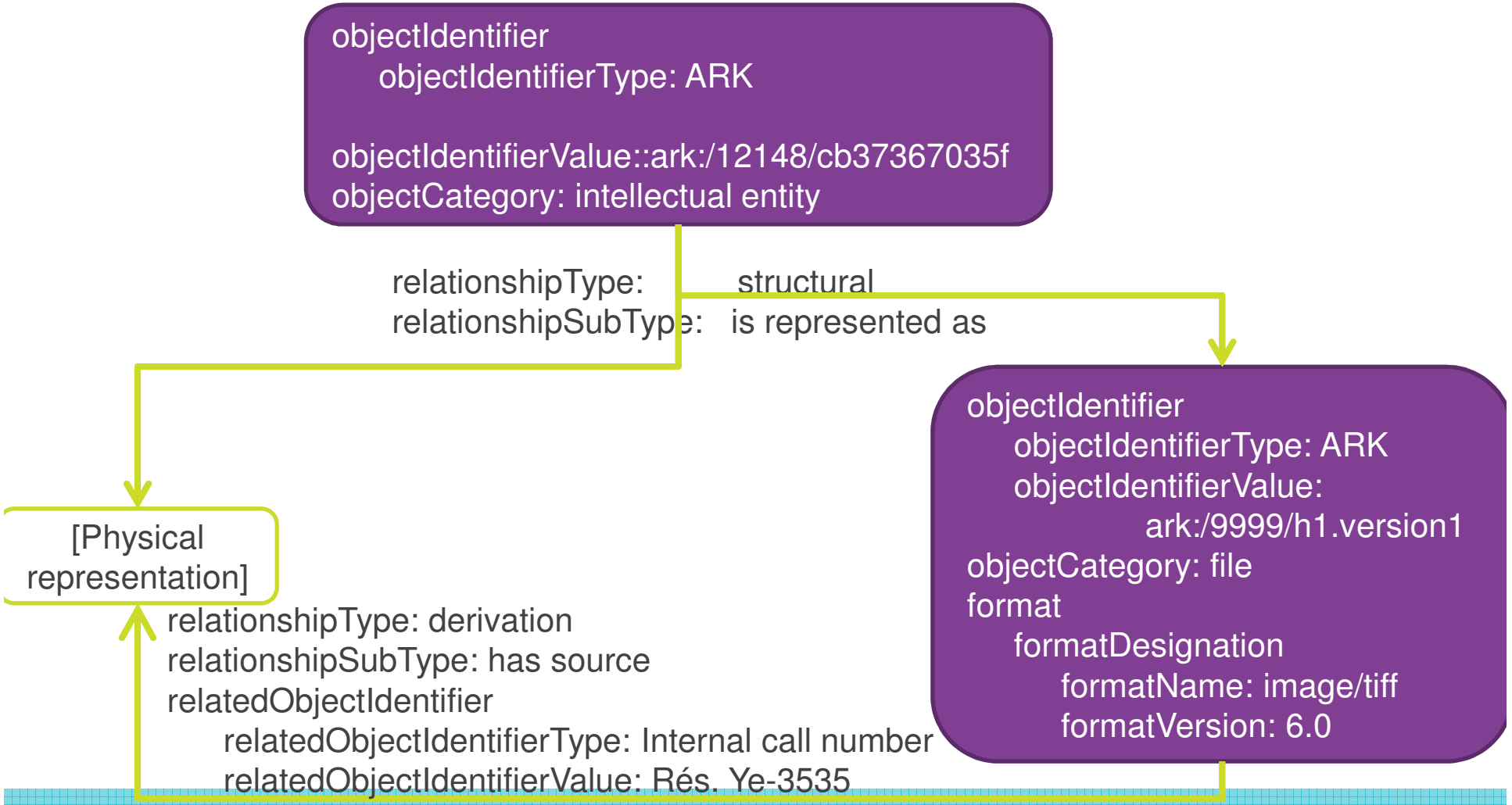


- 1.13 relationship
  - …
  - relatedEnvironmentPurpose
  - relatedEnvironmentCharacteristic

Add physical Objects

- A physical Object is
  - A content Object, such as a manuscript, or printed document
  - An environment Object, such as a physical hardware device.
- Representation: A digital or physical Object
- Either one instantiates or embodies an Intellectual Entity
- Digital and non-digital Objects can be captured uniformly.
- Physical Objects can relate to digital Objects and other physical Objects.
- In V3 *storage* is applicable to Representations.
  For physical Representations: the physical location, e.g. a shelf location.

Add physical Objects



PREMIS: From V2 to V3 based on user needs

- Add preservationLevelType semantic unit
- Add agentVersion semantic unit
- Add “unknown” values
- Add eventDetailInformation semantic unit

minor

- Add authority for controlled vocabulary

bonus

- Make Intellectual Entity an Object category
- Make Environments independent Objects
- Add physical Objects

major

- Update conformance statement

clarification

http://www.loc.gov/standards/premis/premis-conformance-20150429.pdf

Thank you

Resources:
http://www.loc.gov/standards/premis/
PREMIS Implementors Group Forum:
PIG@listserv.loc.gov