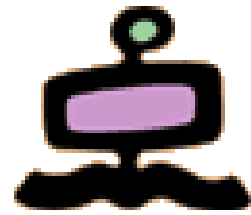# Corso di Biblioteche Digitali

- Vittore Casarosa
  - casarosa@isti.cnr.it
  - Office: 050 621 3115
  - Mobile: 348 397 2168
  - Skype: vittore1201
- "Ricevimento" at the end of the lessons or by appointment
- Final assessment
  - 70% oral examination
  - 30% project (development of a small digital library))
- Reference material:
  - Ian Witten, David Bainbridge, David Nichols, How to build a Digital Library, Morgan Kaufmann, 2010, ISBN 978-0-12-374857-7 (Second edition)
  - Material provided by the teacher
- **http://cloudone.isti.cnr.it/casarosa/BDG/**

Modules

- Computer Fundamentals and Networking
- A conceptual model for Digital Libraries
- Bibliographic records and metadata
- Information Retrieval and Search Engines
- Knowledge representation
- Digital Libraries and the Web
- Hands-on laboratory: the Greenstone system

Parallel evolution

Libraries	The Web
• Description (documents)	• Description (instances)
– Bibliographic records	– Metadata
– MARC	– Dublin Core
• Interoperability	• Interoperability
– Z39-50	– OAI-PMH
• Conceptual model (classes)	• Conceptual model (classes)
– **FRBR – LRM for Works, Expr., Manif.**	– RDF and RDF Schema for all resources (ontologies)
• Information Retrieval	• Information Retrieval
– Full text (catalogue and documents)	– Full text (web pages and resources)

Knowledge representation

- **FRBR: Functional Requirements for Bibliographic Records**
  - LRM: Library Reference Model

- **RDF: Resource Description Framework**
- **RDF Schema**
  - LOD: Linked Open Data
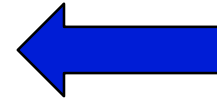
Representation of knowledge

- Description of Information (resources) is an exercise in "knowledge representation"

- Description with metadata (such as Dublin Core) only captures the "syntactic" aspects of the resource

- It would be good to have also a "semantic" description (capturing our "knowledge" of the resource)

- Knowledge representation is the "Holy Graal" of Computer Science (Artificial Intelligence, Expert Systems, Ontologies, etc.)

- Many models/languages proposed in the last 50 years
  - Basically all of them are based on the **entity-relationship model**

- Most of the advances available today are due to "brute force" methods (computing power, **today "neural networks"**)

What is a Conceptual Model ?

- A conceptual model is an abstract framework for identifying and describing significant properties and relationships among the entities of interest in some (small) "part of the universe"

- A conceptual model should be based on a small number of unifying concepts
  - entities, the classes which are the focus of interest
  - attributes, the data which characterizes instances of entities
  - relationships, the properties which link instances of entities

- Two conceptual models are of interest to Digital Libraries
  - FRBR - LRM
  - RDF Schema - LOD

FRBR

Functional Requirements for Bibliographic Records

- Approved by IFLA in 1997, published in 1998
- An abstract conceptual model of the 'bibliographic universe'
- It is based on the entity-relationship model (Entities, Attributes, Relationships)
- FRBR was defined having in mind the main "User Tasks"
  - Find
  - Identify
  - Select
  - Obtain

User tasks

- **find** entities that correspond to the user's stated search criteria (i.e., to locate either a single entity or a set of entities in a file or database as the result of a search using an attribute or relationship of the entity)

- **identify** an entity (i.e., to confirm that the entity described corresponds to the entity sought, or to distinguish between two or more entities with similar characteristics)

- **select** an entity that is appropriate to the user's needs (i.e., to choose an entity that meets the user's requirements with respect to content, physical format, etc., or to reject an entity as being inappropriate to the user's needs)

- acquire or **obtain** access to the entity described (i.e., to acquire an entity through purchase, loan, etc., or to access an entity electronically through an online connection to a remote computer
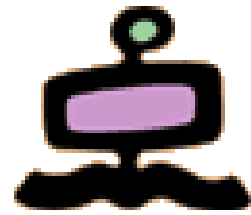
FRBR Entities

Group 1	Work, Expression, Manifestation, Item
Group 2	Person, Corporate body
Group 3	Concept, Object, Event, Place

Group 1 entities

- WORK
  a distinct intellectual or artistic creation

- EXPRESSION
  the intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of such forms

- MANIFESTATION
  the physical embodiment of an expression of a work.
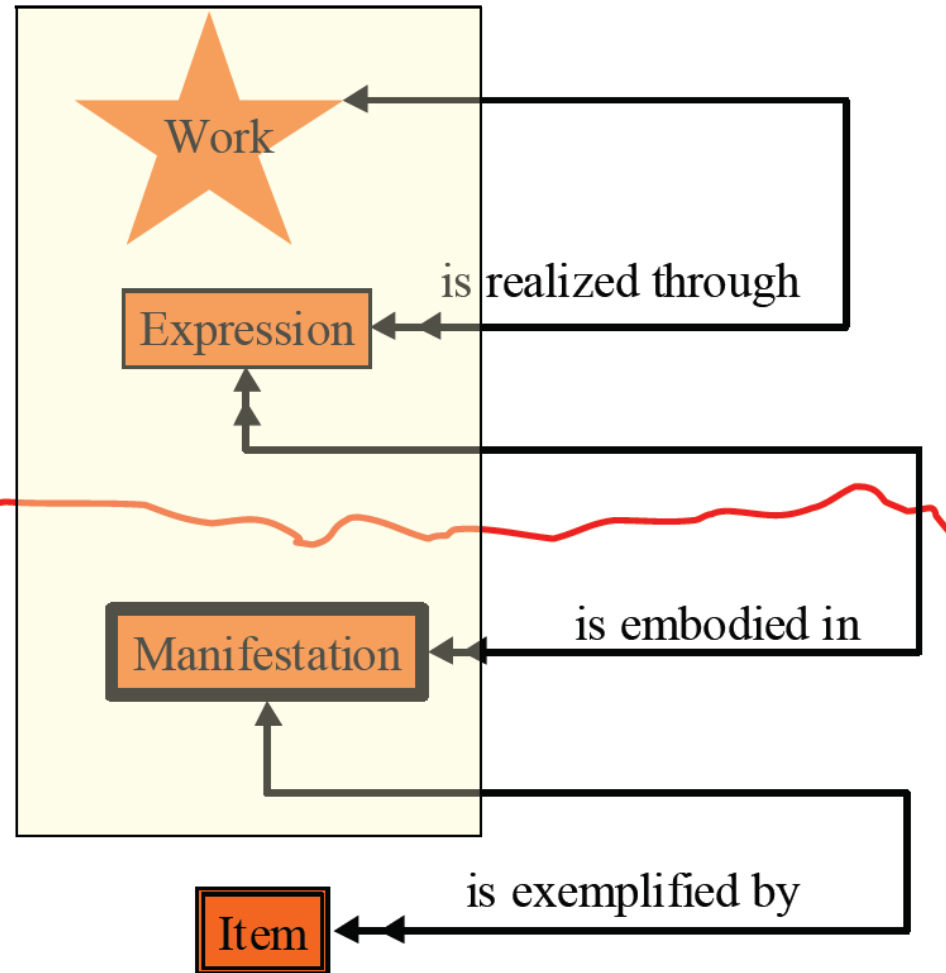
- ITEM
  a single exemplar of a manifestation.

FRBR – Group 1 entities

products of intellectual or artistic endeavour

Intellectual/
artistic content

Physical -
recording of
content

Work

Expression

is realized through

Manifestation
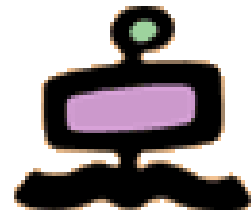
is embodied in

Item

is exemplified by

Examples of Work and Expressions

- w1 Henry Gray's Anatomy of the human body
  - e1 text and illustrations for the first edition
  - e2 text and illustrations for the second edition
  - e3 text and illustrations for the third edition
  - ….
- w1 J. S. Bach's The art of the fugue
  - e1 the composer's score for organ
  - e2 an arrangement for chamber orchestra by Anthony Lewis
  - ….
- w1 Jules et Jim (motion picture)
  - e1 the original French language version
  - e2 the original with English subtitles added
  - ….

Examples of <different> Works

- William Shakespeare's Romeo and Juliet

- Franco Zeffirelli's motion picture Romeo and Juliet

- Baz Lurhmann's motion picture William Shakespeare's Romeo and Juliet

- The theatre play West side story

- The movie West side story

- ….

Examples of Expressions

- **Franz Schubert's Trout quintet (Work)**
  - e1 the composer's notated music
  - e2 the musical work as performed by Rosina Lhevinne, piano, Stuart Sankey, double bass, and members of the Juilliard String Quartet
  - e3 the musical work as performed by Jörg Demus, piano, and the members of the Collegium Aureum
  - e4 the musical work as performed by Emanuel Ax, piano, members of the Guarneri String Quartet, and Julius Levine, double bass
 - ….

Examples of Manifestations

- **J. S. Bach's Six suites for unaccompanied cello (Work)**
  - e1 performances by Janos Starker recorded partly in 1963 and completed in 1965
    - m1 recordings released on 33 1/3 rpm sound discs in 1966 by Mercury
    - m2 recordings re-released on compact disc in 1991 by Mercury
  - e2 performances by Yo-Yo Ma recorded in 1983
    - m1 recordings released on 33 1/3 rpm sound discs in 1983 by CBS Records
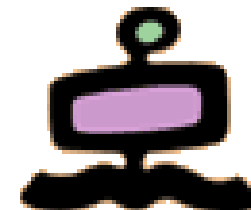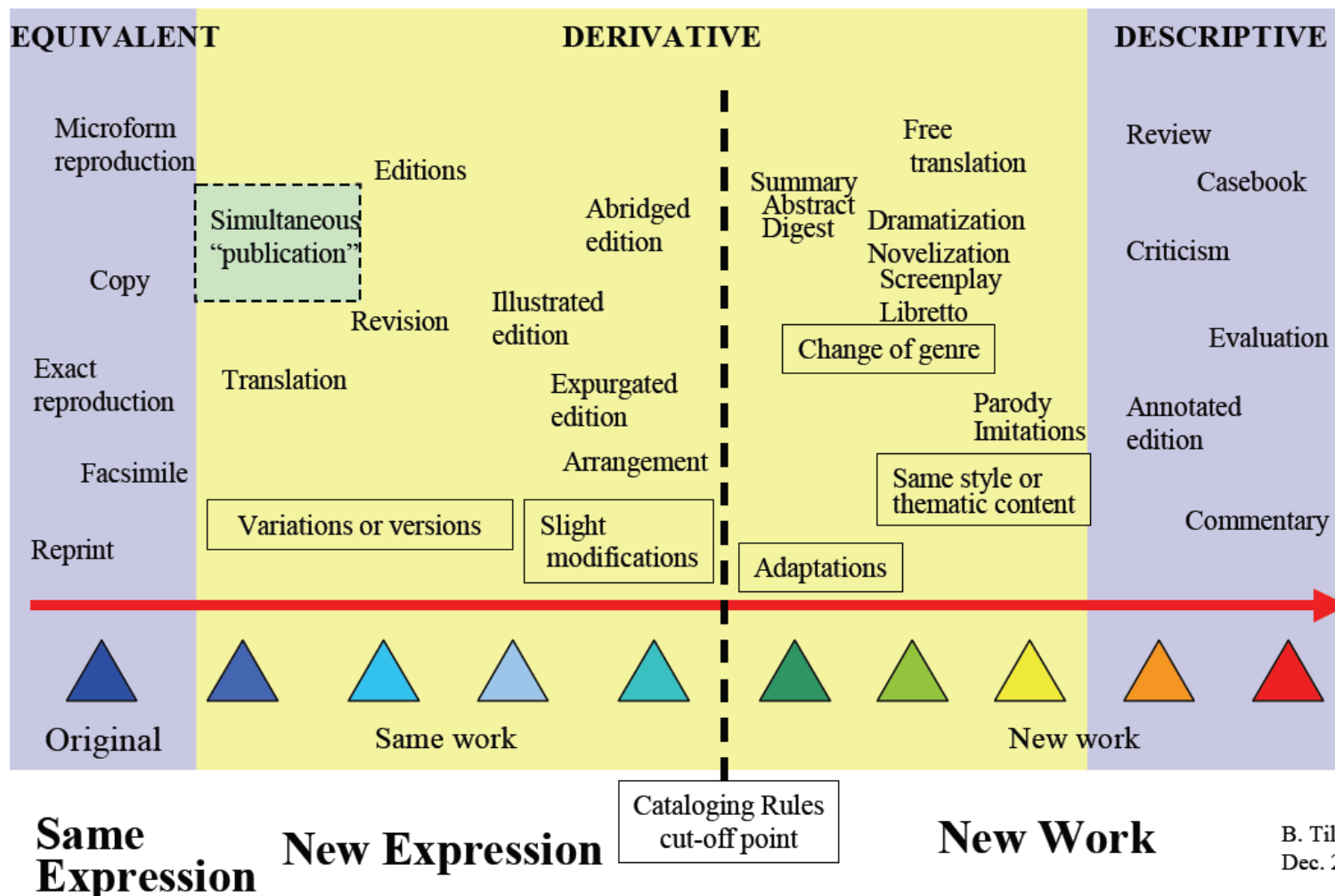   - m2 recordings re-released on compact disc in 1992 by CBS Records

Example of Item

- **w1 Ronald Hayman's Playback**
  - e1 the author's text edited for publication
    - m1 the book published in 1973 by Davis-Poynter
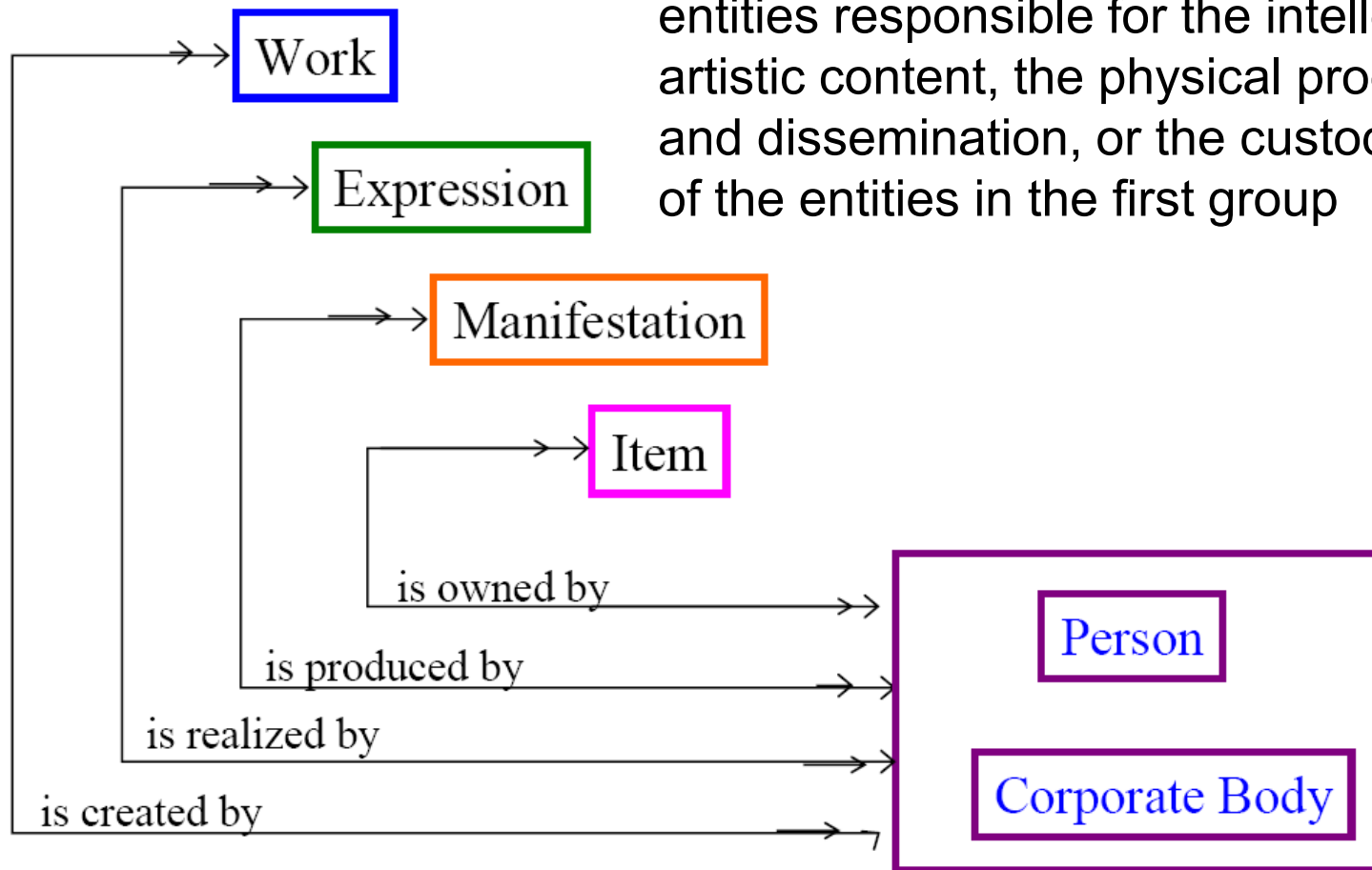   - I 1 copy autographed by the author
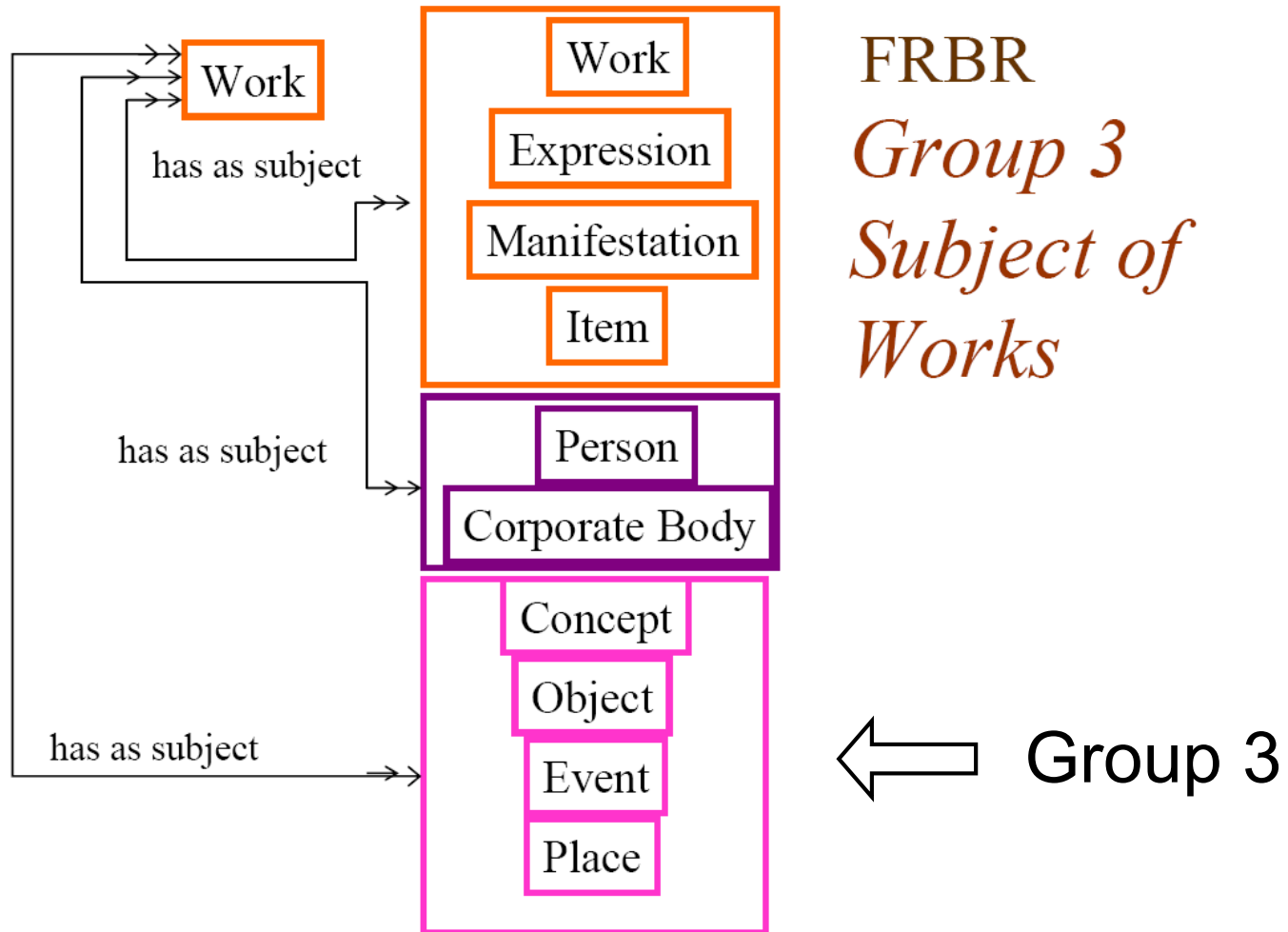
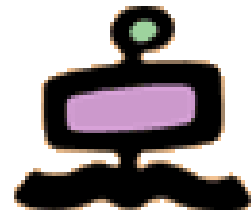# From Work to Item
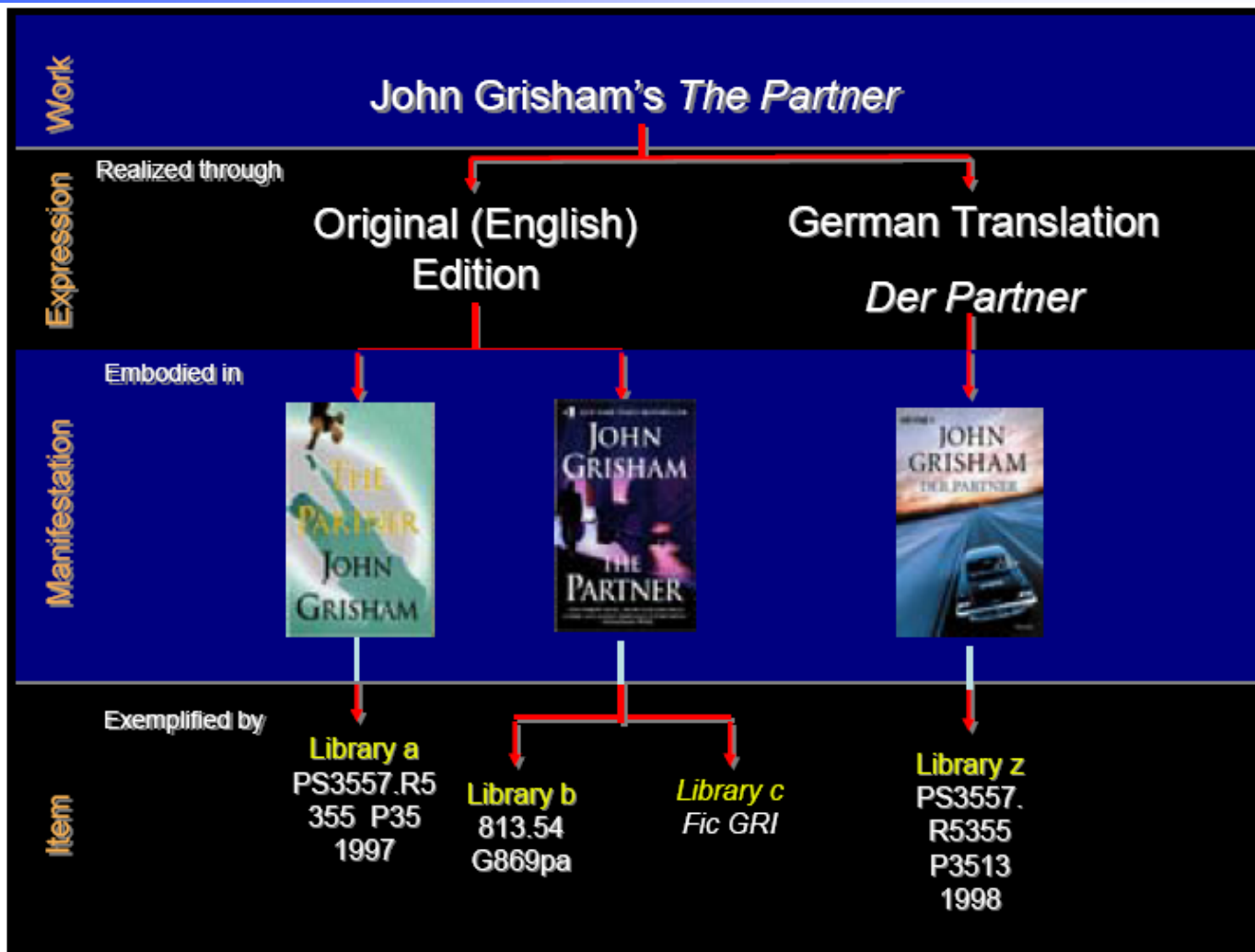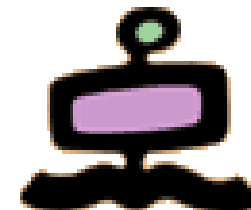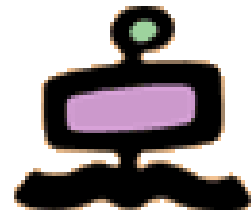
Family of Works

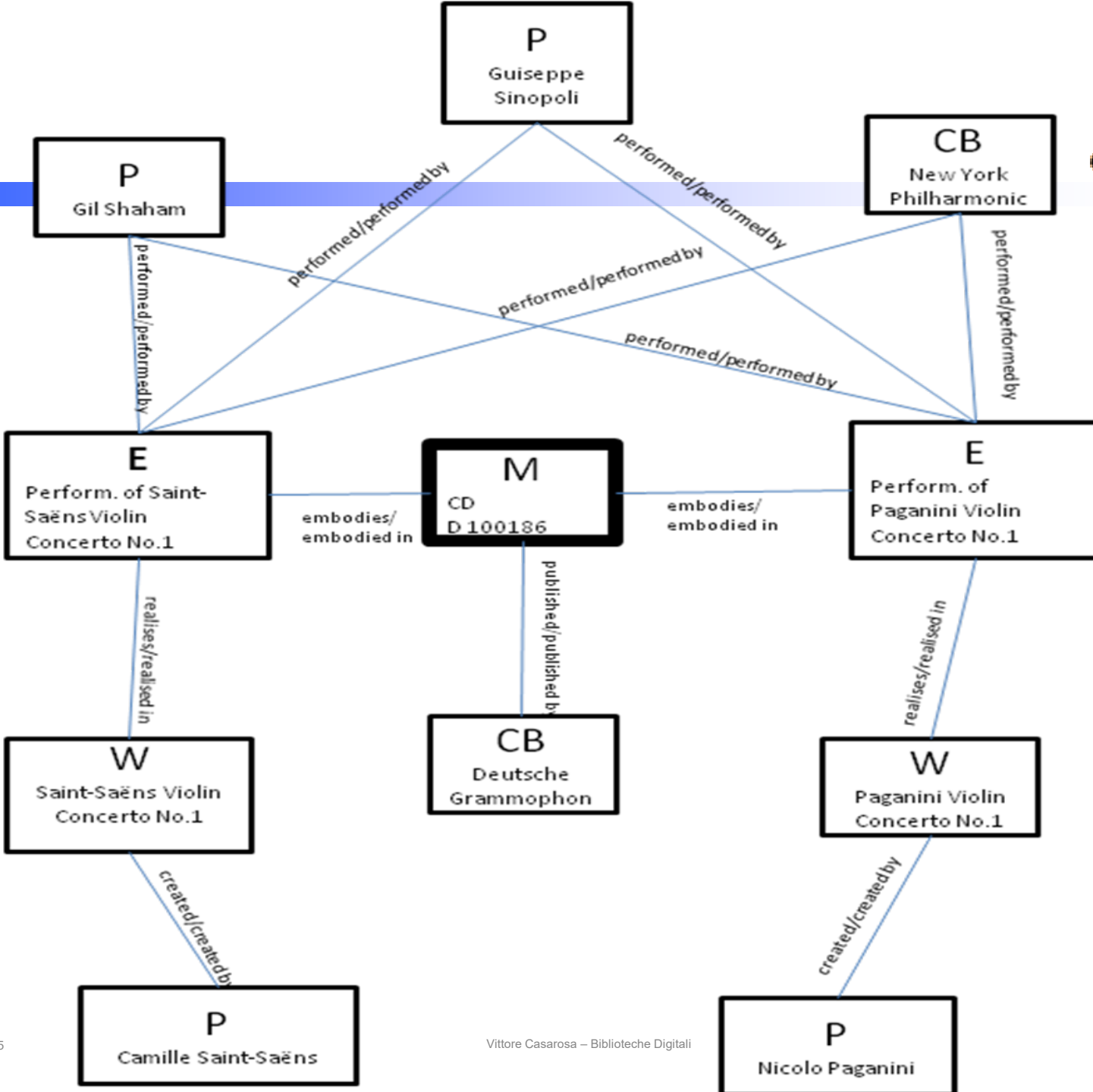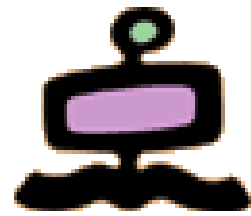# FRBR – Group 2 entities

entities responsible for the intellectual or artistic content, the physical production and dissemination, or the custodianship of the entities in the first group

Work

Expression

Manifestation

Item

is owned by

is produced by

is realized by

is created by

Person

Corporate Body

FRBR – Group 3 entities

FRBR
Group 3
Subject of
Works
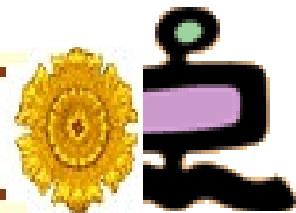
⇐ Group 3

FRBR example

Shakespeare, William, 1564-1616. Hamlet. French.

LC Control No. : 47023612

LCCN Permalink :

Type of Material : Book (Print, Microform, Electronic, etc.)

Personal Name :

Main Title : ... Hamlet, traduit par André Gide.

Published/Created : [Paris] Gallimard [1946]

Description : 2 p. l., 7-237, [2] p. 17 cm.

CALL NUMBER :   Copy 1

-- Request in : Jefferson or Adams Bldg General or
Area Studies Reading Rms

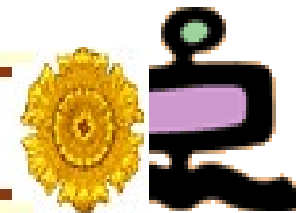**Shakespeare, William, 1564-1616. Hamlet. French.**

Work

LC Control No.	:	47023612
LCCN Permalink	:	
Type of Material	:	Book (Print, Microform, Electronic, etc.)
Personal Name	:	
Main Title	:	... Hamlet, traduit par André Gide.
Published/Created	:	[Paris] Gallimard [1946]
Description	:	2 p. l., 7-237, [2] p. 17 cm.

CALL NUMBER	:	Copy 1
-- Request in	:	Jefferson or Adams Bldg General or Area Studies Reading Rms

Shakespeare, William, 1564-1616. Hamlet. French.

Expression

LC Control No.	:	47023612
LCCN Permalink	:	
Type of Material	:	Book (Print, Microform, Electronic, etc.)
Personal Name	:	
Main Title	:	... Hamlet, traduit par André Gide.
Published/Created	:	[Paris] Gallimard [1946]
Description	:	2 p. l., 7-237, [2] p. 17 cm.

CALL NUMBER	:	Copy 1
-- Request in	:	Jefferson or Adams Bldg General or Area Studies Reading Rms

Shakespeare, William, 1564-1616. Hamlet. French.

LC Control No. :	47023612
LCCN Permalink :	
Type of Material :	Book (Print, Microform, Electronic, etc.)
Personal Name :	
Main Title :	... Hamlet, traduit par André Gide.
Published/Created :	[Paris] Gallimard [1946]
Description :	2 p. l., 7-237, [2] p. 17 cm.

Manifestation

CALL NUMBER :	Copy 1
-- Request in :	Jefferson or Adams Bldg General or Area Studies Reading Rms

Shakespeare, William, 1564-1616. Hamlet. French.

LC Control No. : 47023612
LCCN Permalink :
Type of Material : Book (Print, Microform, Electronic, etc.)
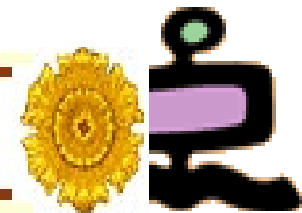**Personal Name** :
Main Title : ... Hamlet, traduit par André Gide.
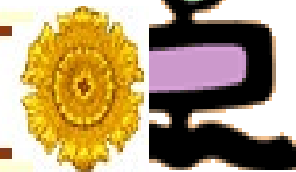**Published/Created** : [Paris] Gallimard [1946]
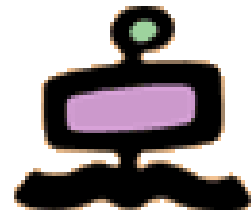**Description** : 2 p. l., 7-237, [2] p. 17 cm.

Item

CALL NUMBER :   Copy 1
-- Request in : Jefferson or Adams Bldg General or
Area Studies Reading Rms

Attributes of Work

- title of the work
- form of work
- date of the work
- other distinguishing characteristic
- intended termination
- intended audience
- context for the work
- medium of performance (musical work)
- numeric designation (musical work)
- key (musical work)
- coordinates (cartographic work)
- equinox (cartographic work)
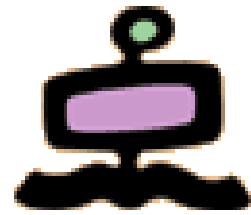
Attributes of Expression

- title of the expression
- form of expression
- date of expression
- language of expression
- other distinguishing characteristic
- extensibility of expression
- revisability of expression
- extent of the expression
- summarization of content
- context for the expression
- critical response to the expression
- use restrictions on the expression
- sequencing pattern (serial)
- expected regularity of issue (serial)

- expected frequency of issue (serial)
- type of score (musical notation)
- medium of performance (musical notation or recorded sound)
- scale (cartographic image/object)
- projection (cartographic image/object)
- presentation technique (cartographic image/object)
- representation of relief (cartographic image/object)
- geodetic, grid, and vertical measurement (cartographic image/object)
- recording technique (remote sensing image)
- special characteristic (remote sensing image)
- technique (graphic or projected image)

Attributes of Manifestation

- title of the manifestation
- statement of responsibility
- edition/issue designation
- place of publication/distribution
- publisher/distributor
- date of publication/distribution
- fabricator/manufacturer
- series statement
- form of carrier
- extent of the carrier
- physical medium
- capture mode
- dimensions of the carrier
- manifestation identifier
- source for acquisition/access authorization
- terms of availability
- access restrictions on the manifestation
- typeface (printed book)
- type size (printed book)
- foliation (hand-printed book)

- collation (hand-printed book)
- publication status (serial)
- numbering (serial)
- playing speed (sound recording)
- groove width (sound recording)
- kind of cutting (sound recording)
- tape configuration (sound recording)
- kind of sound (sound recording)
- special reproduction characteristic (sound recording)
- colour (image)
- reduction ratio (microform)
- polarity (microform or visual projection)
- generation (microform or visual projection)
- presentation format (visual projection)
- system requirements (electronic resource)
- file characteristics (electronic resource)
- mode of access (remote access electronic resource)
- access address (remote access electronic resource)
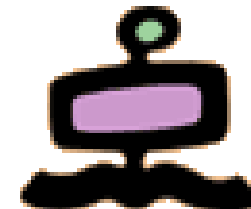
Attributes of Item

- item identifier

- fingerprint

- provenance of the item

- marks/inscriptions

- exhibition history

- condition of the item

- treatment history

- scheduled treatment

- access restrictions on the item

Attributes and user tasks

- The "knowledge" described by means of the entities and their attributes is the base upon which a user can perform his/her tasks

- Verify the usefulness of the attributes for the performance of each task

- User tasks

  – Find

  – Identify

  – Select

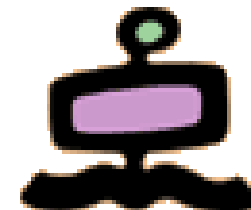 – Obtain

Work and users' tasks

	Find				Identify				Select				Obtain			
	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item
Attributes of a work																
Title of the work	■				■				■							
Form of work	□				□				■							
Date of the work	○				○				○							
Other distinguishing characteristic	○				○				○							
Intended termination	□				■				□							
Intended audience	○				○				□							
Context for the work									○							
Medium of performance (musical work) (note 1)	□				□				□							
Numeric designation (musical work) (note 1)	○				□											
Key (musical work) (note 1)	○				□											
Coordinates (cartographic work)	□				■				■							
Equinox (cartographic work)	□				□				□							

■ = high value   □ = medium value   ○ = low value

Manifestation and users' tasks

|	Find				Identity				Select				Obtain			
	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item
Attributes of a Manifestation																
Title of the manifestation	■	■	■		■	■	■		□	□	□				■	
Statement of responsibility	□	□	□		□	■	■		□	■	■				■	
Edition/issue designation						■	■			■	■				■	
Place of publication/distribution							○		○	○	○				■	
Publisher/distributor			○			□	■			○	□				■	
Date of publication/distribution	□	□	○		□	□	■		□	□	■				■	
Fabricator/manufacturer (note 1)			○				○				○					
Series statement			□				■				□				■	
Form of carrier			□				■				■				■	
Extent of the carrier (note 2)							□			□	□					
Physical medium (note 3)							□				○					
Capture mode							○				□					
Dimensions of the carrier (note 4)							○				□				□	
Manifestation identifier			■			□	■								■	
Source for acquisition/access authorization (note 5)											○				○	
Terms of availability											○				○	
Access restrictions on the manifestation											□				■	

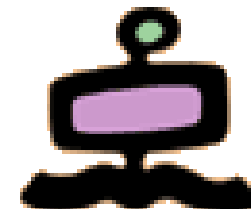■ = high value   □ = medium value   ○ = low value
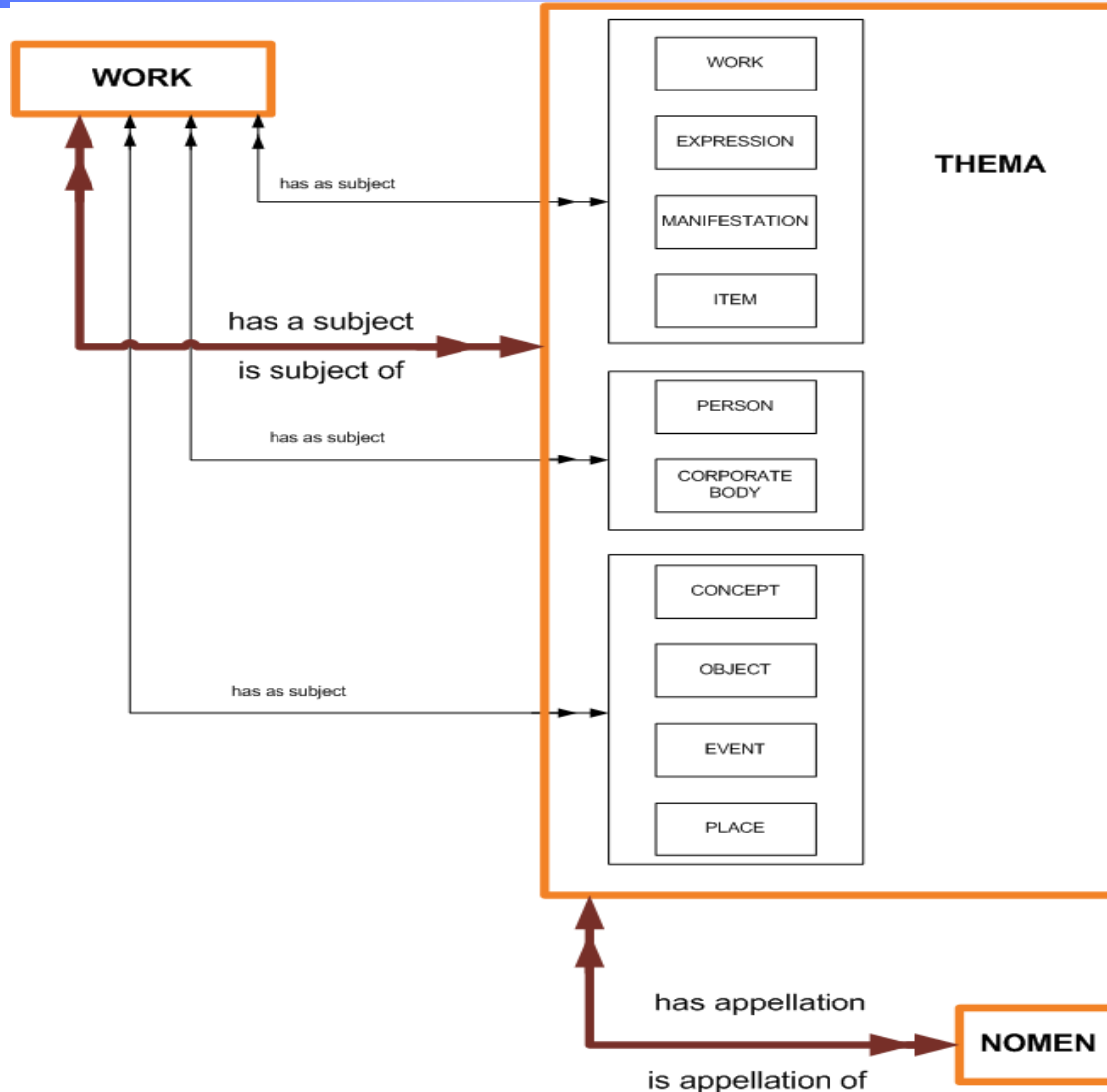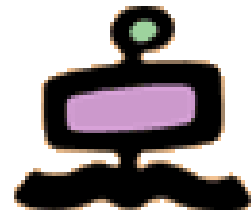
The "FRBR family"

- **FRBR: the original framework**
  - All entities, focusing on Group 1
- **FRAD: Functional Requirements for Authority Data (names)**
  - Focus on Group 2
  - Published in 2009
- **FRSAD: Functional Requirements for Subject Authority Data (subjects)**
  - Focus on 'aboutness' (Group 3)
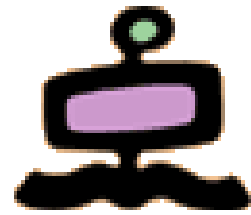 - In revision after IFLA review, published in 2010
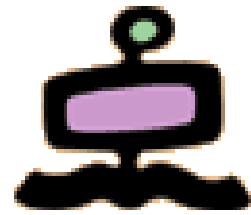
FRAD (names)
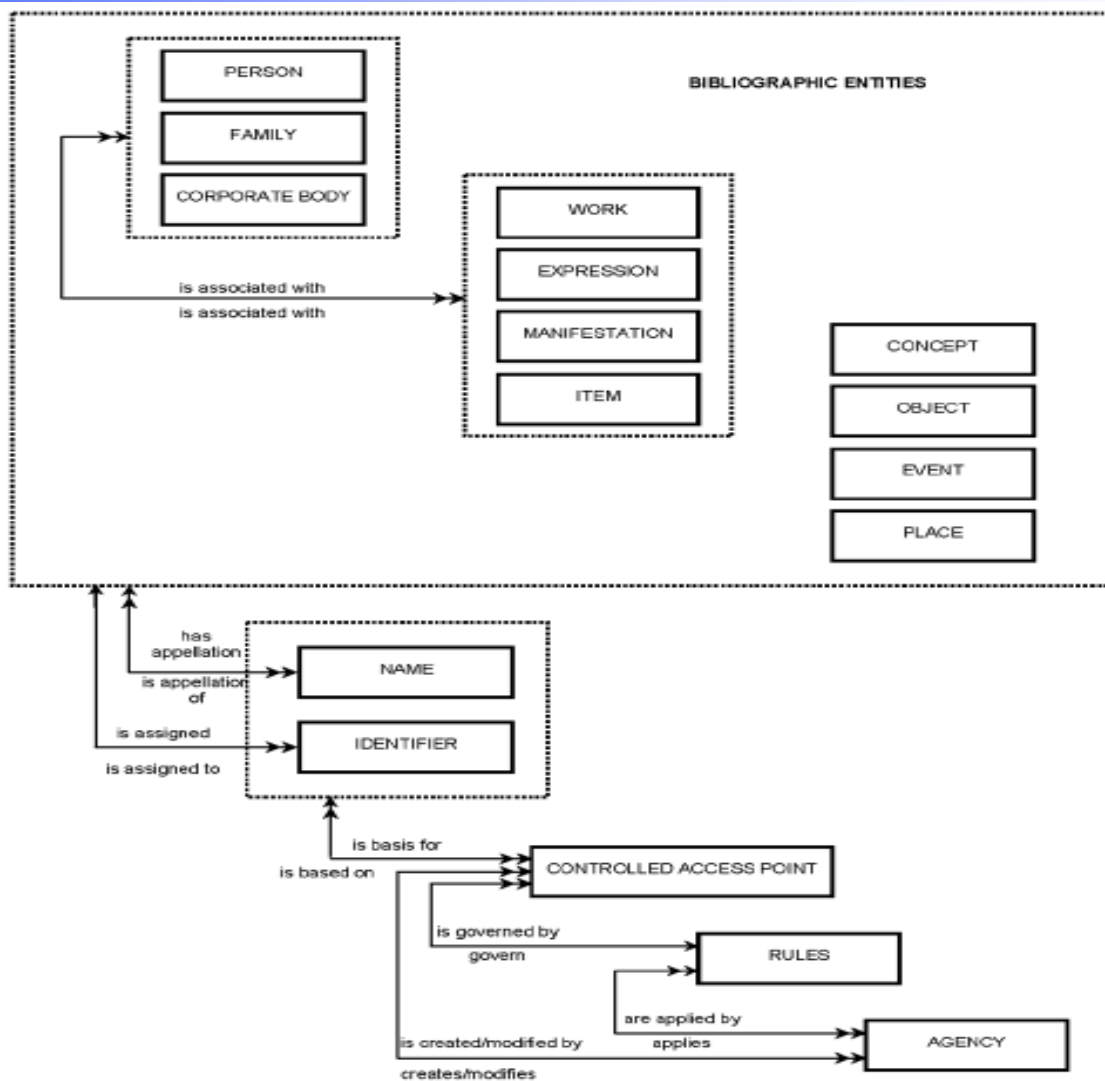
FRSAD (subjects)

Knowledge representation

- **FRBR: Functional Requirements for Bibliographic Records**
  - LRM: Library Reference Model ⬅

- **RDF: Resource Description Framework**
- **RDF Schema**
  - LOD: Linked Open Data

Library Reference Model (LRM)

IFLA LRM (published in August 2017) is a high-level conceptual reference model consolidating the three separately developed IFLA conceptual models: FRBR, FRAD, FRSAD

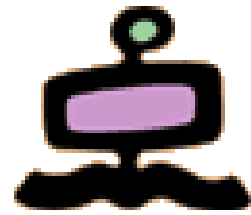| Top Level	Second Level	Third Level
LRM-E1 Res		
--	LRM-E2 Work	
--	LRM-E3 Expression	
--	LRM-E4 Manifestation	
--	LRM-E5 Item	
--	LRM-E6 Agent	
--	--	LRM-E7 Person
--	--	LRM-E8 Collective Agent
--	LRM-E9 Nomen	
--	LRM-E10 Place	
--	LRM-E11 Time-span	

Library Reference Model (LRM)
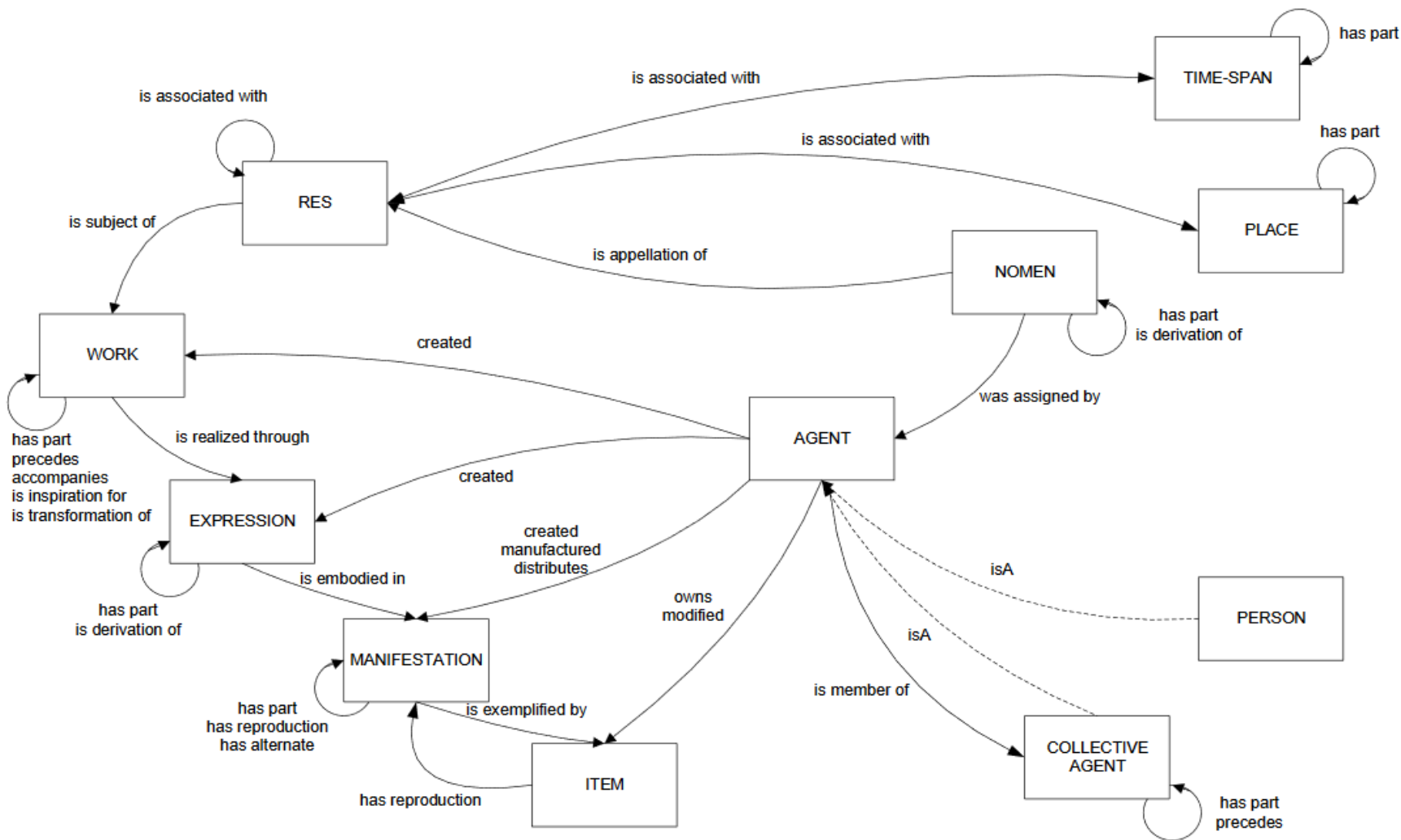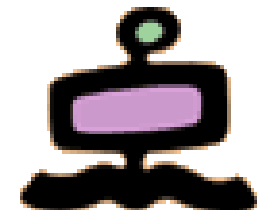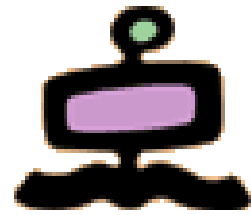
- **Find (FRBR 1997)**

  - To find entities that correspond to the user's stated search criteria (i.e., to locate either a single entity or a set of entities in a file or database as the result of a search using an attribute or relationship of the entity)

- **Find (LRM 2017)**

  - To bring together information about one or more resources of interest by searching on any relevant criteria

User Tasks - Identify

- **Identify (FRBR 1997)**
  - To identify an entity (i.e., to confirm that the entity described corresponds to the entity sought, or to distinguish between two or more entities with similar characteristics)

- **Identify (LRM 2017)**
  - To clearly understand the nature of the resources found and to distinguish between similar resources

User Tasks - Select

- ● Select (FRBR 1997)
  - – To select an entity that is appropriate to the user's needs (i.e., to choose an entity that meets the user's requirements with respect to content, physical format, etc., or to reject an entity as being inappropriate to the user's needs)

- ● Select (LRM 2017)
  - – To determine the suitability of the resources found, and to be enabled to either accept or reject specific resources

User Tasks - Obtain

- ## Obtain (FRBR 1997)

  - To acquire or obtain access to the entity described (i.e., to acquire an entity through purchase, loan, etc., or to access an entity electronically through an online connection to a remote computer)

- ## Obtain (LRM 2017)

  - To access the content of the resource

User Tasks - Explore

- **Contextualize (FRAD 2009)**
  - Place a person, corporate body, work, etc., in context; clarify the relationship between two or more persons, corporate bodies, works, etc.; or clarify the relationship between a person, corporate body, etc., and a name by which that person, corporate body, etc., is known (e.g., name used in religion versus secular name)

- **Explore (FRSAD 2010)**
  - Explore relationships between subjects and/or their appellations (e.g., to explore relationships in order to understand the structure of a subject domain and its terminology)

- **Explore (LRM 2017)**
  - To discover resources using the relationships between them and thus place the resources in a context