



HAL
open science

Modélisation sémantique des annotations pour l'exploitation critique des sources iconographiques

Thomas Bottini, Rebecca Bristow, Louise Quentel

► To cite this version:

Thomas Bottini, Rebecca Bristow, Louise Quentel. Modélisation sémantique des annotations pour l'exploitation critique des sources iconographiques. Journée d'étude travailler avec les images, L'unité de service CERES (Centre d'expérimentation en méthodes numériques pour les recherches en Sciences Humaines et Sociales) de la Faculté des Lettres de Sorbonne Université, Jun 2022, Paris, France. halshs-03950394

HAL Id: halshs-03950394

<https://shs.hal.science/halshs-03950394>

Submitted on 21 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

2022.06.09 — CERES, journée d'étude travailler avec les images

Figure 1: téléchargement.gif

Modélisation sémantique des annotations pour l'exploitation critique des sources iconographiques

Thomas Bottini IReMus UMR 8223 CNRS + CERES

Rebecca Bristow IReMus UMR 8223 CNRS + CERES

Louise Quentel UR Histoire & Critique des Arts, Université Rennes 2

Des opérations mentales & techniques

Produire des savoirs sur des corpus iconographiques renvoie à un ensemble d'opérations à la fois mentales et matérielles :

- Constituer un corpus d'images
- Explorer un corpus
- Documenter une image et son contexte de production
- Accéder au contexte éditorial de l'image
- Indexer une image avec des catégories analytiques
- Annoter une image
- ...

Notion de « Scholarly primitives » dans la littérature DH anglo-saxonne (Unsworth, 2000)

Un cas

L'étude du corpus d'estampes du *Mercure Galant*, périodique mensuel de l'époque de Louis XIV (1672).

→ Le *Mercure Galant* (1672-1710) :

- fondé en 1672 à l'initiative de l'homme de lettres Jean Donneau de Visé
- actualité sociale, littéraire et artistique
- 476 estampes, toutes gravées sur cuivre, associées à la publication entre 1677 et 1710

Situation de travail

- Regarder les estampes du corpus



Figure 2: https://www.iremus.cnrs.fr/sites/default/files/mercure_lecture_dair.jpeg

- Les analyser
- Reporter et formaliser ces analyses dans une feuille de saisie

Propriétés génériques

- date d'édition
- titres
- acteurs de l'édition
- techniques
- dimensions

Lieux, personnes et objets

- lieux et personnes
- objets

Thématique et type de représentation

- thématique

EXTRAORDINAIRE
DU MERCURE
GALANT.

QUARTIER D'OCTOBRE 1678.

DEDIE' AU ROY.

TOME IV.



G. 1563. A PARIS.

28

AV PALAIS



Source gallica.bnf.fr / Bibliothèque nationale de France

Figure 3: Extraordinaire_du_Mercure_galant__bpt6k6373165g_5.jpeg



Source gallica.bnf.fr / Bibliothèque nationale de France

Figure 4: Mercure_galant__bpt6k6220151w_141.jpeg



Figure 5: 1684-06b_385 La Planche « Luxembourg ».jpg

- type de représentation

Étude de cas

→ Créer et utiliser des référentiels

- Vocabulaire d'indexation des estampes

https://docs.google.com/spreadsheets/d/1wS9punldFYlqZpkRgEsDXUSPRH1CDhdPKJYwjMXq_sw/edit#gid=0

- S'aligner sur Iconclass
- Les référentiels de l'Ancien Régime

→ Modélisation sémantique

Notre approche

- Saisir la profondeur de l'organisation interne des sources et du travail analytique
- Notion de graphe de connaissances (sources, contexte, annotations, ressources analytiques...)



Figure 6: 1689-05a_254.jpg



Source gallica.bnf.fr / Bibliothèque nationale de France

Figure 7: Mercure_galant_7_bpt6k6228302k_215.jpeg



Source gallica.bnf.fr / Bibliothèque nationale de France

Figure 8: Mercure_galant__bpt6k6319342c_275.jpeg

A-BE-CE-DE-RES

POUR MON SEIGNEUR LE DUC DE BOURGOGNE

Alphabets des Langues

Orien = Par M^r. Comiers. *tales. d'Ambrun. 1684.*

Hebreu	Rabin	Samaritan	Siriacq	Grec	Arabe	Turc	Armenien	Caracteres Chinois
A	A	A	A	A A.a	A	A	A	天
B	B	B	B	B V B.B	B	B	B	王
Gh	Ch	Gh	Gh	G G.I. γ	T	T	T	康
Dh	Dh	Dh	D	D Δ. Δ. δ	Th	Tf	Th	熙
H	H	H	H	E E. ε	G	Ch	Le	爾
V	V	V	V	Z ζ. ζ	Ch	D	Th	彌
Z	Z	Z	Z	I H. η	Dh	Dz	L	格
Hh	Hh	Hh	Hh	Th Θ. θ	R	R	Hh	尔
T	T	T	T	I I. ι	Z	Z	Gh	沈
I	I	I	Y	K K. κ	S	S	H	
Ch	Ch	Ch	K	L Λ. λ	Sch	Sch	Zz	
L	L	L	L	M M. μ	Sf	Sf	K	
M	M	M	M	N N. ν	Tdh	Dz	G	po mince
N	N	N	N	X Ξ. ξ	T	T	M	po verre
S	S	S	S	O Ο. ο	Trdh	D	I	po ample
Aspiratio	Aa	Aa	A	P Π. π	Aa	A	N	po Deuine
Ph	Ph	Ph	F	R Ρ. ρ	Gh	G	V	po non
Tf	Tf	Tf	Sf	S Σ. σ	Cq	Cq	B	po vielle
Q	Q	PK	Cq	T Τ. τ	K	K	Gg	po apudus
R	R	R	R	Y Υ. υ	L	L	Rr	po fleuve
S	S	Sc	Ch	Ph Φ. φ	M	M	S	po Egalor
Th	T	T	T	L Ch X. χ	N	N	V	po Rompre
Motions		Sur les notes monoyes du sanctuaire I.e. LEAV	Voyelles a e i o u	Pf Ψ. ψ	Voyelles a e ai oeu	Voyelles a e i u	F	Voyelles w t b n

Source gallica.bnf.fr / Bibliothèque nationale de France

Figure 9: Mercure_galant__bpt6k6222729r_143.jpeg



Source gallica.bnf.fr / Bibliothèque nationale de France

Figure 10: Mercure_galant__bpt6k63706517_143.jpeg

Choix méthodologique

- Le choix du CIDOC-CRM : <https://www.cidoc-crm.org/>
- Pourquoi est-ce rassurant ?
- Fonction heuristique (par rapport à un modèle custom)

→ Le CIDOC-CRM en 30 secondes

- Classes & propriétés
- Modèle centré événement

Les entités du CIDOC-CRM

→ Le web sémantique

- Notion de triplet RDF
- Notion d'IRI
- Notion de sémantique partagée

Les images

E36 Visual Item

- Serveur IIIF (Cantaloupe)

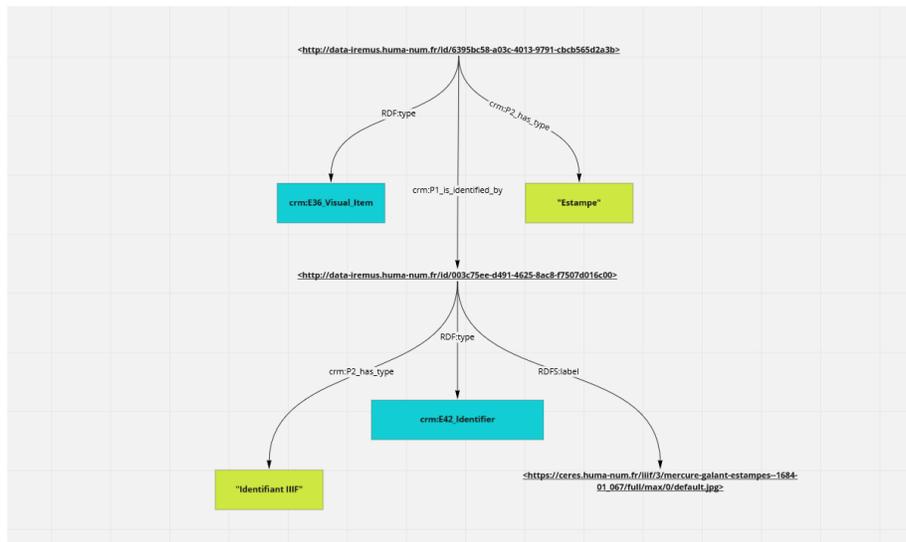


Figure 13: mg-graph-E36-E42.png

https://ceres.huma-num.fr/iiif/3/mercure-galant-estampes-1684-01_067/full/max/0/default.jpg

https://ceres.huma-num.fr/iiif/3/mercure-galant-estampes-1684-01_067/info.json

https://ceres.huma-num.fr/iiif/3/mercure-galant-estampes-1684-01_067/550,1100,200,300/max/0/default.jpg

- L'organisation interne d'une image

Structure d'une annotation

- Utilisation de la classe E13 Attribute Assignment, et bénéfiques pour l'expression du contexte (signature + horodatage + autres éléments contextuels)
- Une annotation se décompose en :
 - Une zone d'image (sous E36, cf. *supra*) → la cible de l'annotation
 - Un contenu ajouté (ce qu'on dit de la zone d'image) → le corps de l'annotation
 - Une propriété qui connecte la cible et le corps
- Rapport entre triplet et E13 (une instance d'E13 assimilable à la réification d'un triplet avec expression du contexte)

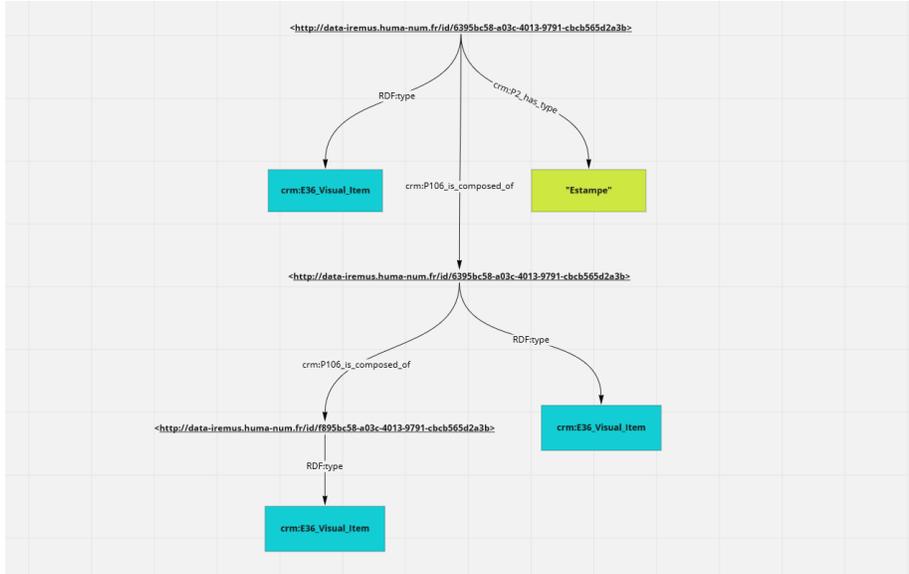


Figure 14: mg-graph-E36-P106.png

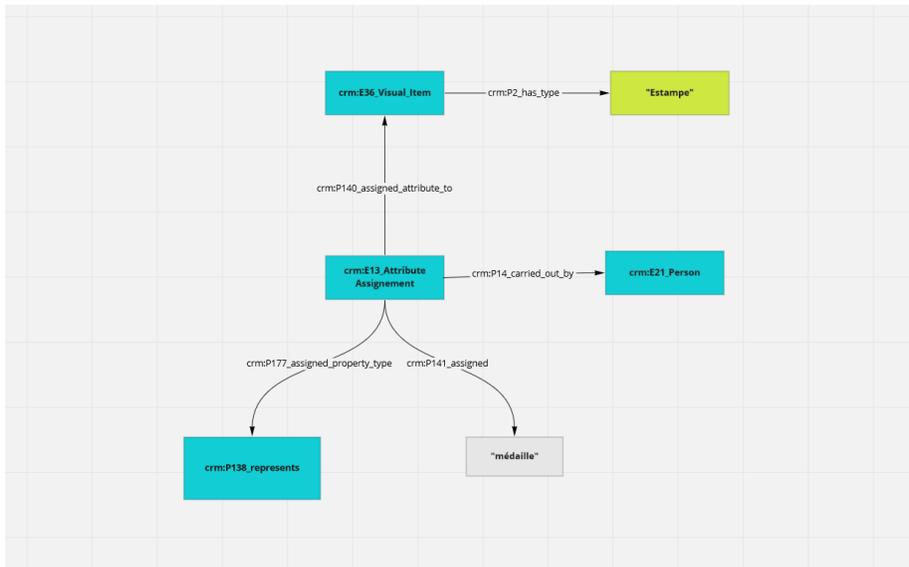


Figure 15: mg-graph-E36-E13.png

- Un triplet vient sans son contexte, donc ça ne convient pas pour exprimer des connaissances scientifiques !
- Le CRM encourage à représenter une connaissance comme l'événement duquel émane cette connaissance, et donc favorise l'expression du contexte.

Annoter avec des concepts

- Notion de thésaurus (= / = ontologie les mots vs. les choses)
- L'exemple du vocabulaire des estampes : types de représentation, thématiques, objets, techniques.

Ex : « Cette image contient une zone qui représente une médaille »

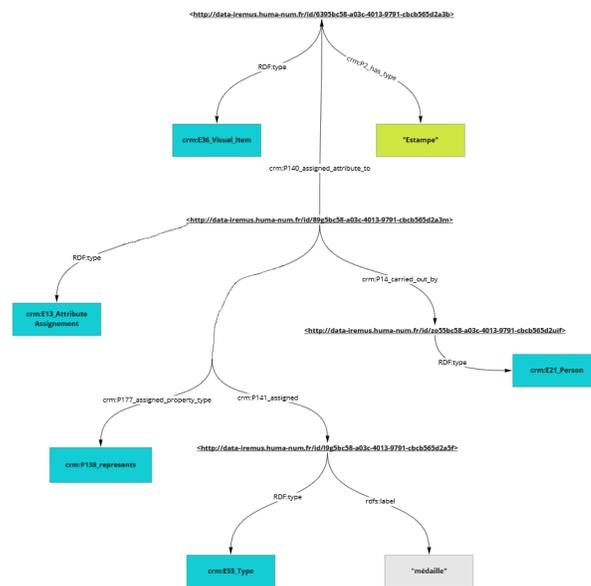


Figure 16: mg-graph-E36-E13-E55_2.png

Exprimer le contexte de production

- Il est intéressant de le représenter autour d'événements (E12)
- Ex : *invenit, sculpsit...*

Le contexte éditorial des estampes

- Un périodique, des livraisons, des articles
- LRMoo

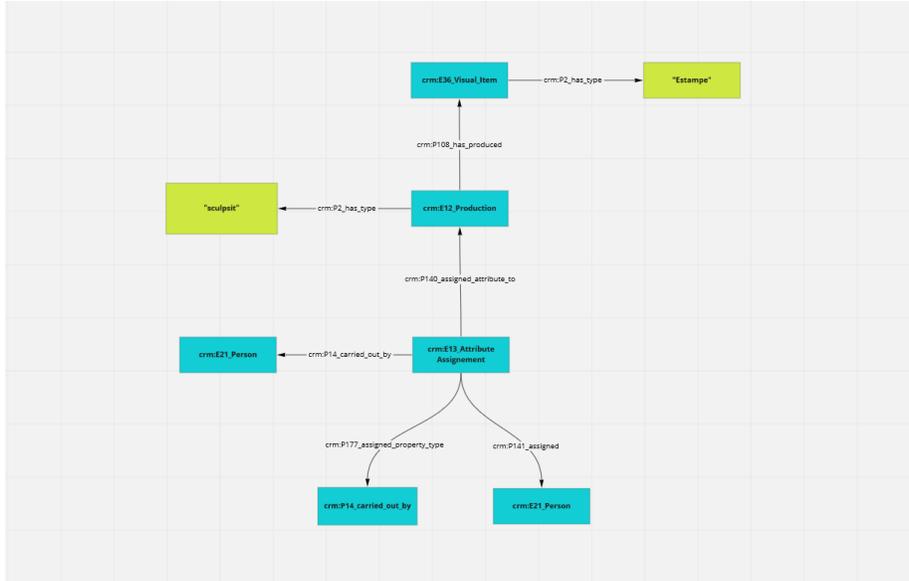


Figure 17: mg-graph-E36-E12.png



Figure 18: mg-graph-E36-F2.png

Science ouverte ?

- SPARQL endpoint
- Rôle du CIDOC-CRM : présenter le contexte de production des données et non seulement les données

Vers une interface d'exploration & d'annotation

- Notre approche : dégager les invariants fonctionnels entre différents corpus et pratiques scientifiques ayant pour point commun le fait de travailler sur des images.
- Plusieurs projets pilotes (nous recherchons des corpus différents pour éprouver notre hypothèse de généralité)
- Trois niveaux :
 - Fonctions génériques non spécifiques aux corpus iconographiques
 - * Naviguer dans les images par leur contexte documentaire/éditorial (livraisons du MG, tweets...)
 - * Visualiser les annotations attachées à une ressource
 - * Effectuer des recherches sémantiques
 - * Annoter une ressource
 - * Créer des corpus et des projets analytiques
 - * Constituer des sélections de ressources ou de fragments de ressources dans un panier et les transformer en projet/corpus
 - * Explorer le corpus par les projets analytiques (<https://www.notion.so/Contexte-des-annotations-6d858b3b7e8c4a6bb4c1f5e630249d58>)
 - * Obtenir des statistiques et les convoquer pour bâtir des stratégies d'exploration des corpus
 - * Elastic Search (sources + contexte + annotations)
 - * Naviguer dans les triplets
 - Fonctions génériques aux corpus iconographiques
 - * Afficher une image et zoomer
 - * Délimiter et identifier des zones d'une image
 - * Bâtir des vues synoptiques de type « galerie »
 - Fonctions spécifiques par projet
 - * Accès au contexte éditorial