

MonumenTAL

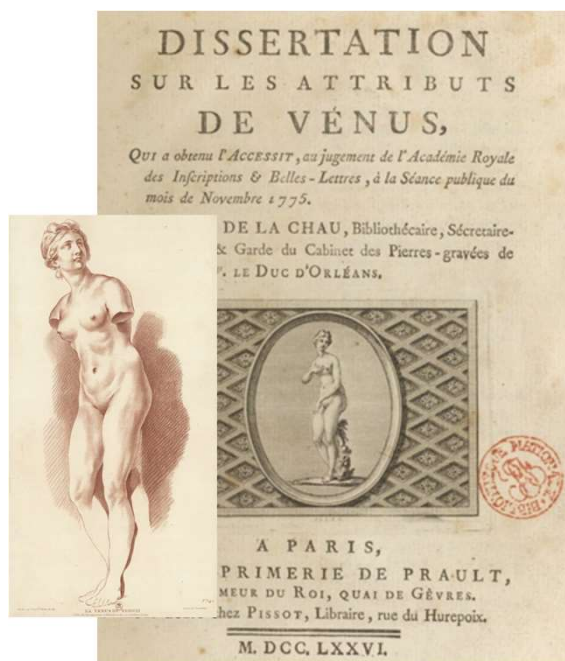
Ressources pour l'étude des appellations d'œuvres visuelles antiques

corpus, dictionnaires et outil de reconnaissance automatique

Présentation : Aurore Lessieux, Anne-Violaine Szabados,
Iris Eshkol-Taravella, Marlène Nazarian



Travail réalisé dans le cadre du labex Les passés dans le présent
(programme Investissements d'avenir réf. ANR-11-LABX-0026-01)



Vénus de Médicis (documents BnF, Gallica)

Présentation de la première phase du projet MonumenTAL, un projet pluridisciplinaire soutenu par le LabEx *les passés dans le présent*, et qui est à mi-parcours de son déroulement.

Il rassemble, des :

- historiens de l'art, archéologues et philologues du LIMC, du laboratoire ArScAn- UMR7041 (Archéologies & Sciences de l'Antiquité),
- de MoDyco (Modèles, Dynamiques, Corpus - UMR 7114)
- et des conservateurs et bibliothécaires de la BnF,

Appellations d'œuvre visuelle antique

- Œuvres visuelles antiques = statues ou scènes représentées sur des artefacts
- Facteurs de variation des appellations d'œuvre :
 - Principes de nommage selon les époques
 - Erudition de l'auteur
 - Substitution d'un nom de personnage grec <> latin

Aphrodite Vénus



- « Zéthos, Amphion, Dircé, un taureau »
(époque antique)
- Taureau Farnèse
- Amphion et Zéthos
- Supplice de Dircé
(époque moderne)



- Vénus de Milo
(statue grecque)
→ Aphrodite ...

MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

Les historiens de l'art sont confrontés à un volume de sources textuelles numériques qui ne cesse de croître et dans lesquelles les œuvres et les images sont désignées avec une grande variabilité. Les œuvres visuelles de l'Antiquité classique, telles que les statues ou les scènes représentées sur des artefacts et des décors architecturaux, sont souvent connues sous plusieurs titres attribués au fil des siècles.

La répétition de motifs iconographiques est une constante dans la création des œuvres figurées relevant de la culture classique - et cela de l'Antiquité à nos jours.

C'est cette répétition de motifs, qui a permis :

- la diffusion d'une culture visuelle dans laquelle les spectateurs reconnaissent, à travers les multiples copies, des types iconographiques célèbres (tel que l'Aphrodite de Cnide) ainsi que
- l'émergence de titres d'œuvres, plus ou moins fixes et plus ou moins acceptés, en raison des commentateurs et amateurs d'art qui cherchent à nommer les œuvres et les types iconographiques qu'ils reconnaissent,

Ainsi, les œuvres visuelles de l'Antiquité classique sont souvent connues sous plusieurs titres attribués :

- au fil des siècles, certains dès la période antique,
- et selon des principes de nommage qui témoignent de divers contextes de transmission textuelle de la connaissance,
- à cela s'ajoute la fréquente substitution, selon les époques, d'un nom de personnage grec par celui de son équivalent latin.

Les ouvrages anciens sont riches d'informations or, la notion de titre d'œuvre, comme elle est pratiquée aujourd'hui, n'est pas fixée dans l'antiquité, ce qui ne facilite pas l'identification des objets et des images dans les textes.

Le projet **MonumenTAL** (2020-2023) : Monuments antiques & TAL

- Traitement Automatique de la Langue appliqué aux *appellations d'œuvres visuelles*
- Projet pluridisciplinaire :
 - Historiens de l'art, archéologues, philologues
 - Linguistes-TAListes
 - Conservateurs, bibliothécaires



Hercule Farnèse
(F.Perrier, BnF, Gallica)

Objectifs :

- Constituer des référentiels (normalisation)
 - appellations d'œuvre
 - liaison objet + image + texte
- Enrichir, indexer

Site Web: <https://heurist.huma-num.fr/heurist/?db=MonumenTAL&website>

MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

La présentation porte sur la recherche réalisée dans le cadre du projet Monumental et qui vise à repérer, collecter et exploiter les titres d'œuvres visuelles de l'Antiquité classique portant une représentation figurée.

Il s'agit d'un projet pluridisciplinaire qui réunit historiens de l'art, archéologues et philologues avec des linguistes spécialistes du Traitement automatique de la langue, ainsi que des conservateurs et bibliothécaires,

Pour atteindre ces objectifs, le projet s'appuie sur les techniques et méthodes du TAL afin d'élaborer, notamment, un référentiel d'appellations d'œuvres et de sujets thématiques, de type thésaurus, exploitable pour de l'indexation de corpus ou de l'annotation des textes.

Corpus textuel

- publications du 18^e s. au 21^e s. en français + textes antiques traduits
- issus de Gallica (océrésés et numérisés par la BNF)
- textes de nature variée

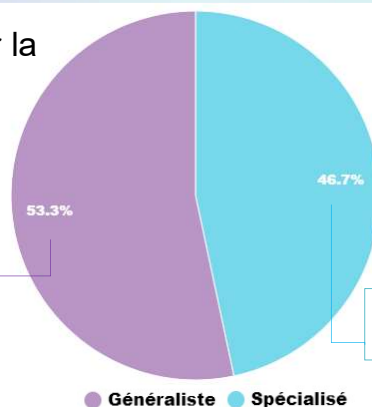
+ généraliste

plus de 1000 références actuellement dans Zotero

+ spécialisé

- 30 textes sélectionnés pour la recherche (591 711 tokens)

2 manuels
11 articles encyclopédiques
3 guides de musée



● Généraliste ● Spécialisé

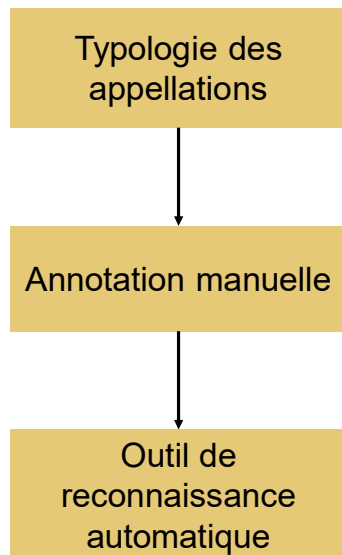
MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

Pour répondre aux objectifs du projet et couvrir les différentes périodes, dans une perspective d'étude diachronique des appellations, le corpus constitué comprend les textes entre le 18^e s. et le 21^e s. issus principalement de Gallica, la bibliothèque numérique de la Bibliothèque nationale de France ; Dont 30 textes sélectionnés pour cette étude.

Pour atteindre la représentativité des données, c'est-à-dire pour rendre compte des différentes pratiques d'appellation, les textes récupérés sont de nature variée (allant des publications très spécialisées orientés vers le public érudit jusqu'aux textes généralistes destinés au grand public) tel que les catalogues de collection, des essais, des articles de périodiques ou d'encyclopédies, des guides, auxquels s'ajoutent traductions des textes antiques grecs et latins qui restituent la façon des auteurs anciens d'évoquer les œuvres et dans lesquels prend racine plus d'un titre. Plus de mille documents sont aujourd'hui référencés dans la bibliographie Zotero du programme,

Le travail a commencé sur des publications traitant majoritairement de la statuaire, pour laquelle il y a beaucoup d'œuvres titrées, et des gemmes gravées, dont les sujets sont simples, puis il a été élargi aux autres domaines artistiques (peinture, mosaïque, céramique) qui sont plus difficiles à récupérer à cause de traditions d'écriture.

Méthodologie



5

La méthodologie.

Typologie des appellations d'œuvre

▪ Appellation Structurée (AS)

- Vénus de Milo ; Vase Cesi-Albani-Torlonia
personnage lieu support collections
- statue / copie de l'Aphrodite de Cnide de Praxitèle
support personnage lieu artiste
- Hercule Farnèse (copie)
personnage collection support

▪ Appellation Thématique (TM)

- Jugement de Pâris ; Amazonomachie
- Commode en Hercule
- Persée délivrant Andromède / Andromède délivrée par Persée

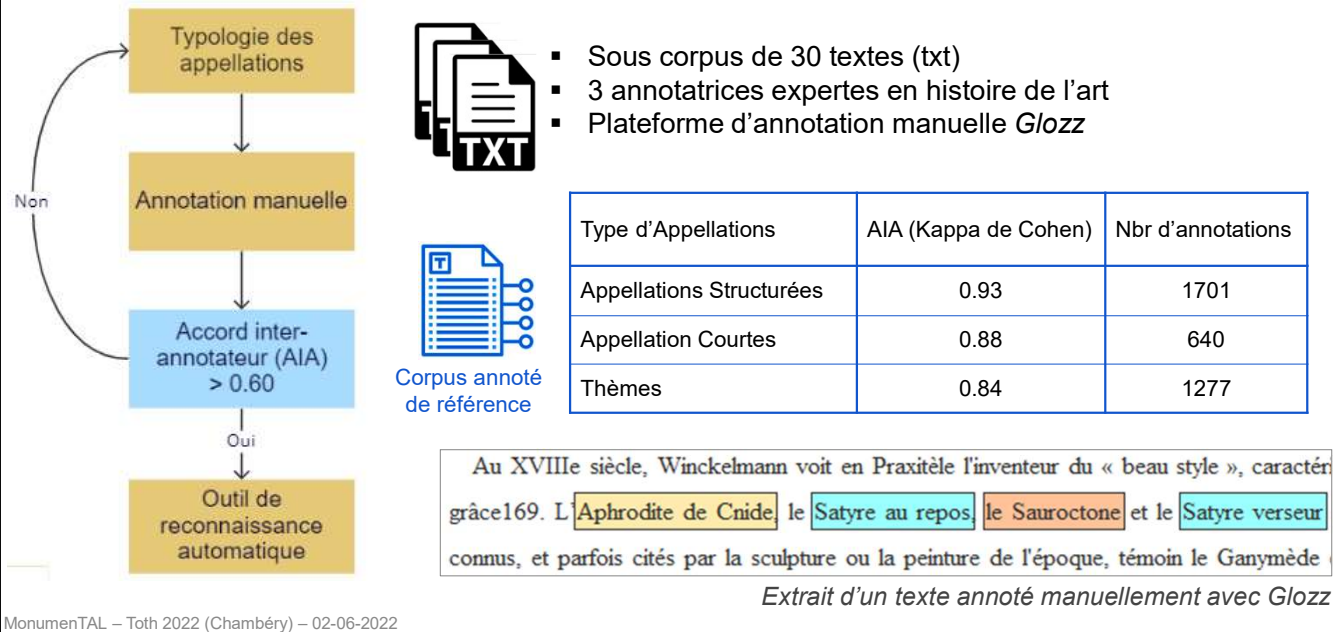
▪ Appellation Courte (AC)

- L'Apollon ; L' Hermès
- Le Sauroctone ; La Cnidienne

MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

- Appellation Structurée (AS) : les plus courantes, suffisent pour reconnaître l'image, l'artefact ou le type iconographique et sont structurées de manière régulière et récurrente.
- Appellation Thématique (TM) : thèmes iconographiques
- Appellation Courte (AC) : mention d'une œuvre uniquement par le personnage représenté.

Annotation manuelle



Une partie du corpus – 30 textes – a été annotée manuellement, selon cette typologie, par les chercheurs en histoire de l'art du projet, en utilisant le logiciel Glozz. L'annotation manuelle a permis de produire un corpus annoté de référence comprenant 1701 Appellations Structurées, 1277 Thèmes et 640 Appellations Courtes.

L'annotation humaine étant un processus d'interprétation, sa validité ne peut pas être réellement évaluée. On se réfère alors, le plus souvent, au calcul d'un accord inter-annotateurs qui permet, entre autre chose, de mesurer la fiabilité et la qualité des annotations produites.

Le Kappa de Coen obtenu, comme le montre le tableau, donne une AIA tout à fait satisfaisante. Quelques mots sur ces résultats :

- L'Appellation Structurée a le meilleur score inter-annotateur car il s'agit de la catégorie pour laquelle les constructions sont les plus identifiables
- La catégorie Thème, qui comprend les descriptions iconographiques, est plus difficile à annoter car elle se confond facilement avec une partie de description complète et détaillée d'une scène d'œuvre d'art figurée. La différenciation nécessite une solide connaissance des expressions traditionnellement employées, et donc du domaine artistique, qui peut varier d'un annotateur à l'autre.

Ce corpus annoté a servi de corpus de référence pour le développement et l'évaluation d'un outil de reconnaissance automatique des appellations d'œuvres antiques fondé sur des méthodes symboliques.

Outil de reconnaissance automatique des appellations d'œuvre antique

- Appellations d'œuvres antiques ≈ Entités nommées (Nouvel et al., 2016)
 - REN
 - les méthodes symboliques (Diez Platas et al., 2020; Collin & Guerraz, 2015)
 - l'apprentissage supervisé de surface (Brandesen et al., 2020)
 - L'apprentissage profond (Brandesen et al., 2021).
 - Méthodes symboliques
- | | |
|--|---|
| <ul style="list-style-type: none">▪ Pourquoi ?<ul style="list-style-type: none">→ optique pluridisciplinaire→ précédents travaux en histoire de l'art et philologie (PRIOUX E. (2011))→ taille restreinte du corpus | <ul style="list-style-type: none">▪ Comment ?<ul style="list-style-type: none">→ patrons morpho-syntaxiques→ dictionnaires électroniques (PERSONNAGE, ARTISTE, COLLECTION, LIEU, etc.)→ outil : Unitex |
|--|---|

MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

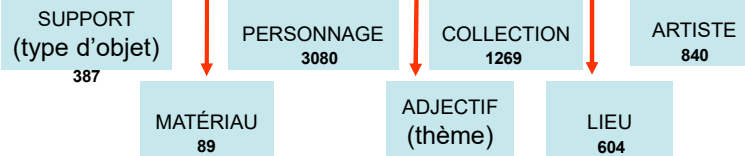
L'annotation des appellations d'œuvres antiques se rapproche de l'annotation des entités nommées dans le sens où les appellations d'œuvres antiques et les entités nommées ont des critères définitoires similaires.

La tâche de reconnaissance des entités nommées exploite des méthodes et techniques variées, dont les méthodes symboliques, qui ont été préférées dans ce projet en raison de la taille restreinte du corpus annoté mais aussi dans une optique de pluridisciplinarité : car ces méthodes permettent de connaître l'état d'avancement de la reconnaissance automatique et de savoir où effectuer les modifications pour en améliorer les résultats ; et cela, par l'ensemble des participants du projet.

Dictionnaires

- statue en marbre de l'Aphrodite de Cnide de Praxitèle
- la Vénus Colonna
- Vénus accroupie

compatible **SKOS**
 ○ prefLabel : *Athéna Médicis*
 ○ altLabel : *Minerve Ingres*
 ○ hiddenLabel : *Athena Medicis*
 Alignement : wikidata



Jeune fille d'Anzio
 (Wikipedia, ph. M.H.Cuesta)

+ ANIMAL (la *Vache de Myron*), THÈME (*Amazonomachie*) et ... appellations des ŒUVRES (575) → la *Fanciulla d'Anzio*

- Import du thésaurus *TheA*: SUPPORT, MATÉRIAU
- Choix d'une sélection à partir des textes (pour les variantes): PERSONNAGE, COLLECTION, ARTISTE, LIEU et ŒUVRES

PERSONNAGE (AC): *l'Apollon* ≠ le *Satyre* / une *Muse* / *Victoire* / le *prêtre* → **individu** ≠ pers. générique/ groupe / allégorie fr / fonction

COLLECTION: notion d'ensemble → personnes physique et morale (collectionneur, collection), édifice. *Cardinal Albani* ≠ *Villa Albani*

SUPPORT: y compris des termes relatifs à la représentation (*image, copie, réplique...*)

→ ajout d'attributs pour différencier et pour scinder en « sous-dictionnaires »

MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

Les dictionnaires utilisés (PERSONNAGE, ARTISTE, COLLECTION, LIEU, etc.) correspondent aux différents composants de l'appellation. Les chiffres indiquent le nombre de termes collectés à ce jour. S'y ajoutent des listes d'adjectifs, voire d'épiclèses pour les personnages. Les épicleses sont des épithètes précisant une spécificité d'une divinité.

Les plus petits vocabulaires ont été importés du thésaurus ThéA (Thésaurus-Antiquité). Pour les listes longues ou susceptibles d'avoir beaucoup de variantes, le vocabulaire a été commencé à partir des textes annotés.

Quelques particularités :

- Les Personnages réunissent des personnages mythologiques et historiques qui font l'objet de la représentation. S'y ajoutent les allégories, les gentilés, les fonctions... L'ensemble fonctionne avec les AS mais seuls les individus à référent unique peuvent être utilisés pour la recherche des AC. Par conséquent des attributs permettent de différencier le type de personnages pour s'adapter aux patrons de reconnaissance et pour créer au besoin des sous-dictionnaires.

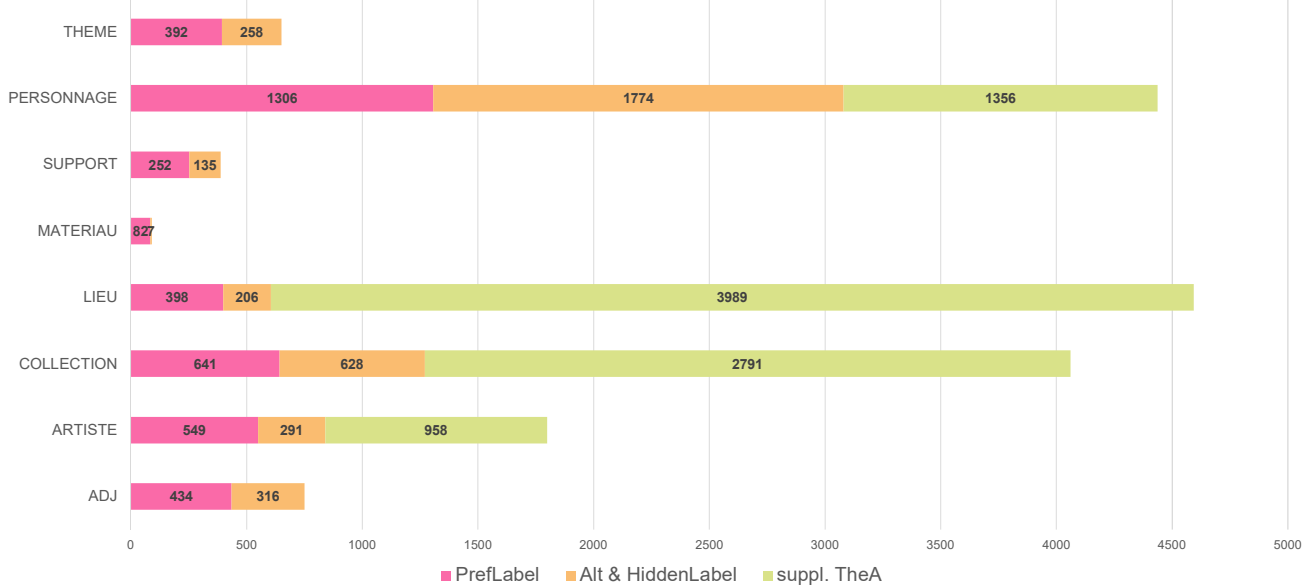
Ce principe est aussi appliqué à :

- COLLECTION qui réunit des personnes physiques et morales (musée) ainsi que des édifices antiques (comme le Parthénon) car en considérant qu'une collection correspond à un ensemble, ici d'objets ou de décors, une formule efficace a été trouvée pour tirer parti de ces composants. Personnes et édifices sont différenciés au sein de ce groupe (attributs).
- Quant à SUPPORT, c'est-à-dire type d'objet, il inclut aussi des termes évoquant la représentation comme image, copie, réplique pour améliorer la reconnaissance.

On retrouve dans cet ensemble de dictionnaires, en cours d'alignement manuel avec Wikidata, et à la structure compatible SKOS (standard utilisé pour les thésaurus numériques), des grandes notions exploitées par plusieurs vocabulaires et ontologies des SHS et du patrimoine culturel : support ou type d'objet, matériau, personne (personnage, artistes...), lieu.

Les dictionnaires en quelques chiffres

nombre de termes par dictionnaire (juin 2022)



MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

10

Histogramme restituant le nombre actuels des termes de nos dictionnaires, en rose les appellations préférées (skos.prefLabel) et en orange les variantes (altLabel et hiddenLabel). Le vert correspond à d'autres termes disponibles dans TheA.

Outil de reconnaissance automatique des appellations d'œuvre antique : résultats

- 40% pour le développement
- 30% pour le test
- 30% pour l'évaluation

| Type d'appellation | Précision | Rappel | F-mesure |
|-----------------------------|-----------|--------|-------------|
| Appellation Structurée (AS) | 0,72 | 0,84 | 0,78 |
| Appellation Courte (AC) | 0,87 | 0,84 | 0,89 |
| Thème (TM) | 0,85 | 0,72 | 0,78 |

Exemple d'AS :

[<AP><PERSO>Apollon</PERSO> <COLL>Borghèse</COLL></AP>](#) , ainsi que nous
[<AP><PERSO>Apollon</PERSO> du <LIEU>Vatican</LIEU></AP>](#) (i) ; {S}
[<AP><PERSO>Apollon</PERSO> <NEW>Dydiméen</NEW><AP>colossal qu'on dû c:](#)
[<AP><PERSO>Apollon</PERSO> <ADJ>Lycien</ADJ> à <LIEU>Argos</LIEU></AP>](#)
[<AP><PERSO>Apollon</PERSO> <NEW>Parnopien</NEW></AP>](#) , partagea , pour

Repérage de nouveaux termes

Exemple de TM :

[<TM>Ganymède enlevé par l'aigle</TM>](#) , placée à
[<TM>Hermès porteur de bélier</TM>](#) , placée sans
[<TM>Héraklès domptant le taureau de Crète</TM>](#)

MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

Les patrons Unitex ont été développés sur 40% des annotations du corpus annoté manuellement, puis testés sur 30% et enfin évalués sur les 30% restants.

L'évaluation de cette reconnaissance automatique a été faite en utilisant des mesures de Précision et de Rappel pour chaque type d'appellation. La mesure de Précision peut être associée à une mesure d'exactitude et celle de Rappel à de l'exhaustivité.

L'Appellation Courte est la catégorie la mieux reconnue avec une F-mesure de 0,89. Ces résultats dépendent des termes contenus dans le dictionnaire des PERSONNAGE et reflètent le choix de rechercher uniquement appellations Courtes avec des noms de personnage bien individualisé tel que l'Apollon.

Les Thèmes, quant à eux, obtiennent une F-mesure de 0,78 due à leur ambiguïté avec les descriptions détaillées de scène, une forme d'expression spécifique des textes sur l'art (comme les descriptions d'œuvres d'art dans un catalogue). Le choix a été fait de préférer la Précision au Rappel. Ainsi, les patrons pour les Thèmes sont les plus stricts possible.

Enfin, concernant les Appellations Structurées, la Précision reflète notre volonté de détecter des appellations en se détachant des limites des dictionnaires. A cette fin, des patrons de reconnaissance plus libres ont été mis en place. A titre d'exemple, dans les exemples d'appellations Structurée "l'Apollon Didyméen" (en graphie erronée) et "l'Apollon Parnopien" qui sont deux appellations reconnues avec un PERSONNAGE suivi d'un Nom Propre (soit une MAJuscule à l'initiale) absent des dictionnaires.

Bilan des ressources produites

- Proposition de la typologie des appellations
- Constitution du corpus de référence annoté en appellations

| Type d'Appellations | AIA (Kappa de Cohen) | Nbr d'annotations |
|-------------------------------|----------------------|-------------------|
| Appellations Structurées (AS) | 0.93 | 1701 |
| Appellation Courtes (AC) | 0.88 | 640 |
| Thèmes (TM) | 0.84 | 1277 |

- Constitution de ressources lexicales (dictionnaires)
 - 9 dictionnaires
 - 7724 termes (actuellement)
- Développement du module de l'annotation automatique
 - Méthodes symboliques
 - F-mesure : AS = 0,78 ; AC = 0,89 ; TM = 0,78

Perspectives histoire de l'art : référentiel des œuvres et thèmes figurés

- Thésaurus ŒUVRE des appellations d'œuvres antiques
→ Exploitation pour un référentiel reliant « objet + image + texte »
(corpus numérique d'objets LIMC-France)

Carthage (photo wikipedia)



(wikipedia, ph. Roccz, 2006)

Pausanias (V, 17, 3):

« Quelque temps après, on fit d'autres consécration à l'Héraion, un **Hermès de marbre portant Dionysos enfant, œuvre de Praxitèle** »

MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

- Gérer les appellations multiples
- Proposer une appellation préférée

- **Hermès d'Olympie**
- **Hermès de Praxitèle**
- **Hermès portant Dionysos enfant**

→ « **Hermès portant Dionysos enfant d'Olympie** » ?



- « **Ganymède enlevé par Zeus métamorphosé en aigle** »
- **Ganymède enlevé par l'aigle**
- **Ganymède et l'aigle**
- **Enlèvement de Ganymède**
- **Rapt de Ganymède**

Le Thésaurus ŒUVRE, réutilisable par la communauté, que ce travail permet de créer actuellement servira à renforcer le référentiel reliant « objet + image + texte », évolution du Corpus d'objet LIMC (LIMC-France)

Ce thésaurus permet de rassembler les différentes variantes d'une appellation d'œuvre et de proposer une appellation préférée :

Voici 2 cas d'exemples :

* **Hermès** : pour une appellation préférée précise : éviter « Hermès de Praxitèle » pour ne pas ajouter du sens en affirmant que c'est un original de cet artiste (rapproché d'un passage de Pausanias, l'attribution est discutée, incertaine) ; éviter Hermès d'Olympie, trop ambigu (plus d'un Hermès à Olympie et problème du mot *hermès* qui désigne un type d'objet (pilier hermaïque) ; quand au thème, bien qu'il semble convenir qu'il s'agit d'un exemplaire statuaire unique, le thème peut être figuré ailleurs. On pourrait par exemple proposer « Hermès portant Dionysos enfant d'Olympie »,

* Le cas de **Ganymède** illustre la simplification au fil du temps (du plus détaillé et explicatif au plus réduit) et la diversité des variantes (y compris pour la peinture néo-classique ou classique française : titres souvent construit avec nom + personnage : Enlèvement de Ganymède...); et polysémie de « rapt » (mot aussi employé pour des actions érotiques et contraintes). **Ganymède et l'aigle** est une appellation courante pour les images antiques mais la « paire » (Ganymède + et + aigle) et l'énumération génèrent du bruit lors de la reconnaissance automatique,

Perspectives : référentiel des œuvres et thèmes figurés

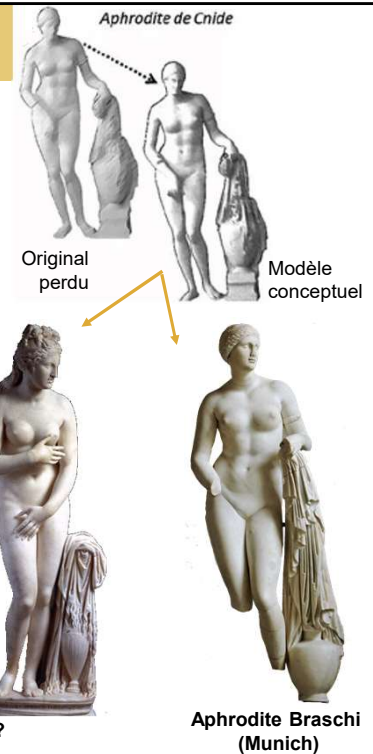
- Différencier les œuvres originales (réfèrent unique) et les types iconographiques (phénomène des copies romaines)

l'Aphrodite de Cnide (par Praxitèle) / La Cnidienne...
 ≠ une *Aphrodite de Cnide* (type générique)

**Reproduction de
 l'original de
 l'Aphrodite de
 Cnide sur une
 monnaie de Cnide
 (Caracalla)
 (BnF)**



- original
- conceptuel
- copies
- typologies



MonumenTAL – Toth 2022 (Chambéry) – 02-06-2022

Il est important, en histoire de l'art et notamment pour le corpus numérique LIMC, de différencier les appellations d'œuvres originales, les types iconographiques et les répliques, comme ici pour l'Aphrodite de Cnide, œuvre perdue mais connue par des copies d'époque romaine plus ou moins fidèles qui elles-mêmes peuvent devenir des types de référence (par ex. l'Aphrodite Braschi comme type probablement proche de l'original ; celle du Capitole, plus pudique avec ses deux mains devant son corps).

La « Vénus du Capitole », telle qu'on l'appelle depuis longtemps, est souvent renommée « Aphrodite du Capitole » par des spécialistes contemporains qui préfèrent mettre en avant son lien avec la statue de Praxitèle et avec le fait probable que l'artiste ou le commanditaire avait ce lien en tête. D'autres « Vénus » sont renommées pareillement.

Perspectives TAL

- Analyse de l'évolution des appellations d'œuvres

- Application de la méthodologie sur :
 - corpus en latin
 - corpus en grec ancien

- Expérimentation sur des textes d'histoire de l'art (français) traitant de différents courants et périodes artistiques

Ce phénomène de substitution d'un nom de personnage latin par son équivalent grec, et inversement, fait l'objet d'analyses en lien avec l'évolution du discours sur l'art et l'iconographie antiques en fonction de l'époque et du contexte de publication et du public visé. C'est analyses ont été commencées dans MonumentAL.

La méthodologie présentée a porté ses fruits sur le français et elle est à présent en cours d'application sur deux corpus textuels, l'un en latin, l'autre en grec ancien.

Également expérimentée sur des textes d'histoire de l'art en français traitant de différents courants et périodes artistiques, la reconnaissance a montré des résultats prometteurs pour AS

- Nougaret, Pierre-Jean-Baptiste, *Anecdotes des beaux-arts. Tome 2, 1776-1780* : art européen
- Geffroy, *Les musées d'Europe. 7, La sculpture au Louvre, 1908* : arts classique, égyptien, proche-oriental ; asiatique...