DataCatalogue : rétro-structuration automatique des catalogues de vente

Hugo Scheithauer, Sarah Bénière, Jean-Philippe Moreux, Laurent Romary

DataCatalogue : rétro-structuration automatique des catalogues de vente

Hugo Scheithauer[1], **Sarah Bénière**[1], **Jean-Philippe Moreux**[2], **Laurent Romary**[1]

[1] ALMAnaCH - Automatic Language Modelling and ANAlysis & Computational Humanities, Inria Paris
[2] Bibliothèque Nationale de France

Webinaire Culture-Inria - 29-11-23

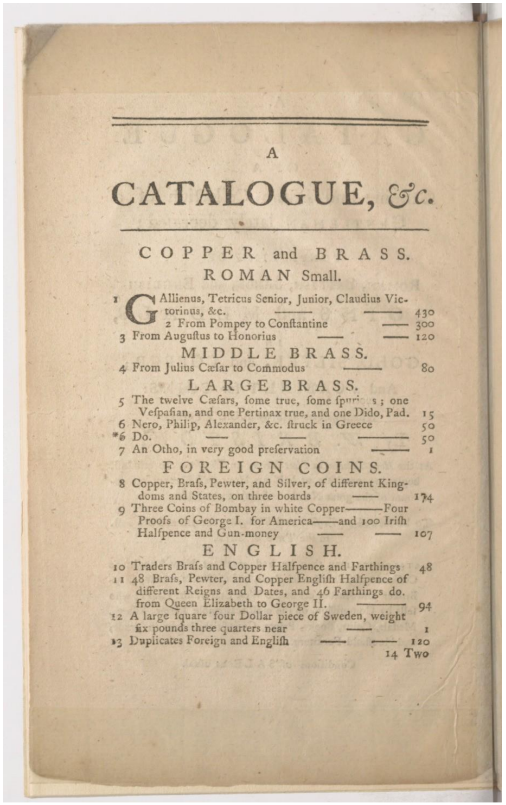
BnF Inria

Institut national d'histoire de l'art INHA

DataCatalogue : objectifs du projet

- Passer d'une **numérisation** à une **base de données textuelle et requêtable**.
- **Segmenter** les catalogues de vente et attribuer à chaque niveau d'information une **étiquette** : entrées de catalogues, numéro des entrées, description des objets, matériaux, sommes monétaires, etc.
- Produire, à partir des zones segmentées, un **encodage XML-TEI** des catalogues de vente.
- Mettre à disposition des chercheurs le corpus encodé dans une **interface de publication** permettant de **requêter** sur les zones segmentées.

Les catalogues de vente : homogénéité de mise en page



A

CATALOGUE, &c.

COPPER and BRASS.

ROMAN Small.

1 Gallienus, Tetricus Senior, Junior, Claudius Victorinus, &c. —— —— —— 430
2 From Pompey to Conſtantine —— —— 300
3 From Auguſtus to Honorius —— —— 120

MIDDLE BRASS.

4 From Julius Cæſar to Commodus —— —— 80

LARGE BRASS.

5 The twelve Cæſars, ſome true, ſome ſpuri[illegible] 1; one Veſpaſian, and one Pertinax true, and one Dido, Pad. 15
6 Nero, Philip, Alexander, &c. ſtruck in Greece 50
*6 Do. —— —— —— 50
7 An Otho, in very good preſervation —— —— 1

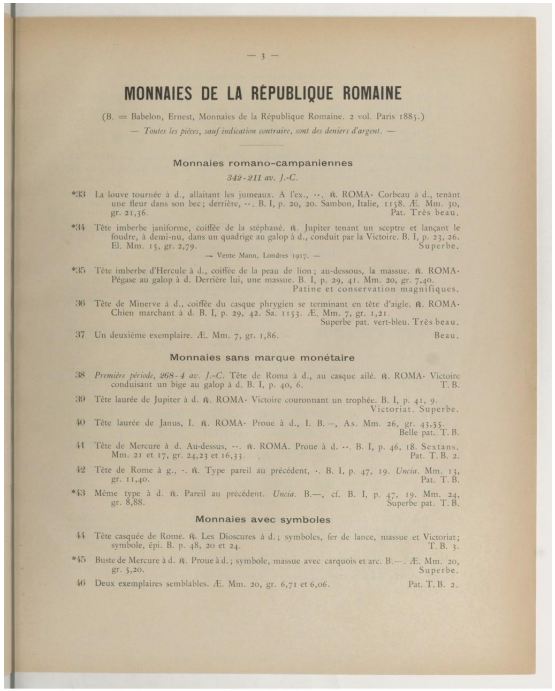
FOREIGN COINS.

8 Copper, Braſs, Pewter, and Silver, of different Kingdoms and States, on three boards —— —— 174
9 Three Coins of Bombay in white Copper——Four Proofs of George I. for America——and 100 Iriſh Halfpence and Gun-money —— —— 107

ENGLISH.

10 Traders Braſs and Copper Halfpence and Farthings 48
11 48 Braſs, Pewter, and Copper Engliſh Halfpence of different Reigns and Dates, and 46 Farthings do. from Queen Elizabeth to George II. —— —— 94
12 A large ſquare four Dollar piece of Sweden, weight ſix pounds three quarters near —— —— 1
13 Duplicates Foreign and Engliſh —— —— —— 120
14 Two

Whiston Bristow, 1762.



— 3 —

MONNAIES DE LA RÉPUBLIQUE ROMAINE

(B. = Babelon, Ernest, Monnaies de la République Romaine. 2 vol. Paris 1885.)

— *Toutes les pièces, sauf indication contraire, sont des deniers d'argent.* —

Monnaies romano-campaniennes

349-211 av. J.-C.

*33 La louve tournée à d., allaitant les jumeaux. A l'ex., ··. ℞. ROMA- Corbeau à d., tenant une fleur dans son bec; derrière, ··. B. I, p. 20, 20. Sambon, Italie, 1158. Æ. Mm. 30, gr. 21,36. Pat. Très beau.

*34 Tête imberbe janiforme, coiffée de la stéphané. ℞. Jupiter tenant un sceptre et lançant le foudre, à demi-nu, dans un quadrige au galop à d., conduit par la Victoire. B. I, p. 25, 26. El. Mm. 15, gr. 2,79. Superbe.

→ Vente Mann, Londres 1917. —

*35 Tête imberbe d'Hercule à d., coiffée de la peau de lion; au-dessous, la massue. ℞. ROMA- Pégase au galop à d. Derrière lui, une massue. B. I, p. 29, 41. Mm. 20, gr. 7,40. Patine et conservation magnifiques.

36 Tête de Minerve à d., coiffée du casque phrygien se terminant en tête d'aigle. ℞. ROMA- Chien marchant à d. B. I, p. 29, 42. Su. 1153. Æ. Mm. 7, gr. 1,21. Superbe pat. vert-bleu. Très beau.

37 Un deuxième exemplaire. Æ. Mm. 7, gr. 1,86. Beau.

Monnaies sans marque monétaire

*38 *Première période, 268-4 av. J.-C.* Tête de Roma à d., au casque ailé. ℞. ROMA- Victoire conduisant un bige au galop à d. B. I, p. 40, 6. T. B.

39 Tête laurée de Jupiter à d. ℞. ROMA- Victoire couronnant un trophée. B. I, p. 41, 9. Victoriat. Superbe.

40 Tête laurée de Janus, I. ℞. ROMA- Proue à d., I. B.—, As. Mm. 26, gr. 43,55. Belle pat. T. B.

41 Tête de Mercure à d. Au-dessous, ··. ℞. ROMA. Proue à d. ··. B. I, p. 46, 18. Sextans. Mm. 21 et 17, gr. 24,23 et 16,33. Pat. T. B. 2.

42 Tête de Rome à g., ·. ℞. Type pareil au précédent, ·. B. I, p. 47, 19. *Uncia*. Mm. 13, gr. 11,40. Pat. T. B.

*43 Même type à d. ℞. Pareil au précédent. *Uncia*. B.—, cf. B. I, p. 47, 19. Mm. 24, gr. 8,88. Superbe pat. T. B.

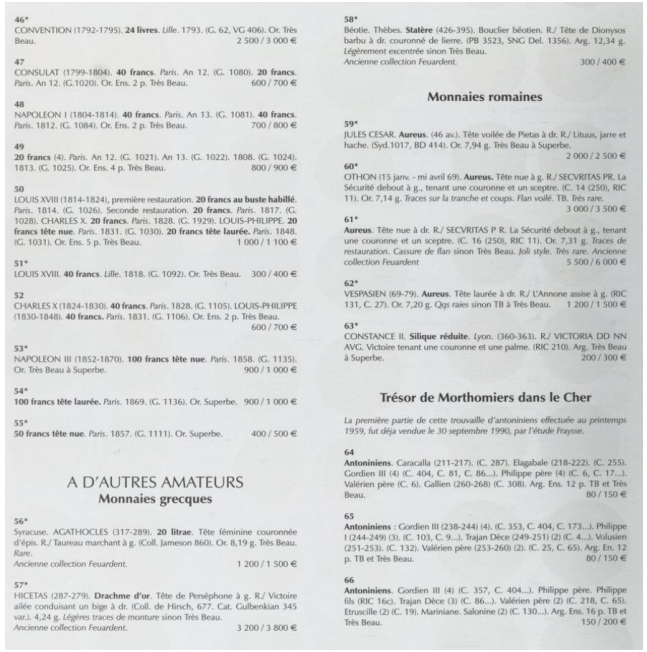
Monnaies avec symboles

44 Tête casquée de Rome. ℞. Les Dioscures à d.; symboles, fer de lance, massue et Victoriat; symbole, épi. B. p. 48, 20 et 24. T. B. 3.

*45 Buste de Mercure à d. ℞. Proue à d.; symbole, massue avec carquois et arc. B.—; Æ. Mm. 20, gr. 5,20. Superbe.

46 Deux exemplaires semblables. Æ. Mm. 20, gr. 6,71 et 6,06. Pat. T. B. 2.

Lucien Naville, 1924.



46*
CONVENTION (1792-1795). **24 livres**. *Lille*. 1793. (G. 62, VG 406). Or. Très Beau. 2 500 / 3 000 €

47
CONSULAT (1799-1804). **40 francs**. *Paris*. An 12. (G. 1080). **20 francs**. *Paris*. An 12. (G.1020). Or. Ens. 2 p. Très Beau. 600 / 700 €

48
NAPOLEON I (1804-1814). **40 francs**. *Paris*. An 13. (G. 1081). **40 francs**. *Paris*. 1812. (G. 1084). Or. Ens. 2 p. Très Beau. 700 / 800 €

49
20 francs (4). *Paris*. An 12. (G. 1021). An 13. (G. 1022). 1808. (G. 1024). 1813. (G. 1025). Or. Ens. 4 p. Très Beau. 800 / 900 €

50
LOUIS XVIII (1814-1824), première restauration. **20 francs au buste habillé**. *Paris*. 1814. (G. 1026). Seconde restauration. **20 francs**. *Paris*. 1817. (G. 1028). CHARLES X. **20 francs**. *Paris*. 1828. (G. 1929). LOUIS-PHILIPPE. **20 francs tête nue**. *Paris*. 1831. (G. 1030). **20 francs tête laurée**. *Paris*. 1848. (G. 1031). Or. Ens. 5 p. Très Beau. 1 000 / 1 100 €

51*
LOUIS XVIII. **40 francs**. *Lille*. 1818. (G. 1092). Or. Très Beau. 300 / 400 €

52
CHARLES X (1824-1830). **40 francs**. *Paris*. 1828. (G. 1105). LOUIS-PHILIPPE (1830-1848). **40 francs.** *Paris*. 1831. (G. 1106). Or. Ens. 2 p. Très Beau. 600 / 700 €

53*
NAPOLEON III (1852-1870). **100 francs tête nue**. *Paris*. 1858. (G. 1135). Or. Très Beau à Superbe. 900 / 1 000 €

54*
100 francs tête laurée. *Paris*. 1869. (G. 1136). Or. Superbe. 900 / 1 000 €

55*
50 francs tête nue. *Paris*. 1857. (G. 1111). Or. Superbe. 400 / 500 €

A D'AUTRES AMATEURS
Monnaies grecques

56*
Syracuse. ACATHOCLES (317-289). **20 litrae**. Tête féminine couronnée d'épis. R./ Taureau marchant à g. (Coll. Jameson 860). Or. 8,19 g. Très Beau. *Rare.*
Ancienne collection Feuardent. 1 200 / 1 500 €

57*
HICETAS (287-279). **Drachme d'or**. Tête de Perséphone à g. R./ Victoire ailée conduisant un bige à dr. (Coll. de Hirsch, 677. Cat. Gulbenkian 345 var.). 4,24 g. *Légères traces de monture* sinon Très Beau.
Ancienne collection Feuardent. 3 200 / 3 800 €

58*
Béotie. Thèbes. **Statère** (426-395). Bouclier béotien. R./ Tête de Dionysos barbu à dr. couronné de lierre. (PB 3523, SNG Del. 1356). Arg. 12,34 g. *Légèrement excentrée* sinon Très Beau.
Ancienne collection Feuardent. 300 / 400 €

Monnaies romaines

59*
JULES CESAR. **Aureus**. (46 av.). Tête voilée de Pietas à dr. R./ Lituus, jarre et hache. (Syd.1017, BD 414). Or. 7,94 g. Très Beau à Superbe. 2 000 / 2 500 €

60*
OTHON (15 janv. - mi avril 69). **Aureus.** Tête nue à g. R./ SECVRITAS PR. La Sécurité debout à g., tenant une couronne et un sceptre. (C. 14 (250), RIC 11). Or. 7,14 g. *Traces sur la tranche et coups. Flan voilé.* TB. *Très rare.* 3 000 / 3 500 €

61*
Aureus. Tête nue à dr. R./ SECVRITAS P R. La Sécurité debout à g., tenant une couronne et un sceptre. (C. 16 (250), RIC 11). Or. 7,31 g. *Traces de restauration. Cassure de flan* sinon Très Beau. *Joli style. Très rare. Ancienne collection Feuardent* 5 500 / 6 000 €

62*
VESPASIEN (69-79). **Aureus**. Tête laurée à dr. R./ L'Annone assise à g. (RIC 131, C. 27). Or. 7,20 g. *Qqs raies* sinon TB à Très Beau. 1 200 / 1 500 €

63*
CONSTANCE II. **Silique réduite**. *Lyon*. (360-363). R./ VICTORIA DD NN AVG. Victoire tenant une couronne et une palme. (RIC 210). Arg. Très Beau à Superbe. 200 / 300 €

Trésor de Morthomiers dans le Cher

La première partie de cette trouvaille d'antoniniens effectuée au printemps 1959, fut déjà vendue le 30 septembre 1990, par l'étude Fraysse.

64
Antoniniens. Caracalla (211-217). (C. 287). Elagabale (218-222). (C. 255). Gordien III (4) (C. 404, C. 81, C. 86...). Philippe père (4) (C. 6, C. 17...). Valérien père (C. 6). Gallien (260-268) (C. 308). Arg. Ens. 12 p. TB et Très Beau. 80 / 150 €

65
Antoniniens : Gordien III (238-244) (4). (C. 353, C. 404, C. 173...). Philippe I (244-249) (3). (C. 103, C. 9...). Trajan Dèce (249-251) (2) (C. 4...). Volusien (251-253). (C. 132). Valérien père (253-260) (2). (C. 25, C. 65). Arg. Ens. 12 p. TB et Très Beau. 80 / 150 €

66
Antoniniens. Gordien III (4) (C. 357, C. 404...). Philippe père. Philippe fils (RIC 16c). Trajan Dèce (3) (C. 86...). Valérien père (2) (C. 218, C. 65). Etruscille (2) (C. 19). Mariniane. Salonine (2) (C. 130...). Arg. Ens. 16 p. TB et Très Beau. 150 / 200 €

Fraysse & Associés, 2011.

Les catalogues de vente : homogénéité de mise en page (vraiment ?)

COLIN (Gustave-Henri)

68 — Gitana.

Elle est debout, relevant de sa main droite sa jupe bleue brodée de fleurs; elle découvre sa jambe aux espadrilles blanches lacées sur un bas de couleur cachou. Son torse est enveloppé d'un châle bariolé à franges. Elle tourne légèrement la tête vers sa droite, une fleur rouge amuse ses cheveux châtains. Elle est d'une beauté calme, avec, dans ses yeux noirs, une flamme. De sa main gauche, elle tient appuyée contre le sol, sa guitare. Près d'elle, un gamin déscheni sur un fond de mur blanc, marqué à gauche par une porte verte.

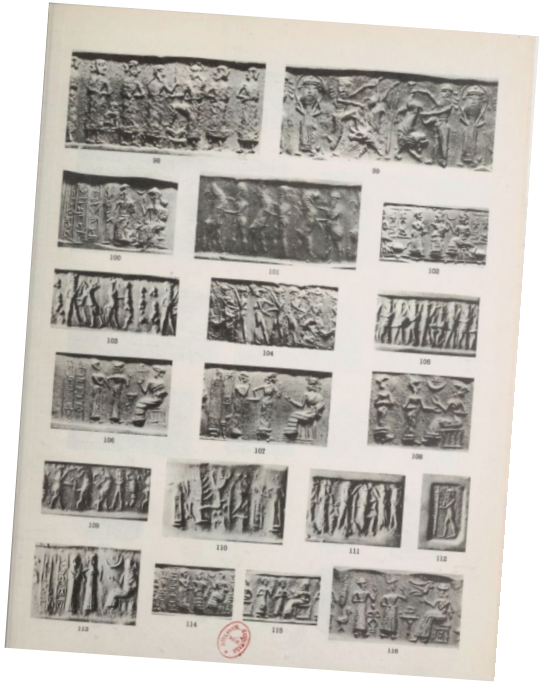
Signé à gauche, en bas : *Gustave Colin.*

Toile. Haut., 2 m. 20; larg., 1 m. 41.

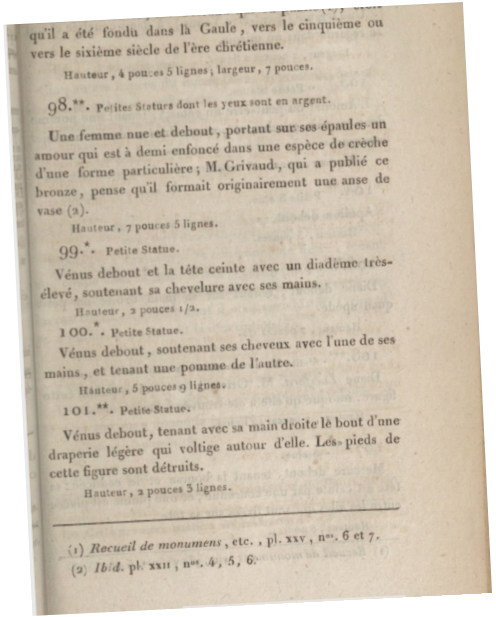
Salon de 1904. Société nationale des Beaux-Arts.

68 — COLIN (Gustave) Gitana.

Lair-Dubreuil, 1919 (INHA).



Bourgey, 1962 (BnF).



qu'il a été fondu dans la Gaule, vers le cinquième ou vers le sixième siècle de l'ère chrétienne.

Hauteur, 4 pouces 5 lignes; largeur, 7 pouces.

98.**. Petites Statues dont les yeux sont en argent.

Une femme nue et debout, portant sur ses épaules un amour qui est à demi enfoncé dans une espèce de crèche d'une forme particulière; M. Grivaud, qui a publié ce bronze, pense qu'il formait originairement une anse de vase (2).

Hauteur, 7 pouces 5 lignes.

99.*. Petite Statue.

Vénus debout et la tête ceinte avec un diadème très-élevé, soutenant sa chevelure avec ses mains.

Hauteur, 2 pouces 1/2.

100.*. Petite Statue.

Vénus debout, soutenant ses cheveux avec l'une de ses mains, et tenant une pomme de l'autre.

Hauteur, 5 pouces 9 lignes.

101.**. Petite Statue.

Vénus debout, tenant avec sa main droite le bout d'une draperie légère qui voltige autour d'elle. Les pieds de cette figure sont détruits.

Hauteur, 2 pouces 5 lignes.

(1) *Recueil de monumens*, etc., pl. xxv, nos. 6 et 7.
(2) *Ibid.* pl. xxii, nos. 4, 5, 6.

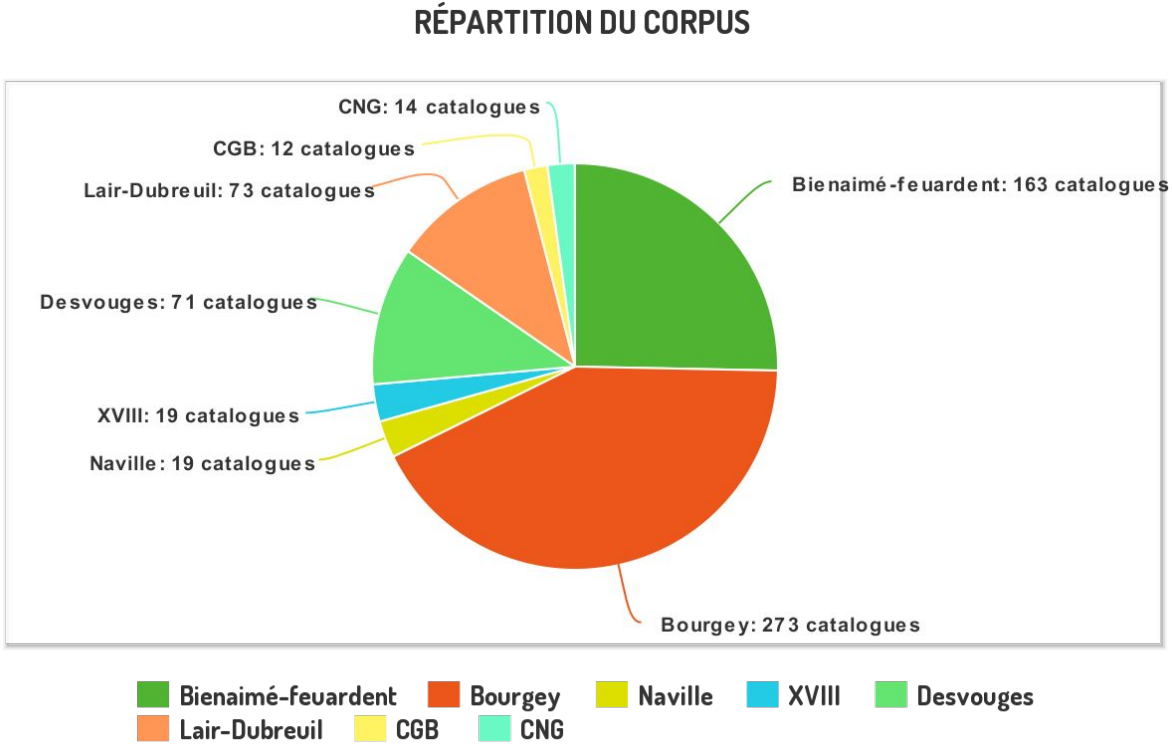
Dubois, 1820 (BnF).

Un premier échantillonnage des catalogues de vente de la BnF et de l'INHA : 4 siècles de contenus structurés

Siècles représentés : XVIIIe, XIXe, XXe, XXIe.

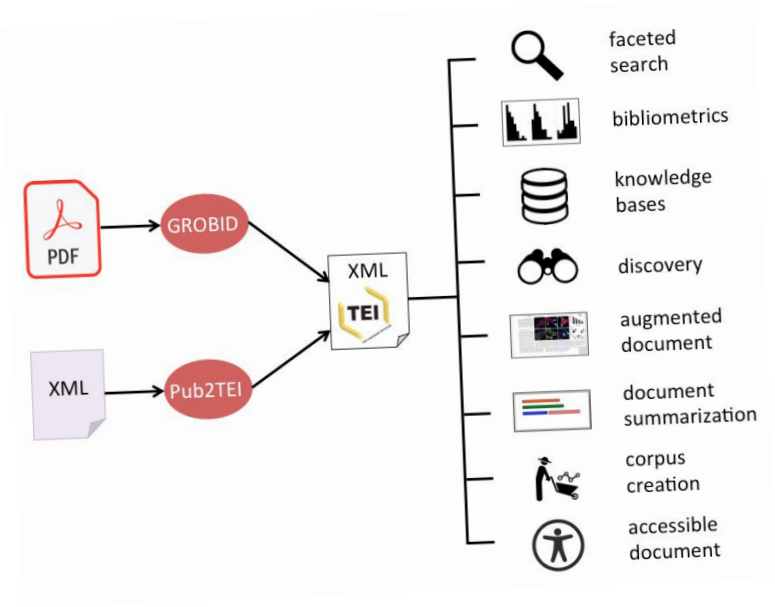
Langues : ~95% en français.

Types de vente : numismatique, livres, antiquités, objets d'art, objets de luxe

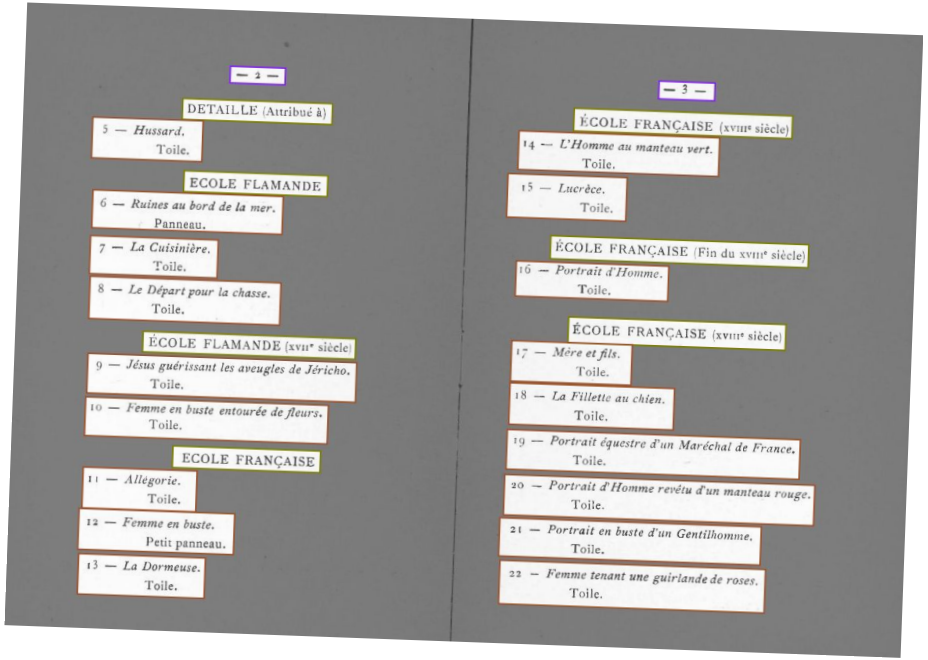


Quelles technologies pour gérer l'hétérogénéité de contenu et la diachronie des catalogues de vente ?

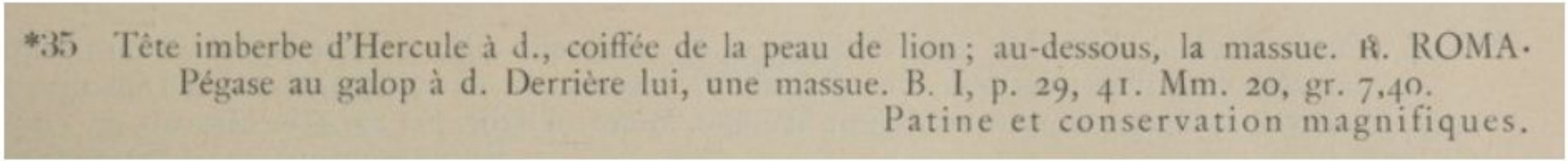
GROBID (GeneRation of BIbliographic Data)



Détection automatique d'objet



Quelles technologies pour gérer l'hétérogénéité de contenu et la diachronie des catalogues de vente ?



*35 Tête imberbe d'Hercule à d., coiffée de la peau de lion ; au-dessous, la massue. ℞. ROMA· Pégase au galop à d. Derrière lui, une massue. B. I, p. 29, 41. Mm. 20, gr. 7,40. Patine et conservation magnifiques.

numéro d'items

*35

description de l'objet mis en vente

Tête imberbe d'Hercule à d., coiffée de la peau de lion ; au-dessous, la massue. ℞. ROMA· Pégase au galop à d. Derrière lui, une massue.

informations bibliographiques

B. I, p. 29, 41.

quantités et mesures

Mm. 20, gr. 7,40.

observations de conservation

Patine et conservation magnifiques.

Quelles technologies pour gérer l'hétérogénéité de contenu et la diachronie des catalogues de vente ?

147 — Grande miniature ovale : portrait d'homme à cheveux blancs avec manteau de fourrure ; signée : *Eudelot*. Cadre doré.

Lair-Dubreuil, 1926 (INHA).

418 — F. 103. Peint. r. et bl. — Vulci. — Ext. *Les vaisseaux d'Ulysse passant devant les sirènes.* On voit deux vaisseaux voguant de front à pleines voiles. La proue est ornée d'une tête de sanglier, et à la poupe est une tête de cygne. Sur chaque vaisseau est représenté un homme debout et drapé, placé à la proue; un rameur est à l'arrière du navire; vers l'anse est une sirène placée sur un rocher, et retournant la tête vers les vaisseaux; plus loin est un dauphin. NIKOSΘENES EΠOIE... *Nicosthènes a fait.*

R. Le même sujet, si ce n'est que les proues sont ornées chacune d'un œil, et qu'aucun personnage ne paraît sur le devant des vaisseaux. Les rameurs sont dans la même position; seulement l'un d'eux lève la main comme surpris à l'aspect de la sirène placée sur un rocher, comme sur l'autre face, et avec un poisson plus loin.

Les voiles des vaisseaux sont peintes en blanc.

Diam., 10 pouces 3 lignes. Haut., 4 pouces.

Rollin, 1836 (BnF).

1	English, Iriſh, Manx, American, Swediſh copper and leaden caſts	45	0. 3. 0
2	Ditto	65	0. 6. 6
3	Ancient ſeals and caſts	22	1. 11. 6
4	Four drawers, middle roman braſs	140	0. 14. 0
5	Ditto, large and ſmall, from Cæſar to Conſtantine	140	1. 14. 0

Langford, 1772 (BnF).

86 — **Drachme.** Mêmes types variés.- (166-88). **Drachme.** Tête radiée à dr. R/. Rose.- **CARIE.** HALICARNASSE. (1er siècle av. J.-C.). **Obole.** Tête d'Hélios de face. R/. Tête d'Athéna à dr. (BMC. 153 sq. var., 247 ; P. 2610 var.). 2g71, 2g74, 0g89. Ens. 3 p. T.B. et Très Beau. *1.000/1.200*

87 — **ILES DE CARIE.** RHODES. (166-88). **Bronze.** Tête radiée d'Hélios à dr. R/. Rose avec 2 branches. (BMC. 324 sq. - SNG. von Aulock 2835). Br. lg38. Patine brune. Très Beau. *1.000/1.500*

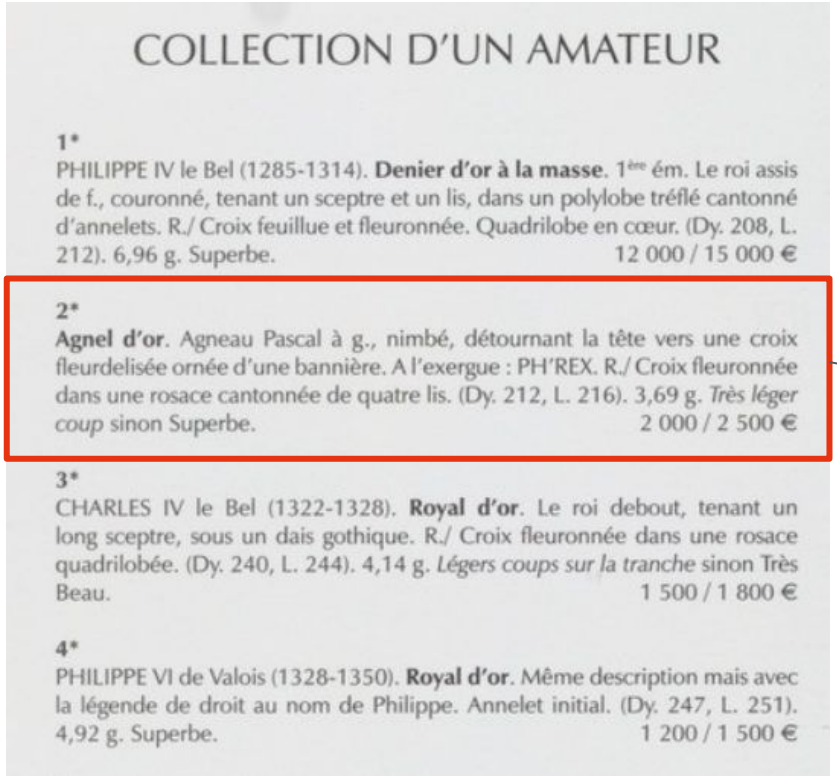
Bourgey, 1995 (BnF).

Standardiser la structuration des catalogues de vente

- Standard XML pour l'**édition de texte**, permet de rendre **lisible** et **compréhensible** un texte par un ordinateur.
- **Structuration** des catalogues en XML TEI, basé sur des **éléments existants** et standardisés, ainsi que sur des **nouveaux éléments** créés pour modéliser précisément l'objet catalogue, notamment la **notice**. Une notice structurée doit rendre compte de la matérialité de l'objet vendu.
- Une **notice** de catalogue décrit un objet matériel mis à la vente selon un ensemble de normes appartenant au monde de la **vente aux enchères**, ainsi que de la **discipline** dont provient l'objet (numismatique, archéologie, etc.).
- Le module TEI DataCatalogue se base sur les travaux de modélisation du projet **Artl@s** et du projet **Ledoux**, et cherche à **affiner** la **représentation des catalogues**.

- Simon Gabay, Barbara Topalov, Caroline Corbières, Lucie Rondeau Du Noyer, Béatrice Joyeux-Prunel, et al.. Automating Artl@s – extracting data from exhibition catalogues. EADH 2021 - *Second International Conference of the European Association for Digital Humanities*, Sep 2021, Krasnoyarsk, Russia. ⟨hal-03331838⟩
- Emmanuel Chateau Dutier, Caroline Corbières. A broader <object> content model for art history. Next Gen TEI, 2021 - TEI Conference and Members' Meeting, Oct 2021, Virtual, United States. ⟨hal-03654979⟩

D’une image à une base de données XML



COLLECTION D’UN AMATEUR

1*
PHILIPPE IV le Bel (1285-1314). **Denier d’or à la masse**. 1ère ém. Le roi assis de f., couronné, tenant un sceptre et un lis, dans un polylobe tréflé cantonné d’annelets. R./ Croix feuillue et fleuronnée. Quadrilobe en cœur. (Dy. 208, L. 212). 6,96 g. Superbe. 12 000 / 15 000 €

2*
Agnel d’or. Agneau Pascal à g., nimbé, détournant la tête vers une croix fleurdelisée ornée d’une bannière. A l’exergue : PH’REX. R./ Croix fleuronnée dans une rosace cantonnée de quatre lis. (Dy. 212, L. 216). 3,69 g. *Très léger coup* sinon Superbe. 2 000 / 2 500 €

3*
CHARLES IV le Bel (1322-1328). **Royal d’or**. Le roi debout, tenant un long sceptre, sous un dais gothique. R./ Croix fleuronnée dans une rosace quadrilobée. (Dy. 240, L. 244). 4,14 g. *Légers coups sur la tranche* sinon Très Beau. 1 500 / 1 800 €

4*
PHILIPPE VI de Valois (1328-1350). **Royal d’or**. Même description mais avec la légende de droit au nom de Philippe. Annelet initial. (Dy. 247, L. 251). 4,92 g. Superbe. 1 200 / 1 500 €

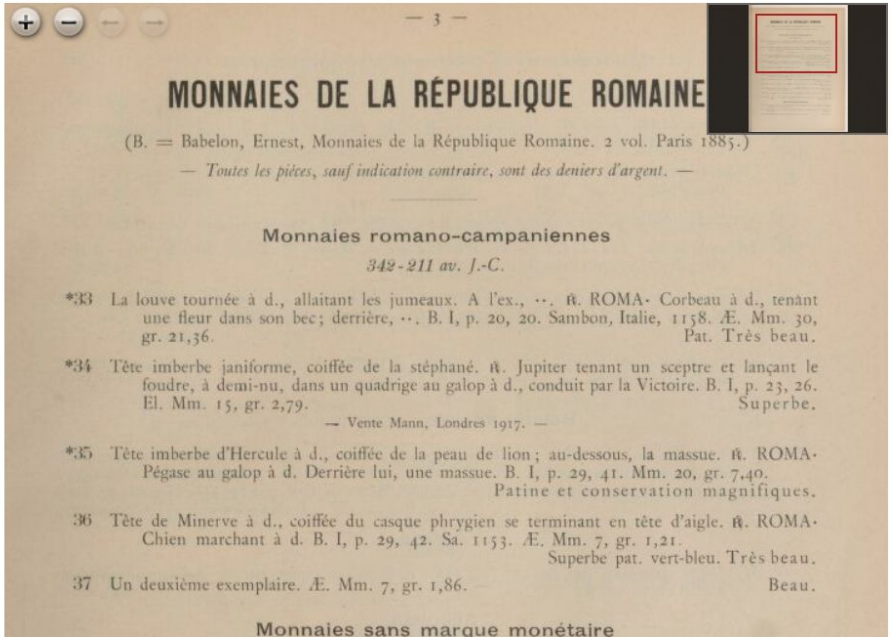
```xml
<catalogueEntry>
    <catalogueDesc>
        <head>Collection d'un amateur</head>
    </catalogueDesc>
    <!-- ... -->
    <catalogueItem>
        <altIdentifier>
            <idno>2</idno>
        </altIdentifier>
        <metamark>*</metamark>
        <objectDesc>
            <supportDesc>
                <support>Agnel d'or.</support>
            </supportDesc>
        </objectDesc>
        <decoDesc>
            <ab>Agneau Pascal à g., nimbé, détournant la tête vers une
                croix fleurdelisée ornée d'une bannière. A l'exergue:
                PH'REX R./ Croix fleuronnée dans une rosace cantonnée de quatre
                lis.</ab>
        </decoDesc>
        <objectDesc>
            <supportDesc>
                <support>(<measure>Dy. 212</measure>,
                    <measure>L. 216</measure>).
                    <measure>3,69 g.</measure></support>
                <condition>Très léger coup sinon Superbe<p>.</p></condition>
            </supportDesc>
        </objectDesc>
        <num type="currency">2000 / 2500 euros</num>
    </catalogueItem>
    <!-- ... -->
</catalogueEntry>
```

Publier les catalogues de vente structurés

- Plateforme de **publication de fichier TEI**
- *Open source*
- Permet d'**indexer** un corpus, de le **visualiser**, et de le rendre **requêtable**
- Interface entièrement **personnalisable**
- Expertise au sein de l'équipe ALMAnaCH, Inria

tei Publisher Start Documentation Actualités Télécharger Fonctions avancées rechercher... Langue Français Connecté en tant que hscheith

AJOUTER UNE VUE



— 3 —

MONNAIES DE LA RÉPUBLIQUE ROMAINE

(B. = Babelon, Ernest, Monnaies de la République Romaine. 2 vol. Paris 1885.)

— *Toutes les pièces, sauf indication contraire, sont des deniers d'argent.* —

Monnaies romano-campaniennes

342-211 av. J.-C.

*33 La louve tournée à d., allaitant les jumeaux. A l'ex., ··. ℞. ROMA· Corbeau à d., tenant une fleur dans son bec; derrière, ··. B. I, p. 20, 20. Sambon, Italie, 1158. Æ. Mm. 30, gr. 21,36. Pat. Très beau.

*34 Tête imberbe janiforme, coiffée de la stéphané. ℞. Jupiter tenant un sceptre et lançant le foudre, à demi-nu, dans un quadrige au galop à d., conduit par la Victoire. B. I, p. 23, 26. El. Mm. 15, gr. 2,79. Superbe.

— Vente Mann, Londres 1917. —

*35 Tête imberbe d'Hercule à d., coiffée de la peau de lion; au-dessous, la massue. ℞. ROMA· Pégase au galop à d. Derrière lui, une massue. B. I, p. 29, 41. Mm. 20, gr. 7,40. Patine et conservation magnifiques.

36 Tête de Minerve à d., coiffée du casque phrygien se terminant en tête d'aigle. ℞. ROMA· Chien marchant à d. B. I, p. 29, 42. Sa. 1153. Æ. Mm. 7, gr. 1,21. Superbe pat. vert-bleu. Très beau.

37 Un deuxième exemplaire. Æ. Mm. 7, gr. 1,86. Beau.

Monnaies sans marque monétaire

MONNAIES DE LA REPUBLIQUE ROMAINE

(B. = Babelon, Ernest, Monnaies de la République Romaine. 2 vol. Paris 1885.) - Toutes les pièces, sauf indication contraire, sont des deniers d'argent. -

Monnaies romano-campaniennes

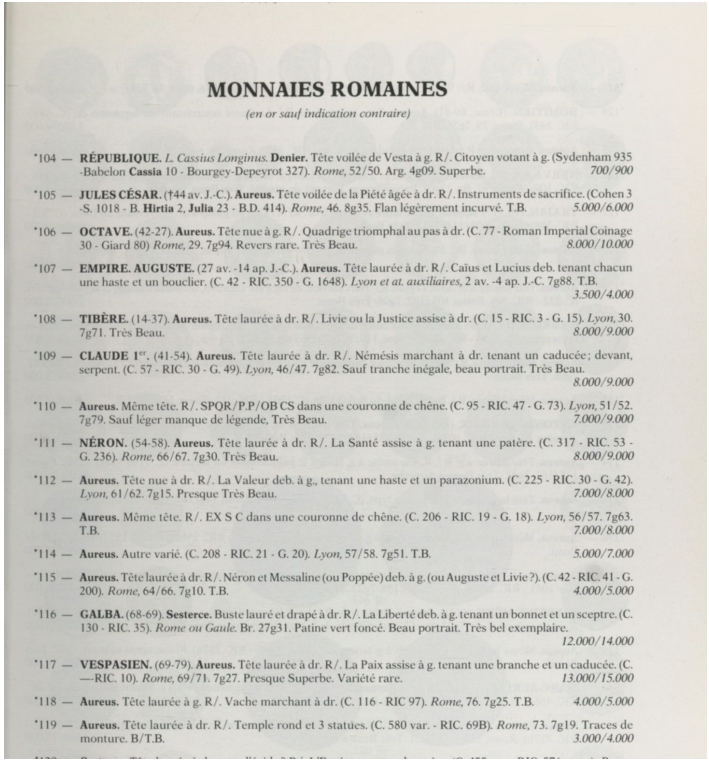
342-211 av. J.-C.

- *33 La louve tournée à d., allaitant les jumeaux. A l'ex., -. ROMA. Corbeau à d., tenant une fleur dans son bec; derrière, B. I, p. 20, 20. Sambon, Italie, 1158. AE. Mm. 30, gr. 21,36. Pat. Très beau.
- *34 Tête imberbe janiforme, coiffée de la stéphané. -. Jupiter tenant un sceptre et lançant le foudre, à demi-nu, dans un quadrige au galop à d., conduit par la Victoire. B.

Exemple de visualisation basique avec TEI Publisher d'un catalogue de vente encodé en TEI

DataCatalogue - v1 : bilan

- Modèle de données TEI pour représenter les catalogues de vente
- Annotation et entraînement de modèles GROBID pour la macro-structuration des catalogues, et la segmentation des notices
- Premières réflexions sur un modèle d'édition pour les catalogues de vente avec TEI Publisher
- Identification des défis à relever pour passer à l'échelle avec GROBID



MONNAIES ROMAINES

(en or sauf indication contraire)

*104 — **RÉPUBLIQUE.** *L. Cassius Longinus.* **Denier.** Tête voilée de Vesta à g. R/. Citoyen votant à g. (Sydenham 935 -Babelon **Cassia** 10 - Bourgey-Depeyrot 327). *Rome,* 52/50. Arg. 4g09. Superbe. *700/900*

*105 — **JULES CÉSAR.** (†44 av. J.-C.). **Aureus.** Tête voilée de la Piété âgée à dr. R/. Instruments de sacrifice. (Cohen 3 -S. 1018 - B. **Hirtia 2, Julia** 23 - B.D. 414). *Rome,* 46. 8g35. Flan légèrement incurvé. T.B. *5.000/6.000*

*106 — **OCTAVE.** (42-27). **Aureus.** Tête nue à g. R/. Quadrige triomphal au pas à dr. (C. 77 - Roman Imperial Coinage 30 - Giard 80) *Rome,* 29. 7g94. Revers rare. Très Beau. *8.000/10.000*

*107 — **EMPIRE. AUGUSTE.** (27 av. -14 ap. J.-C.). **Aureus.** Tête laurée à dr. R/. Caïus et Lucius deb. tenant chacun une haste et un bouclier. (C. 42 - RIC. 350 - G. 1648). *Lyon et at. auxiliaires,* 2 av. -4 ap. J.-C. 7g88. T.B. *3.500/4.000*

*108 — **TIBÈRE.** (14-37). **Aureus.** Tête laurée à dr. R/. Livie ou la Justice assise à dr. (C. 15 - RIC. 3 - G. 15). *Lyon,* 30. 7g71. Très Beau. *8.000/9.000*

*109 — **CLAUDE 1er.** (41-54). **Aureus.** Tête laurée à dr. R/. Némésis marchant à dr. tenant un caducée; devant, serpent. (C. 57 - RIC. 30 - G. 49). *Lyon,* 46/47. 7g82. Sauf tranche inégale, beau portrait. Très Beau. *8.000/9.000*

*110 — **Aureus.** Même tête. R/. SPQR/P.P/OB CS dans une couronne de chêne. (C. 95 - RIC. 47 - G. 73). *Lyon,* 51/52. 7g79. Sauf léger manque de légende, Très Beau. *7.000/9.000*

*111 — **NÉRON.** (54-58). **Aureus.** Tête laurée à dr. R/. La Santé assise à g. tenant une patère. (C. 317 - RIC. 53 - G. 236). *Rome,* 66/67. 7g30. Très Beau. *8.000/9.000*

*112 — **Aureus.** Tête nue à dr. R/. La Valeur deb. à g., tenant une haste et un parazonium. (C. 225 - RIC. 30 - G. 42). *Lyon,* 61/62. 7g15. Presque Très Beau. *7.000/8.000*

*113 — **Aureus.** Même tête. R/. EX S C dans une couronne de chêne. (C. 206 - RIC. 19 - G. 18). *Lyon,* 56/57. 7g63. T.B. *7.000/8.000*

*114 — **Aureus.** Autre varié. (C. 208 - RIC. 21 - G. 20). *Lyon,* 57/58. 7g51. T.B. *5.000/7.000*

*115 — **Aureus.** Tête laurée à dr. R/. Néron et Messaline (ou Poppée) deb. à g. (ou Auguste et Livie ?). (C. 42 - RIC. 41 - G. 200). *Rome,* 64/66. 7g10. T.B. *4.000/5.000*

*116 — **GALBA.** (68-69). **Sesterce.** Buste lauré et drapé à dr. R/. La Liberté deb. à g. tenant un bonnet et un sceptre. (C. 130 - RIC. 35). *Rome ou Gaule.* Br. 27g31. Patine vert foncé. Beau portrait. Très bel exemplaire. *12.000/14.000*

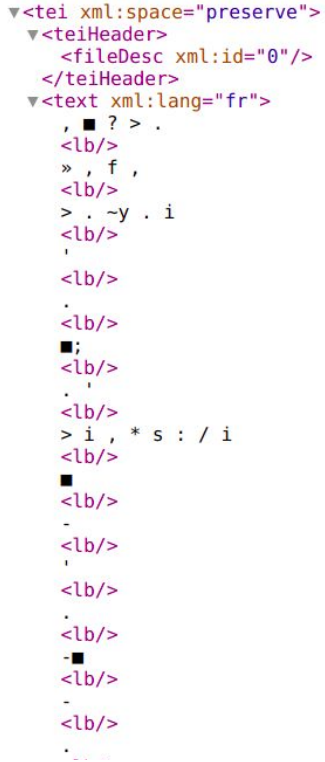
*117 — **VESPASIEN.** (69-79). **Aureus.** Tête laurée à dr. R/. La Paix assise à g. tenant une branche et un caducée. (C. —RIC. 10). *Rome,* 69/71. 7g27. Presque Superbe. Variété rare. *13.000/15.000*

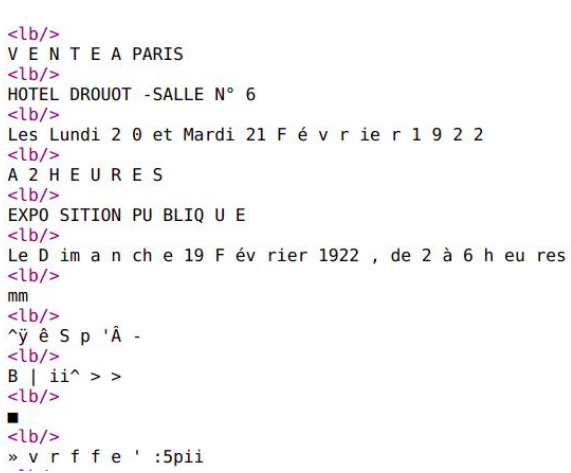
*118 — **Aureus.** Tête laurée à g. R/. Vache marchant à dr. (C. 116 - RIC 97). *Rome,* 76. 7g25. T.B. *4.000/5.000*

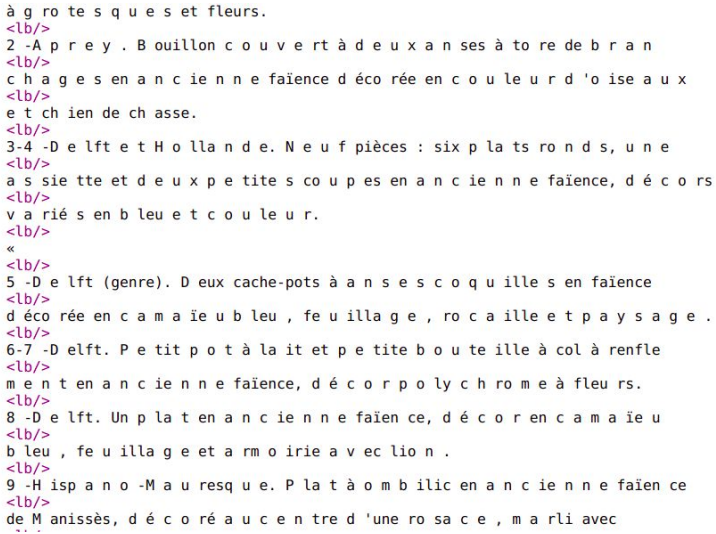
*119 — **Aureus.** Tête laurée à dr. R/. Temple rond et 3 statues. (C. 580 var. - RIC. 69B). *Rome,* 73. 7g19. Traces de monture. B/T.B. *3.000/4.000*

Bourgey, 1992 (BnF).

Des défis à relever pour DataCatalogue v2







Exemples “extrêmes” des scories d’OCR rencontrés dans les PDF numérisés : **caractères non existants dans les documents originaux, mots décomposées.** Ces erreurs ont des effets de cascades sur l’entraînement et l’inférence des modèles.

Des défis à relever pour DataCatalogue v2

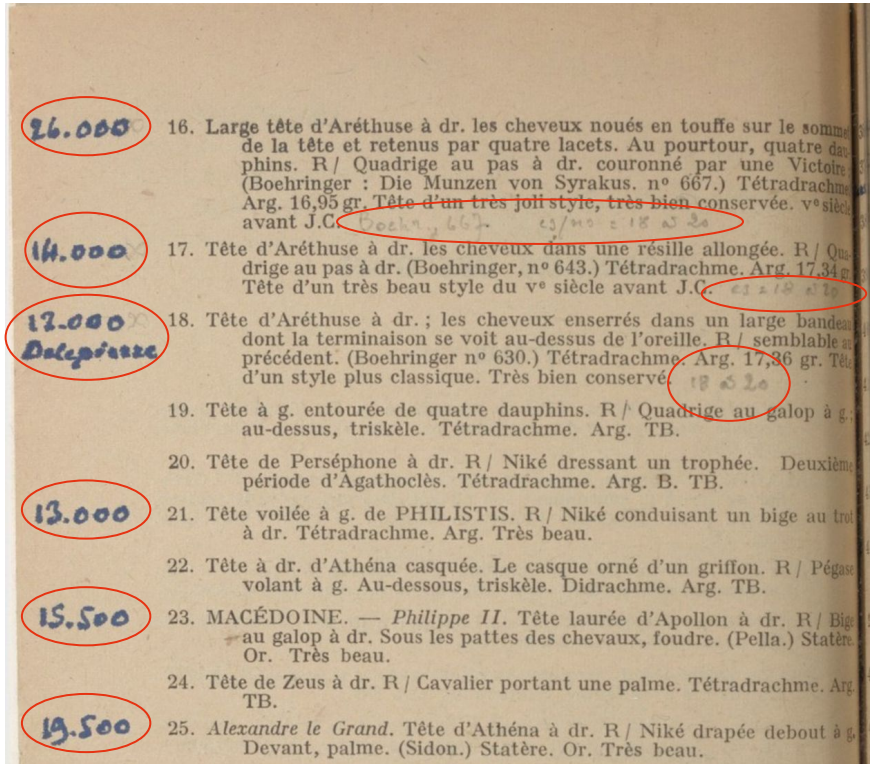
*124 — **DOMITIEN.** (César, 69-81). **Aureus.** Tête laurée à dr. R/. La Santé nourrissant un serpent à dr. (C. 383 - RIC. 243). *Rome*, 79. 7g25. B/T.B. *3.500/4.000*

Bourgey, 1992, **p. 28** (BnF).

Bourgey, 1992, **p. 29** (BnF).

→ Connecter les notices avec leurs illustrations respectives et en rendre compte dans l'encodage TEI

Des défis à relever pour DataCatalogue v2



26.000 16. Large tête d'Aréthuse à dr. les cheveux noués en touffe sur le somme de la tête et retenus par quatre lacets. Au pourtour, quatre dau phins. R/ Quadrige au pas à dr. couronné par une Victoire. (Boehringer : Die Munzen von Syrakus. nº 667.) Tétradrachme Arg. 16,95 gr. Tête d'un très joli style, très bien conservée. Vº siècle avant J.C.

14.000 17. Tête d'Aréthuse à dr. les cheveux dans une résille allongée. R/ Qua drige au pas à dr. (Boehringer, nº 643.) Tétradrachme. Arg. 17,34 gr. Tête d'un très beau style du Vº siècle avant J.C.

12.000 Delepierre 18. Tête d'Aréthuse à dr. ; les cheveux enserrés dans un large bandeau dont la terminaison se voit au-dessus de l'oreille. R/ semblable au précédent. (Boehringer nº 630.) Tétradrachme. Arg. 17,36 gr. Tête d'un style plus classique. Très bien conservé.

19. Tête à g. entourée de quatre dauphins. R/ Quadrige au galop à g. ; au-dessus, triskèle. Tétradrachme. Arg. TB.

20. Tête de Perséphone à dr. R/ Niké dressant un trophée. Deuxième période d'Agathoclès. Tétradrachme. Arg. B. TB.

13.000 21. Tête voilée à g. de PHILISTIS. R/ Niké conduisant un bige au trot à dr. Tétradrachme. Arg. Très beau.

22. Tête à dr. d'Athéna casquée. Le casque orné d'un griffon. R/ Pégase volant à g. Au-dessous, triskèle. Didrachme. Arg. TB.

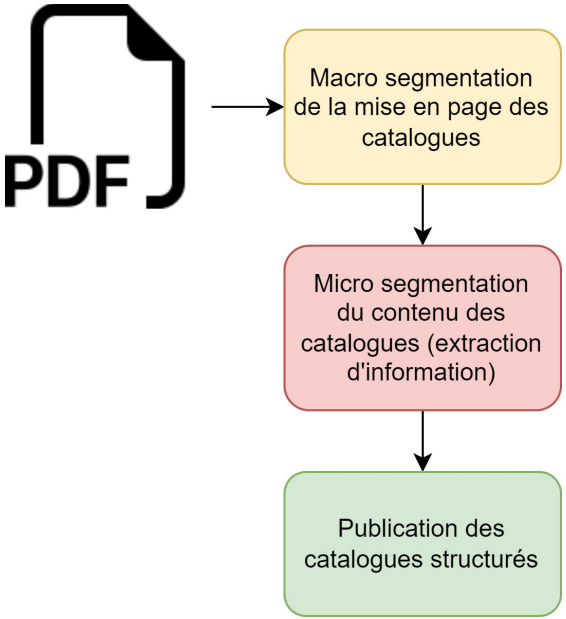
15.500 23. MACÉDOINE. — *Philippe II*. Tête laurée d'Apollon à dr. R/ Bige au galop à dr. Sous les pattes des chevaux, foudre. (Pella.) Statère. Or. Très beau.

24. Tête de Zeus à dr. R/ Cavalier portant une palme. Tétradrachme. Arg. TB.

19.500 25. *Alexandre le Grand*. Tête d'Athéna à dr. R/ Niké drapée debout à g. Devant, palme. (Sidon.) Statère. Or. Très beau.

Segmentation, transcription et structuration des segments de texte manuscrits

DataCatalogue v2 - 2023-2024



- Recrutement de Sarah Bénière dans l’ équipe ALMAnaCH, Inria, en tant qu’ingénieure recherche et développement
- Création d’un projet doctoral sur l’analyse automatique de mise en page mené par Hugo Scheithauer (Paris PhD)

Collaboration et réutilisation des données

- Campagne d'annotation de la macro-structure en coordination avec le projet **COlaF** (Inria, Multispeech)
- Schéma d'annotation suivant l'ontologie **SegmOnto**
- Corpus annotés et développements mis à disposition en ligne sur GitHub et Roboflow, selon les principes de la science ouverte
- Collaboration pluri-institutionnelles entre Inria, la BnF et l'INHA. L'idée est également de créer des corpus annotés et des outils facilement réutilisables par d'autres institutions

https://github.com/DataCatalogue

https://segmonto.github.io/

https://colaf.huma-num.fr/

https://htr-united.github.io/

Breaking news: #prixscienceouverte #donneesrecherche Catégorie Prix jeunes chercheurs - 2023-11-29

Merci pour votre attention !

Ressources :

- GitHub DataCatalogue : https://github.com/DataCatalogue
- https://huggingface.co/spaces/HugoSchtr/DataCat_Yolov5
- Corpus de macro-segmentation des catalogues : https://app.roboflow.com/datacatalogue/macro-segmentation/overview

Contacts :

sarah.beniere[at]inria.fr
hugo.scheithauer[at]inria.fr