

# CONCEPTION ET DÉVELOPPEMENT D'UNE APPLICATION D'ANNOTATION THÉMATIQUE

dans l'environnement Gate

---

Charles Follet   Roland Bary

May 16, 2015

Université de Pau et Pays de l'Adour

- Extension du web vers un web de données: le web sémantique

- Extension du web vers un web de données: le web sémantique
- Recherche sémantique: activité du web sémantique

- Extension du web vers un web de données: le web sémantique
- Recherche sémantique: activité du web sémantique
- Annotation sémantique: Étape du processus de recherche sémantique

- Extension du web vers un web de données: le web sémantique
- Recherche sémantique: activité du web sémantique
- Annotation sémantique: Étape du processus de recherche sémantique
- Problème: Comment annoter sémantiquement un document texte non-structuré ?

- Extension du web vers un web de données: le web sémantique
- Recherche sémantique: activité du web sémantique
- Annotation sémantique: Étape du processus de recherche sémantique
- Problème: Comment annoter sémantiquement un document texte non-structuré ?
- Quelles outils pour l'annotation ?

- Extension du web vers un web de données: le web sémantique
- Recherche sémantique: activité du web sémantique
- Annotation sémantique: Étape du processus de recherche sémantique
- Problème: Comment annoter sémantiquement un document texte non-structuré ?
- Quelles outils pour l'annotation ?
- Quelle approche de conception ?

Le cahier des charges

Cadre d'analyse

    L'environnement GATE

    Les Gazetiers

    Le formalisme JAPE

Structure du numéraire

Annotations

Construction d'une règle JAPE

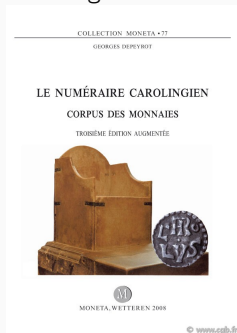
Visualisation



## LE CAHIER DES CHARGES

---

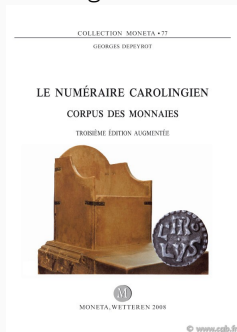
Contexte: Recherche d'informations sur le Numéraire Carolingien de Georges DEPEYROT.



Exemples de recherches possibles:

- Temporelle: Quelles étaient les pièces en circulation de l'an 859 à l'an 865 ?
- Spatiale: Dans quels ateliers, les pièces de type Obole de Charlemagne ont été produites ?
- Thématique: Combien d'exemplaires de la monnaie d'or de Charles le Chauve ont été étudiés ?

Contexte: Recherche d'informations sur le Numéraire Carolingien de Georges DEPEYROT.



Exemples de recherches possibles:

- Temporelle: Quelles étaient les pièces en circulation de l'an 859 à l'an 865 ?
- Spatiale: Dans quels ateliers, les pièces de type Obole de Charlemagne ont été produites ?
- Thématique: Combien d'exemplaires de la monnaie d'or de Charles le Chauve ont été étudiés ?

Solution: Utilisation d'un outil de TAL: GATE

Caractéristiques du projets:

- apprentissage et compréhension du domaine considéré

Caractéristiques du projets:

- apprentissage et compréhension du domaine considéré
- étude des principes d'annotation de documents

Caractéristiques du projets:

- apprentissage et compréhension du domaine considéré
- étude des principes d'annotation de documents
- développement d'une chaîne d'annotation dans GATE

Caractéristiques du projets:

- apprentissage et compréhension du domaine considéré
- étude des principes d'annotation de documents
- développement d'une chaîne d'annotation dans GATE
- mise en place d'une interface web de visualisation des résultats

## Scénario attendu: Texte en entrée de traitement

Auxerre (Yonne)

- Type de 840-864: Lothaire I (817-855), Pépin II, roi d'Aquitaine (839-865), Charles le Chauve (840-877), Lothaire II roi de Lorraine (855-869), Charles l'Enfant roi d'Aquitaine (vers 860)
- 96 Denier de Charles le Chauve (43 exemplaires étudiés)  
+ CAROLVS REX FR croix, 4 globes AVTISIODERO CIV temple  
Gariel XXI 1, 2; Morrison-Grunthal 987-990; Depeyrot (I) (2) 96; Nombreuses variétés de légende, de ponctuation et de gravure. Collections: Assen 1,54; Berlin: MEC 842 (1,59); New York 1,49; Prou 582 (1,63), 583 (1,46). Trésors: Groningen-1890 (vers 840), 1 ex. (MG 38; H. 39, daté de 845-846); Auxerre (840-845), 1 ex. (D. 24); Tzummarum (840-845), 1 ex. (H. 22 daté de 840, ne mentionne pas cette monnaie dans le catalogue); Cosne (840-845), 30 ex. (D. 109; H. 45, daté de 845-846, voir p. 93); Midlaren (855), 1 ex. (MG 55; H. 57 daté de 864); Nourray (869-875), 1 ex. (D. 243; H. 64 daté de 869-870); Roswinkel (875), 1 ex. (MG 66; H. 75 daté de 877) (1,55); Trouvailles: Auxerre, 1 ex. (H. 511)
- Type de 864-875: Charles le Chauve (840-877), Louis le Germanique (840-876), Louis II le Bègue roi d'Aquitaine (867-877), de France (877-879)
- 97 Denier de Charles le Chauve (22 exemplaires étudiés)  
+ GRATIA DEI REX monogramme + AVTISIODER CIVITAS croix  
Gariel XXV 20-23; Morrison-Grunthal 991, 994, 996-1001; Depeyrot (1) (2) 97; Nombreuses variétés de légende, de ponctuation et de gravure (AVTISIODER CIV). Collections: Berlin 1,77, 1,70, 1,59, 1,55; MEC 853 (1,78); Monnaie de Paris 105 (1,63); Prou 584 (1,58), 585 (1,69), 586 (1,79), 587 (1,72), 588 (1,77). Trésors: Fontaines/Mercurey (840-875), 1 ex. (D. 147 ne mentionne pas cette monnaie; H. 66 daté de 875); Cosne (875-877), 2 ex. (D. 110; H. 67 daté de 875) (1,74, 1,73); Montrieux-en-Sologne (887-898), 4 ex. (D. 231; H. 88 daté de 888); Glisy (890), 3 ex. (D. 154; H. 90 daté de 888); Monchy-au-Bois (895-898), 1 ex. (D. 217; H. 101 daté de 897-898); Montrieux-en-Sologne (900), 1 ex. (D. 229; H. 87 daté de 888).
- 98 Denier de Charles le Chauve (1 exemplaire étudié)  
+ GRATIA DEI REX monogramme + ALTIEI CIVITVS (rétro.) croix  
Gariel XXV 24; Morrison-Grunthal 991; Depeyrot (1) (2) 98; Collections: Prou 589 (1,68).



## Scénario attendu: Texte en sortie de traitement

### Auxerre (Yonne)

Type de 840-864: Lothaire I (817-855), Pépin II, roi d'Aquitaine (839-865), Charles le Chauve (840-877), Lothaire II roi de Lorraine (855-869), Charles l'Enfant roi d'Aquitaine (vers 860)

96 Denier de Charles le Chauve (43 exemplaires étudiés)

+ CAROLVS REX FR croix, 4 globes AVTISIODERO CIV temple

Gariel XX 1, 2; Morrison-Grunthal 987-990; Depeyrot (I) (2) 96; Nombreuses variétés de légende, de ponctuation et de gravure. Collections: Assen 1,54; Berlin MEC 842 (1,59); New York 1,49; Prou 582 (1,63), 583 (1,46); Trésors: Groningen-1890 (vers 840), 1 ex. (MG 38; H. 39, daté de 845-846); Auxerre (840-845), 1 ex. (D. 24); Tzummarum (840-845), 1 ex. (H. 22 daté de 840, ne mentionne pas cette monnaie dans le catalogue); Cosne (840-845), 30 ex. (D. 109; H. 45, daté de 845-846, voir p. 93); Midlaren (855), 1 ex. (MG 55; H. 57 daté de 864); Nourray (869-875), 1 ex. (D. 243; H. 64 daté de 869-870); Roswinkel (875), 1 ex. (MG 66; H. 75 daté de 877) (1,55); Trouvailles: Auxerre, 1 ex. (H. 511)

Type de 864-875: Charles le Chauve (840-877), Louis le Germanique (840-876), Louis II le Bègue roi d'Aquitaine (867-877), de France (877-879)

97 Denier de Charles le Chauve (22 exemplaires étudiés)

+ GRATIA DEI REX monogramme + AVTISIODER CIVITAS croix

Gariel XXV 20-23; Morrison-Grunthal 991, 994, 996-1001; Depeyrot (I) (2) 97; Nombreuses variétés de légende, de ponctuation et de gravure (AVTISIODER CIV). Collections: Berlin 1,77, 1,70, 1,59, 1,55; MEC 853 (1,78); Monnaie de Paris 105 (1,63); Prou 584 (1,58), 585 (1,69), 586 (1,79), 587 (1,72), 588 (1,77); Trésors: Fontaines/Mercurey (840-875), 1 ex. (D. 147 ne mentionne pas cette monnaie; H. 66 daté de 875); Cosne (875-877), 2 ex. (D. 110; H. 67 daté de 875) (1,74, 1,73); Montrieux-en-Sologne (887-898), 4 ex. (D. 231; H. 88 daté de 888); Glisy (890), 3 ex. (D. 154; H. 90 daté de 888); Monchy-au-Bois (895-898), 1 ex. (D. 217; H. 101 daté de 897-898); Montrieux-en-Sologne (900), 1 ex. (D. 229; H. 87 daté de 888).

98 Denier de Charles le Chauve (1 exemplaire étudié)

+ GRATIA DEI REX monogramme + ALTIEI CIVITVS (rétro.) croix

Gariel XXV 24; Morrison-Grunthal 991; Depeyrot (I) (2) 98; Collections: Prou 589 (1,68).

## CADRE D'ANALYSE

---

# CADRE D'ANALYSE - GATE

The screenshot shows the GATE Developer 5.1 interface. On the left, the 'Processing Resources' pane lists various modules including ANNE OrthoMatcher, ANNE NE Transducer, ANNE POS Tagger, ANNE Sentence Splitter, ANNE Gazetteer, and ANNE English Tokeniser. The main text area displays a news article about unemployment benefits, with various words and phrases highlighted in different colors (e.g., '10 hours ago', 'Thursday', 'Congress', 'Obama'). A list of selected annotations is shown at the bottom left, including 'Date', 'Location', 'Person', 'Organization', and 'Date'. A list of annotation types is shown on the right side of the interface.

Chaines de traitement

Textes et corpus à traiter

Modules d'annotation

Liste des annotations sélectionnées

Texte annoté

Liste des types d'annotation

# CADRE D'ANALYSE - GATE

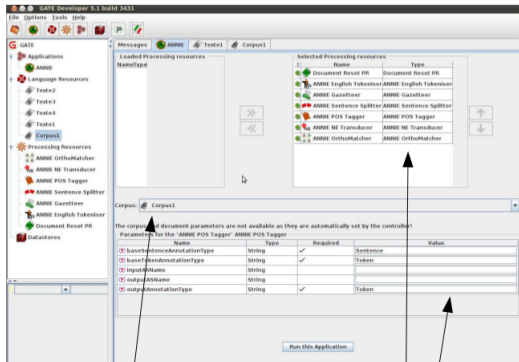
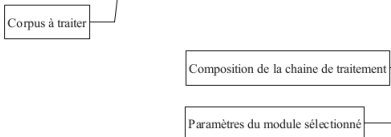


Figure 2.5 : Exemple de chaîne de traitement GATE



# CADRE D'ANALYSE - GAZETIERS

rols.txt	x
1	Pépin le Bref:valeur=Pépin le Bref:periode=752-768
2	Adalbert Lothaire:valeur=Adalbert Lothaire:periode=954-986
3	Anoul roi de Germanie:valeur=Anoul roi de Germanie:periode=887-899
4	Bérenger I:valeur=Bérenger I:periode=888-924
5	Bérenger II:valeur=Bérenger II:periode=950-961
6	Boson roi de Provence:valeur=Boson roi de Provence:periode=879-887
7	Carloman:valeur=Carloman:periode=768-771
8	Carloman:valeur=Carloman:periode=879-884
9	Charlemagne:valeur=Charlemagne:periode=768-814
10	Charles de Provence:valeur=Charles de Provence:periode=876-888
11	Charles l'Enfant roi d'Aquitaine:valeur=Charles l'Enfant roi d'Aquitaine:periode=vers 868
12	Charles le Chauve:valeur=Charles le Chauve:periode=840-877
13	Charles le Gros:valeur=Charles le Gros:periode=884-887
14	Charles le Simple:valeur=Charles le Simple:periode=898-929
15	Conrad le Pacifique:valeur=Conrad le Pacifique:periode=937-993
16	Eudes:valeur=Eudes:periode=887-898
17	Gul de Spolète:valeur=Gul de Spolète:periode=891-894
18	Henri l'Oiseleur:valeur=Henri l'Oiseleur:periode=918-947
19	Hugues de Provence:valeur=Hugues de Provence:periode=926-947
20	Lambert empereur:valeur=Lambert empereur:periode=894-898
21	Lothaire d'Italie:valeur=Lothaire d'Italie:periode=931-950
22	Lothaire I:valeur=Lothaire I:periode=817-855
23	Lothaire II roi de Lorraine:valeur=Lothaire II roi de Lorraine:periode=855-869
24	Louis II le Bègue:valeur=Louis II le Bègue:periode=877-879
25	Louis II le Bègue roi d'Aquitaine:valeur=Louis II le Bègue roi d'Aquitaine:periode=867-877), de France (877-879
26	Louis III:valeur=Louis III:periode=879-882
27	Louis III:valeur=Louis III:periode=890-929
28	Louis IV d'Outre-Mer:valeur=Louis IV d'Outre-Mer:periode=936-954
29	Louis IV l'Enfant:valeur=Louis IV l'Enfant:periode=899-912
30	Louis l'Aveugle roi de Provence:valeur=Louis l'Aveugle roi de Provence:periode=890-929
31	Louis le Germanique:valeur=Louis le Germanique:periode=840-876
32	Louis le Germanique:valeur=Louis le Germanique:periode=876-876
33	Louis le Jeune:valeur=Louis le Jeune:periode=879-882
34	Louis le Pieux:valeur=Louis le Pieux:periode=814-840
35	Louis le Pieux roi d'Aquitaine:valeur=Louis le Pieux roi d'Aquitaine:periode=781-814
36	Louis V:valeur=Louis V:periode=900-907
37	Otton I:valeur=Otton I:periode=936-973) ou Otton II (973-983) (frappes germaniques pour mémoire
38	Pépin I roi d'Aquitaine:valeur=Pépin I roi d'Aquitaine:periode=817-838
39	Pépin II, roi d'Aquitaine:valeur=Pépin II, roi d'Aquitaine:periode=839-865
40	Raoul:valeur=Raoul:periode=923-936

```
Phase: pourcent
Input: Lookup Token
Options: control = appelle
Macro: AMOUNT_NUMBER
(({Token.kind == number}
(
({Token.string == ",", " " | {Token.string == "."})
{Token.kind == number}
)?
({Token.string == "-"}{Token.kind == number}))?)
Macro: PERCENT(
{Token.string == "%"} |
{Token.string == "pourcent"} |
({Token.string == "pour"}{Token.string == "cent"})
14)
Macro: LIAISON_PERCENT_DE_ET(
{Token.string == "de"} |
{Token.string == "et"} |
{Token.string == ", "})
Rule: Percent_number
(((AMOUNT_NUMBER)
({Token.string == "a"}
)?(AMOUNT_NUMBER)(PERCENT))
: pourcent -->: pourcent.Percent = {Kind="pourcent", Rule="Percent_number"}
```

## STRUCTURE DU NUMÉRAIRE

---

### Agen (Lot-et-Garonne)

Type de 793/4-812: Charlemagne (768-814)

Denier de Charlemagne (24 exemplaires étudiés)

+ CARLVS REX FR croix      + AGINNO monogramme

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50) Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex. (MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);

Trouvailles: Bolsward, 1 ex. (H. 528) (1,26).



## Agen (Lot-et-Garonne)

Type de 793/4-812: Charlemagne (768-814)

Denier de Charlemagne (24 exemplaires étudiés)

+ CARLVS REX FR croix      + AGINNO monogramme

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50) Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex. (MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);

Trouvailles: Bolsward, 1 ex. (H. 528) (1,26).

Agen (Lot-et-Garonne)

## **Type de 793/4-812: Charlemagne (768-814)**

Denier de Charlemagne (24 exemplaires étudiés)

+ CARLVS REX FR croix      + AGINNO monogramme

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50) Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex. (MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);

Trouvailles: Bolsward, 1 ex. (H. 528) (1,26).

Agen (Lot-et-Garonne)

Type de 793/4-812: Charlemagne (768-814)

**Denier de Charlemagne (24 exemplaires étudiés)**

+ CARLVS REX FR croix      + AGINNO monogramme

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50) Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex. (MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);

Trouvailles: Bolsward, 1 ex. (H. 528) (1,26).

Agen (Lot-et-Garonne)

Type de 793/4-812: Charlemagne (768-814)

Denier de Charlemagne (24 exemplaires étudiés)

**+ CARLVS REX FR croix      + AGINNO monogramme**

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50) Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex. (MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);

Trouvailles: Bolsward, 1 ex. (H. 528) (1,26).

Agen (Lot-et-Garonne)

Type de 793/4-812: Charlemagne (768-814)

Denier de Charlemagne (24 exemplaires étudiés)

+ CARLVS REX FR croix      + AGINNO monogramme

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; **Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50)** Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex. (MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);

Trouvailles: Bolsward, 1 ex. (H. 528) (1,26).

Agen (Lot-et-Garonne)

Type de 793/4-812: Charlemagne (768-814)

Denier de Charlemagne (24 exemplaires étudiés)

+ CARLVS REX FR croix      + AGINNO monogramme

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50) **Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex. (MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);**

Trouvailles: Bolsward, 1 ex. (H. 528) (1,26).

Agen (Lot-et-Garonne)

Type de 793/4-812: Charlemagne (768-814)

Denier de Charlemagne (24 exemplaires étudiés)

+ CARLVS REX FR croix      + AGINNO monogramme

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50) Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex. (MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);

**Trouvailles: Bolsward, 1 ex. (H. 528) (1,26).**

Agen (Lot-et-Garonne)<sup>ATELIER</sup>

Type de 793/4-812: Charlemagne (768-814)<sup>TYPE</sup>

Denier de Charlemagne (24 exemplaires étudiés)<sup>NATURE</sup>

+ CARLVX REX FR croix + AGINNO monogramme<sup>LÉGENDE</sup>

Gariel XXII 26; Morrison-Grunthal 177-179; Depeyrot (1) (2) 1; Collections: Berlin 1,64, 1,45; Bruxelles 1,25, 1,25; Charleville-Mézières 1,56; Copenhague 1,31; Garrett 1,73; Grenoble; MEC 735 (1,57), 736 (1,24); Monnaie de Paris 88 (1,55); New York 1,64, 1,57; Prou 792 (1,36), 793 (1,60), 794 (1,50)<sup>COLLECTIONS</sup> Trésors: Biebrich (790-814), 2 ex. (MG 11, Völkers, p. 182) (1,60); Ibersheim (814), 1 ex.

(MG 13, Völkers, p. 186) (1,64); Dorestad (822), 4 ex. (MG 18, Völkers, p. 139; H. 7) (1,35);<sup>TRÉSORS</sup>

Trouvailles: Bolsward, 1 ex. (H. 528) (1,26)<sup>TROUVAILLES</sup>.



## ANNOTATIONS

---

## GAZETIER

liste de valeurs.

- Type de 754/5-768
- Type de 768-771
- Type de 771-793/4
- Type de 793/4-812
- Type de 812-814
- ...

## RÈGLE JAPE

règle basée sur une expression régulière.

```
Macro: CHAINE_DEBUT
(
  ({Token.string == "Type"})({SpaceToken})
  ({Token.string == "de"})({SpaceToken})
)

// Exemple : Type de 758/9-762
Rule: PeriodeEmissionRule
(
  CHAINE_DEBUT ({Periode}):p
):PeriodeEmission
-->:PeriodeEmission.PeriodeEmission = {
  Kind = "PeriodeEmission" ,
  D1 = :p.Periode.D1, D2 = :p.Periode.D2
}
```

## GAZETIER

liste de valeurs.

- Type de 754/5-768
- Type de 768-771
- Type de 771-793/4
- Type de 793/4-812
- Type de 812-814
- ...

## RÈGLE JAPE

règle basée sur une expression régulière.

```
Macro: CHAINE_DEBUT
(
  ({Token.string == "Type"})({SpaceToken})
  ({Token.string == "de"})({SpaceToken})
)

// Exemple : Type de 758/9-762
Rule: PeriodeEmissionRule
(
  CHAINE_DEBUT ({Periode}):p
):PeriodeEmission
-->:PeriodeEmission.PeriodeEmission = {
  Kind = "PeriodeEmission" ,
  D1 = :p.Periode.D1, D2 = :p.Periode.D2
}
```

## CONSTRUCTION D'UNE RÈGLE JAPE

---

### 1. Langage naturel

1. Langage naturel
2. Langage régulier

1. Langage naturel
2. Langage régulier
3. Langage de programmation

1. Langage naturel
2. Langage régulier
3. Langage de programmation

français



1. Langage naturel
2. Langage régulier
3. Langage de programmation

français  
expressions régulières

1. Langage naturel
2. Langage régulier
3. Langage de programmation

français  
expressions régulières  
*JAPE*

Annotation : La période d'émission

Exemple : *Type de 793/4-812*

Type de 793/4-812

### 1. LANGAGE NATUREL

Une période est un intervalle entre deux dates. Dans notre travail, les dates sont constituées de trois chiffres. En cas d'ambiguïté, une date peut être suivie d'un / et d'un chiffre traduisant l'indétermination de la date.

Il faut assurer la capture de la période d'émission en ajoutant la contrainte précédée de "Type de" afin de ne pas récupérer toutes les périodes du document.

Type de 793/4-812

### 1. LANGAGE NATUREL

Une période est un intervalle entre deux dates. Dans notre travail, les dates sont constituées de trois chiffres. En cas d'ambiguïté, une date peut être suivie d'un / et d'un chiffre traduisant l'indétermination de la date.

Il faut assurer la capture de la période d'émission en ajoutant la contrainte précédée de "Type de" afin de ne pas récupérer toutes les périodes du document.

Type de 793/4-812

### 2. LANGAGE RÉGULIER

Type de :  $([0-9]\{3\}(/[0-9])?)-([0-9]\{3\}(/[0-9])?)$

Type de 793/4-812

### 2. LANGAGE RÉGULIER

Type de :  $([0-9]\{3\}(/[0-9])?)-([0-9]\{3\}(/[0-9])?)$

Type de 793/4-812

## 2. LANGAGE RÉGULIER

Type de :  $([0-9]\{3\}(/[0-9])?)-([0-9]\{3\}(/[0-9])?)$



Type de 793/4-812

### 2. LANGAGE RÉGULIER

Type de :  $([0-9]\{3\}(/[0-9])?)-([0-9]\{3\}(/[0-9])?)$

Type de 793/4-812

## 3. LANGAGE DE PROGRAMMATION

Macro: CHAINE\_DEBUT

```
(  
  ({Token.string == "Type"})({SpaceToken})  
  ({Token.string == "de"})({SpaceToken})  
)
```

Macro: TROIS\_NOMBRES

```
({Token.kind==number,Token.length == 3})
```

Macro: UN\_NOMBRE

```
({Token.kind==number,Token.length == 1})
```

Macro: SLASH

```
({Token.string=="/"})
```

Macro: DATE\_PRECISE

```
(TROIS_NOMBRES)
```

Macro: DATE\_IMPRECISE

```
(TROIS_NOMBRES SLASH UN_NOMBRE)
```

Macro: DATE

```
(DATE_PRECISE | DATE_IMPRECISE)
```

Rule: PeriodeRule

```
((DATE):d1({Token.string == "-"})(DATE):d2)  
:Periode -->:Periode{"/..."/}
```

Rule: PeriodeEmissionRule

```
(CHAINE_DEBUT ({Periode}):p)
```

:PeriodeEmission

```
-->:PeriodeEmission.PeriodeEmission = {"/..."/}
```

## Type de 793/4-812

### 3. LANGAGE DE PROGRAMMATION

Macro: CHAINE\_DEBUT

```
(  
  ({Token.string == "Type"})({SpaceToken})  
  ({Token.string == "de"})({SpaceToken})  
)
```

Macro: TROIS\_NOMBRES

```
({Token.kind==number,Token.length == 3})
```

Macro: UN\_NOMBRE

```
({Token.kind==number,Token.length == 1})
```

Macro: SLASH

```
({Token.string=="/"})
```

Macro: DATE\_PRECISE

```
(TROIS_NOMBRES)
```

Macro: DATE\_IMPRECISE

```
(TROIS_NOMBRES SLASH UN_NOMBRE)
```

Macro: DATE

```
(DATE_PRECISE | DATE_IMPRECISE)
```

Rule: PeriodeRule

```
((DATE):d1({Token.string == "-"})(DATE):d2)  
:Periode -->:Periode{"/.../"}
```

Rule: PeriodeEmissionRule

```
(CHAINE_DEBUT ({Periode}):p)  
:PeriodeEmission  
-->:PeriodeEmission.PeriodeEmission = {"/.../"}
```

## Type de 793/4-812

### 3. LANGAGE DE PROGRAMMATION

```
Macro: CHAINE_DEBUT
(
  ({Token.string == "Type"})({SpaceToken})
  ({Token.string == "de"})({SpaceToken})
)
```

```
Macro: TROIS_NOMBRES
({Token.kind==number,Token.length == 3})
Macro: UN_NOMBRE
({Token.kind==number,Token.length == 1})
Macro: SLASH
({Token.string=="/"})
Macro: DATE_PRECISE
(TROIS_NOMBRES)
Macro: DATE_IMPRECISE
(TROIS_NOMBRES SLASH UN_NOMBRE)
Macro: DATE
(DATE_PRECISE | DATE_IMPRECISE)
```

```
Rule: PeriodeRule
((DATE):d1({Token.string == "-"})(DATE):d2)
:Periode -->:Periode{/"..."/}
```

```
Rule: PeriodeEmissionRule
(CHAINE_DEBUT ({Periode}):p)
:PeriodeEmission
-->:PeriodeEmission.PeriodeEmission = {/"..."/}
```

## Type de 793/4-812

### 3. LANGAGE DE PROGRAMMATION

```
Macro: CHAINE_DEBUT
(
  ({Token.string == "Type"})({SpaceToken})
  ({Token.string == "de"})({SpaceToken})
)
```

```
Macro: TROIS_NOMBRES
({Token.kind==number,Token.length == 3})
Macro: UN_NOMBRE
({Token.kind==number,Token.length == 1})
Macro: SLASH
({Token.string=="/"})
Macro: DATE_PRECISE
(TROIS_NOMBRES)
Macro: DATE_IMPRECISE
(TROIS_NOMBRES SLASH UN_NOMBRE)
Macro: DATE
(DATE_PRECISE | DATE_IMPRECISE)
```

```
Rule: PeriodeRule
((DATE):d1({Token.string == "-"})(DATE):d2)
:Periode -->:Periode{ /*...*/ }
```

```
Rule: PeriodeEmissionRule
(CHAINE_DEBUT ({Periode}):p)
:PeriodeEmission
-->:PeriodeEmission.PeriodeEmission = { /*...*/ }
```

### Résultat d'annotation

<PeriodeEmission>Type de 793/4-812</PeriodeEmission>

### Résultat d'annotation

<PeriodeEmission>Type de 793/4-812</PeriodeEmission>

*Comment le visualiser de façon élégante ?*

### Résultat d'annotation

`<PeriodeEmission>Type de 793/4-812</PeriodeEmission>`

*Comment le visualiser de façon élégante ?*

XSLT



## VISUALISATION

---