HAL
open science

Du corpus archivistique au corpus numérique : les soubassements du web sémantique. L'exemple des sources relatives au parlement de Flandre

Renaud Limelette

To cite this version:

Renaud Limelette. Du corpus archivistique au corpus numérique : les soubassements du web sémantique. L'exemple des sources relatives au parlement de Flandre. 2019. hal-02901298v1

HAL Id: hal-02901298

https://hal.univ-lille.fr/hal-02901298v1

Preprint submitted on 19 Jan 2021 (v1), last revised 1 Jul 2021 (v3)

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.

Du corpus archivistique au corpus numérique : les soubassements du web sémantique. L'exemple des sources relatives au parlement de Flandre

Version preprint

Renaud Limelette

Depuis une dizaine d'années le Centre d'Histoire Judiciaire (CHJ) s'est lancé dans la valorisation des fonds d'archives liés au parlement de Flandre. Si cette cour souveraine de justice, ancêtre de la Cour d'appel de Douai, a connu une existence relativement brève, de 1667 à 1790, ses archives offrent aux chercheurs de différentes disciplines des sources de questionnement multiples.

Cette valorisation se décline aujourd'hui en deux projets connexes. Le plus ancien s'incarne autour de la constitution d'une base de données [1] regroupant les 30 000 dossiers de procédure civile et criminelle conservés dans la sous-série 8 B 1 des Archives départementales du Nord. Le développement informatique et technique a entièrement été réalisé au sein du CHJ en s'appuyant sur MySQL et le framework Yii[2].

Le second exploite et expose toute la littérature juridique et historique laissée soit directement par les magistrats de la cour soit par les premiers historiens de celle-ci. On y trouve en bonne place les recueils d'arrêts, les arrêts de règlement et la biographie des magistrats. L'exposition des données se fait à travers les outils proposés par la TGIR Huma-Num[3], Nakala[4] et Nakalona[5].

Malgré tous les efforts menés par les ingénieurs du CHJ, aucun chercheur, quelle que soit sa discipline, ne peut lire l'intégralité des sources diffusées, pour tirer de celles-ci celles qui serviront à sa propre recherche. Pour lever cet obstacle, à savoir la sélection de données pertinente, l'informatique offre des moyens de structuration des données pour faciliter leur échange et leur compréhension entre machines.

Nous proposons d'expliquer les choix opérés dans les deux projets autour de la granularité des données et des métadonnées.

---

1 Pour accéder à la base et y lancer une recherche : http://parleflandre.univ-lille2.fr/index.php/recherche.

2 Yii est un cadre d'application (ou infrastructure logicielle) destiné au développement d'application Web. Il s'appuie sur un Modèle-vue-contrôleur. Le modèle est la base de données, la vue est la représentation graphique (visuelle) de l'utilisateur de la base et le contrôleur fait respecter la concordance entre le modèle et la vue lorsque l'utilisateur agit sur les données (création, suppression, modification, recherche, etc.).

3 https://www.huma-num.fr.

4 https://www.huma-num.fr/services-et-outils/exposer.

5 https://www.huma-num.fr/services-et-outils/diffuser#nakalona.

1. La granularité des données ou le choix d'une efficience

Concernant la granularité des données [6], on remarquera que les deux projets touchent au même objet de recherche, le parlement de Flandre, mais la nature des sources utilisées est différente. Or c'est la nature des sources qui conditionnent la manière dont elles vont être exploitées ou valorisées par les humanités numériques.

Notons ici que cette démarche, qui part des sources, est inductive et non déductive. On part des sources pour aboutir à un moyen de les valoriser, et non d'une idée de valorisation pour y soumettre des sources [7]. Cette démarche ne s'entend que si l'on veut, à travers les humanités numériques, expliquer, c'est-à-dire déplier, les composantes d'un corpus, pour en présenter le contenu. Au contraire, si l'on veut d'abord utiliser un média numérique bien précis, pour valoriser un corpus, on choisira une sélection des contenus du corpus qui coïncide assez bien avec le média, mais qui ne reflète pas tous les aspects du corpus. Par exemple, si on choisit d'utiliser la vidéo pour présenter le fonds du parlement de Flandre, on choisira des affaires judiciaires qui ont un potentiel visuel assez fort, et on laissera de côté l'énorme masse de la banalité des affaires.

Cette digression posant le sens de déductif par rapport à inductif, revenons au choix opéré de manière inductive pour le parlement de Flandre et à la granularité des données. Nous verrons que c'est la nature du corpus qui conditionne la solution numérique.

1. 1. 30 000 dossiers de procédure ou 450 mètres linéaires d'archives manuscrites : le projet ParleFlandre

Le premier corpus concernant le parlement de Flandre est la sous-série 8 B 1 des Archives départementales du Nord. Matériellement, cette sous-série contient près de 30 000 dossiers et s'étale sur 450 mètres linéaires. Y sont conservés tous les dossiers de procédures.

Pour mettre en valeur ces dossiers, le Centre d'Histoire Judiciaire (UMR 8025) déposa un projet à l'Agence Nationale de la Recherche sous l'acronyme ParleFlandre. Le but du projet était de présenter au public ces 30 000 dossiers de procédure.

---

6 Toute cette partie repose sur la notion de granularité, la granularité s'entend par le contenu du corpus mis en avant et la manière dont le corpus est sectionné en unités logiques, v. par exemple, STÉPHANIE DELMOTTE, « Publications scientifiques en sciences humaines. L'argumentation dans l'accès aux savoirs dans l'édition numérique », *Les Cahiers du numérique*, 2009/2 (Vol. 5), p. 53-84. URL : https://www.cairn.info/revue-les-cahiers-du-numerique-2009-2-page-53.htm.

7 Pour une plus longue explication de cette démarche et une mise en pratique, on se reportera à RENAUD LIMELETTE, « La gouvernance du service de santé des hôpitaux militaires, de la réforme de 1747 à 1789 », dans Gouvernance, Justice et Santé, à paraître très prochainement, dont on peut trouver la pré-publication dans les archives ouvertes de HAL : https://hal.archives-ouvertes.fr/hal-01688606.

Devant à ce genre de projet, où la masse documentaire est importante, il faut commencer par analyser un échantillon des dossiers pour concevoir leur exploitation numérique.

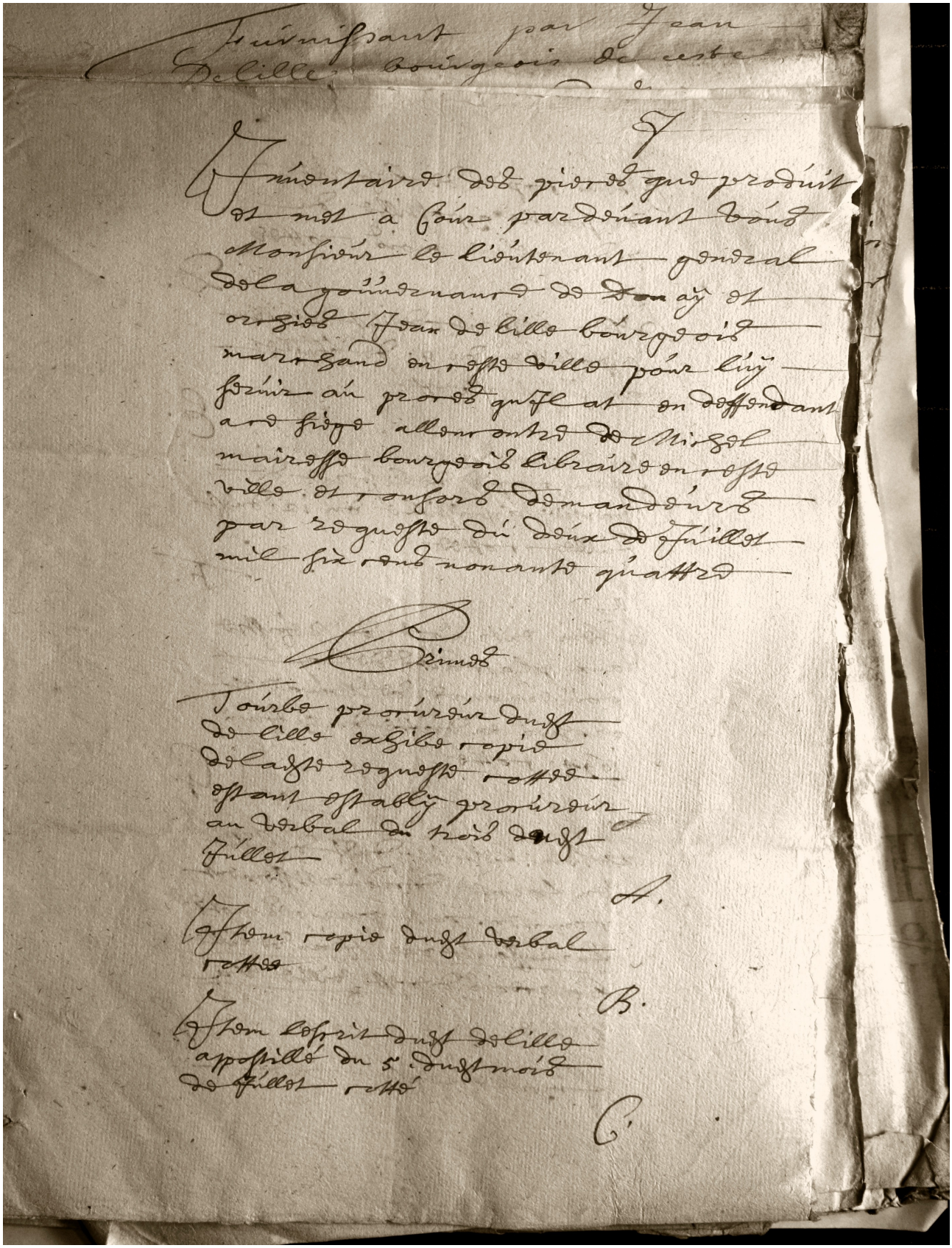
Si aux Archives départementales du Nord, il existe encore aujourd'hui un inventaire des 30 000 dossiers sous forme de fiches analytiques, ces dernières sont tellement lacunaires qu'elles restent quasiment inexploitables. On y trouve seulement, quand elles sont bien remplies, un ou deux noms de famille ou d'institution, une date et un mot clé.

L'échantillon a vite révélé une diversité de données tant sur leur forme que sur leur nature. Néanmoins un dossier type est apparu. Si celui-ci était constitué de nombreuses pièces, elles se présentaient toujours de la même manière. Ainsi ce dossier type contient un inventaire des pièces produites par le procureur des parties. Cet acte est adressé au conseiller-commissaire [8] du parlement ou son équivalent dans une juridiction inférieure, ce qui permet d'identifier rapidement le niveau hiérarchique juridictionnel. Le procureur prend soin de coter d'une lettre, A pour la première, B pour la deuxième, et ainsi de suite pour toutes les pièces transmises. Chaque pièce de l'inventaire est nommée selon le vocabulaire procédural[9] et judiciaire de l'époque. On trouve donc souvent une « commission de procureur », c'est-à-dire l'acte par lequel une partie choisit son procureur et lui donne mandat d'agir pour elle. Puis, dans l'ordre procédural et selon que la partie soit demanderesse ou défenderesse, on trouve la requête introductive, un placet[10], des contredits, des répliques et enfin des dupliques.

8 RENAUD LIMELETTE, « Recherche sur le conseiller-commissaire au parlement de Flandre », *C@hiers du CRHIDI* [En ligne], Vol. 41 – 2018, URL : https://popups.uliege.be:443/1370-2262/index.php?id=595.

9 Pour l'analyse de cette procédure, v. RENAUD LIMELETTE, *op. cit.*, Requête à la cour : causes privilégiées ou instruction sommaire, https://popups.uliege.be/1370-2262/index.php?id=595#tocto3n5.

10 Acte par lequel le requérant demandait au conseiller-commissaire de fixer un jour de comparution.



Extrait de l'inventaire du dossier 8 B 1 / 1846 des Archives départementales du Nord

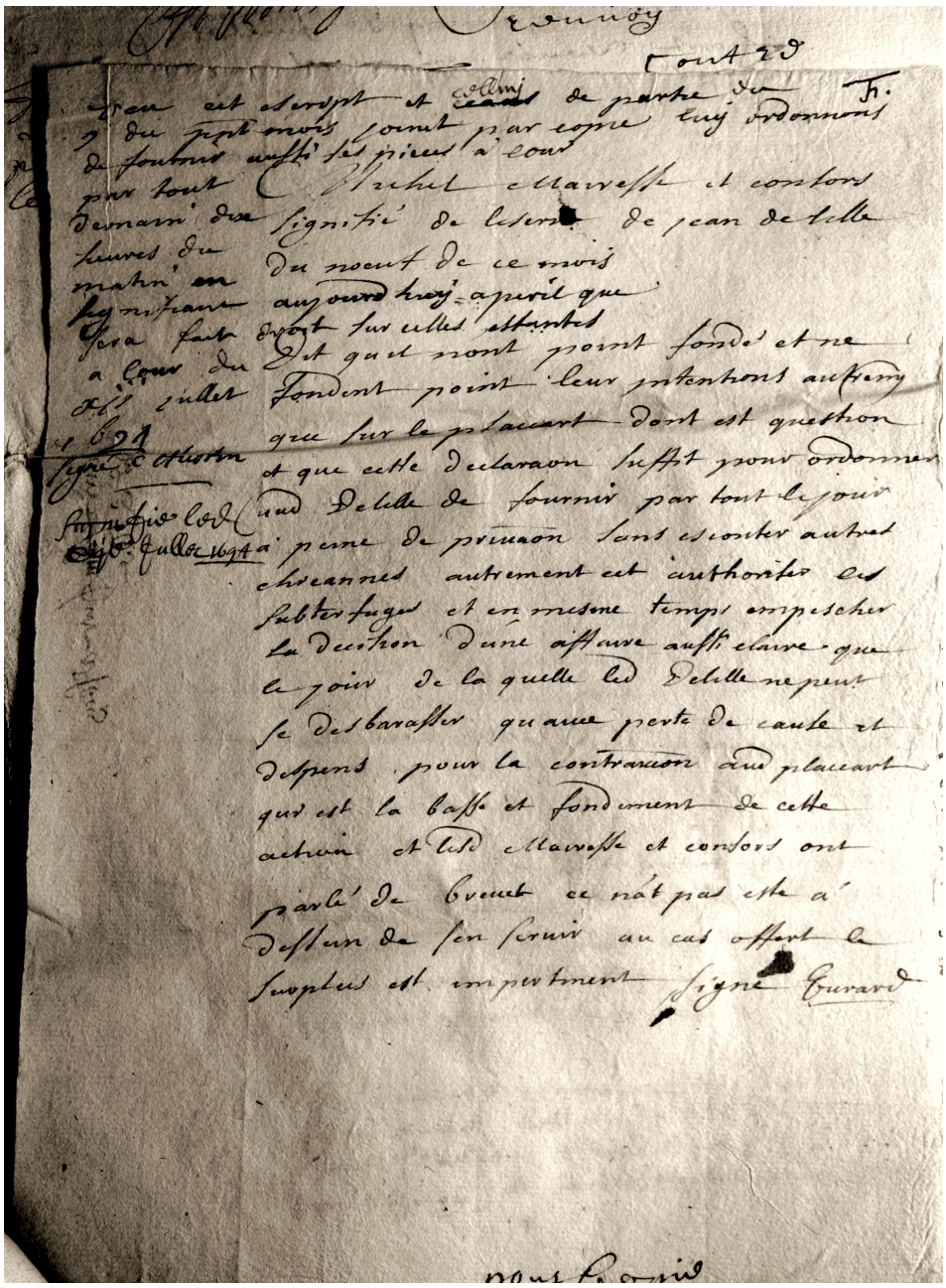
À côté de ces pièces procédurales, on trouve également bien souvent des pièces probatoires, servant d'appui à l'argumentation des parties pour soutenir juridiquement leurs dires. Là, la nature des pièces est très variable, elle dépend du type de contentieux. C'est ici qu'on trouve, sans être exhaustif, des testa-

ments, des contrats de toutes espèces et/ou des registres d'imposition. Bref tout ce qui traduit la vie ordinaire aux XVIIe et XVIIIe siècles.

En marge de ce dossier type d'autres dossiers sont beaucoup moins structurés. Soit ils ne renferment qu'une à plusieurs pièces, parmi celles précédemment évoquées, et on peut alors les rattacher à un autre dossier mieux structuré, soit les pièces sont éparses et parfois forts anciennes, si bien qu'il est difficile d'identifier dans quel autre dossier elles s'insèrent.

Pour être complet, on ajoutera que parfois une pièce peut en contenir une autre intrinsèquement. C'est le cas des apostilles, qui sont la marque, et en même temps un acte procédural distinct, que la cour a pris en compte la pièce fournie. Par exemple, le conseiller peut écrire en marge « *Reçu le* [*…*] *transmis le* », ou un greffier atteste du paiement des frais de procédure[11].

11 LAURIE FRÉGER, « Les épices au parlement de Flandre : pratiques singulières ? », *Revue du Nord*, 2009/4 (n° 382), p. 847-866. DOI : 10.3917/rdn.382.0847. URL : https://www.cairn.info/revue-du-nord-2009-4-page-847.htm.



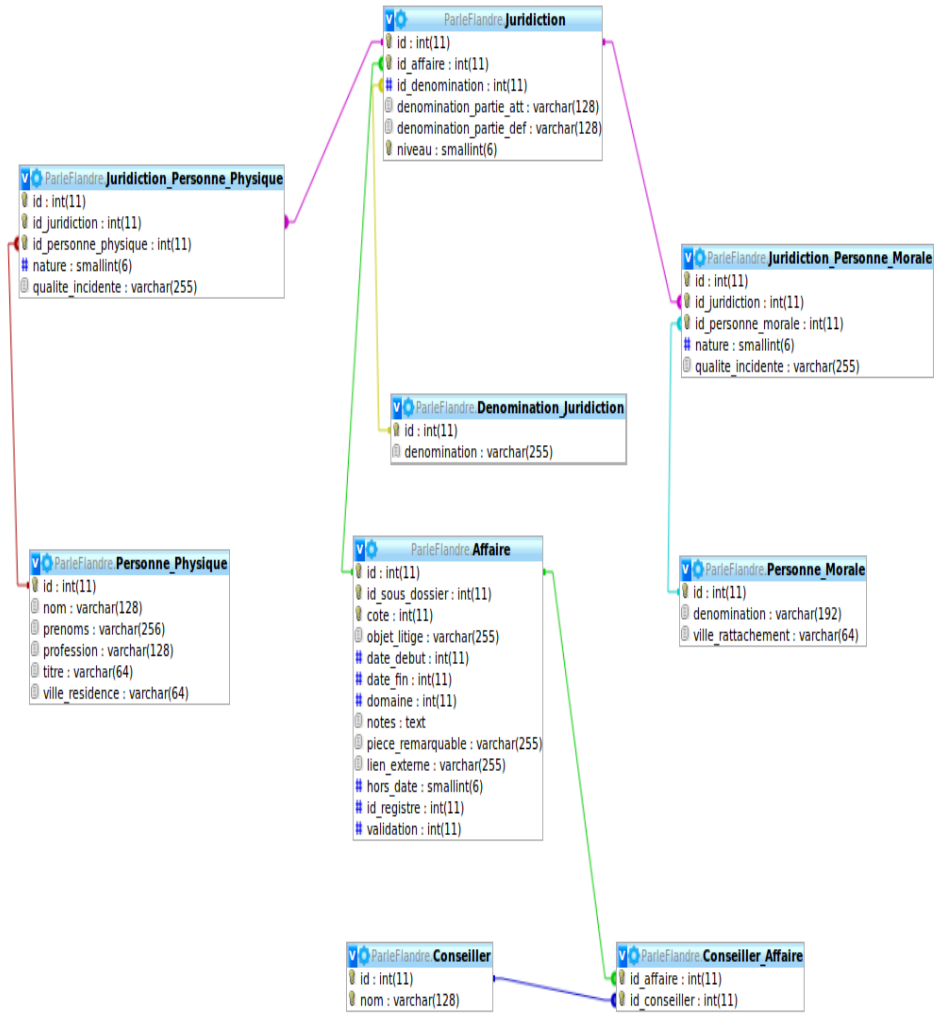
Exemple d'apostille en marge d'un acte, dossier 8 B 1 / 1846 des Archives départementales du Nord

Ainsi donc, c'est un écheveau de pièces plus ou moins structurées, et dont la nature est plus ou moins disparate, qu'il a fallu présenter numériquement. C'est le but du projet ParleFlandre, sous-titré « constitution d'un outil adapté à une exploitation scientifique [12] ». Très vite la solution qui s'est dégagée fut la constitution d'une base de données, dont il fallait dresser le schéma.

Celui-ci, outre ses aspects informatiques, traduit la granularité qui a été retenue. Pour des raisons de temps de dépouillement, d'analyse et de mise en forme numérique, le parti a été pris de concevoir les tables et les champs de la

12 RENAUD LIMELETTE et SABRINA MICHEL, « L'affaire est dans la base ! L'exemple du contentieux du parlement de Flandre (1668-1790) », dans HARALD DECEULAER, SÉBASTIEN DUBOIS et LAETIZIA PUCCIO (dir.), *L'affaire est dans le sac ! Dossiers de procès d'Ancien Régime et perspectives de recherche historique*, Bruxelles, Archives générales du Royaume, Studia, 2014, p.131-152.

base à partir d’un dossier type, en laissant de côté les apostilles mais en conservant la possibilité d’insérer dans la base les dossiers moins structurés qu’un dossier type.



Graphe des dépendances réalisé par François Périchon, IR au CRI de l'Université Lille 2, avec le concours de Mehdi Aksil, AI contractuel au CRI, et Renaud Limelette, IE en analyse de sources au CHJ.

Graphe des dépendances | Base de données ParleFlandre

Détails de l'affaire portant la cote 3224

Cote	3224
Numéro de sous-dossier	0
Objet du litige	Liquidation de dettes, tutelle
Année de début	1693
Année de fin	1694
Notes	Dossier de 21 pièces fournies par Jean Joyeux en appel dont des motifs de droits et pièces de première instance
Pièces remarquables	
Lien externe	

Conseillers

Nom

Aucun résultat trouvé.

Juridictions

Première Instance: Echevinage de la seigneurie de Renescure

Partie 1 - Qualité dans le procès: *Demandeur*

Total de 1 résultat(s).

Dénomination	Qualité incidente	profession / fonction	titre	domicile
Joyeux Jean		Marchand	Bourgeois	Saint-Omer

Partie 2 - Qualité dans le procès: *Défendeur*

Total de 1 résultat(s).

Dénomination	Qualité incidente	profession / fonction	titre	domicile
Fernagut Jacques		Laboureur		Resnescure

Deuxième Instance: Noble cour de Cassel

Partie 1 - Qualité dans le procès: *Appelant*

Total de 1 résultat(s).

Dénomination	Qualité incidente	profession / fonction	titre	domicile
Fernagut Jacques		Laboureur		Resnescure

Partie 2 - Qualité dans le procès: *Intimé*

Total de 1 résultat(s).

Dénomination	Qualité incidente	profession / fonction	titre	domicile
Joyeux Jean		Marchand	Bourgeois	Saint-Omer

Exemple de vue d'une affaire de la base de données ParleFlandre, 8 B 1 / 3224

Reste sur ce premier projet à dresser deux remarques conclusives. Un financement par L'Agence Nationale de la Recherche ne dure que quelques années. Or, pour ce projet, la seule analyse de l'échantillon et la création de la base de données ont épuisé ces maigres années. Pourtant un nombre conséquent de personnel[13] a été mobilisé, mais l'exploitation d'un tel corpus dépasse l'entendement, on comprend mieux pourquoi ces archives ont dormi pendant une cen-

13 Sur la genèse du projet et les difficultés rencontrées par le personnel affecté, v. STÉPHANE BAUDENS et ALEXANDRE JEANNIN, « Les pièces de procédure des archives du parlement de Flandre : rapport d'activité », *Revue du Nord*, 2009/4 (n° 382), p. 739-744. DOI : 10.3917/rdn.382.0739. URL : https://www.cairn.info/revue-du-nord-2009-4-page-739.htm.

taine d'années[14]. La dernière remarque porte sur la nature des pièces et l'objet de la base de données. La base de données n'expose pas les arrêts prononcés par le parlement de Flandre, car les pièces conservées dans la sous-série 8 B 1 sont antérieures au prononcé de l'arrêt, et ces arrêts sont conservés dans la sous-série 8 B 2.

1. 2. De plusieurs dizaines de livres papier à plusieurs milliers de ressources numériques : le projet Meta-parleFlandre

Le second projet reposant sur le parlement de Flandre s'intitule Meta-parleFlandre[15]. Il touche à l'exploitation de la littérature juridique imprimée et produite par les magistrats du parlement. Cette littérature juridique se décline en divers recueils. On pense de suite, pour les habitués du monde judiciaire moderne aux recueils d'arrêts[16]. À côté de ceux-ci, il est aussi tous les recueils sur l'enregistrement des édits royaux et tous les arrêts de règlement produits par le parlement de Flandre.

Par rapport au projet ParleFlandre, dans le projet Meta-parleFlandre les sources mises en valeur sont livresques. Ainsi la première étape de ce projet a été d'identifier les ouvrages afin de les numériser. Or en 2009, la Bibliothèque nationale de France menait ses deuxièmes journées professionnelles sur la numérisation et la valorisation concertées en sciences juridiques. Elle proposait, pour compléter ses propres collections, de financer une part de la numérisation de fonds locaux. Ainsi à l'époque, la bibliothèque numérique scientifique de la réserve patrimoniale des ex-universités de Lille, nommée PôLiB, avait obtenu le financement de la numérisation de 43 ouvrages conservés dans ses fonds et ceux de l'Institut catholique de Lille. Parmi ces ouvrages repérés et numérisés, de nombreux ouvrages touchent au parlement de Flandre et forment le corpus du projet Meta-parleFlandre[17]. Une phase essentielle dans ce travail de numéri-

14 ROSINE CLEYET-MICHAUD, « Le fonds du parlement de Flandre : historique de la conservation », *Revue du Nord*, 2009/4 (n° 382), p. 683-685. DOI : 10.3917/rdn.382.0683. URL : https://www.cairn.info/revue-du-nord-2009-4-page-683.htm ; VÉRONIQUE DEMARS-SION, « « Heurts » et malheurs d'un fonds : les tribulations des archives du parlement de Flandre », *Revue du Nord*, 2014/4 (n° 407), p. 829-858. DOI : 10.3917/rdn.407.0829. URL : https://www.cairn.info/revue-du-nord-2014-4-page-829.htm.

15 https://parleflandre.nakalona.fr.

16 On lira à propos sur ces recueils d'arrêt GERALDINE CAZALS, *L'arrestographie flamande : jurisprudence et littérature juridique à la fin de l'Ancien Régime (1668-1789)*, Genève, Droz, 2018.

17 *Recueil des édits déclarations arrests et réglemens qui sont propres et particuliers aux Provinces du ressort du Parlement de Flandres*, Douai, Jacques-François Willerval, 1730. URI : http://nordnum.univ-lille.fr/ark:/72505/a0115224001366EQToQ/from/a011522400136C4fuAQ. PIERRE-ANTOINE-SAMUEL-JOSEPH PLOUVAIN et PHILIPPE-JOSSE-AUGUSTE SIX, *Recueil des Édits, Déclarations, Lettres-Patentes, etc. enregistrés au parlement de Flandres, des arrêts du Conseil d'État particuliers à son ressort ; ensemble des règlemens rendus par cette cour depuis son érection en conseil souverain à Tournay*, Douai, Derbaix, 1785-1790, 12 tomes. URI : http://nordnum.univ-lille.fr/ark:/72505/a011522400136C4fuAQ. JACQUES POLLET, *Arrests du Parlement de Flandre sur diverses questions de droit, de coutume, et de pratique*, Lille, Liévin Danel, 1716. URI : http://nordnum.univ-lille.fr/ark:/72505/a0115217019356v106i/from/a011522400136C4fuAQ. *Recueil d'Arrêts du parlement de Flandres*, Lille, Jean-Baptiste-Joseph Henry, 1773, 2 tomes. LOUIS PILLOT, *Histoire du Parlement de Flandres*, Douai, Adam d'Aubers, 1849, 2 tomes. URI :

sation fût de coupler la numérisation en mode image avec la possibilité d'interroger le texte de chaque document. Pour ce faire « l'océrisation » a été complétée par une correction manuelle pour obtenir une version texte fiable à 99 %. En effet, les graphies particulières des imprimeurs de l'époque sont mal prises en compte par une simple « océrisation », et si l'on veut un grain assez fin pour rechercher dans ce corpus d'ouvrages il faut passer du temps à reprendre chaque mot pour avoir une version texte fiable. Présenté d'abord sur PôLiB, ce corpus sur les sources du droit en Flandre est dorénavant accessible sur NordNum[18] et sur Gallica.

Si la numérisation des ouvrages était une étape essentielle, elle n'en était pas l'aboutissement. En effet, naviguer à l'intérieur d'un recueil d'arrêts ou d'édits de plusieurs centaines de pages n'est pas chose aisée, même si le site de consultation dispose d'un bon moteur de recherche. En fait, l'intérêt de Meta-parleFlandre est de faciliter la recherche pour le lecteur. Au lieu de consulter un ouvrage dans son entièreté, l'ouvrage est divisé en unités logiques selon la nature des données. Ainsi on trouve sur Meta-parleFlandre des arrêts rendus par le parlement de Flandre, des édits enregistrés au parlement de Flandre et des arrêts de règlement pris par le parlement de Flandre. En d'autres termes, la numérisation des ouvrages avait abouti à constituer des grains numériques trop gros, il a fallu moudre ce grain pour séparer chaque arrêt, chaque édit et chaque règlement de l'ouvrage initial dans lequel ils étaient rassemblés. Ainsi profitant de la numérisation de chaque ouvrage en PDF, le projet Meta-parleFlandre se construit autour de la dissociation des pages des PDF pour les réassembler en unités juridiques ou judiciaires[19].

Une fois le lourd travail de définition de la granularité des données terminé une autre tâche est à aborder, celle de la conception de la diffusion des données.

2. Les métadonnées : enjeu de diffusion

L'ancienneté relative du projet ParleFlandre (2007) n'a pas poussé les porteurs de ce projet ANR à se questionner sur la nécessité de mettre en place des métadonnées, si bien qu'à ce jour aucune métadonnée n'est associée aux affaires judiciaires alimentant la base de données.

À l'inverse le projet Meta-parleFlandre a de suite pris en compte cette problématique. S'il y a encore quelques années les solutions trouvées pour les projets en Sciences Humaines Sociales étaient souvent particulières à chaque projet, aujourd'hui depuis le déploiement de la Très Grande Infrastructure de Recherche Huma-Num les solutions sont plus universelles.

---

http://nordnum.univ-lille.fr/ark:/72505/a011544092129DYRL7H.

18 L'intégralité du projet négocié avec la BnF est accessible sous l'URI suivant : http://nordnum.univ-lille.fr/ark:/72505/a011521028444a7MYpO/from/a011522400136C4fuAQ.

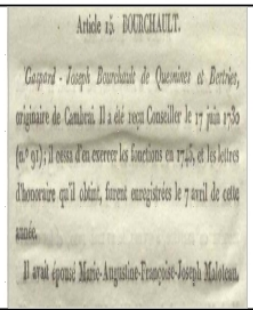
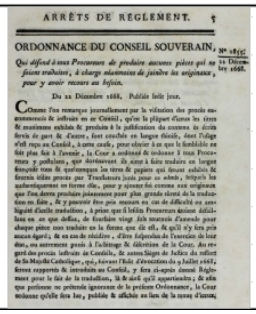
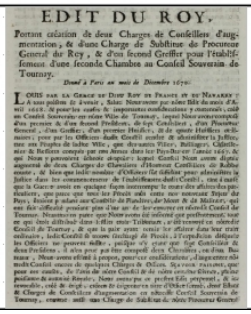
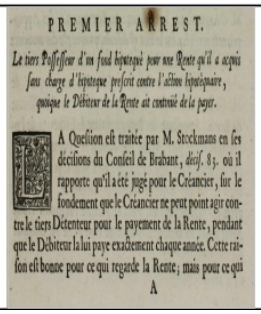
19 C'est-à-dire que maintenant il y a autant de PDF que d'arrêt, d'édits et de règlements.

Si la grille de service a permis de solutionner les questions relatives au stockage, à la diffusion, au signalement et à l'exposition des données issues, au centre de toutes ces problématiques il est un dénominateur commun : la description des données selon un référentiel de métadonnées.

Pour Meta-parleFlandre nous verrons qu'à chaque donnée sont associées des métadonnées au format Dublin Core, et que ces métadonnées sont au cœur de la synergie du projet.

2. 1. Le vocabulaire Dublin Core et les données de Meta-parleFlandre

Comme je l'ai déjà énoncé précédemment [20], quatre natures de données alimentent le site Meta-parleFlandre : des fiches biographiques des magistrats, des arrêts de règlement produits par le parlement de Flandre, des édits ou autres normes enregistrés au parlement, et de la jurisprudence commentée.

Fiches biographiques de magistrats	Arrêts de règlement	Édits ou autres normes	Jurisprudence commentée
			

Exemple de représentation des quatre natures de données dans Meta-parleFlandre

Pour un humain, un lecteur commun du site, il est presque instinctif de reconnaître la nature de chaque document : une lecture rapide, même sans grande attention, des caractères gras suffit souvent pour identifier un document. Mais dès que l'on commence à vouloir affiner l'analyse, l'humain n'est plus capable d'embrasser d'un seul coup toutes les informations, et plus il y a de données plus l'affaire se complique.

C'est ici que le passage au numérique prend tout son intérêt. Encore faut-il ne pas se limiter à numériser les ouvrages anciens accueillants les fiches biographiques, les arrêts de règlement, les édits et la jurisprudence de la cour. En plus de « l'océrisation » et de la division des ouvrages en autant d'unités logiques[21] qu'il y a d'arrêts, d'édits ou de magistrats, la description précise de chaque unité logique par un référentiel de métadonnées permet à la machine d'interpréter très rapidement chaque donnée[22].

20 *Supra*, 1. 2. D'une dizaine de livres papier à plusieurs milliers de documents numériques : le projet Meta-parleFlandre.

21 *Supra*, note 2.

22 Sur le sujet de l'intérêt des métadonnées et des vocabulaires contrôlés, v. VANDA BROUGHTON, « Emergent vocabulary control in Web 2.0. Comparisons with conventional LIS theory and

Ainsi dans le projet Meta-parleFlandre chaque donnée est décrite selon le vocabulaire Dublin Core. Parmi les éléments et qualificatifs du vocabulaire Dublin Core, nous avons retenu le schéma suivant : titre, sujet, type de ressource, source, couverture, créateur, contributeur, identifiant et date, avec les qualificatifs date de création et date de parution.

Arrêtons-nous maintenant aux résolutions que nous avons faites par ordre de difficulté.

— Pour le « titre », aucune difficulté n'a été rencontrée dans la mesure où chaque donnée avait déjà un titre dans sa version bibliographique[23].

— L'URI de consultation de l'ouvrage d'où est tirée la ressource sur la plateforme NordNum a servi à remplir l'élément source[24].

— Comme toutes les données sont tirées de livres publiés, l'élément « type de ressource » n'a pas donné lieu à discussion[25].

— Nous avons choisi d'utiliser l'élément « contributeur » pour y mettre le nom du laboratoire porteur du projet, sous la forme [26] « Centre d'Histoire Judiciaire (CHJ) – UMR 8025 » ; le but est de bien dissocier dans la vie de la donnée les créateurs orignaux de la version bibliographique des créateurs secondaires qui en font une édition numérique.

— La distinction de l'élément « date » par les raffinements « date de création » et « date de parution » permet de différencier la date originelle de la ressource avec la date de mise en ligne[27].

— Si l'élément « couverture » pouvait poser quelques problèmes, compte tenu de la variabilité du ressort du parlement de Flandre [28], la difficulté a été vaincue en retenant simplement « Parlement de Flandre » comme le permet le Dublin Core Metadata Initiative[29].

— L'élément « créateur » a pu sembler facile à utiliser puisque chacun des ouvrages d'où proviennent les données est identifiable par des auteurs. Pour autant la manière de référencer les auteurs restait incertaine, le Dublin Core Metadata Initiative laissant beaucoup de place à l'interprétation [30]. Pour identifier chaque auteur nous avons retenu la dénomination retenue par l'International

---

practice », *Les Cahiers du numérique*, 2010/3 (Vol. 6), p. 49-75. URL : https://www.cairn.info/revue-les-cahiers-du-numerique-2010-3-page-49.htm.

23 http://purl.org/dc/terms/title.

24 Le Dublin Core Metadata Initiative (ci-après DCMI) recommande d'identifier la ressource au moyen d'une chaîne de caractères conforme à un système d'identification formel, v. http://purl.org/dc/elements/1.1/source. Or NordNum remplit cette recommandation en utilisant pour chaque ressource des URIs (Uniforme Ressource Identifier) de type ARK (Archival Ressource Key), v. http://n2t.net/e/ark_ids.html.

25 Nous avons retenu à chaque fois l'étiquette « Text » comme le recommande le DCMI à propos des livres, v. http://purl.org/dc/dcmitype/Text.

26 Le DCMI reconnaissant qu'il faut entendre comme contributeur « l'entité chargée d'apporter des contributions à la ressource », v. http://purl.org/dc/terms/contributor.

27 http://purl.org/dc/terms/date.

28 RENAUD LIMELETTE, À la recherche de son juge dans le ressort du parlement de Flandre, SERGE DAUCHY (dir.), *Les cahiers du Centre de Recherches en Histoire du Droit et des Institutions*, Bruxelles, Presses de l'Université Saint-Louis, 2009, 31, p. 29-46.

29 http://purl.org/dc/terms/coverage.

30 http://purl.org/dc/elements/1.1/creator.

Standard Name identifier [31] que suit la Bibliothèque nationale de France dans ses notices d'autorité[32]. Ainsi chaque auteur est identifié par son nom, son prénom et ses dates de vie et de mort entre parenthèses.

— L'un des éléments qui a demandé le plus de réflexion est l'élément « sujet ». Instinctivement on comprend que sous l'élément « sujet » se trouvent des mots-clés[33]. Or, on conviendra, que si on laisse à plusieurs spécialistes de la même discipline le soin de définir par des mots-clés les pans de leur discipline, on ne trouvera pas au final une liste de mots-clés arrêtée[34] (figée). Au contraire, chaque spécialiste ayant sa conception et sa représentation de sa discipline, on tombera inévitablement dans une cacophonie littérale, une logorrhée plus ou moins subjective.

Prenons à dessein un exemple juridique, que chacun comprendra aisément. Si on vous demande de décrire par un seul mot-clé la situation pénale suivante, « M. X, après mûre réflexion, a tué sa mère », il y a de grandes chances que les réponses fournies soient très différentes les unes des autres. J'aurais sans doute, meurtre, homicide, assassinat, homicide volontaire, crime, matricide. C'est la même chose avec les visiteurs d'un site internet, quand ils utilisent un moteur de recherche ils n'ont pas tous le même référentiel de mots, et donc faute d'avoir rendu public ce référentiel de mots-clés, les résultats proposés seront incertains ou incomplet. Il en est de même pour une machine qui doit interpréter et classer des données du web.

À travers cet élément « sujet » c'est tout le web sémantique qui transparaît. C'est pourquoi les recommandations du DCMI prennent soin d'insister sur la nécessité d'utiliser un vocabulaire contrôlé [35]. Reste ici à préciser qu'un vocabulaire contrôlé ne se développe pas seul, au contraire il doit être mis en place par des spécialistes chargés de choisir les termes descripteurs [36] les plus représentatifs d'un domaine. Il fallait donc pour le projet Meta-parleFlandre trouver un vocabulaire contrôlé le plus adéquat. Après de nombreuses recherches, nous avons retenu celui proposé par le Ministère de la culture pour décrire et indexer les archives historiques locales[37].

— Enfin vient l'élément « identifiant » qui permet d'identifier la donnée. La DCMI recommande ici [38] d'utiliser une chaîne de caractère reconnue dans un

31 http://www.isni.org.

32 https://catalogue.bnf.fr/recherche-autorite.do.

33 http://purl.org/dc/terms/subject.

34 Nous reprenons à notre compte la réflexion de Michel Villey sur le sens des mots dans ses *Carnets* : *Nous voici placés devant le monde des abominables mots qui poursuivent à travers le temps leur tourbillon désordonné. La plupart de nos contemporains nagent à leur aise dans cette extrême confusion des mots, dont le sens n'est plus définissable. Je ne puis construire sur cette glaise, et sens le besoin de restituer aux mots leur sens qui est objectif, assis sur l'étymologie qui ets le témoin de leur formation spontanée conforme à l'ordre naturel*. V. M. VILLEY, *Réflexions sur la philosophie et le droit. Les Carnets*, textes préparés et indexés par Anne-Marie FRISON-ROCHE et CHRISTOPHE JAMIN, Paris, Presses Universitaires de France, 1995, p. 176 et 177.

35 *Supra* note 24.

36 Pour une approche synthétique on se reportera à MICKAËL COUSTATY, NORBERT TSOPZE, KARELL BERTET *et al.*, « Traitement des documents anciens à l'aide d'ontologie », *Les Cahiers du numérique*, 2012/3 (Vol. 8), p. 91-118. DOI : 10.3166/LCN.8.3.91-118. URL : https://www.cairn.info/revue-les-cahiers-du-numerique-2012-3-page-91.htm.

37 https://francearchives.fr/article/37828#.

38 http://purl.org/dc/elements/1.1/identifier.

système d'identification formel. Pour Meta-parleFlandre l'identifiant retenu est la fin d'un URI (Uniforme Ressource Identifier) de type handle [39]. L'URI provient du réservoir de données de Meta-parleflandre qu'est Nakala. En effet chaque donnée enregistrée dans Nakala se voit attribuer une référence unique et pérenne. L'intérêt est de pouvoir consulter la donnée sans aller dans les pages publiques de Nakala ou de l'exposition du projet. Ainsi chaque donnée de Meta-parleFlandre a un URI construit de cette façon :

— http://www.nakala.fr/data/identifiant de la donnée

— http://hdl.handle.net/identifiant de la donnée

Ces deux URL·s permettent d'accéder directement à la donnée, c'est-à-dire au PDF décrit à travers le vocabulaire Dublin Core.

On peut se demander quel est le but de toutes ses métadonnées. L'un des avantages du Dublin Core est de rendre le corpus de données interopérable [40], mais avec le moteur de recherche nativement présent dans Omeka il permet des recherches beaucoup plus ciblées. En effet, l'effort investi par l'utilisation de vocabulaires contrôlés donne ici toute son efficacité. La recherche avancée sur les contenus enregistrés dans Omeka permet de lancer une requête à travers les champs du Dublin Core retenus en configurant la recherche avec des opérateurs de type : « contient », « ne contient pas », « est exactement », « est vide », « n'est pas vide », « commence par » ou « fini par ». On peut bien entendu croiser la recherche avec un autre champ du Dublin Core en appliquant un opérateur booléen de type « et ou « ou [41] ». C'est là un des gros avantage des métadonnées quand leur schéma est bien conçu[42].

2. 2. La synergie du projet Meta-parleFlandre et le web sémantique

Quand on construit un projet numérique [43], outre les aspects de recherche scientifique attachés au projet, on désire que celui-ci soit visible sur la toile, sans avoir à fouiller le tréfonds des pages de résultats de son moteur de recherche. Ainsi pour remonter dans les pages de résultats, il fallait, il y a encore quelques années, multiplier les liens hypertextes pointant sur son propre site.

---

39 https://www.handle.net/index.html.

40 *Infra*, note 51.

41 https://omeka.org/classic/docs/GettingStarted/Searching.

42 À l'inverse, à défaut de se conformer à au moins un vocabulaire contrôlé, une recherche dans la base de données parleflandre renvoie des résultats partiellement inexploitables si on n'y prend pas garde. Par exemple, la requête « dette » dans le champ « Objet du litige » agrège tous les objets du litige où il a été saisi soit « dette » soit « dettes », alors que la requête « dettes » ne renvoie que les objets du litige où seul « dettes » a été saisi.

43 Un guide pratique accompagne les chercheurs et les ingénieurs dans les étapes à suivre pour mener à bien un projet numérique, M AUD INGARAO et SAMANTHA SAÏD, *Guide pratique pour la production de corpus numérique* : http://mutec.huma-num.fr/sites/www.mutec-shs.fr/files/Guide%20pratique%20pour%20la%20production%20de%20corpus%20num%C3%A9rique_0.pdf.

Aujourd'hui, on conviendra que dans nos disciplines scientifiques ce n'est pas le PageRank[44] qui devrait compter, mais la qualité scientifique du site. Ainsi, dès lors que le PageRank n'apporte qu'une vue quantitative des personnes visitant un site il ne devrait pas être un motif d'évaluation d'un projet numérique. Autrement dit, ce n'est pas le nombre de visiteurs qui compte mais quelle motivation les a poussés à visiter le site[45].

Pour faire valoir la qualité scientifique de son projet numérique, encore faut-il avoir les outils adéquats.

Dès l'initiation du projet Meta-parleFlandre, l'interopérabilité des données a été prise en compte. Pour cette raison, nous avons opté pour le pack Nakalona proposé par la TGIR Huma-Num.
Le système de gestion de documents numériques Omeka permet de nous concentrer sur l'édition des données.

Déjà sous Omeka les données peuvent être affichées sous des formats facilitant la compréhension du corpus par une machine et l'indexation dans les moteurs de recherche[46]. En effet, Omeka propose des sorties dans différents formats[47] dont on mesurera l'intérêt.


```
▼ items:
  ▼ 0:
      item_type_id:    null
      collection_id:   15
      featured:        0
      public:          1
      added:           "2019-07-03 10:55:49"
      modified:        "2019-07-04 14:37:00"
      owner_id:        2
      id:              707
```


Exemple de données au format JSON exposées dans Omeka, relatif à un Arrêt du parlement, concernant les salaires des Huissiers pour les devoirs qu'ils font au profit de leur bourse commune[48]

Sans être du web sémantique JSON est bien connu comme format d'échange de données[49]. Dans le cas ci-dessus, rien n'est très explicite sans savoir que l'id 707 est l'item intitulé « Arrêt du parlement, concernant les salaires des Huis-

44 Sur le rapport entre le PageRank et le référencement des moteurs de recherche, v. MATHIEU ANDRO, MARION CHAIGNE et FRANCK SMITH, « Valoriser une bibliothèque numérique par des choix de référencement et de diffusion. L'expérience de la bibliothèque Sainte-Geneviève », *Les Cahiers du numérique*, 2012/3 (Vol. 8), p. 75-90. DOI : 10.3166/LCN.8.3.75-90. URL : https://www.cairn.info/revue-les-cahiers-du-numerique-2012-3-page-75.htm, notamment le paragraphe 2.2 – Tenir compte des algorithmes de référencement des moteurs de recherche pour améliorer la visibilité de la bibliothèque numérique.

45 On lira ici avantageusement, DOMINIQUE CARDON, « Dans l'esprit du PageRank. Une enquête sur l'algorithme de Google », *Réseaux*, 2013/1 (n° 177), p. 63-95. DOI : 10.3917/res.177.0063. URL : https://www.cairn.info/revue-reseaux-2013-1-page-63.htm.

46 Bien qu'Omeka s'appuie essentiellement sur le Dublin Core, un article récent montre l'écart qu'il peut y avoir dans la conception du Dublin Core entre créateurs de la norme et utilisateurs de celle-ci, v. DEBORAH MARON et MELANIE FEINBERG, « What does it mean to adopt a metadata standard ? A case study of Omeka and the Dublin Core », dans *Journal of Documentation*, vol. 74, n° 4, 2018, p. 674-691, https://doi.org/10.1108/JD-06-2017-0095.

47 C'est grâce à la fonction ContextSwitch du framework Zend framework, sur lequel s'appuie Omeka, qu'une ressource peut avoir plusieurs représentations. V. JULIEN PAULI et GUILLAUME PONÇON, *Zend Framewok. Bien développer en PHP*, Les cahiers du programmeur, Eyrolles, 2008, p. 136 *sqq*.

48 https://parleflandre.nakalona.fr/items/show/707?output=json.

siers pour les devoirs qu'ils font au profit de leur bourse commune ». Pour plus d'expressivité et dans une approche de web sémantique on préférerait une exposition en JSON-LD[50], le nouveau standard du W3C[51].


```
−<rdf:RDF>
 −<rdf:Description rdf:about="https://parleflandre.nakalona.fr/items/show/707">
  −<dcterms:title>
     Arrêt du parlement, concernant les salaires des Huissiers pour les devoirs qu'ils font au profit de leur bourse commune
   </dcterms:title>
  −<dcterms:subject>
     <a href="/items/browse?advanced%5B0%5D%5Belement_id%5D=49&advanced%5B0%5D%5Btype%5D=is+exactly&advanced%5B0%5D%5Bterms%5D=Huissier+de+justice">Huissier de justice</a>
   </dcterms:subject>
  −<dcterms:subject>
     <a href="/items/browse?advanced%5B0%5D%5Belement_id%5D=49&advanced%5B0%5D%5Btype%5D=is+exactly&advanced%5B0%5D%5Bterms%5D=Organisation+judiciaire">Organisation judiciaire</a>
   </dcterms:subject>
  −<dcterms:subject>
     <a href="/items/browse?advanced%5B0%5D%5Belement_id%5D=49&advanced%5B0%5D%5Btype%5D=is+exactly&advanced%5B0%5D%5Bterms%5D=R%C3%A9mun%C3%A9ration">Rémunération</a>
   </dcterms:subject>
   <dcterms:creator>Plouvain, Pierre-Antoine-Samuel-Joseph (1754-1832)</dcterms:creator>
   <dcterms:creator>Six, Philippe-Josse-Auguste-Joseph (1732-1793)</dcterms:creator>
  −<dcterms:source>
     http://nordnum.univ-lille.fr/ark:/72505/a0115224001360im5OI/from/a011522400136ypxWI2
   </dcterms:source>
  −<dcterms:created>
     <a href="/items/browse?advanced%5B0%5D%5Belement_id%5D=56&advanced%5B0%5D%5Btype%5D=is+exactly&advanced%5B0%5D%5Bterms%5D=14%2F05%2F1725">14/05/1725</a>
   </dcterms:created>
  −<dcterms:issued>
     <a href="/items/browse?advanced%5B0%5D%5Belement_id%5D=60&advanced%5B0%5D%5Btype%5D=is+exactly&advanced%5B0%5D%5Bterms%5D=01%2F07%2F2019">01/07/2019</a>
   </dcterms:issued>
  −<dcterms:issued>
     <a href="/items/browse?advanced%5B0%5D%5Belement_id%5D=60&advanced%5B0%5D%5Btype%5D=is+exactly&advanced%5B0%5D%5Bterms%5D=1790">1790</a>
   </dcterms:issued>
   <dcterms:contributor>Centre d'Histoire Judiciaire (CHJ) - UMR 8025</dcterms:contributor>
   <dcterms:type>Text</dcterms:type>
   <dcterms:identifier>11280/afe3b32f</dcterms:identifier>
   <dcterms:coverage>Parlement de Flandre</dcterms:coverage>
 </rdf:Description>
```


Exemple de données exposées au format DC-RDF, relatif à un Arrêt du parlement, concernant les salaires des Huissiers pour les devoirs qu'ils font au profit de leur bourse commune[52]

On reconnaît ici parfaitement les éléments du Dublin Core repérables à travers les balises forgées sur <dcterms : élément-Dublin-Core></dcterms : élément-Dublin-Core>. En d'autres termes, c'est à la fois compréhensible pour un humain, déjà initié au Dublin Core, et facilement interprétable par une machine[53].


```
−<itemContainer xsi:schemaLocation="http://omeka.org/schemas/omeka-xml/v5 http://omeka.org/schemas/omeka-xml/v5/omeka-xml-5-0.xsd" uri="https://parleflandre.nakalona.fr/items/browse?output=omeka-xml"
 accessDate="2019-10-03T14:52:17+02:00">
  −<miscellaneousContainer>
   −<pagination>
      <pageNumber>1</pageNumber>
      <perPage>50</perPage>
      <totalResults>486</totalResults>
    </pagination>
  </miscellaneousContainer>
 −<item itemId="707" public="1" featured="0">
  −<fileContainer>
   −<file fileId="432">
    −<src>
       https://parleflandre.nakalona.fr/files/original/f05e055089c302280c6d26c88290f0d7.pdf
     </src>
     <authentication>db2722c362fd519f5b8bd66a98e3a3d6</authentication>
    </file>
   </fileContainer>
  −<collection collectionId="15">
   −<elementSetContainer>
    −<elementSet elementSetId="1">
       <name>Dublin Core</name>
     −<description>
        The Dublin Core metadata element set is common to all Omeka records, including items, files, and collections. For more information see, http://dublincore.org/documents/dces/.
      </description>
     −<elementContainer>
      −<element elementId="50">
         <name>Title</name>
         <description>A name given to the resource</description>
       −<elementTextContainer>
        −<elementText elementTextId="9471">
          −<text>
             Ordonnances et arrêts de règlement du parlement de Flandre
           </text>
         </elementText>
       </elementTextContainer>
      </element>
     −<element elementId="43">
        <name>Identifier</name>
       −<description>
          An unambiguous reference to the resource within a given context
        </description>
```


49 Aurélie Sobréro, *Développer des services REST en Java – Échanger des données au format JSON*, ENI, 2014.

50 Markus Lanthaler et Christian Gütl, « On using JSON-LD to create evolvable RESTful services », dans *WS-REST '12 Proceedings of the Third International Workshop on RESTful Design*, Lyon, 2012, https://dl.acm.org/citation.cfm?doid=2307819.2307827. L'exposition en JSON-LD est disponible néanmoins avec Omeka S, https://omeka.org/s/docs/user-manual.

51 https://www.w3.org/2018/jsonld-cg-reports/json-ld.

52 https://parleflandre.nakalona.fr/items/show/707?output=dc-rdf.

53 C'est le sens du RDF, il facilite « l'exploitation et le traitement automatique », v. *Le guide des bonnes pratiques numériques*, http://www.huma-num.fr/ressources/guides.

Exemple de données au format Omeka-XML, relatif à un Arrêt du parlement, concernant les salaires des Huissiers pour les devoirs qu'ils font au profit de leur bourse commune[54]

Ici le format est plus explicite, car il détaille pour chaque terme du Dublin Core sa définition entre les balises <description></description>. C'est plus compréhensible pour un humain, même non initié au Dublin Core, mais trop volubile pour une machine.

Ainsi donc, Omeka fournit des formats de sortie qui selon leur fonction auront un intérêt spécifique. Par exemple, la sortie en XML diffère de la sortie RDF. La première met en avant la structure des données alors que la seconde s'attache aux liaisons logiques des données[55].

Grâce aux extensions Import et Export Omeka et Nakala sont reliés entre eux. Ainsi avec Nakala les métadonnées sont présentées publiquement sous un modèle RDF avec une sortie Turtle ou RDF/XML.

```
@prefix ore:    <http://www.openarchives.org/ore/terms/> .
@prefix geonames: <http://www.geonames.org/ontology#> .
@prefix void:   <http://rdfs.org/ns/void#> .
@prefix owl:    <http://www.w3.org/2002/07/owl#> .
@prefix xsd:    <http://www.w3.org/2001/XMLSchema#> .
@prefix yago:   <http://localhost:8080/class/yago/> .
@prefix skos:   <http://www.w3.org/2004/02/skos/core#> .
@prefix ir:     <http://www.ontologydesignpatterns.org/cp/owl/informationrealization.owl#> .
@prefix rdfs:   <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dbpedia: <http://localhost:8080/resource/> .
@prefix units: <http://dbpedia.org/units/> .
@prefix nkl:    <http://localhost:8080/pubby/> .
@prefix geo:    <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix p:      <http://localhost:8080/property/> .
@prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix prv:    <http://purl.org/net/provenance/ns#> .
@prefix doap:   <http://usefulinc.com/ns/doap#> .
@prefix meta:   <http://example.org/metadata#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix prvTypes: <http://purl.org/net/provenance/types#> .
@prefix foaf:   <http://xmlns.com/foaf/0.1/> .
@prefix dc:     <http://purl.org/dc/elements/1.1/> .

<https://www.nakala.fr/data/resource/11280/afe3b32f?output=ttl>
        rdfs:label          "RDF description of Afe3b32f" ;
        foaf:primaryTopic   <https://www.nakala.fr/resource/11280/afe3b32f> .

<https://www.nakala.fr/resource/11280/afe3b32f>
        a                   foaf:Document ;
        dcterms:creator     <https://www.nakala.fr/user/98839290-95d0-4ae7-8dee-f0c0b566efd1> ;
        dcterms:extent      "891179"^^xsd:long ;
        dcterms:format      "PDF" ;
        dcterms:identifier  "11280/afe3b32f" ;
        dcterms:issued      "2019-07-03T10:48:11.096+02:00"^^xsd:dateTime ;
        dcterms:modified    "2019-07-03T11:02:11.865+02:00"^^xsd:dateTime ;
        dcterms:publisher   <https://www.nakala.fr/account/11280/5f539340> ;
        ore:isAggregatedBy  <https://www.nakala.fr/collection/11280/7681d5cf> ;
        skos:altLabel       "2056.pdf" ;
        foaf:primaryTopic   <https://www.nakala.fr/data/11280/afe3b32f> ;
        foaf:sha1           "adba0df29e78a1249b7adea51c2b9d7f00ece70f" .

<https://www.nakala.fr/collection/11280/7681d5cf>
        ore:aggregates  <https://www.nakala.fr/resource/11280/afe3b32f> .
```

Exemple de métadonnées publiques exposées dans Nakala au format Turtle, relatif à un Arrêt du parlement, concernant les salaires des Huissiers pour les devoirs qu'ils font au profit de leur bourse commune

54 https://parleflandre.nakalona.fr/items/show/707?output=omeka-xml.

55 GAUTIER POUPEAU, « XML vs RDF : logique structurelle contre logique des données », *Les petites cases*, 29 août 2010, https://www.lespetitescases.net/xml-vs-rdf.

```
−<rdf:RDF>
 −<rdf:Description rdf:about="https://www.nakala.fr/data/resource/11280/afe3b32f?output=xml">
    <rdfs:label>RDF description of Afe3b32f</rdfs:label>
   −<foaf:primaryTopic>
    −<foaf:Document rdf:about="https://www.nakala.fr/resource/11280/afe3b32f">
       <dcterms:publisher rdf:resource="https://www.nakala.fr/account/11280/5f539340"/>
       <foaf:sha1>adba0df29e78a1249b7adea51c2b9d7f00ece70f</foaf:sha1>
       <dcterms:identifier>11280/afe3b32f</dcterms:identifier>
       <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2019-07-03T11:02:11.865+02:00</dcterms:modified>
      −<ore:isAggregatedBy>
       −<rdf:Description rdf:about="https://www.nakala.fr/collection/11280/7681d5cf">
          <ore:aggregates rdf:resource="https://www.nakala.fr/resource/11280/afe3b32f"/>
        </rdf:Description>
       </ore:isAggregatedBy>
       <dcterms:format>PDF</dcterms:format>
       <skos:altLabel>2056.pdf</skos:altLabel>
       <dcterms:extent rdf:datatype="http://www.w3.org/2001/XMLSchema#long">891179</dcterms:extent>
       <dcterms:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2019-07-03T10:48:11.096+02:00</dcterms:issued>
       <foaf:primaryTopic rdf:resource="https://www.nakala.fr/data/11280/afe3b32f"/>
       <dcterms:creator rdf:resource="https://www.nakala.fr/user/98839290-95d0-4ae7-8dee-f0c0b566efd1"/>
      </foaf:Document>
    </foaf:primaryTopic>
   </rdf:Description>
  </rdf:RDF>
```


Exemple de métadonnées publiques exposées dans Nakala au format RDF_XML, relatif à un Arrêt du parlement, concernant les salaires des Huissiers pour les devoirs qu'ils font au profit de leur bourse commune[56]

Toutes ces vues ou expositions de formats différents nous permettent de comprendre comment une machine, un ordinateur, arrange [57] les données. Certes, le porteur d'un projet numérique n'est peut-être pas amené à mettre ses mains dans cet « arrangement », mais aujourd'hui, il doit, semble-t-il, comprendre l'intérêt de ces formats[58].

Au-delà de ces expositions de données dans différents formats, Omeka propose un entrepôt des métadonnées interopérable grâce au protocole OAI-PMH.[59] Cette solution donne la possibilité de diffuser les enregistrements du projet Meta-parleFlandre sur la plateforme Isidore. Ainsi Meta-parleflandre est aussi une collection à part entière dans Isidore[60].

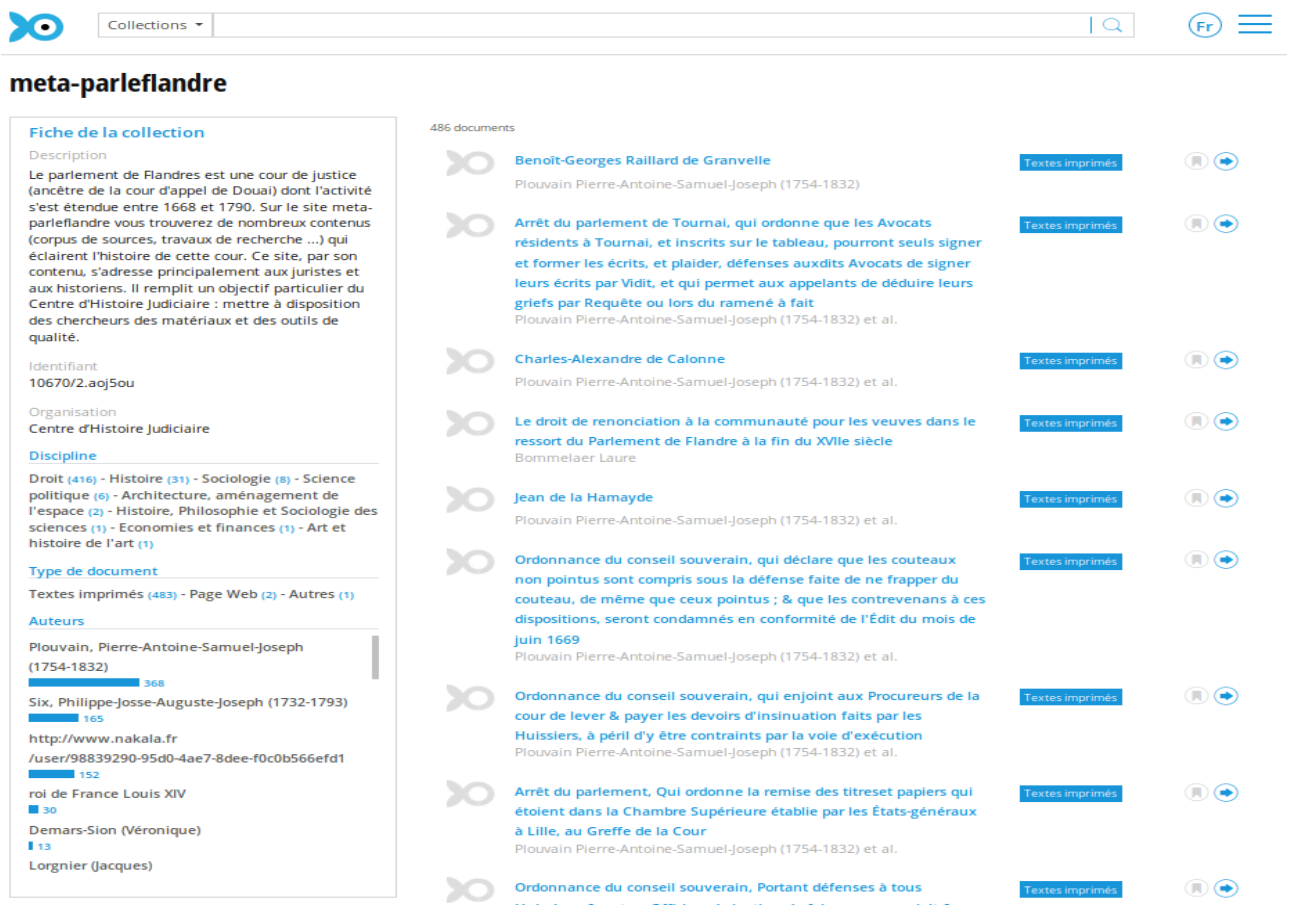
56 https://www.nakala.fr/data/resource/11280/afe3b32f?output=xml.

57 C'est à dessein que nous utilisons ce verbe, plus représentatif que « structurer » car il introduit l'idée de passage d'un état à un autre, d'une forme à une autre, tout en faisant référence à l'idée d'organisation. Il englobe alors à la fois la structuration d'un document XML et la logique du web sémantique, v. *supra*, note 46.

58 Sur ce questionnement on se reportera à KATHRYN LA BARRE, « A Semantic (Faceted) Web ? », *Les Cahiers du numérique*, 2010/3 (Vol. 6), p. 103-131. URL : https://www.cairn.info/revue-les-cahiers-du-numerique-2010-3-page-103.htm et à BRUNO BACHIMONT, FABIEN GANDON, GAUTIER POUPEAU *et al.*, « Enjeux et technologies : des données au sens », *Documentaliste-Sciences de l'Information*, 2011/4 (Vol. 48), p. 24-41. DOI : 10.3917/docsi.484.0024. URL : https://www-cairn-info.ressources-electroniques.univ-lille.fr/revue-documentaliste-sciences-de-l-information-2011-4-page-24.htm.

59 CAMILLE PRIME-CLAVERIE et ANNAÏG MAHÉ, « Le défi de l'interopérabilité entre plates-formes pour la construction de savoirs augmentés en sciences humaines et social », dans ISAD KEMBELLEC et EVELYNE BROUDOUX (dir.), *Ecrilecture augmentée dans les communautés scientifiques*, *Humanités numériques et construction des savoirs*, ISTE édition, 2017.

60 Collection disponible sur https://isidore.science/collection/10670/2.aoj5ou.



Vue de la collection Meta-parleFlandre sur Isidore

Outre la diffusion du corpus numérique vers une communauté spécifique aux Sciences Humaines et Sociales, le projet Isidore enrichit les données à travers plusieurs référentiels comme HAL[61], OpenEdition[62] ou Rameau[63].

Conclusion

En fin de compte, à partir de deux projets connexes sur le fonds, en tant qu'ils touchent à valoriser la connaissance du parlement de Flandre, les réalisations numériques diffèrent largement. Dans un court laps de temps, moins d'une dizaine d'années, les outils numériques disponibles en Sciences Humaines et Sociales, et par là les possibilités techniques pouvant être facilement mises en œuvre, ont largement permis aux porteurs de projets de se libérer de contraintes techniques en s'appuyant sur des solutions évaluées par la Très Grande Infrastructure de Recherche qu'est Huma-Num.

Cette constatation faite, ceci nous amène à la réflexion suivante. Le porteur d'un projet numérique doit rester vigilant sur l'évolution des Humanités Numé-

61 Le référentiel de Hal est disponible ici : https://halshs.archives-ouvertes.fr/browse/domain.

62 Le référentiel d'OpenEdition est disponible ici : https://journals.openedition.org/index/49.

63 Le référentiel Rameau est disponible ici : http://rameau.bnf.fr.

riques. Aujourd’hui, tout bouge très vite et il ne faudrait pas que les choix fait dans la précipitation de la conception d’un projet soient bloquants ou obsolètes très peu de temps après.

Pour se prémunir de telles déconvenues nous ne pouvons rappeler que le suivi de bonnes pratiques [64], étape par étape, et l’utilisation de solutions libres sont des éléments essentiels dans la conduite de projet.

64 *Supra*, note 45.