



Construction d'une ontologie pour le patrimoine culturel et historique : les sceaux byzantins

Mémoire

présenté par

Asmae CHELAULAU
et **Alexandra DEMKOVIC**

Master 1

Mention « *Sciences du Langage* »

Parcours « *Langue et Informatique* »

dirigé par

Victoria EYHARABIDE

soutenu le

18 septembre 2023

Dédicace

Ce mémoire est dédié à nos familles et amis.

Ainsi qu'à Stefan qui est un véritable pilier et une source de motivation au quotidien.

Remerciements

Nous tenons à remercier tout d’abord l’équipe enseignante qui a été présente de manière significative pour partager le savoir et les connaissances nécessaires à la réalisation de ce mémoire et cela tout au long de l’année.

Ainsi qu’à Madame Victoria Eyharabide, notre encadrante, qui nous a accompagnées tout au long du semestre pour nous apporter conseils, soutien et disponibilité dans l’avancement des étapes pour l’élaboration de notre travail. Son expertise et son expérience dans le domaine des ontologies a pu nous guider et permettre de mieux saisir les aboutissants de ce travail enrichissant et à la thématique original.

Enfin, nous souhaitons remercier nos proches qui ont su être présents tout au long de nos études pour nous permettre de réaliser ce mémoire sur un thème passionnant.

Résumé

Le présent mémoire a pour objectif de contribuer à la gestion et à l'analyse des informations extraites à partir des sceaux byzantins qui sont des objets historiques d'une grande importance tant historique que culturelle. La finalité principale de cette étude est la création d'une ontologie riche en données paramétrées de manière systématique sur ces sceaux appartenant à la collection Zacos présentée au Musée d'art et d'histoire de Genève. Cette ontologie vise à établir un cadre conceptuel clair et une méthodologie rigoureuse, permettant ainsi un accès précis aux détails des sceaux et ouvrant de nouvelles perspectives pour la recherche dans ce domaine passionnant.

Pour atteindre ces objectifs multiples et nécessaires, l'usage multiple de techniques et d'outils informatiques s'est montré essentiel. L'utilisation de bibliothèques avancées, notamment spaCy, sert à renforcer le balisage du fichier XML initial, enrichissant ainsi la source de données ou encore l'élaboration de règles Jape ou du framework GATE pour automatiser davantage le processus d'annotation des éléments pertinents et de la collecte des données. De manière concrète, ces méthodes employées d'enrichissement garantissent une meilleure délimitation et exploitation des informations extraites des sceaux.

Le peuplement de l'ontologie repose sur l'utilisation de l'outil Protégé, un environnement de gestion d'ontologie, pour concevoir, éditer et examiner la structure ontologique, garantissant ainsi sa solidité et sa flexibilité, notamment grâce à l'utilisation du moteur de raisonnement PELLET. De plus, pour faciliter l'interrogation et l'extraction d'informations spécifiques de l'ontologie, des requêtes SPARQL sont élaborées.

Ce mémoire participe donc à la combinaison de l'expertise en ontologie avec les outils informatiques d'automatisation mis à disposition pour répondre aux besoins de préservation et de recherche dans ce domaine complexe, ainsi qu'à une meilleure compréhension du contexte historique et culturel entourant ces ressources historiques.

Mots clés : Sceaux byzantins, Ontologie, Gestion de l'information, Traitement des données, RDF, OWL, Balisage XML, Jape, GATE, Protégé, Reasoner PELLET, SPARQL, Patrimoine culturel, SWRL.

Abstract

This thesis aims to contribute to the management and analysis of information extracted from Byzantine seals, which are historically and culturally significant artifacts. The primary objective of this study is to create an ontology rich in systematically parameterized data about these seals, part of the Zacos collection exhibited at the Museum of Art and History in Geneva. This ontology seeks to establish a clear conceptual framework and rigorous methodology, thereby enabling precise access to seal details and opening up new avenues for research in this fascinating field.

To achieve these multiple and essential goals, the multifaceted use of computer techniques and tools has proven essential. The use of advanced libraries, notably spaCy, serves to enhance the markup of the initial XML file, enriching the data source, or the development of Jape rules and the GATE framework to further automate the annotation of relevant elements and data collection. In concrete terms, these employed enrichment methods ensure a better delineation and exploitation of the information extracted from the seals.

The population of the ontology relies on the use of the Protégé tool, an ontology management environment, to design, edit, and review the ontological structure, ensuring its robustness and flexibility, notably through the use of the PELLET reasoning engine. Furthermore, to facilitate querying and extracting specific information from the ontology, SPARQL queries are developed.

This thesis participates in the combination of ontology expertise with the automation tools available to meet the preservation and research needs in this complex field, as well as to a better understanding of the historical and cultural context surrounding these historical resources.

Keywords : Byzantine Seals, Ontology, Semantic Web, Data Processing, RDF, OWL, XML Markup, Jape, GATE, Protégé, PELLET Reasoner, SPARQL, Cultural Heritage, SWRL.

Table des matières

Résumé	iii
Abstract	iv
Table des figures	4
Liste des abréviations	6
Introduction générale	7
1 Cadre historique de la création des sceaux byzantins	9
1.1 Approche historique de l'apparition des sceaux sous l'Empire Byzantin	9
1.2 Histoire et création des sceaux byzantins	11
1.2.1 Processus de création des sceaux	11
1.2.2 Utilisation des sceaux byzantins	12
1.2.3 Signification symbolique des sceaux byzantins	12
1.3 La collection Zacos	13
1.4 Questions de compétences pour l'exploration des informations à représenter ontologiquement	15
2 Etat de l'art	17
2.1 Les fondements des ontologies	17
2.2 Les ontologies pour le patrimoine culturel	18
2.3 Systématisation des connaissances pour les méthodes d'apprentissage d'onto- logies	21
3 Méthodes de représentation ontologique des données sémantiques et matériels	23
3.1 Exploration des données brutes et numérisées sur les sceaux byzantins de la collection Zacos	23
3.1.1 Observation des données brutes des sceaux byzantins	23
3.1.2 Structure des données numérisés des sceaux de la collection Zacos	24

3.2	Etapes de la construction de l'ontologie	26
3.2.1	Parcours du fichier XML à l'aide de bibliothèque	26
3.2.2	Ensemble des méthodes de codification de l'ontologie des sceaux by- zantins de la collection Zacos	27
3.2.2.1	Import de la bibliothèque 'rdflib' pour manipuler les graphes RDF	27
3.2.2.2	Définition des espaces de noms RDF	27
3.2.3	Création du contenu du graphe contenant les classes, instances et rela- tions de l'ontologie	28
3.2.3.1	Création du graphe RDF	28
3.2.3.2	Définition de fonctions pour créer des classes et des relations d'objets	28
3.2.3.3	Création des classes dans l'ontologie	29
3.2.3.4	Création des relations entre les classes	30
3.2.4	Sérialisation du graphe RDF par l'ajout des instances et classes des sceaux byzantins	31
3.2.4.1	Instanciation des attributs paramétrés	31
3.2.4.2	Sérialisation du graphe RDF	31
3.2.5	Enrichissement par la modification des balises de date dans un fichier XML	32
3.2.5.1	Utilisation d'un fichier texte regroupant les appellations et les conversions des dates	32
3.2.5.2	Utilisation d'un fichier JSON regroupant les appellations et les conversions des dates	33
3.2.6	Amélioration par l'enrichissement des balises du fichier XML avec spaCy	34
3.2.6.1	Enrichissement du balisage à l'aide du postagging de la bi- bliothèque spaCy	34
3.2.6.2	Enrichissement du balisage à l'aide de l'Universal Dependen- cies de la bibliothèque spaCy	35
3.2.6.3	Enrichissement du balisage à l'aide de la reconnaissance d'en- tités nommées de la bibliothèque spaCy	36
3.2.6.4	Enrichissement du balisage à l'aide de la lemmatisation de la bibliothèque spaCy	36
3.2.7	Amélioration par l'enrichissement des balises du fichier XML avec GATE	37
3.2.7.1	Ciblage des informations à l'aide de règles Jape	37
3.2.7.2	Balisage du contenu textuel depuis le framework GATE	40
3.2.8	Peuplement de l'ontologie grâce aux annotations effectuées	42

3.2.8.1	Enrichissement des domaines, ranges, et autres caractéristiques des classes	42
3.2.8.2	Enrichissement des instances	43
3.2.8.3	Enrichissement des prédicats	43
4	Evaluation	46
4.1	Vérification de la consistance avec le moteur de raisonnement Pellet	46
4.2	Exploitation des informations paramétrées par le requêtage en SPARQL	46
5	Comparaison des résultats obtenus et des limites rencontrées	49
5.1	Résultats et limites des outils de spaCy	50
5.2	Résultats de l'utilisation des règles Jape et du framework GATE	52
5.3	Difficultés dans l'élaboration du code pour convertir les dates	53
	Conclusion	54
A	Annexe	57
	Bibliographie	60

Table des figures

1.1	Le boullôtèrion	14
3.1	Sceau Byzantin en plomb datant du XIIe siècle et appartenant à Joseph	25
3.2	Sceau en or datant du XIe siècle et appartenant à l'empereur Constantin IX Monomaque	25
3.3	Exemple d'étiquetage de l'analyse des dépendances sur le sceau byzantin CdN 2004-0313	35
3.4	Exemple de lemmatisation des verbes sur le sceau byzantin CdN 2004-0313	37
3.5	Illustration de la grammaire Jape utilisée pour annoter toutes les occurrences correspondantes à la Vierge Marie	38
3.6	Liste des annotations obtenues pour le mot lôros à l'aide de Spacy.	40
3.7	Liste des annotations obtenues pour le mot bouclier à l'aide de Spacy.	40
3.8	Fichier XML contenant les nouvelles annotations.	41
3.9	Fichier XML contenant les nouvelles annotations.	42
3.10	Illustration de la règle SWRL.	44
3.11	Résultat de la règle SWRL avec les inférences obtenues à l'aide de Pellet.	45
3.12	Résultat de la règle SWRL traitant la propriété tient_des_deux_mains avec les inférences obtenues à l'aide de Pellet.	45
5.1	Représentation du graphe pour les sceaux byzantins de la collection Zacos	50
5.2	Exemple d'erreur obtenu avec le postagging pour le sceau byzantin CdN 2004-0551	51
5.3	Exemple 1 d'erreur de lemmatisation d'un nom commun en verbe pour le sceau byzantin CdN 2004-0313	52
5.4	Exemple 2 d'erreur de lemmatisation d'un nom commun en verbe pour le sceau byzantin CdN 2004-0373	52
5.5	Résumé des différentes étapes du mémoire.	56
A.1	Exemple de visualisation du graphe enrichi dans l'environnement Protégé	57
A.2	Exemple de création de classes, instances et relations dans l'ontologie	58

A.3	Exemple de visualisation du graphe enrichi dans l'environnement Protégé . . .	58
A.4	Exemple de résultat obtenu par requêtage SPARQL depuis le site SPARQL - Playground	59

Liste des abréviations

FOAF	Friend of a Friend
GATE	General Architecture for Text Engineering
IRI	Internationalized Resource Identifier
Jape	Java Annotation Patterns Engine
OWL	Web Ontology Language
OWL DL	Web Ontology Language Description Logic
PDF	Portable Document Format
RDF	Resource Description Framework
RDFLib	Resource Description Framework Library
RDFS	RDF Schema
rule[-	
0.2cm]0cm0.7cm	
Regex	
SPARQL	SPARQL Protocol and RDF Query Language
spaCy	Space-efficient Natural Language Processing in Python
SWRL	Semantic Web Rule Language
TALN	Traitement automatique du langage naturel
URI	Uniform Resource Identifier
XML	eXtensible Markup Language

Introduction générale

Dans un article intitulé «Toward Principles for the Design of Ontologies Used for Knowledge Sharing» publié en 1993 dans le journal «International Journal of Human-Computer Studies» [R.95], l’informaticien Thomas R. Gruber définit l’ontologie comme étant « la spécification d’une conceptualisation. [...] Une conceptualisation est une vue abstraite et simplifiée du monde que l’on veut représenter ». Il s’appuie sur deux dimensions qui sont qu’une ontologie représente une vue abstraite et simplifiée du monde réel dans ce domaine, en identifiant les concepts, les relations et les propriétés, appelés les triplets. La spécification formelle de l’ontologie implique une description précise en utilisant des langages formels et des méthodes rigoureuses pour permettre la précision, la cohérence et la fiabilité des représentations conceptuelles. Cela permet ainsi l’utilisation des connaissances dans le domaine concerné et de diffuser, de cette manière, des connaissances à une plus grande échelle, notamment lorsqu’il ces connaissances servent à la compréhension de notre histoire et de notre identité collective.

Parmi les biens historiques catégorisés comme précieux, les sceaux byzantins se distinguent par leur valeur historique et culturelle. Ces sceaux, utilisés dans l’Empire byzantin pendant des siècles, fournissent des informations uniques sur les institutions, les individus et les pratiques sociales de cette période. L’étude et la préservation du patrimoine culturel revêtent donc d’une importance particulière pour saisir l’évolution des civilisations et des sociétés humaines.

La difficulté rencontrée se trouve dans l’exploitation et la diffusion des connaissances liées aux sceaux byzantins qui nécessitent une organisation structurée des données. La construction d’ontologies et l’établissement de standards de données spécifiques au patrimoine culturel permet de structurer et sauvegarder les informations propres à ces objets historiques et fragiles. La construction d’une ontologie pour les sceaux byzantins favorise la compréhension, l’analyse et la valorisation de ce patrimoine historique. Une ontologie des sceaux byzantins permettrait de structurer les données relatives à ces objets culturels en identifiant et en décrivant les entités et les relations pertinentes, telles que les périodes, les matériaux, les représentations iconographiques et les inscriptions. Il s’agira, pour cela, de recourir à diverses méthodologies et approches, allant de l’encodage sémiotique détaillé des caractéristiques visuelles et sémiotiques des sceaux à l’utilisation du modèle RDF (Resource Description Framework) pour représenter les connaissances de manière structurée et interopérable.

Ainsi, comment construire une ontologie pour préserver et faciliter l'accès aux connaissances du patrimoine culturel et historique des sceaux byzantins ?

La sauvegarde du patrimoine culturel est possible non seulement grâce aux ontologies, qui proposent un cadre conceptuel et une représentation formelle permettant de modéliser et d'interconnecter les connaissances sur un domaine spécifique, mais aussi les standards de données, qui garantissent une cohérence et une interopérabilité des informations.

Nous examinerons les différentes méthodes utilisées pour la construction d'une ontologie des sceaux byzantin, et, les standards de données au domaine spécifique des sceaux byzantins, en passant par les avantages, les difficultés et les contributions que cette approche peut apporter à la compréhension et à la préservation de cet héritage culturel et historique.

Chapitre 1

Cadre historique de la création des sceaux byzantins

Dans ce chapitre, nous allons voir le cadre historique et sociale de l'Empire byzantin, en passant par le processus de création de ces objets historiques, la collection qui les réunit, de même que les questions de compétences pour l'exploration des données.

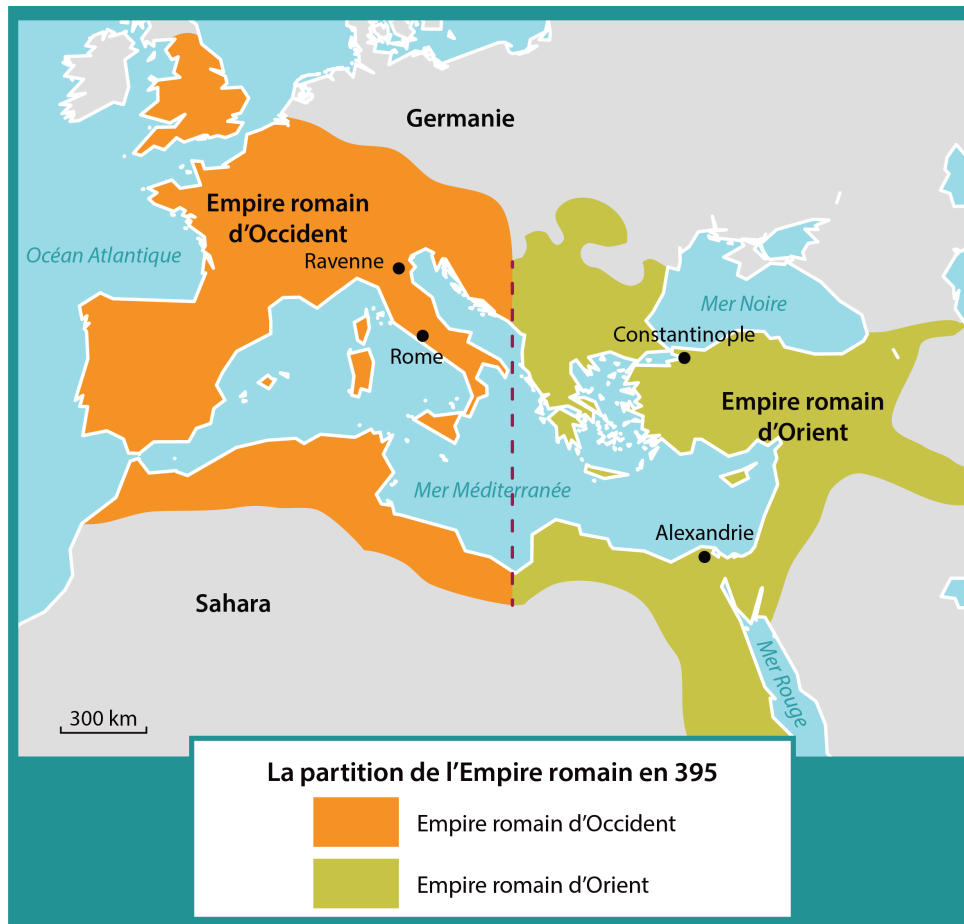
1.1 Approche historique de l'apparition des sceaux sous l'Empire Byzantin

L'Empire Byzantin est apparu lors de la séparation de l'empire romain au IV^e siècle. Cet empire s'étend tout au long des côtes de la mer Méditerranée, la mer Rouge et la mer Noire. Constantin le Grand choisit donc Byzance, qui était un port grec séparant la Turquie actuelle de l'Europe, pour en faire la capitale Constantinople. Cela permit un développement économique et sociétal caractérisé par un lien entre l'Eglise et les autorités impériales et un système monétaire durable.

Au VI^e siècle, la chute de l'Empire romain d'Occident et la chute du côté de l'Empire byzantin ont permis à ce dernier de tenter de reconquérir les territoires qui étaient auparavant occupés par l'Empire romain (l'Afrique du Nord, l'Italie et une partie de l'Espagne).

Etant donné le contexte historique, la société est plongée dans une culture chrétienne et la langue utilisée est la langue latine. Ce n'est qu'au VIII^e siècle que l'empire byzantin s'affaiblit suite aux attaques des Slaves, des Avars et des Perses mais les réformes introduites par l'empereur Héraclius permirent de sauver l'empire. Les relations entre les différents peuples tant dans les Balkans que sur le pourtour de la Méditerranée a évolué au gré des invasions, conflits de nature différente, et d'alliance. C'est pourquoi le christianisme orthodoxe s'est répandu bien au-delà

des terres grecques [VS77]. La société byzantine est désormais composée par un empereur au sommet, d'une classe dirigeante formée des membres du Sénat, d'une bureaucratie civile et militaire ainsi que d'une aristocratie provinciale gouvernant les principales villes de l'empire. La culture et la langue grecques deviennent centrale à cette période et se diffuse largement.



La société de l'empire byzantin est subdivisée en trois grandes parties. Au sommet de l'échelle se trouve l'empereur et il n'est possible de s'élever socialement entre ces trois grands groupes que dans des cas particuliers (grâce à la richesse par exemple). En effet, au début de la naissance de l'empire byzantin il était possible qu'un individu s'élève dans l'échelle sociale en devenant plus riche ou en évoluant à l'armée. Cependant, au cours des siècles cette mobilité a été réduite. Chaque citoyen possède des droits différents en fonction de son rang.

La position d'un individu au sein de la société byzantine était largement influencée par sa richesse en ce qui concerne l'acquisition de terres ou son rang dans l'armée.

Plus un homme possède de terres, plus il sera dans une bonne position au sein de la société. Nous pouvons constater ce même phénomène chez les soldats car un homme ayant un rang élevé au sein de l'armée aura une position importante dans l'échelle sociale.

Dans ce mémoire, nous étudierons notamment la classe sociale qui regroupe les aristocrates, les

fonctionnaires, les officiers et les propriétaires terriens car grâce aux reliques tels que les sceaux il est possible de retracer l'histoire de l'empire byzantin et de la structure et le fonctionnement de la société byzantine.

1.2 Histoire et création des sceaux byzantins

Les sceaux byzantins sont des objets d'une grande importance historique et culturelle. Ils ont été utilisés par l'Empire byzantin pour authentifier des documents et garantir leur sécurité. Dans cette présentation, nous allons explorer l'histoire des sceaux byzantins, leur création, leur utilisation et leur signification symbolique.

Les premiers sceaux byzantins ont été créés pendant l'Empire romain d'Orient, également connu sous le nom d'Empire byzantin. Ils ont été utilisés pour authentifier des documents officiels et garantir leur sécurité. Le sceau était une marque de propriété, de puissance et d'autorité. Il était également utilisé pour marquer les objets et les animaux.

Le premier sceau connu de l'Empire byzantin a été créé sous le règne de l'empereur Justinien en 527 après J.-C. Il s'agissait d'un petit disque en bronze avec une image de l'empereur. Les sceaux byzantins ont ensuite évolué pour inclure des images de saints, de personnages bibliques, de croix et d'autres symboles religieux. Les sceaux étaient également gravés avec des inscriptions principalement en grec, peu importe la région ou le pays d'origine [Iva01], et indiquaient le nom et le titre de la personne qui possédait le sceau, ainsi que des prières.

1.2.1 Processus de création des sceaux

Ces sceaux étaient fabriqués en différents matériaux, tels que l'or, l'argent, le bronze et le plomb, mais pouvaient être également fabriqués à partir de pierres précieuses, comme le jade, la cornaline, l'agate, le lapis-lazuli et la chrysoprase. Etant généralement de petite taille, d'environ un à deux centimètres de diamètre, ils étaient utilisés pour apposer une marque distinctive sur des documents officiels, et souvent étaient portés autour du cou ou attachés à une ceinture par une chaîne. Ils étaient alors de différentes tailles et formes, la plupart étant de forme circulaire ou ovale.

Le processus de création d'un sceau byzantin impliquait l'utilisation d'un poinçon pour graver une image sur un disque de métal ou de pierre. Les images étaient souvent détaillées et complexes, nécessitant une grande habileté et une grande précision. Les sceaux étaient ensuite utilisés pour apposer une empreinte sur de la cire chaude ou du plomb fondu pour sceller les documents officiels.

Concernant les dimensions des sceaux byzantins celles-ci peuvent varier considérablement. Les plus petits sont généralement de la taille d'une pièce de monnaie, avec un diamètre d'environ deux à trois centimètres. Les plus grands peuvent atteindre plusieurs centimètres de diamètre ou être de forme rectangulaire avec des dimensions plus importantes. Les formes arrondies ou angulaires ne sont pas parfaitement façonnées.

1.2.2 Utilisation des sceaux byzantins

Les sceaux byzantins étaient utilisés pour authentifier des documents officiels, tels que des contrats, des actes de propriété, des testaments et des lettres. Les sceaux garantissaient l'authenticité des documents et empêchaient leur falsification. Les sceaux étaient également utilisés pour protéger la vie privée des individus en garantissant la confidentialité des lettres et des documents personnels.

Utilisés également à des fins religieuses, les images présentes sur les sceaux sont des figures religieuses gravées qui étaient considérées comme sacrées et leur utilisation garantissait la protection divine. Les saints et personnages bibliques étaient utilisés comme saint patron propre à chaque famille ou encore selon la fonction du propriétaire du sceau. Les sceaux étaient également utilisés pour marquer les objets religieux, tels que les reliques et les icônes.

Enfin, les sceaux étaient utilisés par ces importantes personnes pour des correspondances privées. En effet, les sceaux étaient attachés à l'aide d'une cordelette pour assurer la confidentialité des envois officiels. Communément en plomb et en or pour l'empereur, cette petite pièce permettaient d'identifier son propriétaire car elle était créée à l'aide d'une pince appelée boullôtèrion ou d'une bague sigillaire qui étaient strictement personnelles. Ces petits objets étaient personnalisés car chaque individu possédait une bague sigillaire ou un boullôtèrion pour pouvoir estamper communément des prières, des figures religieuses et des citations.

1.2.3 Signification symbolique des sceaux byzantins

Les sceaux servaient essentiellement à identifier son propriétaire qui posséder une autorité particulière à une époque et à un lieu donnés. Similaire au système des cartes d'identité, de part les différentes informations inscrites sur les sceaux, les personnages choisies sur l'avvers d'un sceau ou encore les inscriptions gravées sur les surfaces de ceux-ci, permettent de retracer des liens et événements dans le temps durant l'époque byzantine et établir la structure sociale avec ses acteurs d'autant. Les images impériales représentaient l'empereur et sa famille, ainsi

que des symboles de la puissance impériale, tels que des aigles et des couronnes. Les sceaux étaient donc un moyen de communiquer des messages symboliques importants sur la puissance et l'autorité[Teo21].

Leurs surfaces étaient gravées avec des images et des inscriptions qui reflétaient le statut et l'autorité de leur propriétaire. Les images les plus courantes représentaient l'empereur ou l'impératrice assis sur un trône, entouré d'anges et de symboles impériaux tels que des aigles, des lions et des licornes. Outre les images et les inscriptions, les sceaux byzantins pouvaient également inclure des symboles tels que des croix, des étoiles et des cercles, ainsi que des motifs géométriques et floraux. Ces motifs étaient souvent utilisés pour décorer le pourtour du sceau.

Par conséquent, ces objets étaient considérés comme des symboles de pouvoir et de statut, et leur utilisation était strictement réglementée par l'État byzantin. Les sceaux byzantins étaient des objets d'une grande importance historique et culturelle. Ils ont été utilisés par l'Empire byzantin pour authentifier des documents et garantir leur sécurité. Les sceaux étaient fabriqués à partir de différents matériaux, tels que l'or, l'argent, le bronze et le plomb, et étaient gravés avec des images religieuses et impériales détaillées et complexes. Les sceaux avaient une signification symbolique importante, représentant la puissance et l'autorité de l'empereur et de l'Église. Aujourd'hui, les sceaux byzantins sont considérés comme des objets d'art précieux et sont étudiés pour leur valeur historique et culturelle.

1.3 La collection Zacos

Débutée dans les années 1950, la collection Zacos est une importante collection privée qui a été rassemblée au fil des ans par Nicolas et Jeanne Zacos, deux collectionneurs experts en art byzantin. Au fil des ans, la collection Zacos s'est considérablement développée, atteignant plus de 20 000 sceaux, ce qui en fait l'une des plus grandes collections privées de sceaux byzantins au monde, tout en offrant une richesse de détails sur la vie quotidienne, la religion et la culture de l'Empire byzantin.

Etant des objets d'une grande valeur historique et culturelle qui permettent de mieux comprendre la société et la culture byzantines, les sceaux byzantins de la collection Zacos sont donc une source précieuse d'informations pour les historiens, les archéologues, les numismates et les étudiants de l'art byzantin. La collection comprend des sceaux rares et uniques couvrant une période de l'Empire byzantin allant du IV^e au XV^e siècle et qui fournissent des informations

précieuses sur la vie quotidienne, la culture et la religion de l'Empire byzantin. C'est pourquoi, les sceaux de la collection Zacos ont été étudiés et publiés dans de nombreux ouvrages académiques, contribuant ainsi à l'avancement de la recherche sur l'art et la culture byzantins. Exposés dans de nombreux musées et galeries d'art dans le monde, ces sceaux font partis du patrimoine culturel de l'Empire byzantin témoignant de la richesse et de la sophistication de cette culture.

La qualité de conservation de cette collection a permis à Jean-Claude Cheynet d'effectuer un travail de sigillographie (discipline qui a pour objet l'étude des sceaux sous tous les aspects) et de répertorier pour chaque sceau un propriétaire, une date, la description textuelle de l'avvers et du revers, son diamètre, un numéro d'inventaire ainsi que le contexte de sa création. Ces informations ont été collectées par Maria Campagnolo-Pothitou dans l'ouvrage, que nous avons cité auparavant, intitulé *Sceaux de la collection Georges Zacos au musée d'art et d'histoire de Genève* [dedCPC⁺16] qui a été publié en 2016.

Leur place dans des musées s'explique du fait que selon l'ICOM (Conseil international des musées), « Un musée est une institution permanente à but non lucratif, au service de la société et de son développement, ouverte au public, qui acquiert, conserve, recherche, communique et expose, à des fins d'étude, d'éducation et de plaisir, des éléments matériels témoignant des êtres humains et de leur environnement » [Def01]. De plus, « Les musées détiennent des preuves essentielles pour établir et approfondir les connaissances » [Def06]. L'informatique vient donc assister cette volonté de sauvegarder ce patrimoine culturel et à la diffusion de ces données historiques.



FIGURE 1.1 – Le boullôtèrion

1.4 Questions de compétences pour l'exploration des informations à représenter ontologiquement

« »

Pour pouvoir établir l'ontologie des sceaux byzantins, il est nécessaire d'établir en amont les différents objectifs visés pour tester la pertinence et l'accessibilité des différentes informations paramétrées. Les questions de compétences ci-dessous, serviront à détecter différents éléments qui peuvent intéresser toute personne souhaitant accéder à des informations ciblées. Ces questions portent sur les différents personnages présents comme Jésus-Christ, la Vierge Marie ou encore des figures politiques comme les empereurs, ecclésiastiques comme les patriarches, mais aussi des éléments temporels comme les dates, les périodes, ou encore les matériaux utilisés, les symboles et écritures apparaissant sur ces sceaux.

Huit questions de compétences ont été choisies pour ce travail.

La première question « Quelles sont les personnages bibliques dans la collection présents sur l'avvers du sceau et selon les dates de début des sceaux associés ? » pour cibler une surface d'un sceau, les personnages bibliques apparaissant dessus et selon la datation.

La deuxième question « Quelles sont la description et la date de fin des sceaux liés à des personnages ecclésiastiques ? » vise la recherche par date de fin et selon la description ajoutée par le chercheur à propos du sceau donné quicomporte des personnages ecclésiastiques.

La troisième question « Quelles bibliographies sont associées aux sceaux de la collection ? » permet de rechercher des informations paramétrées autour du sceau si le chercheur souhaite obtenir directement les sources.

La quatrième question « Quelles sont les mentions obligatoires à propos des sceaux de type "Objet de décor" présents sur l'avvers des sceaux ? » permet d'utiliser le listage pour les mentions obligatoires des sceaux comportant un objet de décoration sur son avers.

la cinquième question « Quelles sont les descriptions des sceaux appartenant à la période "Moyen_âge" selon leur date ? » permet de trier dans un ordre choisi, selon la date de début et la description de ceux-ci.

La sixième question « Quels sont les avers, les revers et les dates de création des sceaux associés à des armes dans la base de données ? » est utilisée pour obtenir les informations sur

les surfaces des sceaux et les dates de leur création afin d'obtenir directement leur contenu selon la date et selon s'ils contiennent une arme.

La septième question « Quelles sont les personnages de la classe "Figure_ecclesiastique" qui sont nimbés ? » permet de savoir si, notamment, des figures ecclésiastiques sont nimbés.

La huitième question « Quelles sont les titre, matériel et période sceaux en or ou en plomb ? » permet d'obtenir certains types de sceau selon des données ciblées.

Etant donné la variété de sceaux byzantins rassemblés dans la collection Zacos, c'est à travers cette série de questions de compétences qu'il est possible d'explorer sous différents aspects les sceaux byzantins par des méthodes automatisées. Au regard de l'objectif principal qu'est l'élaboration de l'ontologie, cette démarche est cruciale pour l'appréhension de l'organisation et de l'exploitation de ces données riches en histoire et en culture.

Chapitre 2

Etat de l'art

Dans les années 90, les recherches dans le domaine de la représentation des connaissances ont montré qu'il était possible de modéliser de manière explicite et consensuelle les caractéristiques structurelles et descriptives d'un domaine de connaissance, dans le but de faciliter le partage de ces connaissances. Ces efforts ont conduit à la création de modèles conçus pour clarifier la signification des données, que l'on appelle couramment des ontologies.

2.1 Les fondements des ontologies

Tout d'abord, le document «Toward Principles for the Design of Ontologies Used for Knowledge Sharing»[\[R.95\]](#) publié en 1995 présente des principes directeurs pour la conception d'ontologies dans le contexte du partage de connaissances. Ces principes sont cruciaux pour garantir l'efficacité et la qualité des ontologies. Ils comprennent la clarté conceptuelle, la modularité et la réutilisabilité, ainsi que l'extensibilité. Une ontologie doit être clairement définie, utilisant des concepts bien définis et des relations explicites pour faciliter la compréhension. La modularité et la réutilisabilité permettent de construire des ontologies flexibles et adaptables, en intégrant des parties existantes pour répondre à de nouveaux besoins. L'extensibilité est également essentielle, permettant d'étendre l'ontologie pour inclure de nouveaux concepts et relations sans perturber sa structure.

Un autre aspect crucial abordé dans le document est la cohérence de l'ontologie. Il est essentiel d'éviter les contradictions et les incohérences internes dans la définition des concepts et des relations. Une ontologie cohérente garantit une représentation fiable et précise des connaissances partagées.

Enfin, la formalité et l'expressivité sont des principes clés. L'utilisation de langages formels et structurés, tels que l'OWL, permet une représentation précise et une manipulation informa-

tique des connaissances. Une ontologie formelle et expressive facilite ainsi l'interprétation et la transmission des connaissances dans différents domaines d'application.

De plus, l'ouvrage «Ontologies : Principles, Methods and Applications»[UG96], publié en 1996, explore les fondements, les méthodes et les applications des ontologies, qui sont des structures conceptuelles utilisées pour représenter les connaissances et les relations entre les concepts dans différents domaines. L'ouvrage examine en profondeur les principes sous-jacents à la construction d'ontologies, les techniques de développement et les diverses applications pratiques.

Les auteurs explorent les concepts fondamentaux des ontologies, notamment leur rôle dans la représentation sémantique des connaissances et la manière dont elles favorisent l'interopérabilité entre les systèmes informatiques. Ils examinent également les méthodes de création d'ontologies, allant de l'identification des concepts et des relations à la formalisation de ces éléments à l'aide de langages de représentation sémantique tels que RDF et OWL.

L'ouvrage met en lumière les différentes applications des ontologies dans des domaines variés tels que la recherche d'informations, la gestion des connaissances, la bioinformatique, l'ingénierie logicielle et bien d'autres. Il explique comment les ontologies facilitent la recherche efficace d'informations, la résolution de problèmes complexes et l'amélioration de la compréhension des données dans ces domaines.

En somme, «Ontologies : Principles, Methods and Applications»[UG96] offre un aperçu complet des concepts fondamentaux, des méthodes de développement et des multiples utilisations des ontologies. Cet ouvrage est une ressource précieuse pour quiconque s'intéresse à la modélisation sémantique, à la représentation des connaissances et à l'application des ontologies dans divers secteurs.

2.2 Les ontologies pour le patrimoine culturel

Maintenant que nous avons une compréhension solide des ontologies, examinons comment elles peuvent être mises en pratique dans le contexte historique.

Tout d'abord, dans l'article intitulé «The Digital Enhancement of a Discipline : Byzantine Sigillography and Digital Humanities»[SFS20], les auteurs explorent l'application des méthodes et des outils des sciences humaines numériques (Digital Humanities) dans le domaine de la Sigillographie Byzantine, qui étudie les sceaux, en particulier dans leur contexte historique et culturel.

Les auteurs soulignent que «Les sciences humaines numériques offrent de nouvelles possibilités passionnantes pour enrichir notre compréhension de la Sigillographie Byzantine». Les techniques numériques telles que la numérisation, la modélisation 3D, l'analyse d'images et la construction de bases de données sont discutées en détail. L'article met en évidence les avantages de ces approches, notamment une meilleure accessibilité aux données, leur préservation et leur analyse.

Enfin, l'article envisage les futurs développements possibles et les directions de recherche émergentes dans ce domaine passionnant, tout en résumant l'importance de l'utilisation des sciences humaines numériques pour améliorer l'étude des sceaux byzantins.

D'un autre côté, l'article «The CIDOC Conceptual Reference Module : An Ontological Approach to Semantic Interoperability of Metadata»[\[Doe03\]](#) présente le module de référence conceptuel CIDOC, qui adopte une approche ontologique pour assurer l'interopérabilité sémantique des métadonnées. Le module de référence conceptuel CIDOC est développé par le Comité international pour la documentation (CIDOC) de l'ICOM (Conseil international des musées) et vise à créer une structure conceptuelle commune pour décrire les objets culturels et le patrimoine.

L'article explore comment le module de référence conceptuel CIDOC utilise les principes ontologiques pour modéliser les concepts clés dans le domaine de la documentation culturelle. Il s'appuie sur des langages de modélisation sémantique tels que RDF et OWL pour définir des relations sémantiques entre les concepts. Cette approche permet de créer un cadre commun pour décrire les informations liées aux objets culturels, aux événements historiques, aux lieux et aux personnes, en garantissant la cohérence et l'interopérabilité des métadonnées à travers les différentes institutions et systèmes.

L'article met en évidence comment le module de référence conceptuel CIDOC facilite l'échange d'informations et la compréhension mutuelle entre les organisations culturelles et patrimoniales, en créant un langage commun pour décrire et interconnecter les données. Il explique également comment cette approche ontologique peut contribuer à surmonter les défis liés à la diversité des terminologies et des structures de données dans le domaine de la documentation culturelle.

En résumé, l'article met en avant le rôle essentiel du module de référence conceptuel CIDOC en tant qu'outil ontologique pour atteindre l'interopérabilité sémantique des métadonnées liées au patrimoine culturel, permettant ainsi une meilleure collaboration et une utilisation plus efficace des données dans ce domaine.

Bien qu'il n'y ait pas de recherche spécifique sur les ontologies appliquées aux sceaux by-

zantins, parmi les différentes études qui ont examiné la manière dont les ontologies peuvent être utilisées pour représenter des connaissances historiques et archéologiques, nous allons vous en présenter deux.

Tout d'abord, nous allons introduire le travail effectué avec le projet «OntoCoins». En effet, «OntoCoins : Données ouvertes liées pour la numismatique et le patrimoine culturel»[CLBJA⁺17] est une initiative visant à fournir des données ouvertes liées dans le domaine de la numismatique et du patrimoine culturel. Ce projet consiste à créer un ensemble de données structurées qui relie divers aspects de l'information numismatique et du patrimoine culturel, le rendant ainsi accessible et utile aux chercheurs, aux historiens et aux institutions culturelles.

L'idée fondamentale derrière OntoCoins est d'utiliser les technologies et les principes du web sémantique pour relier et organiser les données relatives aux pièces de monnaie, aux artefacts historiques et au patrimoine culturel. Cette approche de données liées améliore l'interopérabilité et la découvrabilité de l'information numismatique et des ressources du patrimoine culturel. En structurant et en reliant ces données, OntoCoins facilite la recherche, l'analyse et le partage des connaissances dans ces domaines.

Le projet se concentre sur la création d'une ontologie et d'un vocabulaire normalisés pour décrire les concepts et les artefacts numismatiques, permettant ainsi une représentation et un échange cohérents des données. Cette approche normalisée améliore la qualité des données et favorise l'intégration de l'information numismatique avec des ensembles de données plus larges du patrimoine culturel.

En résumé, ce projet est une initiative qui tire parti des principes des données ouvertes liées pour améliorer l'accessibilité et l'utilité de l'information numismatique et des ressources du patrimoine culturel. Il fournit un ensemble de données structurées et interconnectées qui bénéficie aux chercheurs et aux institutions culturelles dans leurs efforts pour étudier et préserver notre patrimoine culturel.

Pour finir, l'article «MusicKG : Representations of Sound and Music in the Middle Ages as Linked Open Data»[ELM19] publié en 2019 présente le projet MusicKG qui a pour objectif de représenter les connaissances sur la musique et les sons du Moyen âge sous forme de données liées ouvertes. L'utilisation de données liées pour organiser les connaissances sur la musique peut également être appliquée à l'organisation des connaissances sur les sceaux byzantins, permettant ainsi une meilleure compréhension de leur rôle et de leur signification dans la société byzantine. Les données sont organisées en plusieurs catégories, notamment les manuscrits mu-

sicaux, les instruments de musique et les théories musicales.

Le projet MusicKG utilise des outils numériques tels que la reconnaissance optique de caractères, la modélisation 3D et la classification automatisée pour rendre ces données plus accessibles et plus facilement compréhensibles pour les chercheurs et le grand public. Ce processus peut également être appliqué à l'étude des sceaux byzantins pour faciliter l'analyse et la compréhension de ces artefacts importants. En somme, le projet MusicKG est une contribution significative à la représentation de connaissances historiques sous forme de données liées et peut être appliqué à d'autres domaines de l'étude historique, y compris l'étude des sceaux byzantins.

En conclusion, bien qu'il n'y ait pas de recherche spécifique sur les ontologies appliquées aux sceaux byzantins, plusieurs études ont examiné l'utilisation des ontologies pour représenter des connaissances historiques et archéologiques. Celles-ci pourraient être adaptées à la représentation des connaissances sur les sceaux byzantins et pourraient notamment constituer une base pour de futures recherches dans ce domaine.

2.3 Systématisation des connaissances pour les méthodes d'apprentissage d'ontologies

L'étude rédigée par Agnieszka Konys sur la systématisation des connaissances pour les méthodes d'apprentissage d'ontologies[[Agn18](#)] développe l'idée de la nécessité croissante de solutions intelligentes pour garantir une sémantique interopérable au sein des systèmes d'information modernes. Etant donné que la quantité d'informations se multiplie davantage ces dernières années, il devient important d'établir un cadre clair et fonctionnel pour permettre une communication efficace entre les individus et les outils informatiques. Dans le contexte du Web sémantique, la qualité des ontologies est cruciale pour maximiser la richesse d'informations paramétrées tout en assurant une exploitation efficace de ces données. La création d'une ontologie se juge donc à son niveau de pertinence, c'est pourquoi l'utilisation de différentes techniques de traitement automatique du langage naturel (TALN) apporte davantage de ressources et de potentiel d'automatisation des ontologies, notamment par des systèmes d'apprentissage des ontologies à partir de diverses sources de données. Bien que ces techniques d'automatisation favorise la non-intervention humaine tout en garantissant la qualité des résultats obtenus, elles demeurent pour autant partielles ou limitées. L'article liste un ensemble de techniques utilisées dans le processus d'apprentissage des ontologies, appelant ainsi à une systématisation des connaissances dans ce domaine, comme l'analyse des définitions de dictionnaire et la désambiguïsation des sens des mots, l'extraction de connaissances à partir de modèles syntaxiques,

les relations sémantiques, les motifs de verbes, les méthodes d'extraction, les dépendances syntaxiques, les méthodes de programmation logique inductive, l'utilisation de modèles de sujets, et autres.

En réponse à ce défi, l'auteure propose donc une démarche pratique pour guider le processus d'apprentissage des ontologies, à travers une certaine interopérabilité des outils informatiques, tout en visant à apporter une clarté conceptuelle et méthodologique dans un domaine donné afin de faciliter la compréhension et la mise en œuvre de ces méthodes.

Chapitre 3

Méthodes de représentation ontologique des données sémantiques et matériels

Dans ce chapitre, les différentes méthodes employées et le processus des étapes pour appliquer ces méthodes, d'une part, l'exploration des données, et, d'autre part, les méthodes pour construire l'ontologie.

3.1 Exploration des données brutes et numérisées sur les sceaux byzantins de la collection Zacos

En ce qui concerne les données exploitables, celles-ci sont d'abord observées et récoltées par les chercheurs pour seulement ensuite être entrées informatiquement sur le site du Musée de Genève. C'est à partir de là que les données numérisées peuvent être récoltées informatiquement du fait de leur nature numérisée et regroupées dans des fichiers servant de support de travail comme nous allons le développer maintenant.

3.1.1 Observation des données brutes des sceaux byzantins

Le travail de création d'une ontologie pour les sceaux byzantins se fait à partir des sceaux eux-mêmes qui sont répertoriées sur le site du Musée de Genève [[byza](#)]. De manière générale les sceaux comportent des éléments similaires allant des inscriptions aux figures emblématiques représentées sur les surfaces de l'objet antique [[byzb](#)].

Dans la collection Zacos, nous retrouvons de la même manière deux surfaces l'avant et le revers qui comportent effectivement un ensemble de figures emblématiques de l'iconographie byzantine avec des figures religieuses telles que le Christ, la Vierge Marie, les Saints ou encore

des figures ecclésiastiques comme certains patriarches éminents, mais aussi des figures politiques variés tels que les empereurs, les femmes d'empereurs, les impératrices, des membres de la famille impériale ou autres personnages dotés d'une autorité particulière.

Ces différents personnages sont mis en scène sur les surfaces des sceaux, particulièrement la face avers, et peuvent être accompagnés de symboles spécifiques comme des croix et des croixettes, ou encore les Evangiles ou le modèle de l'Eglise de Sainte-Sophie. Les figures peuvent avoir des positions différentes, peuvent être à plusieurs représentés sur une surface comme la Vierge Marie qui tient l'Enfant Jésus, ou encore la Vierge Marie qui est accompagné d'empereurs. Ces représentations sont une mise en scène pouvant parfois représenter des passages bibliques, ou alors représenter une ou plusieurs figures spécifiques dans un décor caractéristique et propre à ces figures, comme la présence d'items propre au soldat (arme, bouclier, monture, etc), des trônes, des vêtements spécifiques à un statut social, entre autres. Ces représentations sont caractéristiques de la Chrétienté et de la culture byzantine où un personnage est doté d'une signification particulière et donc une utilisation particulière.

Ces sceaux peuvent aussi comporter la plupart du temps tant sur l'avvers que le revers, et surtout sur le revers, des inscriptions écrites en langue grecque dans l'alphabet grec ancien, comme les initiales de Jesus-Christ, ou bien des phrases descriptives de la scène et des figures représentées sur les sceaux.

Cependant, du fait de l'ancienneté de ces objets et de la qualité des matériaux utilisés pour les fabriquer, la plupart sont abîmés, et, tant les représentations que les inscriptions peuvent être illisibles ou difficilement descriptibles. Cela apporte d'emblée un obstacle dans le traitement et l'exploitation des informations à extraire et réutiliser pour créer une ontologie claire et rigoureusement paramétrée. Des exemples de sceaux byzantins de la collection sont observables dans les figures 3.1 et 3.2.

3.1.2 Structure des données numérisés des sceaux de la collection Zacos

A partir de ces données brutes extraites sur ces sceaux, les données ont été collectées et publiées dans des catalogues, des publications et d'autres sources d'information. Pour ce mémoire, nos données sont tirées du site internet de la collection Zacos ainsi que de l'ouvrage Sceaux de la collection Georges Zacos au musée d'art et d'histoire de Genève [dedCPC⁺16]. Les informations ont été inscrites sur le site du Musée de Genève [byza] suivant une certaine méthode et une description standard similaire pour chaque sceau pour décrire les différents éléments présents.



FIGURE 3.1 – Sceau Byzantin en plomb datant du XIIe siècle et appartenant à Joseph



FIGURE 3.2 – Sceau en or datant du XIe siècle et appartenant à l'empereur Constantin IX Monomaque

Le fichier XML décrit trente objets de la collection Zacos conservés au Musée d'art et d'histoire de Genève. Le contenu du fichier est organisé de manière structurée dans une format XML où chaque objet est contenu dans une balise mère 'Seal' regroupant d'autres balises filles telles que '<Titre>', '<Date>', '<Person>', '<Materiaux>', '<Inventaire>', '<Collection>', '<Periode>', '<Mention_obligatoire>', '<Avers>', '<Revers>', '<Description>' et 'Bibliographie', dont cette dernière peut contenir des balises '<Article>'. Tous les éléments contenu dans le fichier n'ont pas nécessairement de contenu associé dans chaque balise fille pré-citées. Cette structure permet de casser et d'organiser les informations relatives à chaque sceau de manière claire et cohérente. En plus d'être un texte multilingue regroupant principalement trois langues (français, le grec ancien et le latin), les descriptions des sceaux sont parsemées de mots anciens, techniques et des

mots translittérés non identifiables et non standardisés et qui sont donc peu reconnaissables par une machine.

Les propriétés et caractéristiques de chacun des sceaux de la collection regroupent donc des informations sur la date ou la période associée à l'objet, des informations sur la personne propriétaire de l'objet ou l'entité associée à l'objet, des informations sur le matériau à partir duquel l'objet est fabriqué, des informations sur le numéro d'inventaire attribué à l'objet servant d'identifiant unique, des informations sur la collection à laquelle appartient l'objet, des informations sur la période historique à laquelle l'objet est lié, ou encore une description détaillée ou des informations supplémentaires sur l'objet observable sur les avers et revers des sceaux ou dans une partie descriptive ajoutée par l'observateur. Ces éléments sont enrichis par des bibliographie et des articles ayant rapport avec le sceau et les éléments présents sur celui-ci.

3.2 Etapes de la construction de l'ontologie

La création d'une ontologie des sceaux byzantins de la collection Zacos est une tâche complexe qui nécessite une compréhension approfondie de l'histoire, de l'art et de la culture byzantine. Une ontologie est un ensemble de concepts et de relations qui décrivent un domaine spécifique, tel que les sceaux byzantins. Les étapes nécessaires pour créer une ontologie des sceaux byzantins de la collection Zacos amènent à mettre l'accent sur les méthodes utilisées pour décrire les sceaux et les relations entre eux.

3.2.1 Parcours du fichier XML à l'aide de bibliothèque

Pour pouvoir exploiter le contenu du fichier XML, l'analyse de celui-ci est donc l'étape initiatrice dans le processus de construction de l'ontologie du patrimoine culturel spécifique aux sceaux byzantins.

Dans le code Python généré pour ce travail, la bibliothèque 'xml.etree.ElementTree' est utilisée pour parcourir et analyser le fichier XML qui contient les données des sceaux byzantins. Tout d'abord, la fonction 'parse' de la bibliothèque citée charge le fichier XML et le convertit en une structure d'arbre XML qui permet d'accéder facilement aux différentes parties du document XML. De cette manière, le code parcourt l'arbre XML en utilisant des itérations et des requêtes XPath pour extraire les informations pertinentes. Les données extraites sont ensuite stockées dans des listes correspondant aux différents attributs des sceaux byzantins. Par exemple, si le fichier XML contient des informations sur le titre, les personnages, le matériau, la collection ou

encore la période d'un sceau, alors chaque attribut aura sa propre liste qui contiendra les valeurs extraites du fichier XML.

Cette étape permet de récupérer les données spécifiques des sceaux byzantins qui seront utilisées ultérieurement pour créer les instances de la classe 'Sceau' et les triplets RDF correspondants. En extrayant ces informations du fichier XML, cela facilite la structuration et l'organisation des données dans le graphe RDF, conformément à l'ontologie définie.

3.2.2 Ensemble des méthodes de codification de l'ontologie des sceaux byzantins de la collection Zacos

Dans cette sous-section, nous allons voir la préparation des bibliothèques et espaces sous Python pour créer l'ontologie.

3.2.2.1 Import de la bibliothèque 'rdflib' pour manipuler les graphes RDF

Dans un premier temps, le point de départ du code Python est l'importation des bibliothèques spécialisées nécessaires à la construction de l'ontologie. La seconde bibliothèque exploitée est 'rdflib'. C'est une bibliothèque, au rôle primordial, qui fournit des classes et des méthodes pour représenter, manipuler et sérialiser des graphes RDF contenant des outils spécialisés dans l'organisation et la gestion des données et annotations pour la construction de l'ontologie, en utilisant des concepts et des relations adaptés.

3.2.2.2 Définition des espaces de noms RDF

Dans un second temps, il s'agit d'aborder la définition des espaces de noms RDF. Les espaces de noms RDF permettent de définir les termes, les classes, les propriétés et les relations utilisés dans le modèle RDF, permettant de structurer et apporter une sémantique claires pour les données. Dans le code Python généré, plusieurs espaces de noms RDF sont définis.

L'espace de noms 'OntoSeals' représente l'ontologie spécifique aux sceaux byzantins qui est en cours de construction. Il permet de définir les classes, les relations et les propriétés spécifiques à ce domaine. 'DBpedia' [LIJ⁺14] fait, quant à lui, référence à la base de connaissances DBpedia, une ressource majeure dans le domaine de la représentation des connaissances, permettant d'enrichir les données des sceaux byzantins en les reliant à des informations complémentaires. Aussi, 'SCHEMA' représente le schéma Schema.org, une ontologie largement utilisée pour la structuration des données sur le Web, qui permet d'annoter des ressources Web utiles liées aux sceaux byzantins, et, ainsi mieux intégrer les premières données avec d'autres sources de données. Ensuite, deux types d'espaces de noms RDF d'ontologie sont utilisés. D'une part, l'ontologie 'FOAF' (Friend of a Friend) sert à décrire les personnes et leurs relations, comme relier un sceau à son propriétaire. D'autre part, l'ontologie Dublin Core qui fournit des termes pour

décrire des ressources bibliographiques, et, ainsi, annoter des informations bibliographiques sur les sceaux byzantins, telles que les sources primaires ou les références disponibles dans la balise de la bibliographie par exemple.

Ainsi, la définition de ces espaces de noms RDF en amont sert à la spécification des ontologies utilisées et à la création d'une structure cohérente pour les données des sceaux byzantins. Chaque espace de noms représente une source d'informations ou une ontologie spécifique qui contribue à l'enrichissement des données et à la représentation sémantique précise.

3.2.3 Création du contenu du graphe contenant les classes, instances et relations de l'ontologie

Dans cette sous-section, nous allons voir l'application des méthodes depuis Python pour créer les classes, instances et relations de l'ontologie à partir du fichier XML généré.

3.2.3.1 Création du graphe RDF

Pour la création du graphe RDF, il s'agit d'utiliser la classe 'Graph' fournie par la bibliothèque 'rdflib' précédemment mentionnée. Ce graphe contient l'ensemble des triplets RDF représentant les connaissances relatives aux sceaux byzantins. En effet, le graphe RDF est une structure de données qui permet de représenter des informations sous la forme de triplets, où chaque triplet est composé d'un sujet (l'instance de sceau), d'un prédicat (la relation) et d'un objet (la valeur correspondante).

Chaque triplet RDF sujet-prédicat-objet exprime une assertion ou une relation entre des entités, c'est-à-dire, dans ce contexte, il s'agit de capturer les informations sur les différents aspects des sceaux, tels que leur matériel, leur période, leur collection, etc. Cette classe permet d'ajouter, supprimer et interroger les triplets RDF dans le graphe. Dans le code Python généré, la création du graphe se fait simplement en instanciant un objet de la classe 'Graph', comme suit : 'g = Graph()', et il recevra, au fur et à mesure, les triplets RDF qui seront générés à partir des données des sceaux byzantins, chaque triplet RDF représentant donc une assertion ou une relation spécifique dans le domaine des sceaux byzantins.

3.2.3.2 Définition de fonctions pour créer des classes et des relations d'objets

Une fois le graphe créé, vient l'étape de la définition des classes et des relations d'objets qui représentent les concepts et les liens entre les différentes entités. Pour faciliter cette tâche, le code Python généré met en place des fonctions spécifiques, telles que 'createClass' et 'createObjectRelation', qui permettent de générer les triplets RDF correspondants dans le graphe RDF. Ces deux fonctions utilisent la méthode 'add' du graphe RDF pour ajouter le triplet RDF

correspondant à la relation d'objets. L'utilisation de ces fonctions facilite non seulement la création des classes et des relations dans l'ontologie, en automatisant le processus de génération des triplets RDF, mais aussi, permet d'éviter les erreurs et les duplications, tout en assurant la cohérence et la structure appropriée de l'ontologie.

Cela contribue ainsi à une modélisation précise et cohérente des connaissances sur les sceaux byzantins avec une sémantique riche. En effet, la classe Python 'createClass' sert à créer une classe dans l'ontologie en ajoutant un triplet RDF correspondant au graphe RDF. Cette classe prend en paramètre le nom de la classe à créer. Par exemple, si l'on souhaite créer une classe appelée «Titre», on peut appeler la fonction 'createClass(«Titre»)', qui ajoutera le triplet RDF représentant cette classe au graphe dit.

De même, la fonction 'createObjectRelation' est utilisée pour créer une relation d'objets entre deux classes dans l'ontologie. Cette fonction prend en paramètre le nom de la relation et les classes de domaine entre lesquelles la relation est établie. Par exemple, si l'on souhaite créer une relation appelée «a materiel» entre les classes «Sceau» et «Materiel», la fonction 'createObjectRelation(«a materiel», «Sceau», «Materiel»)' peut être utilisée pour générer le triplet RDF correspondant dans le graphe.

3.2.3.3 Création des classes dans l'ontologie

Les classes créées dans l'ontologie, à partir de la classe 'createClass' et du fichier XML fourni, sont 'Titre', 'Personne', 'Materiel', 'Sceau', 'Collection' et 'Periode'. Chacune de ces classes représente un aspect spécifique des sceaux byzantins et joue un rôle clé dans la représentation des informations relatives à ces objets culturels.

La classe 'Titre' est utilisée pour catégoriser les différents titres ou rang sociaux associés aux sceaux byzantins, selon qu'il s'agisse d'un homme politique comme un Empereur ou une Impératrice, ou sinon d'un homme d'église, comme un ecclésiaste, etc. La classe 'Personne' représente les individus liés aux sceaux byzantins, qu'il s'agisse des propriétaires des sceaux ou d'autres personnes associées à l'objet. Cette classe permet de modéliser les informations relatives aux personnes, telles que leurs noms, leurs rôles et leurs liens avec les sceaux byzantins.

La classe 'Materiel' est utilisée pour représenter les matériaux utilisés dans la fabrication des sceaux byzantins. Elle permet de spécifier les caractéristiques physiques et les propriétés des matériaux utilisés, tels que l'or, l'argent, le bronze, etc. La classe 'Sceau' est au cœur de l'ontologie, représentant les sceaux byzantins eux-mêmes. Elle englobe les informations spécifiques à chaque sceau, comme précédemment mentionnées.

La classe 'Collection' permet de regrouper les sceaux byzantins en collections, qu'elles soient privées, publiques ou institutionnelles. Elle joue un rôle important dans l'organisation et la catégorisation des sceaux byzantins en fonction de leur provenance et de leur appartenance à une collection spécifique. Ici, il ne s'agit que de la collection Zacos.

Enfin, la classe ‘Periode’ est utilisée pour représenter les différentes périodes historiques dans lesquelles les sceaux byzantins ont été créés et utilisés. Cela permet de contextualiser les sceaux et de les relier à des événements historiques spécifiques ou à des changements stylistiques et iconographiques.

En utilisant la classe Python ‘createClass’ du code Python, chaque classe est créée dans l’ontologie en ajoutant les triplets RDF correspondants au graphe RDF. Ces triplets RDF contiennent des informations sur le nom de la classe et l’espace de noms auquel elle appartient. Par exemple, la classe ‘Sceau’ peut être définie avec le triplet RDF suivant : ‘ g.add((OntoSeals.Sceau, RDF.type, OWL.Class)) ‘. Dans ce triplet, ‘OntoSeals.Sceau’ représente le nom de la classe ‘Sceau’ dans l’espace de noms RDF ‘OntoSeals’, et ‘RDF.type’ est une propriété prédéfinie de RDF indiquant que ‘OntoSeals.Sceau’ est une classe. De manière similaire, les autres classes telles que ‘Titre’, ‘Personne’, ‘Materiel’, ‘Collection’ et ‘Periode’ sont créées en utilisant la fonction ‘createClass’, avec leurs noms et espaces de noms respectifs. Cette hiérarchie de classes dans l’ontologie fournit ainsi une structure conceptuelle pour organiser les informations relatives aux sceaux byzantins.

3.2.3.4 Création des relations entre les classes

Pour créer des relations entre les classes, la fonction ‘createObjectRelation’ est utilisée pour créer une relation d’objets entre deux classes en ajoutant un triplet RDF correspondant au graphe RDF. De cette manière, elle établit des liens sémantiques entre les différentes entités représentées par les classes, en définissant les propriétés et les associations qui les caractérisent. Par exemple, la relation ‘est de type’ peut être définie entre la classe ‘Sceau’ et la classe ‘Titre’ de la manière suivante : ‘ g.add((OntoSeals.Sceau, OntoSeals.est de type, OntoSeals.Titre)) ‘. Dans ce triplet, ‘OntoSeals.Sceau’ est la classe de domaine, ‘OntoSeals.est de type’ est la relation d’objets, et ‘OntoSeals.Titre’ est la classe de plage. En ajoutant ce triplet au graphe RDF, on établit la relation «un sceau est de type titre» dans l’ontologie.

De la même manière, d’autres relations peuvent être créées pour décrire les associations et les caractéristiques des sceaux byzantins, comme la relation ‘appartient a’ qui définit la relation entre la classe ‘Sceau’ et la classe ‘Collection’ pour indiquer l’appartenance d’un sceau à une collection spécifique, ou encore la relation ‘a materiel’ qui établit une relation entre la classe ‘Sceau’ et la classe ‘Materiel’ pour représenter le matériau utilisé dans la fabrication d’un sceau donné.

3.2.4 Sérialisation du graphe RDF par l'ajout des instances et classes des sceaux byzantins

3.2.4.1 Instanciation des attributs paramétrés

Suite à l'assignation de chaque valeur d'un sceau donné à l'attribut 'self.X' de l'instance, où 'X' représente une valeur d'un sceau donné, il est possible d'utiliser la méthode 'addInstance' pour générer les triplets RDF correspondants et les ajouter au graphe RDF, car cette méthode prend en compte les attributs de chaque instance de sceau et génère les triplets RDF appropriés en utilisant les espaces de noms RDF définis précédemment.

3.2.4.2 Sérialisation du graphe RDF

La dernière étape du code généré consiste en la sérialisation du graphe RDF pour représenter les données structurées en RDF d'une manière lisible et compatible avec les standards du Web sémantique. Pour effectuer la sérialisation, la méthode 'serialize' de l'objet 'g' est utilisée pour convertir le graphe RDF en une représentation textuelle conforme au format Turtle, un langage de sérialisation RDF utilisé dans le Web sémantique. Ce format utilise une syntaxe simple et humainement lisible, ce qui facilite la compréhension et l'interprétation des données RDF par les humains et les machines.

En effet, concrètement, les triplets RDF sont présentés sous forme de sujets, de prédicats et d'objets, où chaque triplet est séparé par un point-virgule et est affiché sur une ligne distincte. L'utilisation de cette syntaxe apporte la possibilité d'explorer et d'analyser des données RDF, en mettant en évidence les relations entre les différentes instances de sceaux byzantins et en fournissant une vue globale de l'ontologie construite, par exemple sur un site Web spécialisé à partir duquel il est possible de générer une vue du graphe ou encore par une des fonctionnalités de l'outil Protégé.

Ainsi, la représentation textuelle du graphe RDF en format Turtle peut être utilisée aussi ultérieurement pour être intégrée dans des ressources Web pertinentes par l'annotation. Cela enrichit, par la même occasion, le contenu sémantique sur les sceaux byzantins ou encore permettre de répondre à des questions complexes dans le domaine des sceaux byzantins en exploitant les standards et les outils du Web sémantique.

Ainsi, la sérialisation des étapes de construction de l'ontologie des sceaux byzantins de la collection Zacos permet de structurer et annoter de manière cohérente les données relatives à ces objets historiques. Le modèle RDF facilite la gestion et l'exploitation des informations matériels et sémantiques de manière numérique. L'analyse de la structure des sceaux dans le fichier XML permet d'extraire les informations pertinentes et de les stocker dans des listes correspon-

dant aux attributs des sceaux.

L'utilisation de la bibliothèque rdflib et la définition des espaces de noms RDF sont primordiales pour manipuler et organiser les données qui sont paramétrées dans le graphe créée et peuplée de manière structurée. En combinant ces étapes, l'ontologie des sceaux byzantins de la collection Zacos peut être construite de manière rigoureuse, et faciliter ainsi l'exploitation et l'enrichissement des informations sur ce patrimoine culturel.

3.2.5 Enrichissement par la modification des balises de date dans un fichier XML

Le fichier XML d'origine comporte une multitude d'informations, pour la majorité de type textuel, ce qui peut pour certaines données rendre la compréhension et la sauvegarde de l'information plus complexe.

3.2.5.1 Utilisation d'un fichier texte regroupant les appellations et les conversions des dates

Le premier cas traité est celui des dates contenues dans les balises filles '`<Date>`' qui sont écrit non seulement en chiffres, mais pour certaines en toutes lettres et concernant des périodes vagues données en terme de siècles comme «XI^e siècle». Les dates des trente sceaux du fichier XML d'origine doivent donc être remplacées en chiffres pour une facilité d'usage. Pour cela, un code a été généré avec Python.

Tout d'abord, le module '`re`' est importé, permettant l'utilisation d'expressions régulières, comme '`re.findall`', pour effectuer des recherches et des remplacements de motifs dans le texte. Ici, il s'agit de lire les balises '`Date`' dans le fichier XML et stocker son contenu dans une variable, et stocker dans une seconde variable les dates regroupées et converties dans un fichier `.txt`. De cette manière, le code généré trouve leurs correspondances dans le fichier '`dates.txt`', et remplace les balises '`<Date>`' par des balises '`<DateStart>`' et '`<DateEnd>`' contenant les différentes dates correspondantes et extraites du fichier '`dates.txt`'.

Ainsi, les dates présentes dans les balises du nouveau fichier XML seront toutes converties en chiffres numériques et non plus en caractères textuels comme des siècles ou parties de siècles, amenant à davantage de lisibilité et une meilleure lecture et utilisation des informations sur les sceaux byzantins.

3.2.5.2 Utilisation d'un fichier JSON regroupant les appellations et les conversions des dates

Un autre possibilité pour faciliter la manipulation des dates abstraites dans le fichier XML et la création d'un fichier JSON regroupant l'ensemble des années de l'empire byzantin, de 330 à 1453 la chute de l'empire byzantin, où l'année est la clé, et chaque clé à pour valeurs les différentes appellations attribuables à cette année, par exemple «3e quart du (century)e s.» correspondant aux années entre XX75 jusqu'à la fin du siècle donné.

Ensuite, les fichiers JSON et XML sont mis en relation pour effectuer la conversion et le remplacement des années et périodes de siècles plus abstraites par des années plus spécifiques selon l'encadrement que nous avons décidé. Les principales conditions sont au nombre de quatre.

Pour la première condition, il s'agit du remplacement des années de départ dans les balises «<DateStart> </DateStart>». Si la première année contenue dans la balise date contient trois chiffres et correspond à une année présente dans le fichier JSON, elle sera remplacée par les balises «<DateStart> </DateStart>» contenant cette année comme année de départ. De même, si la première année contenue dans la balise date contient quatre chiffres et correspond à une année du fichier JSON, elle sera également remplacée dans les balises «<DateStart> </DateStart>».

Pour la deuxième condition, il en est de même avec le remplacement, cette fois, des années de fin dans les balises «<DateEnd> </DateEnd>». Si la seconde année contenue dans la balise date contient trois chiffres et correspond à une année du fichier JSON, elle sera remplacée par les balises «<DateEnd> </DateEnd>» contenant cette année comme année de fin. De même, si la seconde année contenue dans la balise date contient quatre chiffres et correspond à une année du fichier JSON, elle sera remplacée dans les balises «<DateEnd> </DateEnd>».

S'ensuit le remplacement des mélanges de caractères en lettres et chiffres, pour permettre plus de faciliter dans la manipulation des données qui permettent d'établir le cadre temporel. D'une part, si le contenu de la balise «<Date> </Date>» est un mélange de caractères en lettres et chiffres, il sera recherché dans le fichier JSON pour trouver une date correspondante. La balise «<Date> </Date>» sera remplacée par les balises «<DateStart> </DateStart> - <DateEnd>» contenant la première et la dernière date correspondante du fichier JSON.

D'autre part, étant donné que les dates ne sont pas écrites sous la même forme, il s'agit aussi de traiter des appellations, telles que «avant XXXX». Lorsque les balises «<Date> </Date>» contiennent l'appellation «avant» suivie d'une année, le terme «avant» sera remplacé par «(?)» dans les balises «<DateStart> </DateStart>». L'année spécifiée après le mot «avant» dans le

fichier XML sera placée dans les balises «<DateEnd> </DateEnd>» comme date de fin.

Enfin, en utilisant des opérations de lecture, de traitement et d'écriture de fichiers, nous pouvons parcourir les balises de date du fichier XML et appliquer les conditions mentionnées précédemment pour effectuer les modifications nécessaires.

Par ailleurs, la difficulté réside dans les détails, à savoir comment encadrer les différentes périodes abstraites lorsqu'il s'agit de siècles. Il est alors nécessaire de structurer l'automatisation de l'extraction, de la conversion et du remplacement des données manipulées pour permettre une meilleure manipulation de celle-ci par la suite et faciliter leur exploitation par la standardisation de leur traitement.

3.2.6 Amélioration par l'enrichissement des balises du fichier XML avec spaCy

L'enrichissement du balisage peut se renforcer davantage encore avec des bibliothèques telles que spaCy¹ qui est une bibliothèque open-source de traitement automatique du langage naturel (TALN), possédant plusieurs outils dont le postagging, la lemmatisation ou encore l'Universal Dependencies (UD), et ainsi obtenir des informations riches sur la structure grammaticale et sémantique du contenu textuel de notre fichier XML.

3.2.6.1 Enrichissement du balisage à l'aide du postagging de la bibliothèque spaCy

Pour la première étape, il s'agit de recourir au postagging de la bibliothèque spaCy à chaque mot du contenu textuel, plus précisément les contenus textuels des balises 'Avers', 'Revers', et 'Description'. Comportant des éléments de précision sur le contenu des sceaux, ces balises sont celles qui servent à améliorer la base de données à paramétrer afin d'enrichir l'ontologie sur les sceaux byzantins par davantage de relations, de classes ou encore d'instances. En effet, SpaCy peut identifier les parties du discours, telles que les noms, les verbes, les adjectifs, et autres et donc catégoriser plus précisément les informations extraites des balises XML.

1. <https://spacy.io/>

3.2.6.2 Enrichissement du balisage à l'aide de l'Universal Dependencies de la bibliothèque spaCy

L'Universal Dependencies (UD) ou l'analyse de dépendances permet de comprendre comment les mots sont structurés dans une phrase et comment ils interagissent les uns avec les autres. Cela permet d'identifier les relations entre les entités représentées sur les sceaux, telles que l'empereur, la Vierge Marie, ou des objets présents sur les sceaux byzantins.

Voici un exemple ci-après de différents étiquetages obtenus grâce à l'outil d'analyse des dépendances.

```

<DET>un</DET> <NOUN dep='nmod'>cercle</NOUN> <NUM>d'</NUM> <VERB le
Nicolas</Personnage> <ADP>à</ADP> <NOUN dep='xcomp'>mi-corps</NO
> <NOUN dep='conj'>tenue</NOUN> <ADJ>épiscopale</ADJ> <VERB lemma
age>Livre</Personnage> <ADP>en</ADP> <NOUN dep='nmod'>main</NOU
> <NOUN dep='obl:arg'>droite</NOUN> <PUNCT>.</PUNCT> <ADP>De</AD
mod'>l'</NOUN> <NOUN dep='conj'>effigie</NOUN> <PUNCT>,</PUNCT>
rs><Revers><SPACE> </SPACE> <PROPN dep='nmod'>Légende</PROPN> <A
rmontée</VERB> <NUM>d'</NUM> <DET>un</DET> <NOUN dep='obj'>losang
ADP> <NOUN dep='obl:arg'>tirets</NOUN> <CCONJ>et</CCONJ> <VERB le
<PUNCT>:</PUNCT></Revers><Description><PROPN dep='nsubj'>Sophia<

```

FIGURE 3.3 – Exemple d'étiquetage de l'analyse des dépendances sur le sceau byzantin CdN 2004-0313

Dans cet exemple, nous pouvons voir plusieurs types de composant de la notation effectuée, telle qu'elle est utilisée dans l'analyse linguistique pour décrire les relations grammaticales entre les noms (NOUN) et leurs dépendances respectives dans une phrase ou un texte.

Le premier élément, <NOUN dep='conj'>, représente un nom (NOUN) qui agit comme une conjonction (conj), c'est-à-dire qu'il connecte généralement deux éléments au sein de la phrase. Le deuxième, <NOUN dep='nmod'>, indique un nom (NOUN) servant de modificateur nominal (nmod) à un autre nom, et sert à désigner des informations ou du contexte supplémentaires. Le quatrième élément, <NOUN dep='nsubj'>, indique que le nom (NOUN) joue le rôle de sujet nominal (nsubj) de la phrase, ce qui signifie qu'il effectue l'action du verbe principal. Ensuite, <NOUN dep='obl:arg'> décrit un nom (NOUN) agissant comme un complément oblique (obl:arg), apportant des informations supplémentaires sur une action ou un état dans la phrase. Enfin, <NOUN dep='xcomp'> signifie qu'un nom (NOUN) fonctionne comme un complément clausal ouvert (xcomp), décrivant une action ou un état lié au verbe principal. Ces notations sont utilisées pour définir les rôles grammaticaux et les relations des noms au sein des phrases lors

de l'analyse linguistique. Leur signification précise dépend du contexte et du cadre linguistique employé pour l'analyse.

3.2.6.3 Enrichissement du balisage à l'aide de la reconnaissance d'entités nommées de la bibliothèque spaCy

Un autre outil de la bibliothèque spaCy permet la détection et la catégorisation des entités nommées, telles que les noms de personnes, des lieux, ou encore des objets. Cette fonctionnalité est précieuse pour identifier les acteurs clés, les lieux mentionnés, et les objets représentés sur les sceaux byzantins. Par exemple, spaCy peut identifier les différents personnages apparaissant sur les sceaux byzantins, comme les empereurs ou les saints, mais pour autant, étant donné la particularité de la tâche et du contexte d'utilisation sur un ensemble de données traitant d'une époque révolue et très ancienne, il peut arriver que cet outil détecte le nom d'un saint comme étant un lieu au lieu d'un nom propre, comme c'est le cas dans les résultats avec le nom propre «Sainte-Sophie» qui n'est pas non plus une sainte mais signifie la «Sainte Sagesse».

3.2.6.4 Enrichissement du balisage à l'aide de la lemmatisation de la bibliothèque spaCy

Le processus de lemmatisation a été utilisé pour obtenir le lemme, notamment des verbes, pour pouvoir les exploiter plus facilement par la suite. Cet outil permet de réduire des mots à leur forme de base, ce qui est essentiel pour regrouper les différentes formes d'un mot en une seule entité pour pouvoir faciliter la création de l'ontologie par la suite.

En utilisant cet outil, le but était de faciliter la saisie des informations de manière plus large en ce qui concerne les verbes. Par exemple, les verbes présents dans les balises des sceaux sont assez similaires et répétitifs comme le verbe tenir ou bénir. Le fait de les lemmatiser permet de simplifier leur balisage et extraction par la suite pour permettre la création de triplets, ou capturer les éléments autour comme l'objet qui suit le verbe.

Le but étant de faciliter la création de triplets RDF, où les entités sont liées aux prédicats (verbes) et aux objets, la normalisation des verbes contribue ainsi à la capture d'éléments connexes dans les balises des sceaux pour une extraction d'informations plus riche sur les actions et les attributs associés aux objets historiques représentés par les sceaux, tout en permettant de créer des relations sémantiques plus cohérentes entre les concepts de l'ontologie.

Voici un exemple ci-après d'une lemmatisation optimale des verbes avec le lemme du verbe indiqué dans la balise correspondante.

```
/NOUN> <VERB lemma='indique'>indique</VERB> <ADP>sans</ADP> <NOUN dep='obl:mod'>doi
> <VERB lemma='financer'>financé</VERB> <DET>une</DET> <NOUN dep='obj'>distributio
<NOUN dep='obl:mod'>église</NOUN> <VERB lemma='placer'>placée</VERB> <ADP>sous</ADP>
```

FIGURE 3.4 – Exemple de lemmatisation des verbes sur le sceau byzantin CdN 2004-0313

3.2.7 Amélioration par l'enrichissement des balises du fichier XML avec GATE

Pour améliorer autrement le balisage des informations textuelles pertinentes concernant les sceaux byzantins, il est possible d'utiliser un ensemble de règle Jape (Java Annotation Patterns Engine). Ces règles sont à rédiger et à exploiter à travers le framework GATE (General Architecture for Text Engineering). Il s'agit donc d'allier différents outils permettant la capture et le paramétrage des données isolées.

3.2.7.1 Ciblage des informations à l'aide de règles Jape

A l'aide d'expressions régulières ou de motifs complexes, sous formes de règles syntaxiques, les règles Jape permettent d'identifier les éléments voulus pour les annoter avec précision de manière automatique, que ce soit des entités comme les personnages, des relations comme le fait pour un personnage de tenir un objet dans une certaine main, ou encore des structures grammaticales facilitant la saisie par exemple de noms propres (PROPN), d'adjectifs (ADJ), des verbes (VERB), des noms (NOUN), etc. Pour cela, il faut rigoureusement préciser, dans les règles, le motif pour la cible à annoter, le motif du contexte dans lequel la cible peut être capturée, en notant les contraintes pour plus de précision dans la saisie, et les conditions d'application de l'annotation une fois la cible trouvée. Pour l'application aux cas des sceaux byzantins, ces règles peuvent être très pratique pour capturer les différents personnages et leurs différentes appellations, comme «Jesus-Christ» qui peut aussi être appelé «l'Enfant» ou encore «Christ Emmanuel», mais aussi des objets et catégoriser ces objets selon leur type, des relations spécifiques, des positions, etc.

Dans le cas de la règle Jape pour annoter la Vierge Marie, qui est un personnage redondant et centrale dans l'univers des sceaux byzantins, il faut prendre en compte le fait qu'elle peut être nommée dans le contenu textuel selon différentes façons. C'est pourquoi, comme pour Jesus-Christ, Sainte-Sophie ou encore le Saint-Esprit, des règles Jape ont été spécifiquement rédigées pour ces personnages essentiels et peuvent être enrichies par des Macros réutilisables dans les règles. Les mentions de la Vierge Marie peuvent prendre diverses formes, telles que «La Vierge», «Mère de Dieu», ou d'autres appellations similaires.

Pour rester sur l'exemple de la Vierge Marie, le code Jape associé commence par la définition, dans la phase, les annotations spécifiques qui seront utilisées dans la ou les règles qui suivent. Cette phase prend en général des «Token», ou encore des «SpaceToken» pour les espaces entre les tokens ou mots. Mais aussi les annotations déjà présentes dans le fichier de base, telles que les annotations créées à partir des étiquettes obtenues lors de l'étape du postagging. Nous pouvons citer entre autres les annotations des verbes lemmatisés, les déterminants, les adjectifs ou bien la ponctuation. A partir de là, les règles, telle que «ViergeMarieAp LaViergeEpiskepsis», capture l'appellation de la Vierge Episkepsis. Dans les informations ajoutées à cette annotation, il y a aussi le genre du personnage qui est pour la Vierge le genre féminin. Toutes ces appellations sont reliées par le concept général «La Vierge Marie» qui servira par la suite à relier toutes ces appellations à la même personne qu'est la sous-classe de la Vierge Marie dans la classe «Personnage» de l'ontologie, comme illustre la figure 3.5.

Toutes ces règles sont séparées par le «|» (le «ou» logique) ou caractérisées par d'autres opérateurs tels que «?» (zéro ou une occurrence), permettant ainsi de trouver l'appellation la plus adaptée au nom que nous cherchions. De plus, l'option «control = appell», nous permet de trouver la correspondance la plus longue et éviter de contabiliser chaque correspondance. Nous pouvons reprendre l'exemple cité auparavant de la Vierge Episkepsis qui grâce à ces règles Jape va être contabilisée comme une seule correspondance plutôt que deux telles que «la Vierge» et «la Vierge Episkepsis».

```

Input: Token — Annotations de base
Options: control = appell — Occurrence la plus longue

Macro: VirginMaryMacro
(
  ((Token.string =~ "La|la"))?
  (Token.string =~ "Vierge|vierge")
  ((Token.string =~ "Marie|marie"))?
)

Rule: ViergeMarie| — Nom de la règle
(( (VirginMaryMacro):ViergeMarieAp |
  (
    ((VirginMaryMacro)((Token.string =~ "(Episkepsis|episkepsis)(orante)?"))):ViergeMarieAp_LaViergeEpiskepsis
    ((VirginMaryMacro)((Token.string =~ "Platytera|platytéra"))):ViergeMarieAp_ViergePlatytera
    ((VirginMaryMacro)((Token.string =~ "Hodégéttria|hodégéttria"))):ViergeMarieAp_ViergeHodegetria
    ((VirginMaryMacro)((Token.string =~ "Nikopoios|nikopoios"))):ViergeMarieAp_ViergeNikopoios
    ((Token.string =~ "Mère|mère")(Token.string =~ "de")(Token.string =~ "Dieu|dieu")):ViergeMarieAp_MereDeDieu
    ((VirginMaryMacro)((Token.string =~ "Odigitria|odigitria"))):ViergeMarieAp_ViergeOdigitria
    ((Token.string =~ "(La|la)")(Token.string =~ "(Théotokos|théotokos)")):ViergeMarieAp_LaTheotokos
  )
):Personnage — Nom de l'annotation
-->
:ViergeMarieAp.Personnage = {majorType = "Personnage biblique", type = "Vierge Marie",minorType = "Vierge Marie", Genre = "Feminin"},
:ViergeMarieAp_LaViergeEpiskepsis.Personnage = {majorType = "Personnage biblique", type = "Vierge Marie",minorType = "Vierge Episkepsis",
"Feminin"},
:ViergeMarieAp_ViergePlatytera.Personnage = {majorType = "Personnage biblique", type = "Vierge Marie",minorType = "Vierge Platytera", Genre =
"Feminin"},
:ViergeMarieAp_ViergeHodegetria.Personnage = {majorType = "Personnage biblique", type = "Vierge Marie",minorType = "Vierge Hodégéttria", Genre =
"Feminin"},
:ViergeMarieAp_ViergeNikopoios.Personnage = {majorType = "Personnage biblique", type = "Vierge Marie",minorType = "Vierge Nikopoios", Genre =
"Feminin"},
...

```

Partie gauche de la règle
contient les motifs
d'annotations à identifier

Attributs de
l'annotation

FIGURE 3.5 – Illustration de la grammaire Jape utilisée pour annoter toutes les occurrences correspondantes à la Vierge Marie

De même, les regex permettent de capturer le contenu textuel écrit en lettres grecques et qui n’a pas été balisé au préalable dans la balise ‘Greek’. En effet, bien que le fichier XML soit rédigé principalement en français, il comporte également des noms propres ou des ensembles de tokens qui se suivent en phrases, groupes nominales, groupes verbales entre autres et qui ont leur importance. Etant donné que les données portent sur les sceaux byzantins, ces écritures en langue grecque, plus précisément de l’alphabet grec ancien, renvoie notamment aux noms des personnages représentés sur les sceaux, ou encore une sorte de court résumé inscrit sur le sceau en rapport avec le personnage, comme les initiales de la Vierge Marie, nommée la «Mère de Dieu» ou encore «Jésus-Christ» en grec.

Suite aux annotations des personnages, nous avons donc procédé à la création de grammaires Jape pour trouver les correspondances des positions des personnages à l’intérieur des sceaux. Après analyse de la partie textuelle du fichier XML, nous avons remarqué que les positions des personnages sont soit de face, soit assises, au centre, à droite ou à gauche et debout.

Nous avons alors créé une grammaire Jape en prenant en compte ces informations pour certaines précédée ou suivi d’une annotation «ADP» (adposition) annoté à l’aide du PosTagger. Pour illustrer ces propos nous pouvons prendre en exemple la position «assise», qui peut être annotée tout simplement telle qu’elle ou si elle est suivi de l’adposition «sur», elle sera alors annotée de la manière suivante «position = assis_sur» .

Nous pouvons aussi remarquer la position «de face» qui se décline aussi en «en face», et pour cette raison nous avons décidé de chercher toutes les correspondances d’une adposition suivi d’un token dont la chaîne de caractères est «face».

Nous avons effectué la même démarche pour la position des mains. Cette information nous permettra de faire la relation entre les personnages et les objets car dans la grande majorité des sceaux les personnages sont représentés avec un objet de chaque extrémité.

Ensuite, nous avons avancé à l’étape des objets. Pour automatiser cette étape nous avons fait recours à la bibliothèque Spacy pour obtenir les objets dans la phrase selon leur relation syntaxique à l’aide de l’outil des Universal Dependencies. Nous estimons que les objets correspondent aux complément d’objet direct. L’idée initiale était de pouvoir obtenir le sujet, le verbe et le COD dans la phrase comme par exemple dans la phrase suivante «La Vierge Marie tient le médaillon» où le sujet est la «Vierge Marie», le verbe est «tient» et le COD est le «médaillon». Cependant, notre corpus est issu d’un travail de sigillographie ce qui veut dire que d’un côté il présente des termes spécifiques et en langue étrangère tels que thôkos et labarum, et de l’autre une structure syntaxique assez complexe où les outils du TAL ne sont pas assez précis. Pour clarifier le point précédent, nous pouvons voir dans la figure 4.3 que le mot lôros (écharpe

en grec) est annoté de manière différente à chaque occurrence. Nous, nous sommes alors demandé si c'est un cas particulier car il s'agit d'un mot en latin cependant ce n'est pas le cas car d'autres mot plus communs tels que bouclier sont pas précisément annotés comme nous pouvons le constater dans la figure 4.4.

```
<NOUN dep='obl:arg'>lôros</NOUN>
<NOUN dep='obl:arg'>lôros</NOUN>
<NOUN dep='conj'>lôros</NOUN>
<NOUN dep='nmod'>lôros</NOUN>
<NOUN dep='obl:arg'>lôros</NOUN>
<NOUN dep='nmod'>lôros</NOUN>
```

FIGURE 3.6 – Liste des annotations obtenues pour le mot lôros à l'aide de Spacy.

```
<NOUN dep='conj'>bouclier</NOUN>
<NOUN dep='nmod'>bouclier</NOUN>
<NOUN dep='obj'>bouclier</NOUN>
```

FIGURE 3.7 – Liste des annotations obtenues pour le mot bouclier à l'aide de Spacy.

Afin d'obtenir des valeurs correctes, nous avons alors décidé de créer des dictionnaires dans des fichiers dont l'extension est .lst. Ces dictionnaires vont contenir tous les types d'objets regroupés en fonction de leur champs lexical. Par exemple, les objets dotés d'une assise ont été listés dans un dictionnaire alors que les objets de type arme vont être regroupés dans un autre. Tous ces dictionnaires vont être réunis dans un fichier avec l'extension .def de type gazeeter où les attributs de ces annotations seront de type «Objet» et ils correspondent à leur champs lexical. Par exemple, on obtiendra alors ce type de résultat :

```
<Objet majorType="Objet_d_assise" minorType="Souppédion">souppédion</Objet>
```

Ici, le souppédion est un objet de type objet_d_assise, de type souppédion. A partir de ce résultat nous allons donc créer la classe Objet, la sous classe objet_d_assise et la sous sous classe souppédion et la valeur de la balise Objet sera l'instance de la classe Objet.

3.2.7.2 Balisage du contenu textuel depuis le framework GATE

Une fois toutes les grammaires Jape créées et les dictionnaires renseignés, il faut passer à l'étape de la création d'une application sur Gate afin de l'utiliser sur le corpus créé à partir du fichier xml.

Nous obtenons alors toutes les annotations que nous avons établies au préalable ainsi que les attributs correspondants. Le résultat est présent dans l'image suivante et comme nous pouvons le constater, cette étape nous a permis de trouver précisément tous les personnages, les positions de ces personnages et les objets présents dans la représentation des sceaux.

SCEAU ECCLÉSIASTIQUE
milieu 11e s. (?) - 1ère moitié 12e s. (?)
Trésor de la Grande Eglise
Plomb
CdN 2004-0433a
Byzantine et copte Numismatique
Moyen Age
MAH Musée d'art et d'histoire, Ville de Genève. Don de la Fondation Migore, legs Janet Zakos, 2004
dans un **cercle de grènetis**, **Vierge assise**, de face, sur un **trône bas**, au piètement de perles, garni d'un **coussin**; pieds posés sur un **soppédion**, elle tient devant elle l'**Enfant**, **main droite** sur son épaule, la gauche le long du corps du **Christ** qui **bénit de la droite** et tient un **rouleau de la gauche**. De part et d'autre du **nimbe**, les sigles pour : légende métrique, accentuée, sur six lignes, précédée d'une **croisette** : Bénéficiaire de donations importantes, tant de la part de fidèles de condition modeste que de grandes familles aristocratiques, voire de l'empereur lui-même, l'Eglise avait pour charge principale de venir en aide aux indigents par des oeuvres de charité ponctuelles (distributions de vivres, accès aux bains, rachats de prisonniers de guerre), ou de long terme (entretien d'hospices, d'hôpitaux ou d'orphelinats). La légende « parlante » de ce sceau unique fait allusion à la tâche impartie à l'économe de la Grande Eglise et à son bureau, l'Oikovoμeiov, à savoir seconder le **patriarche** de Constantinople pour gérer au mieux la fortune de l'institution dans l'intérêt des ayants droit. **La Vierge** elle-même, représentée au droit du sceau, se **porte** garante du bon fonctionnement de l'Oikovoμeiov de la Grande Eglise, en sa qualité de patronne de Constantinople et de donataire de **Sainte-Sophie** par **Justinien Ier** (ainsi figure-t-elle sur les sceaux des **ekdikoi**). Les empereurs à ses **côtés**, outre la dévotion qu'ils lui témoignent, tiennent les chrysobulles de donation à la Grande Eglise. Spink, Byzantine Seals from the collection of Georges Zakos, **Part II**, Auction 132, Londres, 1999 (25.5), n° 208Zakos G., Veglery A., Byzantine lead seals, Allemagne, 1973.

Type	Set	Start	End	Id	Features
Personnage	TEST	35515	35526	140576	{Genre=Masculin, majorType=Figure_Politique, minorType=Empereur}
Personnage	TEST	35699	35710	140577	{Genre=Masculin, majorType=Figure_Politique, minorType=Empereur}
Personnage	TEST	45754	45765	140582	{Genre=Masculin, majorType=Figure_Politique, minorType=Empereur}
Personnage	TEST	47205	47214	140583	{Genre=Masculin, majorType=Figure_Politique, minorType=Empereur}
Personnage	TEST	37534	37542	140580	{Genre=Masculin, majorType=Figure_Politique, minorType=Empereur}
Personnage	TEST	45341	45349	140581	{Genre=Masculin, majorType=Figure_Politique, minorType=Empereur}
Personnage	TEST	47534	47543	140586	{Genre=Masculin, majorType=Figure_Politique, minorType=Empereur}
Personnage	TEST	47615	47624	140587	{Genre=Masculin, majorType=Figure_Politique, minorType=Empereur}

Original markups
TEST
☐ Avers
☒ Benir
☐ Bibliographie
☐ Collection
☐ Date
☐ Description
☒ DeuxMains
☐ Geneve
☐ Inventaire
☒ MainDroite
☒ MainGauche
☐ Materiaux
☒ Nimbes
☐ NonNimbes
☒ Objet
☐ Periode
☒ Personnage
☒ Position
☐ Revers
☐ Seal

FIGURE 3.8 – Fichier XML contenant les nouvelles annotations.

Ensuite, afin d'exploiter ces résultats, nous devons sauvegarder ces informations sous format XML afin de pouvoir créer de manière automatique les nouvelles classes, sous classes et les instances. Nous allons détailler ce point par la suite.

Pour sauvegarder ces annotations nous avons sauvegarder le fichier avec la fonction Inline XML présente dans le logiciel Gate. Avec cette fonctionnalité il faut choisir toutes les annotations que nous souhaitons conserver et dans notre cas nous souhaitons conserver les annotations ou bien les balises de base telles que le titre, la date, l'inventaire, l'avvers etc. (sauf les balises créées à l'étape du postagging et des relations syntaxiques) ainsi que les nouvelles annotations. Nous obtenons donc un fichier plutôt complet et potentiellement exploitable pour un peuplement automatique de l'ontologie.


```

<?xml version="1.0" encoding="UTF-8" >
<Geneve>
  <Seal>
    <Titre>SCEAU ECCLESIASTIQUE</Titre>
    <Date>milieu 11e s. (?) - 1ère moitié 12e s. (?)</Date>
    Trésor de la Grande Eglise
    <Matériaux>Plomb</Matériaux>
    <Inventaire>CdN 2004-0433a</Inventaire>
    <Collection>Byzantine et copte Numismatique</Collection>
    <Période>Moyen Age</Période>
    MAH Musée d'art et d'histoire, Ville de Genève. Don de la Fondation Migore, legs Janet Zakos, 2004
    <Avers> dans un <Objet majorType="Objet de decor" language="" minorType="CerclesGrenetis">cercle de grénétis</Objet> , <Personnage
    majorType="Personnage biblique" minorType="Vierge Marie" Genre="Feminin">Vierge</Personnage> <Position majorType="Assise">assise</Position>
    , <Position majorType="De face">de face</Position> , sur un <Objet majorType="Objet d assise" language="" minorType="Trône">trône bas
    </Objet> , au piétement de perles , garni d' un <Objet majorType="Objet d assise" language="" minorType="Souppédion">coussin</Objet> ;
    pieds posés sur un <Objet majorType="Objet d assise" language="" minorType="Souppédion">souppédion</Objet> , elle tient <Position majorType="
    Devant">devant elle</Position> l' <Personnage majorType="Personnage biblique" minorType="Enfant" Genre="Masculin">Enfant</Personnage> ,
    <MainDroite majorType="Main Droite">main droite</MainDroite> sur son épaule , la gauche le long du corps du <Personnage majorType="
    Personnage biblique" minorType="Christ" Genre="Masculin">Christ</Personnage> qui <Benir VERB="Benir">benit</Benir> <MainDroite majorType="
    Main Droite">de la droite</MainDroite> et tient un <Objet majorType="Symboles" language="" minorType="Rouleau">rouleau</Objet> <MainGauche
    majorType="Main Gauche">de la gauche</MainGauche> . De part et d' autre du <Nimbes Nimbe="nimbe">nimbe</Nimbes> , les sigles pour :
    </Avers><Revers> légende métrique , accentuée , sur six lignes , précédée d' une <Objet majorType="Symboles" language="" minorType="
    Croisette">croisette</Objet> :</Revers><Description>Bénéficiaire de donations importantes , tant de la part de fidèles de condition
    modeste que de grandes familles aristocratiques , voire de l' empereur lui-même , l' Eglise avait pour charge principale de venir en aide
    aux indigents par des oeuvres de charité ponctuelles ( distributions de vivres , accès aux bains , rachats de prisonniers de guerre ) , ou
    de long terme ( entretien d' hospices , d' hôpitaux ou d' orphelinats ) . La légende « parlante » de ce sceau unique fait allusion à la
    tâche impartie à l' économ de la Grande Eglise et à son bureau , l' Oikovouros , à savoir seconder le <Personnage majorType="
    Figure Politique" minorType="Patriarche" Genre="Masculin">patriarche</Personnage> de Constantinople pour gérer au mieux la fortune de l'
    institution dans l' intérêt des ayants droit . <Personnage majorType="Personnage biblique" minorType="Vierge Marie" Genre="Feminin">La
    Vierge</Personnage> elle-même , représentée au droit du sceau , se <Objet majorType="Objet de decor" language="" minorType="Porte">porte
    </Objet> garante du bon fonctionnement de l' Oikovouros de la Grande Eglise , en sa qualité de patronne de Constantinople et de donataire
    de Sainte-<Personnage majorType="Figure Politique" minorType="FemmeDeLEmpereur" Genre="Feminin">Sophie </Personnage>par <Personnage
    majorType="Figure Politique" minorType="Empereur" Genre="Masculin">Justinien Ier</Personnage> ( ainsi figure -t- elle sur les sceaux des
    ekkikoi ) . Les empereurs <Position majorType="A côté">à ses côtés</Position> , outre la dévotion qu' ils lui témoignent , tiennent les
    chrysobulles de donation à la Grande Eglise .</Description> <Bibliographie>Spink. Byzantine Seals from the collection of Georges Zakos,
    <Personnage majorType="Figure Politique" minorType="Empereur" Genre="Masculin">Part II</Personnage>, Auction 132. Londres, 1999 (25.5) . , n°
  
```

FIGURE 3.9 – Fichier XML contenant les nouvelles annotations.

3.2.8 Peuplement de l'ontologie grâce aux annotations effectuées

A l'aide des différentes annotations effectuées précédemment d'après différentes méthodes, les différents éléments annotés peuvent alors être exploités pour améliorer le peuplage de l'ontologie initiale.

3.2.8.1 Enrichissement des domaines, ranges, et autres caractéristiques des classes

Il s'agit dans cette étape de préciser pour chaque classe, Object Properties et instances les différentes caractéristiques qui les distinguent. Pour procéder à l'enrichissement, nous avons utilisé des codes pythons venant automatiser certaines de ces étapes, comme l'ajout dans le Domaines et le Range des Object Properties les différentes classes correspondantes. Par exemple, le fait que un Personnage peut être le parent ou l'enfant d'un autre personnage, ou encore cibler des classes pour des relations spécifiques comme le fait qu'un personnage n'est vêtu qu'avec un habit, en reliant donc la classe « Personnage » avec la relation est_vetu_avec et la sous-classe « Habillements » située dans la classe « Objet » ou le fait que seul le Christ et Saint Nicolas bénissent de la main droite dans l'échantillon de sceaux que nous traitons.

De même, ces relations définies par le prédicat sont caractérisables par leur type. En effet, il s'agit de préciser si la relation peut être par exemple réciproque ou non, ce sont les relations dites symétriques ou asymétriques. Lorsqu'une relation précise si une propriété est intrinsèque avec lui-même comme le fait qu'on est quelque chose, dans notre cas, il y a l'exemple de la nimbe, si un personnage est nimbé alors la relation est dite réflexive, sinon la relation est dite

irréflexive. Aussi, Une relation peut être fonctionnelle lorsqu'une propriété ne peut avoir qu'une seule valeur pour chaque individu.

3.2.8.2 Enrichissement des instances

A partir de l'ontologie créée initialement, il a été possible d'ajouter de nouveaux éléments avec une meilleure précision et une amélioration dans l'exploitation de ces données. Pour cela, nous avons pris en compte les différences entre les personnages et les objets ou les actions d'un sceau à l'autre, mais aussi les différences pour un même personnage ou objet d'un sceau à l'autre.

Les personnages et objets ont pu être ciblés et numérotés grâce à un code python pour permettre de distinguer un personnage d'un autre comme un objet d'un autre d'un sceau à l'autre, afin de ne pas confondre ces personnages et ces sceaux d'un sceau à l'autre. En effet, bien qu'un personnage soit le même individu, il faut toutefois prévoir le fait que d'un sceau à l'autre, il n'a pas forcément les mêmes caractéristiques, comme l'âge, la tenue, ou encore l'action effectuée diffère. Par exemple, le Christ peut apparaître comme étant l'Enfant dans tel sceau et apparaître comme étant un adulte dans un autre. Dans un sceau il peut bénir mais pas dans un autre, ou encore tenir les Evangiles dans l'un mais pas dans l'autre.

3.2.8.3 Enrichissement des prédicats

Plusieurs nouvelles relations ont été ajoutées et précisées pour apporter une meilleure distinction et un enrichissement dans la complexité des informations contenues sur les surfaces des sceaux byzantins. En effet, nous avons à partir de l'Object Properties « contient », ajouter avec du code python deux sous-prédicats « contient_personnage », « contient_objet » et « contient_objet_d_assise » pour distinguer les nuances que comporte cette relation. De cette manière, il est aussi possible de délimiter le cadre d'exploitation du prédicat et distinguer la nature d'un objet dans ce cas.

A partir de toutes ces propriétés, nous avons utilisé l'onglet SWLR Tab dans protégé pour rentrer des règles SWRL afin de créer des relations entre les instances. Une règle SWRL, ou Semantic Web Rule Language, est une extension du langage OWL (Web Ontology Language) utilisée pour représenter des règles dans le contexte du Web sémantique. Les règles SWRL permettent d'exprimer des connaissances sous forme de conditions et d'actions, ce qui facilite l'inférence et la déduction de nouvelles informations à partir des données existantes. Les règles SWRL combinent des concepts d'OWL (ontologies) avec des règles de logique.

Une règle SWRL est généralement composée de deux parties : Une partie *antécédent* qui spécifie les conditions qui doivent être remplies pour que la règle s'applique. Les conditions

sont généralement exprimées en utilisant des classes, des propriétés, des individus et des opérations logiques. Une partie *conséquent* qui spécifie les actions à entreprendre si les conditions spécifiées dans la partie antécédent sont satisfaites. Ces actions peuvent inclure l'ajout ou la modification de données dans une ontologie ou la création de nouvelles relations entre les données.

Voici un exemple des règles SWRL que nous avons utilisé dans notre ontologie :

```
OntoSeals :Sceau(?a) ^ OntoSeals :contient_objet_d_assise(?a, ?b)
^ OntoSeals :contient_personnage(?a, ?m) -> OntoSeals :est_assis_sur(?m, ?b)
```

Cette règle SWRL a été conçue pour créer une relation entre les personnages et les objets d'assise lorsque ces derniers sont représentés sur le sceau. Toute règle nécessite tout d'abord d'explicitement le contexte de notre ontologie qui s'appelle OntoSeals. Dans la partie antécédent, si un individu (?a) est marqué du sceau (OntoSeals :Sceau(?a)), si cet individu (?a) contient un objet d'assise (?b) (OntoSeals :contient_objet_d_assise(?a, ?b)) et si cet individu (?a) contient un personnage (?m) (OntoSeals :contient_personnage(?a, ?m)), alors dans la partie conséquent le personnage (?m) est assis sur l'objet d'assise (?b) (OntoSeals :est_assis_sur(?m, ?b)). Cette règle est illustrée dans la figure 3.1.

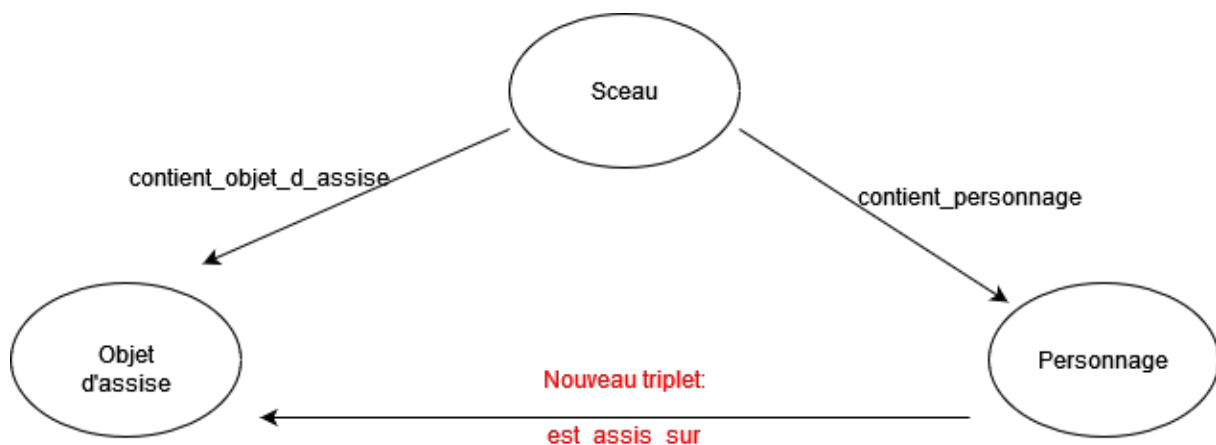


FIGURE 3.10 – Illustration de la règle SWRL.

Nous avons procédé de la même manière pour d'autres relations telles que *tient_des_deux_mains* ou *tient_en_main_gauche* lorsque le personnage bénit de la main droite.

La règle suivante nous a permis d'établir une relation entre les personnages et le fait de tenir un objet des mains. Généralement, les objets tenus dans les deux mains ont été annotés et rangés dans la classe Symboles.

```
OntoSeals :Personnage(?p) ^ OntoSeals :tient(?p, ?s) ^ OntoSeals :Symboles(?s) ->
OntoSeals :tient_des_deux_mains(?p, ?s)
```

Ici on peut voir que si le personnage tient un objet de type Symboles alors il le tient des deux mains. Enfin, pour déterminer toutes les nouvelles relations créées, nous avons choisi d'utiliser le reasoner Pellet pour qu'il puisse appliquer les règles SWRL et définir les relations entre instances. Par inférence nous avons alors obtenus les résultats comme illustrés dans les figures suivantes.

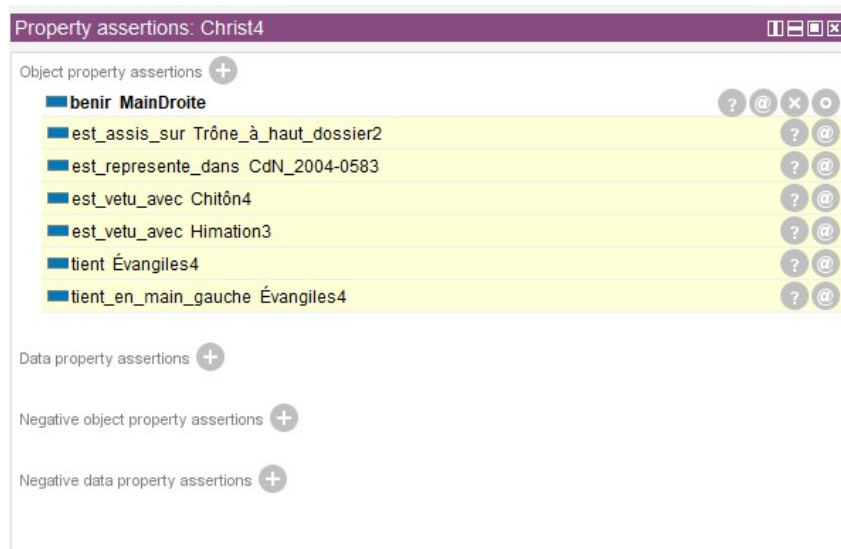


FIGURE 3.11 – Résultat de la règle SWRL avec les inférences obtenues à l'aide de Pellet.



FIGURE 3.12 – Résultat de la règle SWRL traitant la propriété tient_des_deux_mains avec les inférences obtenues à l'aide de Pellet.

Chapitre 4

Evaluation

4.1 Vérification de la consistance avec le moteur de raisonnement Pellet

Au fur et à mesure de l'établissement des prédicats et de leurs caractéristiques, nous avons utilisé le moteur de raisonnement (ou reasoner) Pellet dans l'environnement de développement Protégé. Cette plateforme open-source sert à la création et la gestion d'ontologies basées sur le langage RDF et le standard OWL (Web Ontology Language), et gère OWL_DL qui est une sous-classe plus élaborée dans ses caractéristiques de raisonnement. Il permet d'inférer des informations implicites à partir des informations explicites fournies dans l'ontologie. Ainsi, il est possible de vérifier la consistance de nouvelles relations sémantiques créées grâce à cet outil intégré dans l'environnement Protégé[\[Pel\]](#).

D'une part, le raisonnement s'effectue grâce à la déduction d'informations supplémentaires à partir des axiomes et des faits définis dans l'ontologie, comme le fait que si la Vierge Marie est le parent du Christ alors le Christ est son enfant. De cette manière, il peut aussi classer le Christ comme faisant partie des enfants.

D'autre part, Pellet peut donc détecter et signaler des incohérences dans les résultats obtenus par raisonnement suite à la création des prédicats. De cette manière, il peut aider à corriger des ontologies complexes et conséquentes comme celle créée et peuplée pour les sceaux byzantins et l'ensemble de la collection.

4.2 Exploitation des informations paramétrées par le requête en SPARQL

Pour tester l'exploitation des informations, nous avons établis huit requêtes SPARQL pour obtenir différents éléments de réponses possibles selon les informations qui ont pu être paramé-

trées. Ces questions sont les suivantes :

La première requête vise à lister les personnages bibliques uniquement selon l'avvers des sceaux donnés et la date de départ des sceaux associés. Elle commence par définir un préfixe "ns1" pour simplifier les requêtes ultérieures. Ensuite, elle sélectionne les variables ?Sceau et ?date_start. Le motif WHERE commence par chercher tous les sceaux de type "Sceau" selon l'avvers et leur date de début. Il utilise une clause OPTIONAL pour rechercher les personnages contenus dans chaque sceau et filtre les résultats en vérifiant que le personnage est de type "Personnage_biblique", qui est le nom de la classe correspondante dans l'ontologie utilisée. La clause FILTER vérifie que le personnage n'est pas lié au sceau, en utilisant la fonction BOUND. Ainsi, seuls les sceaux sans personnages bibliques sont renvoyés.

La deuxième requête cherche à obtenir la description et la date de fin des sceaux associés à des personnages ecclésiastiques. Elle utilise le même préfixe "ns1" et sélectionne les variables ?Sceau, ?description, et ?dateFin. Dans la clause WHERE, elle recherche les sceaux de type "Sceau" avec leur description, date de fin et les personnages contenus, en vérifiant que le personnage est de type "Figure_ecclesiastique", qui est la classe dédiée pour ce type de personnage. Cette requête renvoie les informations sur les sceaux associés à des personnages ecclésiastiques.

La troisième requête a pour but de lister les bibliographies associées à des sceaux de la collection. Elle utilise le préfixe "ns1" et sélectionne les variables ?sceau et ?bibliographie. La clause WHERE recherche tous les sceaux de type "Sceau" et, en option, leurs bibliographies associées. Cette requête renvoie les informations sur les bibliographies liées aux sceaux de la collection.

La quatrième requête vise à obtenir la mention obligatoire et l'avvers des sceaux de type "Objet de décor". Elle utilise le préfixe "ns1" et sélectionne les variables ?Sceau, ?mentionObligatoire, et ?nomAvers. Dans la clause WHERE, elle recherche les sceaux de type "Sceau" avec leur mention obligatoire, leur nom d'avvers, et les objets contenus. La clause OPTIONAL recherche les objets de type "Objet_de_decor" et filtre les résultats pour ne renvoyer que les sceaux sans objets de décor associés.

La cinquième requête a pour objectif de lister les sceaux avec leurs descriptions et dates de début, triés par date de début croissante, appartenant à la période "Moyen_age". Elle utilise le préfixe "ns1" et sélectionne les variables ?Sceau, ?Periode, ?Description, et ?date_start. La clause WHERE recherche tous les sceaux de type "Sceau" avec leur période, description, et date de début. La requête utilise ensuite la clause ORDER BY pour trier les résultats par date de début croissante, ce qui permet d'obtenir une liste ordonnée de sceaux de la période "Moyen_age".

La sixième requête vise à obtenir l'avvers, le revers et la date de création des sceaux associés à des armes. Elle utilise plusieurs préfixes, dont "ns1," "rdfs," et "owl." Elle sélectionne les

variables ?Sceau, ?date_start, ?nomAvers, et ?revers. Dans la clause WHERE, elle recherche les sceaux de type "Sceau" associés à des objets de type "Arme" ou à des sous-classes d'armes. La requête renvoie les informations sur les sceaux liés à des armes, y compris leur date de création, leur nom d'avvers, et leur revers.

La septième requête SPARQL vise à extraire la liste des personnages appartenant à la classe "Figure_ecclesiastique" et à afficher leur nimbe s'ils en possèdent un. Pour ce faire, nous utilisons un préfixe pour simplifier l'écriture des URI, puis nous utilisons la clause SELECT pour spécifier les variables ?personnage et ?nimbe que nous souhaitons récupérer. La clause WHERE définit les conditions de la recherche : nous recherchons des triples où ?personnage est de type "Figure_ecclesiastique". La clause OPTIONAL nous permet de récupérer le nimbe associé à chaque personnage s'il existe, tout en incluant les personnages qui n'en ont pas. Cela permet d'obtenir une liste de personnages et de leurs nimbres s'ils sont présents.

La huitième requête SPARQL a pour objectif de lister les sceaux qui sont fabriqués en or ou en plomb, en fournissant leur titre, leur matériau et leur période. De manière similaire à la première requête, nous utilisons un préfixe pour simplifier les URI, puis nous spécifions les variables que nous souhaitons récupérer dans la clause SELECT. Dans la clause WHERE, nous recherchons des triples où ?sceau est de type "titre", possède un matériau (a_materiel), et est associé à une période (a_pperiode). De plus, nous utilisons un filtre (FILTER) pour limiter les résultats aux sceaux en or ou en plomb en vérifiant le matériau de chaque sceau. Cela nous permet d'obtenir une liste de sceaux, chacun accompagné de son titre, de son matériau et de sa période s'il satisfait les critères de matériau en or ou en plomb.

Chapitre 5

Comparaison des résultats obtenus et des limites rencontrées

La création d'une ontologie pour les sceaux byzantins avec le logiciel protégé s'est faite de manière simple et claire où chaque classe principale a pu être créée, suivant de leurs relations avec les instances correspondantes. L'enrichissement du contenu des informations à paramétrer dépend donc des techniques et méthodes rigoureuses à utiliser autour de ce logiciel.

Néanmoins, à fur et à mesure des différentes utilisations des méthodes employées pour le peuplement, l'enrichissement et l'amélioration des informations sur les sceaux byzantins, plusieurs limites et problèmes ont été rencontrés. Ces obstacles de natures différentes concernent tant le paramétrage que le balisage multiple des informations contenues aussi bien dans le fichier XML d'origine que dans les nouvelles versions qui ont suivi.

L'ensemble des étapes a pu permettre la représentation ci-dessous des différents triplets paramétrés dans le graphe à l'aide de ces différentes méthodes employées, en utilisant le balisage présent dans le fichier XML et le codage python pour créer et peupler l'ontologie pour les sceaux byzantins. De cette manière, les éléments principaux qui constituent un sceau byzantin, au travers de tous les éléments qui le caractérisent, ont pu être représentés comme il se doit. La difficulté réside dans la suite des étapes en ce qui concerne tant le choix des méthodes rigoureuses à utiliser selon la tâche à traiter que dans la rédaction des codes et choix de pattern pour créer ces méthodes afin d'automatiser le processus pour des utilisations ultérieures. Ci-dessous, une représentation graphique de l'ontologie peuplée et paramétrée avec les différentes informations ; l'image a été visualisée et téléchargée depuis le site WebVOWL[\[web\]](#).

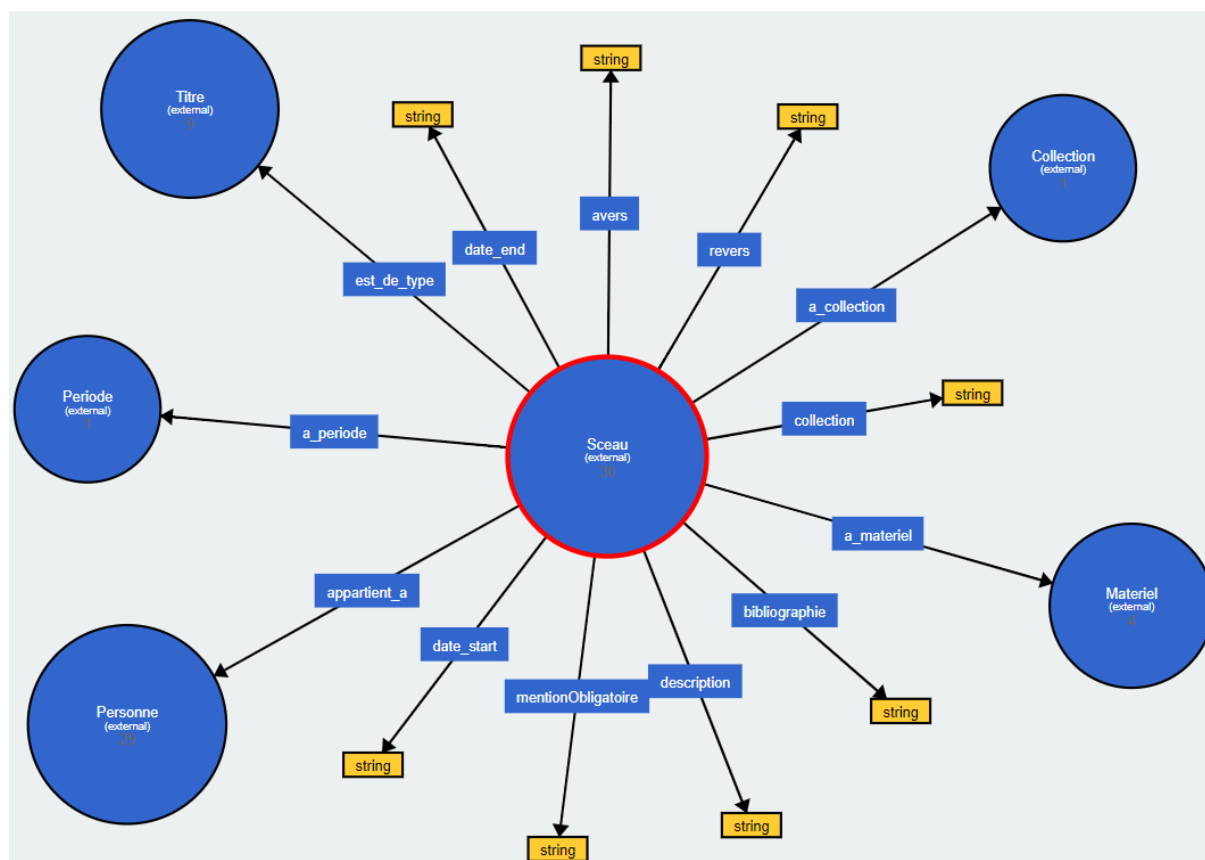


FIGURE 5.1 – Représentation du graphe pour les sceaux byzantins de la collection Zacos

5.1 Résultats et limites des outils de spaCy

Une fois la sélection et l'application des méthodes sélectionnées et présentées précédemment, les résultats obtenus étaient majoritairement convainquants mais des imperfections se sont tout de même immiscées dans le balisage.

Les différents outils apportés par la bibliothèque de spaCy ont permis de procéder à un balisage efficace et sur un ensemble de contenu textuel assez conséquent, ce qui a permis d'emblée d'obtenir un enrichissement du contenu textuel du fichier XML d'origine et très rapidement. L'utilisation de ces outils à l'aide de codes python permet d'être plus rapide dans le processus de balisage qu'avec les règles japes qui demandent davantage de connaissance dans la rédaction des règles Jape.

Il demeure que ces outils ne sont pas parfaits et nécessitent des corrections dans les résultats obtenus. Il y a eu des problèmes techniques comme les soucis de balisage d'un fichier XML

à l'autre où des balises ont été perdues comme la balise '<Greek>' pour les éléments rédigés en grec. Néanmoins, la plupart des outils ne sont pas parfaitement efficaces même dans une tâche simple sans ambiguïté dans le sens d'un mot ou token donné. Pour palier à ce problème de perte des balises '<Greek>', nous avons élaboré un code python qui récupère les balises contenues dans le fichier XML original, ainsi à partir de là, nous avons enrichi le balisage en distinguant ces balises selon la balise où celles-ci apparaissaient. Ainsi, les balises '<Greek>' contenues dans la balise '<Avers>' étaient par exemple changées en '<AversGreek>', de même pour le revers et la partie description. Ces nouvelles balises et sont contenu ont été intégré dans une data Property nommé de la même manière.

En effet, bien que l'enrichissement du balisage à l'aide de la reconnaissance d'entités nommées avec spaCy apporte une meilleure compréhension des éléments clés présents sur les sceaux byzantins, cette subtilité nécessite une attention particulière lors de l'utilisation de la reconnaissance d'entités nommées pour s'assurer de l'exactitude de la catégorisation, comme avec l'exemple de la confusion entre les noms des saints et les lieux géographiques. Cela peut être expliqué par le fait que de nos jours les lieux géographiques sont plus utilisés pour désigner des directions que les noms des saints pour parler de figures religieuses.

Aussi, d'autres problèmes avec le postagging sont survenus pour une confusion dans la nature d'un token donné, comme dans l'exemple ci-après où le verbe « bénissant » a été balisé comme étant un <Personnage>.

```
<PUNCT>,</PUNCT> <NOUN dep='fixed'>debout</NOUN> <ADP>sur</ADP> <DET>ses</DET> <NOUN d
='dep'>genoux</NOUN> <PUNCT>,</PUNCT> <Personnage>l'</Personnage> <Personnage>Enfant
</Personnage> <Personnage>bénissant</Personnage> <PUNCT>.</PUNCT> <ADP>De</ADP> <NOUN
dep='ROOT'>part</NOUN> <CCONJ>et</CCONJ> <NUM>d'</NUM> <ADJ>autre</ADJ> <ADP>de</ADP>
```

FIGURE 5.2 – Exemple d'erreur obtenu avec le postagging pour le sceau byzantin CdN 2004-0551

De même, en ce qui concerne la lemmatisation, un lemme a été inventé pour le nom commun « épais » qui a été lemmatisé en « éper » dans l'exemple 2 ci-dessous.

legs Janet Zakos, 2004</Mention_obligatoire
> <VERB lemma='éper'>épais</VERB> <ADJ>grè
-corps</NOUN> <PUNCT>,</PUNCT> <ADP>de</ADP>
<VERB lemma='accoutumer'>accoutumée</VERB>

FIGURE 5.3 – Exemple 1 d’erreur de lemmatisation d’un nom commun en verbe pour le sceau byzantin CdN 2004-0313

Tandis le nom commun « barbe » a été confondu avec un verbe comme dans l’exemple 1 ci-dessous.

<P> <ADV>plus</ADV> <ADJ>grande</ADJ> <NOUN dep='ob.
<VERB lemma='barbe'>barbe</VERB> <ADJ>courte</ADJ>
<NOUN>deux</NOUN> <ADJ>cent</ADJ> <VERB lemma='lâcher'>

FIGURE 5.4 – Exemple 2 d’erreur de lemmatisation d’un nom commun en verbe pour le sceau byzantin CdN 2004-0373

5.2 Résultats de l’utilisation des règles Jape et du framework GATE

Les règles Jape ont pu permettre le balisage concret de l’ensemble des éléments intéressants à paramétrer tant que les règles étaient efficacement rédigées. L’outil GATE permet une meilleure visualisation du balisage et de la catégorisation de ces balises grâce à ces règles japes contrairement aux différents outils de spaCy qui ne permettent pas un balisage aussi personnalisé et personnalisable par l’ajout d’attribut spécifique comme l’ajout du genre en français (féminin, masculin), ou encore une catégorie qui relie et rassemble plusieurs appellations d’un même personnage.

Toutefois, ces règles doivent être écrites selon la manière dont les données sont présentées et répondre à la complexité de leur orthographe comme la présence de deux langues et deux alphabets dans le contenu textuel, à savoir la langue française et son alphabet latin et, de l’autre, la langue grecque ancienne et son alphabet grec ancien, ou encore la présence de chiffres romains en ce qui concerne notamment les noms des empereurs.

En ce qui concerne l'enrichissement des balises '<Greek>', la regex n'a pas permis de capturer correctement tous les tokens ou groupes de tokens écrits en grec lorsqu'il y avait des caractères particuliers comme des parenthèses dans un token donné. Ainsi, l'enrichissement n'a pas pu être effectué comme cela était souhaité et cette règle jape doit donc être corrigée.

C'est pourquoi, il a été nécessaire de combiner ces différents outils avec des méthodes spécifiques à chacun pour obtenir de meilleurs résultats tout en visant l'automatisation de ces techniques pour de futures utilisations sur d'autres collections de sceaux byzantins par exemple, ou pour un enrichissement des bases de données sur l'ensemble de plusieurs collections de sceaux byzantins et autres objets.

5.3 Difficultés dans l'élaboration du code pour convertir les dates

En ce qui concerne le code Python générant un fichier JSON permettant la conversion des dates rédigées en dates numériques, il a fallu faire face à certaines difficultés dans le code avec l'utilisation des listes créées manuellement pour compenser certains mauvais calibrage dans l'élaboration du code Python qui ne convertissait pas par exemple le 'DateEnd' correctement du fait qu'il ne prenait pas la dernière date pour signifier que la fin de la période était la dernière date du siècle, à savoir si le siècle est par exemple le XIe siècle alors la dernière date devra être l'an 1199 ou 1200 et non pas l'an 1100. Aussi, certains siècles n'ont pas été capturés comme il l'aurait fallu et n'ont donc pas été convertis.

Par ailleurs, pour l'idée de mettre une espace dans le 'DateStart' lorsque la date prend la forme "avant XXXX" vient engendrer un problème dans l'environnement Protégé qui ne peut prendre des valeurs nulles 'None' dans ces 'data Properties'.

Ainsi, ce code comporte encore des erreurs de calibrage, mais qui demeurent exploitables et peuvent être corrigés et améliorés.

Conclusion

La création et la gestion d'une ontologie pour les sceaux byzantins nécessitent une combinaison de compétences techniques, de connaissances historiques et d'expertise en traitement du langage naturel.

Tout au long de la création et du peuplement de l'ontologie des sceaux byzantins de la collection Zacos, il a été question de répondre à la complexité de sa conception et des moyens d'enrichissement du contenu et des connexions entre les différentes informations essentielles que celle-ci contient. La préparation des données en amont demeure une étape essentielle et laborieuse pour permettre le paramétrage rigoureux des données ensuite, en traitant la lecture des informations sélectionnées à l'aide de méthodes adaptées pouvant aller jusqu'à la conversion de celles-ci afin d'obtenir de ces informations des données exploitables et réutilisables tout en assurant leur qualité et leur cohérence. Le nettoyage de ces données favorise leur intégration fluide dans l'ontologie, tout en maintenant leur intégrité sémantique. C'est pourquoi l'automatisation de l'extraction et le paramétrage des données pour de futures utilisations sont des éléments essentiels à la mise en forme de ces données collectées avec soin grâce à des outils comme les règles JAPE, GATE et Python pour établir des méthodes rigoureuses de traitement automatisé et améliorant la cohérence des données collectées.

Par ailleurs, dans le contexte des sceaux byzantins, il est possible d'enrichir les informations disponibles en exploitant des ressources externes telles que les sites Web, les bases de connaissances telles que BBpedia et des encyclopédies en ligne comme Wikipedia qui contiennent des informations complémentaires sur les sceaux, les personnages, les objets, ou encore les symboles, ainsi que sur leurs utilisations historiques et sociales. De même, lorsque les informations complémentaires sont disponibles depuis des documents manuscrits pouvant être anciens ou numérisés, il est possible de recourir à la technique d'ocrisation pour extraire davantage d'informations dans ses ressources textuelles comme les documents PDF, Word de documents académiques ou de rapports historiques ou encore des images ou photographies. Cette technique permettra donc d'étendre davantage les connaissances sur les sceaux byzantins, leurs contextes historiques et culturels, ainsi que sur les personnages et les symboles associés, et, à partir de ces

supports exploitables, procéder à leur numérisation pour une facilité d’exploitation des données collectées. Cela permet une recherche plus avancée et une analyse approfondie des données, tout en élargissant la portée des connaissances capturées par l’ontologie. En effet, une autre ressource exploitable pour ce mémoire est l’échantillon de sceaux publiés sur le site du musée d’art et d’histoire de Genève[Bib23]. Ces sceaux ont été étudiés par Jean-Claude Cheynet qui a fourni le même travail que pour les sceaux de la collection de Zacos. Pour ce corpus, il est possible d’utiliser la méthode du web scraping pour extraire toutes les informations sur ces sceaux qui étaient disponibles sur le site, et ainsi obtenir un fichier XML et json afin de pouvoir les réutiliser pour le reste de notre processus au même titre que les données du recueil de sceaux de Zacos.

De même, étant donné toutes les incohérences avec le texte originel et étant dans l’incapacité de corriger manuellement plus de 480 pages du livre de J.-C. Cheynet, il est possible alors d’utiliser un autre outil d’océrisation plus performant. Nous avons alors fait recours à l’outil d’océrisation hébergé par le site avepdf.com [Doc] qui permet de passer d’un fichier au format pdf à un fichier au format txt.

De cette manière, l’enrichissement des informations paramétrées dans notre ontologie peut être complétées davantage grâce à ses supports disponibles comme sur le site du musée de Genève concernant l’ensemble de la collection Zacos, mais aussi pour d’autres collections de sceaux byzantins qui pourront être reliées entre elles. De même, en ce qui concerne l’amélioration du balisage, il est possible de procéder à des systèmes d’apprentissage supervisé ou non supervisé afin d’améliorer les petits défauts rencontrés avec les outils de spaCy qui ont été utilisés.

La collecte d’informations sur des éléments historiques favorise la préservation de ces témoignages culturels transmissibles au public, tout en facilitant l’exploration et la manipulation dans les domaines de la recherche, de l’éducation et de la préservation du patrimoine culturel et historique. Cette tâche contribue ainsi à la compréhension et à l’appréciation de l’histoire et de la culture byzantines, tout en ouvrant la voie à de nouvelles explorations et découvertes dans ce domaine plutôt méconnu du grand public.

L'ensemble du travail peut être résumé dans la figure 4.5.

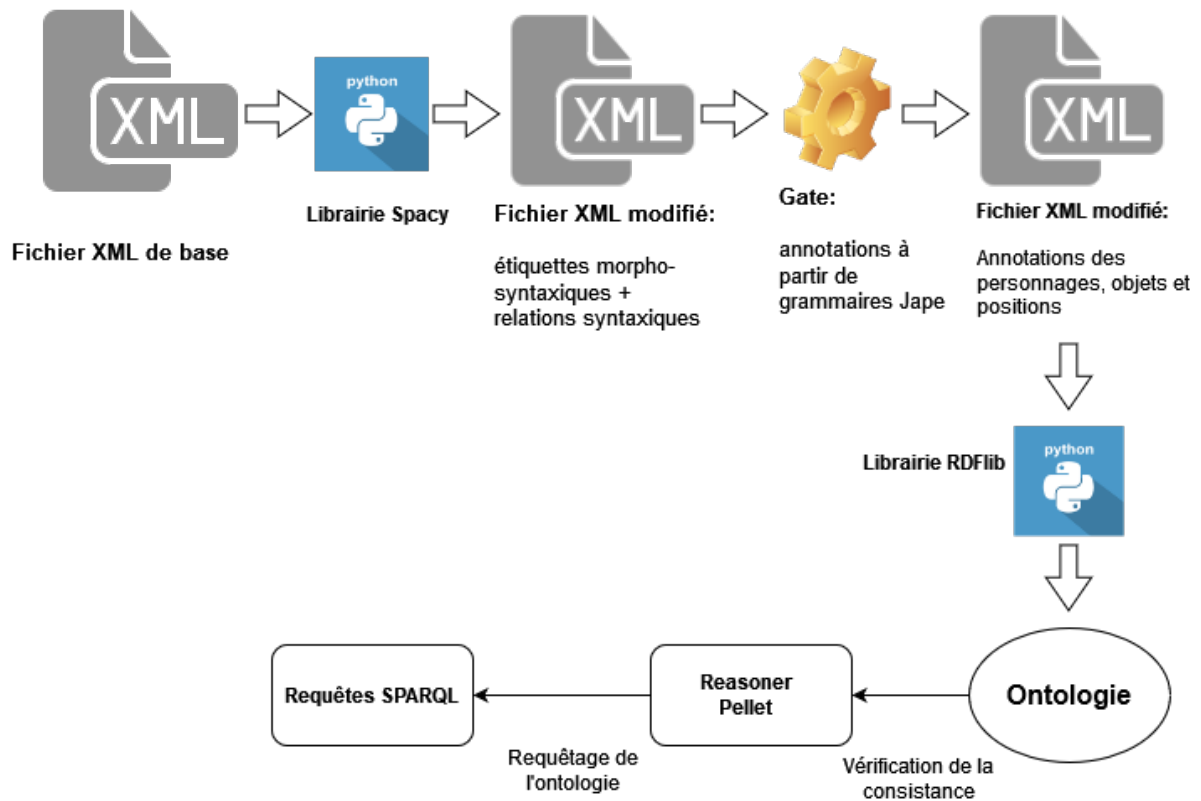


FIGURE 5.5 – Résumé des différentes étapes du mémoire.

Annexe A

Annexe

```
<?xml version="1.0" encoding="utf-8"?>
<Geneve>
  <Seal>
    <Titre>SCEAU ECCLÉSIASTIQUE</Titre>
    <Date>milieu 11e s. (?) / 1ère moitié 12e s. (?)</Date>
    <Person>Trésor de la Grande Eglise</Person>
    <Materiaux>Plomb</Materiaux>
    <Inventaire>CGN 2004-0433a</Inventaire>
    <Collection>Byzantine et copte Numismatique</Collection>
    <Periode>Moyen Age</Periode>
    <Mention_obligatoire>MAH Musée d'art et d'histoire, Ville de Genève. Don de la Fondation Migore, Legs Janet Zakos, 2004</Mention_obligatoire>
    <Avers> dans un cercle de grénétis, Vierge assise, de face, sur un trône bas, au piétement de perles, garni d'un coussin ; pieds posés sur un souppédion, elle tient devant elle l'Enfant, main dr
    </Avers>
    <Revers> légende métrique, accentuée, sur six lignes, précédée d'une croix : <Greek><![CDATA[+ Θ(ρο)ῶ τοιαύτῃ σκηνῇ ὡς ἔστιν ἡ τοῦ ἁγίου Ἰωάννου ἐκ τῆς ἀρχαίας </Greek>+ C'est moi qui
    <Description>Bénéficiaire de donations importantes, tant de la part de fidèles de condition modeste que de grandes familles aristocratiques, voire de l'empereur lui-même, l'Eglise avait pour ch
    <Bibliographie>Spink. Byzantine Seals from the collection of Georges Zakos, Part II, Auction 132. Londres, 1999 (25.5) ., n° 208 Zakos G., Vegliery A., Byzantine lead seals, Allemagne, 1972, planch
  </Seal>
  <Seal>
    <Titre>SCEAU ECCLÉSIASTIQUE</Titre>
    <Date>13e -14e siècles</Date>
    <Person>Ekdikoi de l'église de Sainte-Sophie de Constantinople</Person>
    <Materiaux>Plomb</Materiaux>
    <Inventaire>CGN 2004-0439</Inventaire>
    <Collection>Byzantine et copte Numismatique</Collection>
    <Periode>Moyen Age</Periode>
    <Mention_obligatoire>MAH Musée d'art et d'histoire, Ville de Genève. Don de la Fondation Migore, Legs Janet Zakos, 2004</Mention_obligatoire>
    <Avers> La Vierge et l'empereur Justinien Ier, non nimbé, tiennent un modèle de l'église de Sainte-Sophie. Le dôme de l'église repose sur trois arcatures, portées par quatre colonnes, fermées à
    <Revers> légende sur cinq lignes, surmontée d'une croix et suivie d'une perle accolée de fleurons : <Greek><![CDATA[+ τοῖς ἑκατόντα τοῖς ἀποστόλοις καὶ ἑκατὸν </Greek>+ (Aux
    <Description>Outre le fondateur de Sainte-Sophie, Justinien Ier est l'instituteur du tribunal de la Grande Eglise, dont la mission était de contribuer à mieux défendre les criminels face à l'in
    <Bibliographie>Zakos, G., John Nesbitt. Byzantine lead seals. Berne, 1984., n° 73a Marielle Martiniani-Beber (dir.), Byzance en Suisse, cat. exp. Genève, Musée Rath, 4 décembre 2015 - 13 mars 20
  </Seal>
</Geneve>
```

FIGURE A.1 – Exemple de visualisation du graphe enrichi dans l’environnement Protégé

```
def createObjectRelation(name, domainType, rangeType):
    rel = URIRef(OntoSeals[name])
    g.add((rel, RDF.type, OWL.ObjectProperty))
    if domainType is not None:
        dom = OntoSeals[domainType]
        g.add((rel, RDFS.domain, dom))
    if rangeType is not None:
        ran = OntoSeals[rangeType]
        g.add((rel, RDFS.range, ran))
    return rel

# Créer la classe "Main"
g.add((OntoSeals["Main"], RDF.type, OWL.Class))

# Créer la sous-classe "MainDroite" avec une relation 'est sous-classe de'
g.add((OntoSeals["MainDroite"], RDF.type, OWL.Class))
g.add((OntoSeals["MainDroite"], RDFS.subClassOf, OntoSeals["Main"]))

# Créer l'instance "de_la_main_droite" qui est une sous-classe de "MainDroite"
g.add((OntoSeals["de_la_main_droite"], RDF.type, OntoSeals["MainDroite"]))

# Créer la sous-sous-classe "Nimbe" dans la sous-classe "Symboles"
g.add((OntoSeals["Nimbe"], RDF.type, OWL.Class))
g.add((OntoSeals["Nimbe"], RDFS.subClassOf, OntoSeals["Symboles"]))

# Créer l'instance "nimbé" à mettre dans la sous-sous-classe "Nimbe"
g.add((OntoSeals["nimbé"], RDF.type, OntoSeals["Nimbe"]))

# Définir une propriété d'objet personnalisée pour la relation 'contient'
contient = OntoSeals["contient"]
g.add((contient, RDF.type, OWL.ObjectProperty))

# Créer les sous-Object Properties "contient_personnage" et "contient_objet"
contient_personnage = OntoSeals["contient_personnage"]
contient_objet = OntoSeals["contient_objet"]
g.add((contient_personnage, RDF.type, OWL.ObjectProperty))
g.add((contient_objet, RDF.type, OWL.ObjectProperty))

# Ce sont des sous-Object Properties de "contient"
g.add((contient_personnage, RDFS.subPropertyOf, contient))
g.add((contient_objet, RDFS.subPropertyOf, contient))

contient_obj = createObjectRelation("contient_objet", "Sceau", "Objet")
contient_pers = createObjectRelation("contient_personnage", "Sceau", "Personnage")
```

FIGURE A.2 – Exemple de création de classes, instances et relations dans l'ontologie

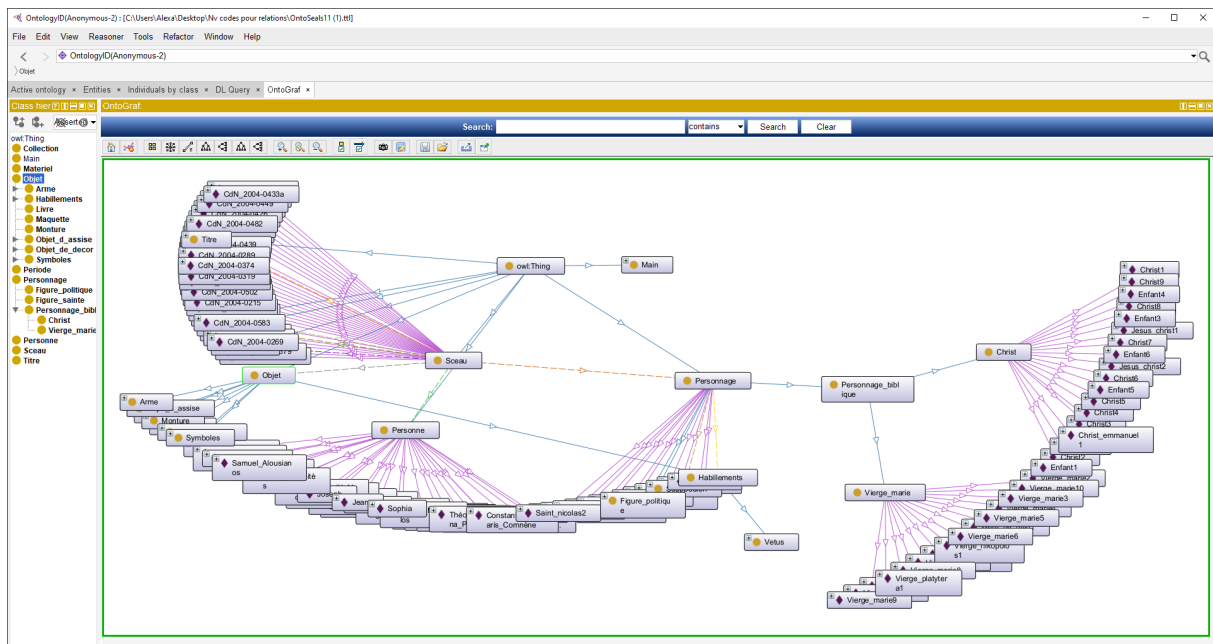


FIGURE A.3 – Exemple de visualisation du graphe enrichi dans l'environnement Protégé

The screenshot shows the SPARQL Playground interface. On the left, the 'RDF input' is set to 'Turtle' and contains several triples. Below it, the 'Output' is set to 'JSON' and shows the resulting JSON-LD data. On the right, the 'SPARQL query' is entered, and the 'Results' section displays a table of the query results.

SPARQL query

```
PREFIX ns1: <http://www.co-ode.org/ontologies/ont.owl#>

SELECT ?Sceau ?date_start
WHERE {
  ?Sceau rdf:type ns1:Sceau .
  ?Sceau ns1:avers ?nomAvers .
  ?Sceau ns1:date_start ?date_start .
  OPTIONAL {
    ?Sceau ns1:contient_personnage ?Personnage .
    ?Personnage rdf:type ns1:Personnage_biblique .
  }
  FILTER ( !BOUND(?Personnage) )
}
```

Results

Sceau	date_start
<http://www.co-ode.org/ontologies/ont.owl#CdH_2004-0215>	"1250"^^<http://www.w3.org/2001/XMLSchema#int>
<http://www.co-ode.org/ontologies/ont.owl#CdH_2004-0322>	"1250"^^<http://www.w3.org/2001/XMLSchema#int>
<http://www.co-ode.org/ontologies/ont.owl#CdH_2004-0526>	"1250"^^<http://www.w3.org/2001/XMLSchema#int>
<http://www.co-ode.org/ontologies/ont.owl#CdH_2004-0248>	"1080"^^<http://www.w3.org/2001/XMLSchema#int>
<http://www.co-ode.org/ontologies/ont.owl#CdH_2004-0289>	"1055"^^<http://www.w3.org/2001/XMLSchema#int>
<http://www.co-ode.org/ontologies/ont.owl#CdH_2004-	"950"^^<http://www.w3.org/2001/XMLSchema#int>

FIGURE A.4 – Exemple de résultat obtenu par requêtage SPARQL depuis le site SPARQL - Playground

Le site SPARQL Playground est un outil en ligne exploitable par les utilisateurs pour tester des bases de données RDF (Resource Description Framework) en utilisant le langage de requête SPARQL (SPARQL Protocol and RDF Query Language)[[SPA](#)].

Bibliographie

- [Agn18] Konys Agnieszka. *Knowledge systematization for ontology learning methods*, volume 126. Elsevier B.V., 2018. (Cité en page 21)
- [Bib23] Sceaux byzantins | Musées d’art et d’histoire de Genève, April 2023. [Online ; accessed 10. Apr. 2023]. (Cité en page 55)
- [byza] Site d’archives des Sceaux byzantins de la Collection Zacos ; dernière consultation le 7 septembre 2023. (Cité en pages 23 et 24)
- [byzb] Site d’archives d’exemples de Sceaux byzantins ; dernière consultation le 7 septembre 2023. (Cité en page 23)
- [CLBJA⁺17] Lopez Cédric, Marie-Laure Le Brazidec, Chevillon Jean-Albert, Francis Couturas, Dominique Hollard, and Aurélien Pierre. Ontocoins : données ouvertes liées pour la numismatique, patrimoine culturel. *Actes IC 2017*, 06 2017. (Cité en page 20)
- [dedCPC⁺16] Geneva (Switzerland). Musée d’art et d’histoire, Maria Campagnolo-Pothitou, Jean-Claude Cheynet, Bettina Jacot-Descombes, and Flora Bevilacqua. *Sceaux de la collection George Zacos au Musée d’art et d’histoire de Genève*. Musée d’art et d’histoire de Genève, 2016. (Cité en pages 14 et 24)
- [Def01] International council of museums icom, 2001. (Cité en page 14)
- [Def06] International council of museums icom, 2006. (Cité en page 14)
- [Doc] Free Online PDF and Document Tools. (Cité en page 55)
- [Doe03] Martin Doerr. The CIDOC Conceptual Reference Module : An Ontological Approach to Semantic Interoperability of Metadata. *AIMag.*, 24(3) :75, September 2003. (Cité en page 19)
- [ELM19] Victoria Eyharabide, Vincent Lully, and Florentin Morel. MusicKG : Representations of Sound and Music in the Middle Ages as Linked Open Data. In *Semantic Systems. The Power of AI and Knowledge Graphs*, pages 57–63. Springer, Cham, Switzerland, November 2019. (Cité en page 20)
- [Iva01] Jordanov Ivan. Byzantine lead seals from the stronghold near dobri dol, plovdiv region. *Revue Numismatique*, (157) :pp. 443–469, 2001. (Cité en page 11)

- [LIJ⁺14] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick Van Kleef, Sören Auer, and Christian Bizer. Dbpedia - a large-scale, multilingual knowledge base extracted from wikipedia. *Semantic Web Journal*, 6, 01 2014. (Cité en page 27)
- [Pel] Introduction à Protégé ; Site de l'Université TÉLUQ ; dernière consultation le 14 septembre 2023. (Cité en page 46)
- [R.95] Gruber Thomas R. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, 43(5-6) :pp. 907–928, 1993-1995. (Cité en pages 7 et 17)
- [SFS20] Alessio Sopracasa, Martina Filosa, and Simona Stoyanova. The digital enhancement of a discipline. byzantine sigillography and digital humanities. *magazén*, 12 2020. (Cité en page 18)
- [SPA] Site SPARQL-Playground ; dernière consultation le 14 septembre 2023. (Cité en page 59)
- [Teo21] Vesevska Irena Teodora. A rare byzantine lead seal from medieval bučín. 74 :pp.183–194, 2021. (Cité en page 13)
- [UG96] Mike Uschold and Michael Gruninger. Ontologies : principles, methods and applications. *The Knowledge Engineering Review*, 11(2) :93–136, 1996. (Cité en page 18)
- [VS77] Jr. Vryonis Speros. *Byzantine Studies : Essays on the Slavic World and Early Medieval Europe*. New Rochelle, N.Y., ©1992, 1977. (Cité en page 10)
- [web] Site de visualisation graphique ; dernière consultation le 7 septembre 2023. (Cité en page 49)