

MÉMOIRE DE RECHERCHE  
RAPPORT TECHNIQUE  
*POUR UN NOUVEAU ROMAN, ÉDITION NUMÉRIQUE*  
11 juin 2023

Mme Florence DE CHALLONGE  
M. Matthieu MARCHAL

MÉVEL Adrien  
Master 2 Lettres Modernes,  
« Éditions numériques et imprimées de textes littéraires »

Année universitaire 2022-2023

# 1 Introduction au projet d'édition numérique

Le présent document se veut un exposé des modalités et difficultés techniques que nous avons pu rencontrer et des moyens mis en œuvre pour les surmonter. Chaque outil utilisé fera l'objet d'une présentation succincte, après quoi leur apport au projet et les choix ayant mené à la réalisation de l'édition numérique de *Pour un nouveau roman* seront détaillés.

## 2 Conception et réalisation d'une base de donnée

### 2.1 Principes généraux

On appelle base de donnée un mode de structuration de l'information qui permet de stocker un grand nombre d'informations sur un petit espace disque. Il existe plusieurs modèles de structuration de ces bases de données, mais pour le présent travail n'est employé que le modèle le plus courant : le modèle relationnel.

Dans une base de donnée relationnelle on ne stocke pas seulement des informations brutes telles « Robbe-Grillet, Une voie pour le roman futur » mais bien des informations mises en relations les unes avec les autres, chacune ayant une nature définie au sein de la base, on aurait donc plutôt : « L'auteur Robbe-Grillet a écrit l'article nommé "Une voie pour le roman futur". ».

Nous nous proposons d'illustrer via une base de données les liens que tisse chacun des textes constituant l'ensemble avec les publications antérieures et les référents (textuels ou autres) qui sous-tendent l'argumentation tout en rendant compte des différents thèmes abordés afin de donner une vue d'ensemble du recueil perçu comme un tissu de textes au sein d'un environnement dont il donne, de manière implicite et/ou explicite, une représentation.

### 2.2 Mise en œuvre

Dès décembre 2022, nous avons produit une première version de cette base de donnée relationnelle dans le cadre de l'évaluation du cours de base données animé par Mme Delphine TRIBOUT. Le modèle soumis alors à évaluation nécessitait quatre entités : la base incluait une entité « SUBJECTS », un recensement des domaines abstraits dont traitaient les ARTICLES. Afin d'être le plus pertinent possible nous nous proposons de produire des valeurs les plus précises possibles pour l'attribut sDomain (« théorie Littéraire XX<sup>e</sup> » plutôt que « littérature »).

Cette entité a depuis été retirée du modèle car elle nous semblait avoir peu d'intérêt tant le choix des domaines à affubler à tel ou tel article était d'une part redondant, d'autre part le fruit d'une appréciation personnelle parfois difficile à objectiver. Si tel outil est toujours le produit d'une recherche particulière et, dès lors, le résultat d'une lecture donnée, le découpage des domaines traités par les articles du recueil nous paraissait au mieux d'un intérêt limité (« À quoi servent les théories » traite de théorie littéraire du XX<sup>e</sup> et de philosophie), au pire, difficilement défendable. Par exemple, nous avons réuni l'ensemble des filiations du nouveau roman tels « Faulkner » ou « Kafka » généralement mentionnés ensemble au sein du domaine « histoire littéraire internationale synchronique » plutôt que de les séparer dans des catégories par siècle et/ou pays car Alain Robbe-Grillet ne fait pas une histoire de la littérature anglaise ou tchèque mais inscrit ses références dans une histoire littéraire internationale ; on aurait pu également considérer que ces références puisqu'elles s'inscrivent dans une volonté de décrire une filiation au nouveau roman devraient être rattachées au domaine « théorie littéraire XX<sup>e</sup> ». De manière générale, il nous semblait que l'attribution de domaine aux références toujours effectuées en fonction de notre lecture du recueil induisait trop de choix problématiques pour être pleinement satisfaisante.

## 2.2.1 Modèle conceptuel

La première étape de constitution d'une base de donnée est l'élaboration d'un modèle conceptuel. Ce modèle conceptuel constitue une représentation sommaire d'une partie restreinte du monde, en l'occurrence une lecture donnée de *Pour un nouveau roman*. Les modèles conceptuels sont constitués :

- d'entité, les objets représentés (par exemple : les premières publications, les articles de *Pour un nouveau roman*, les références transtextuelles)
- chacun des objets d'une entité sont appelés « instances » (ainsi « Une voie pour le roman futur » est une instance de l'entité ARTICLES)
- d'attributs, les qualités de ces objets (par exemple : date de publication, page de début, nature de la référence)
- d'association, les relations qui unissent les entités (par exemple : « correspond à », « mentionne »), elles peuvent également être munies d'attributs
- chacune des associations sont munies de cardinalité qui précisent le nombre de fois minimal et maximal où l'entité est impliquée dans l'association (par exemple : l'entité « ARTICLE » peut ne pas mentionner d'entités de TRANSTEXTS et peut en mentionner un nombre virtuellement infini, la cardinalité de l'association « MENTION » à l'endroit de « ARTICLES » sera donc 0,n).

Par convention on évite l'usage d'accents et d'espace et les noms d'entités et d'associations sont inscrits en majuscule, les entités sont représentées par un rectangle, les associations par un cercle. Par ailleurs, parmi les attributs, notons la nécessité d'utiliser l'un des attributs comme clef primaire soulignée par convention, identifiant de chacun des objets.

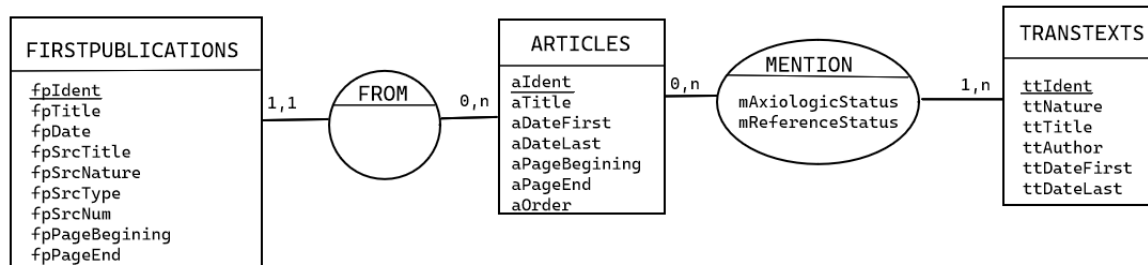


FIGURE 1 – Modèle conceptuel

Afin de permettre une implémentation aisée et rigoureuse nous préfixons nos attributs avec la (ou les) première(s) lettre(s) de l'entité ou de l'association à laquelle ils renvoient ; dans les cas où une association commence par la même lettre qu'une autre entité nous lui substituons les initiales des deux entités mises en relation.

Chacun des articles de *Pour un nouveau roman* constitue une instance de l'entité ARTICLES définie par un identifiant (aIdent), leur titre (aTitle), une ou deux dates (aDateFirst et aDateLast) déclarée(s) par Alain Robbe-Grillet leur place dans le recueil (aOrder) et leur étendue incarnée par deux attributs aPageBeginning et aPageEnd correspondant respectivement à la première et à la dernière page de l'article.

La deuxième entité FIRSTPUBLICATIONS est liée par une association FROM à ARTICLES. Ses attributs préfixés « fp » caractérisent l'instance constituée par la première publication, ceux préfixés « fpSrc » décrivent la source de cette première publication, soit le journal ou la revue dont elles sont issues. Créer une nouvelle entité pour ces sources ne nous parut pas significatif car ces sources ne nous intéressent qu'en ce qu'elles induisent une tonalité (polémique, scientifique, savante, profane) ou une réception particulière aux premières publications.

La troisième entité SUBJECTS est un recensement des domaines abstraits dont traitent (association ABOUT) les ARTICLES. Afin d'être le plus pertinent possible nous nous proposons de produire des valeurs les plus précises possibles pour l'attribut sDomain (« théorie Littéraire XX<sup>e</sup> » plutôt que « littérature »).

Intitulée TRANSTEXTS la quatrième et dernière entité est constituée de toutes les œuvres, auteurs ou concepts (identifiés comme étant de seconde main) mentionné par Alain Robbe-Grillet. La nature diverse (« caricature bien connue » ou simplement « Heidegger ») des instances de cette entité explique le foisonnement d'attributs qui seront, selon les cas, sans valeur ou bel et bien mobilisés.

L'association MENTION illustre les liens qu'entretiennent les ARTICLES avec les TRANSTEXTS, les attributs mAxiologicStatus et mReferenceStatus caractérise le lien que *Pour un nouveau roman* entretient avec telle ou telle référence. Si les valeurs possibles de mAxiologicStatus sont relativement restreintes (« éloge », « blâme », « ambiguë »), les valeurs de mReferenceStatus sont plus difficiles à caractériser simplement. En effet si dans certains cas Alain Robbe-Grillet cite une œuvre de manière explicite en donnant auteur et titre, il s'épargne bien souvent de donner ses références précises ; alors nous faut-il être en mesure de caractériser toutes les nuances de l'implicite : l'auteur est-il cité sans être nommé ? l'emprunt manifeste est-il désigné comme un emprunt d'une source à son tour déclarée ou non ? etc. Aussi optons-nous pour un système similaire à celui mis en œuvre pour l'attribut asImportance. Si nous nous sommes efforcés d'établir un système rigoureux et adapté au texte, telles catégories ne se défont jamais tout à fait d'une appréciation subjective (voir 2.3).

## 2.2.2 Modèle relationnel

La deuxième étape de la constitution d'une base de donnée est la conversion du modèle conceptuel au modèle relationnel qui correspond à une représentation schématique de la manière dont les données seront inscrits dans la base. Le point crucial de cette conversion est la gestion des associations. Entités et associations sont remplacés par des relations ou *table* qui, selon les cas, illustre des relations de dépendances ou non entre elles.

En effet, lorsqu'une seule des entités liées par une association à une autre a une cardinalité maximale de « 1 », cela signifie qu'elle n'a pas d'existence indépendante de l'autre entité. Alors l'association ne devient pas une relation mais n'est plus présente dans le modèle relationnel que par la présence d'une « clef secondaire » dans la relation dépendante de l'autre, cette clef secondaire à la même valeur que la clef primaire de l'instance cible. Au contraire, lorsque les deux entités sont reliées par une association dont les cardinalités maximales sont « n », alors l'association devient une relation contenant deux clefs secondaires : les clefs primaires des deux instances liées par l'association.

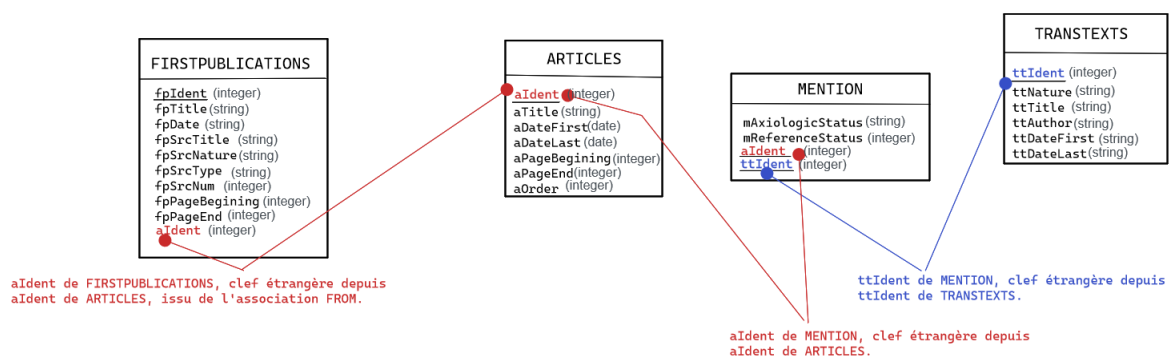


FIGURE 2 – Modèle relationnel

L'association FROM reliant les entités FIRSTPUBLICATIONS et ARTICLES disparaît dans le modèle relationnel car la cardinalité maximal de FIRSTPUBLICATIONS a pour valeur « 1 », laquelle est donc dépendante de ARTICLE dont la cardinalité maximale a pour valeur « n » (un article peut être une compilation ou une réécriture de plusieurs publications premières mais les articles originaux ne correspondent jamais qu'à un seul article du recueil final). Dès lors les instances FIRSTPUBLICATIONS contiennent désormais une clef secondaire qui correspond à la clef primaire d'une instance de ARTICLES.

L'association ABOUT devient une table car les deux entités qu'elle relie ont pour cardinalité maximale « n » (un même SUBJECT peut être traité par plusieurs ARTICLES et chaque ARTICLES peut traiter de plusieurs SUBJECT). ABOUT est dans le modèle relationnel une relation avec pour clef primaire deux clefs secondaires, l'une correspondant à la clef primaire de ARTICLES, l'autre correspondant à la clef primaire de SUBJECTS.

L'association MENTION devient une table car les deux entités qu'elle relie ont pour cardinalité maximale « n » (un même ARTICLES peut faire référence à plusieurs TRANSTEXTS et chaque TRANSTEXTS peut être mentionnée par plusieurs ARTICLES). MENTION est dans le modèle relationnel une relation avec pour clef primaire deux clefs secondaires, l'une correspondant à la clef primaire de ARTICLES, l'autre correspondant à la clef primaire de TRANSTEXTS.

### 2.2.3 Implémentation

Lors de l'implémentation, nous nous connectons à un serveur local (soit un serveur hébergé sur notre propre machine) via un logiciel dédié et rentrons à la main ou grâce à des scripts les données qui prennent dans l'interface de l'application l'apparence de tableaux (on retrouve notre modèle relationnel). Proches de la langue naturel, les scripts servant à la création de la base n'ont en eux-même que peu d'intérêt, on envoie littéralement des chaînes de caractères dans un ordre donné.

## 2.3 Mode d'établissement des valeurs des attributs

### 2.3.1 aOrder, conception de la structure du recueil

Il convient de noter une particularité dans la structure du recueil qui nécessita un choix de notre part : cinq articles sont présentés dans le recueil comme des sous-sections d'un chapitre « Éléments d'une anthologie moderne », dès lors il eût pu paraître nécessaire de prévoir des valeurs de aOrder sur le modèle 5.1, 5.2 etc. dénotant section et sous-section, cependant dans la mesure où l'article enchâssant les cinq critiques littéraires constituant l'ensemble est ajouté *a posteriori* il nous parut préférable de le considérer comme un article à part entière détaché de ses sous-articles qui n'entretiennent aucun lien explicite si ce n'est leur introduction, sorte de propos général ayant une fonction de seuil, ce choix nous paraît d'autant plus déterminant que l'on note l'absence de conclusion achevant de constituer l'ensemble.

De même si la table des matières de *Pour un nouveau roman* présente des sous-sections « personnage », « intrigue », « engagement » de l'article « De quelques notions périmées », ces sous-sections sont bien moins marquées dans le texte et nous semblent constituer davantage des paragraphes titrés issus d'articles fortement réécrits pour s'intégrer comme un tout homogène.

### 2.3.2 Valeurs de mReferenceStatus

Afin de modéliser de manière efficace et rigoureuse le statut des références, nous avons opté pour un système d'entiers inversement proportionnels au degré d'explicite des références dans le texte d'Alain Robbe-Grillet.

- Valeur **0, explicite** citation, du moins segment présenté dans le texte comme telle dont la source (auteur ou œuvre est mentionné).
- Valeur **1, mention** l'entité est mentionnée sans être citée. Il peut s'agir d'une glose interprétée (où l'interprétation de Robbe-Grillet est explicite).
- Valeur **2, mention ambiguë** cette valeur est réservée presque exclusivement à des entités collectives mentionnées sans nécessairement que les signifiés (les auteurs désignés par « les critiques traditionnels ») soient identifiables. Pareille identification étant difficile voire impossible : on constate qu'il s'agit bien souvent d'un procédé rhétorique visant à discréditer sans les nommer des adversaires réels ou imaginaires.
- Valeur **3, emprunt non déclaré fortement suggéré** réservée aux cas où Alain Robbe-Grillet emprunte un concept cite ou glose une référence dont il ne donnera pas la source mais dont la paternité est suffisamment présente à l'esprit de ces lecteurs ou suffisamment appuyé par lui pour être inféré. Ainsi lorsque Robbe-Grillet disserte sur tel élément de Balzac qui « fait vrai » le lecteur compétent reconnaît sans mal la conception de Barthes régulièrement mobilisé par Alain Robbe-Grillet.
- Valeur **4, emprunt ou mention non déclaré(e) non suggéré(e) reconstitué(e)** réservée pour des emprunts qui sont reconnus par l'éditeur sans qu'ils ne soient signalés par l'auteur, ainsi p. 69 lorsque Alain Robbe-Grillet cite des lieux propices à la poésie romantique y glissant « vallon » nous identifions Lamartine. Faisant l'objet d'une intervention de l'éditeur qui peut reconnaître des références non produites par Robbe-Grillet ou au contraire ne pas en reconnaître il nous a semblé nécessaire de différencier cette valeur de « 3 ». Enfin notons que dans ces cas comme dans les cas précédents lorsque la valeur de l'un des attributs est reconstituée par l'éditeur nous les insérons entre crochet, pour les repérer et corriger aisément si besoin mais également par honnêteté intellectuelle si pareil travail devait être amené à intégrer un travail de recherche plus vaste sur *Pour un nouveau roman*.

### 2.3.3 Valeurs de mAxiologicStatus

Pour délimiter les valeurs de axiologicStatus nous partons de deux polarités premières, le blâme et l'éloge constituant le moteur des théories de Alain Robbe-Grillet et le cœur de sa rhétorique, auxquels nous adjoignons deux autres statuts axiologique l'ambiguïté et l'indifférence.

S'il est aisé de reconnaître que la référence Balzac (ses œuvres ou le concept emprunt des conceptions de Barthes) fait l'objet d'un blâme il est plus difficile de juger le statut d'un référent comme Stendhal qui n'est pas mentionné pour lui-même mais comme argument servant à critiquer « un jeune écrivain contemporain qui écrirait comme Stendhal ». Nous avons choisi d'utiliser les valeurs blâme et éloge de manière indifférente lorsque la référence est critiquée ou vantée de manière explicite ou sollicitée comme raison d'une critique ou d'un éloge portant sur une tierce référence.

Le statut « ambiguë » sert lorsque Alain Robbe-Grillet exprime de manière explicite une difficulté à rejeter ou inclure tout à fait une référence comme étant symptomatique de la modernité (objet d'éloge) ou de la tradition (objet de blâme). Cette valeur sert également dans les cas où Alain Robbe-Grillet se sert d'un argument proprement sartrien pour critiquer Sartre (parfois désigné de manière implicite (valeur 3 de mReferenceStatus) par une formule tel « les engagés » ou pléthore de synonyme désignant ce que l'on qualifierait encore aujourd'hui de « staliniste »). Cette valeur marque donc également l'habileté rhétorique, d'aucuns diraient la « mauvaise foi » d'Alain Robbe-Grillet. Notons cependant que le pastiche à valeur de charge tel au sein de « Nouveau roman, homme nouveau » n'est pas considéré comme ambiguë car l'emprunt à Sartre ici ne sert qu'à renforcer encore une opposition frontale à ses thèses.

Enfin la valeur « indifférent » sert à désigner une mention qui n'est pas mobilisée dans l'argu-

mentation semblant avoir une valeur plus neutre de comparaison dénuée du moindre jugement de valeur sur le référent lui-même. En effet dans les premières pages du recueil Alain Robbe-Grillet s'attaque à un « dictionnaire encyclopédique de notre temps » pour la définition que celui-ci propose de Schönberg, si dans ce cas la référence au musicien n'est pas tout à fait neutre (le choix de ce compositeur intellectuel supposé hermétique rappelle le nouveau roman), il est difficile de rattacher le référent « Schönberg » au système axiologique de l'essai ; c'est bien le dictionnaire encyclopédique qui fait l'objet d'une charge et si l'on devine une sympathie pour le musicien de la part de Robbe-Grillet cette sympathie est inférée par le lecteur sans être partie prenante de l'argumentaire.

### 2.3.4 Les valeurs datées de TRANSTEXTS et FIRSTPUBLICATIONS en *string*

Dans le cas des dates de TRANSTEXTS constituées selon les cas d'une seule date `ttDateFirst` (de publication) ou de deux `ttDateFirst` et `ttDateLast` (naissance et mort d'un auteur ou première publication et traduction française antérieure à la publication de l'article au sein duquel la référence est mobilisée), nous avons dû opter pour le domaine *string* car certaines dates étaient antérieures à 1900, limite posée par SQL.

Dans le cas des FIRSTPUBLICATIONS, nous avons choisi dans le modèle relationnel d'utiliser le domaine *string* pour l'attribut `fpDate` car selon la nature de la source (quotidien, mensuel, annuel) la valeur de l'attribut changera de structure (10 octobre 1957, été 1963) ne permettant pas de l'inscrire comme une date (à moins de créer des dates arbitraires, ce qui n'aurait à nos yeux pas de sens).

### 2.3.5 Valeurs de l'attribut `ttTitle` de TRANSTEXTS

Au sein de TRANSTEXTS nous réunissons des instances de nature diverses (les discriminer est le rôle de l'attribut `ttNature`) pour lesquels le choix de l'attribut `ttTitle` peut sembler étrange : nous nous trouvons parfois à employer cet attribut pour inscrire le nom d'un concept dans la base. En effet si la dénomination de « titre » peut paraître saugrenue pour des instances de nature « concept » elle nous paraissait moins arbitraire que n'aurait pu l'être « nom » pour des œuvres, au demeurant l'emploi de « nom » ne nous satisfaisait pas non plus pour désigner les concepts. Par ailleurs considérer un concept comme un titre avait l'avantage de pouvoir plus aisément lui attribuer un auteur, voire appliquer des mises à jour ultérieures sur la base dans le cas où l'usage particulier que Robbe-Grillet fait de tel ou tel concept pourrait être rattaché à une œuvre particulière.

Notons que dans les cas où nous reconstituons une référence non mentionnée et peu sous-entendu (valeur « 4 » de `mReferenceStatus`) ou complétons pour paraphraser des propos allusifs de Robbe-Grillet, nous ajoutons des crochets droits à la valeur afin de noter notre intervention. Dans ces cas nous nous sommes efforcés de suivre une dénomination canonique non ambiguë.

Enfin si nous aurions pu faire des « entités collectives » des concepts en vertu du fait que les « critiques traditionnels » semblent davantage une idée abstraite qu'un groupe nettement délimité, dès lors nous aurions pu faire des valeurs `ttAuthor` « critiques traditionnel » un `ttTitle`, titre du concept. Nous avons privilégié la valeur « entités collectives » car ces instances sont mobilisées par Alain Robbe-Grillet bel et bien en tant que producteurs de textes et l'abstraction relativement élevée de pareille dénomination nous semble un procédé rhétorique permettant de cibler le plus grand nombre de personnes sans jamais les nommer, typique d'un écrit à la charge polémique aussi élevée.

## 3 Encodage

Afin de proposer une édition numérique enrichie de *Pour un nouveau roman*, nous optons pour un encodage en XML-tei. Pensé courant octobre et novembre, l'encodage à proprement parler commence début mars.

### 3.1 Principes généraux

Le langage XML est un langage de balisage : l'encodeur entoure chacun des éléments qu'il juge devoir être encodé (l'ensemble du texte, un paragraphe, un mot, voire une lettre) de balises ouvrantes et fermantes éclairant la nature du passage qu'elles entourent (libre à l'encodeur d'encadrer un paragraphe en tant que paragraphe ou en tant que saut sémantique).

<p><quote type="cit" corresp="pr\_02\_92" cert="0" ana="2"><J'y pense, j'y pense, un livre quelle prétention dans un sens, mais quelle extraordinaire merveille s'il est raté dans les grandes largeurs.></quote> Ainsi <ref type="author" cRef="pinget\_robert" cert="9">Robert Pinget</ref> nous prévient de ses ambitions et cet écrivain honnête (l'espèce n'est pas si répandue), qui met tout son soin, depuis déjà quelques années, à <hi rend="italic">rafer</hi> ses livres dans les grandes largeurs, est en train de passer à peu près inaperçu—même des spécialistes, noyés par profession dans le flot quotidien des récits linéaires, et réussis— alors que ces livres (ceux de <ref type="author" cRef="pinget\_robert" cert="9">Pinget</ref>), apparemment <hi rend="italic">sans queue ni tête,</hi> sont peut-être déjà précisément les <quote type="cit" corresp="unknown1" cert="9" ana="2"><extraordinaires merveilles></quote> annoncées.</p></small>

FIGURE 3 – Exemple d'encodage xml-tei

L'encodeur étant libre de choisir les éléments qu'il encode et comment il les encode, des conventions internationales ont émergées afin de permettre l'interopérabilité des corpus balisés. Le plus utilisé en science humaine demeure sans doute le standard de la *Text Encoding Initiative* (TEI) que nous employons pour la présente édition. Initialement pensé pour la restitution de sources anciennes (tels des manuscrits du XVI<sup>e</sup>) cet encodage nous permet d'identifier des éléments structurel très fin et de mettre en correspondance différents extraits constituant un corpus (voir 3.2.2).

Lors du rapport d'étape soumis courant décembre 2022, nous avions l'intention de produire une ODD. Un fichier permettant la génération d'un schéma et d'une documentation technique sur les choix d'encodage fait pour le présent travail (soit une spécialisation des éléments xml-TEI), le temps nous a finalement manqué pour produire tel travail qui nous parût au demeurant en partie redondant avec le présent document.

### 3.2 Mise en œuvre

#### 3.2.1 Un premier fichier d'encodage

Afin de pouvoir encoder le texte du *Pour un nouveau roman* nous nous procurons une version numérique du texte en achetant la seule édition disponible de l'œuvre au format « .epub ». Nous savons que les fichiers d'un epub se présente sous la forme d'une archive : il suffit de changer l'extension « .epub » en « .zip » pour pouvoir explorer son contenu et en extraire les fichiers contenant le texte.

Ces fichiers se présentent sous la forme de fichiers « .html », un fichier par chapitre du recueil. Grâce à un script perl nous récupérons l'ensemble du contenu textuel au sein d'un seul fichier .xml.

Le même script procède simultanément au nettoyage des fichiers d'origine (nous remplaçons les balises html par des balises xml-tei lorsque cela est pertinent et supprimons toutes les balises inutiles).

Ont été supprimés :

- les éléments <div> vides servant à espacer le corps du texte (une indication des suppressions et des tailles des éléments supprimés est à chaque fois ajoutée en commentaire dans le fichier d'encodage)



- les éléments `<span>` parasite qui redoublait, entre autres, tous les éléments `<p>` sans ajouter d'information de mise en forme
  - les éléments `<b>` et `<a>` muni d'un attribut `@id` marquant le début des chapitres
- Ont été remplacés par les balises xml-tei jugée pertinentes :
- les éléments `<a>` muni d'un attribut `@id` signalant les débuts de page par des éléments `<pb>` muni d'un attribut `@n`
  - les éléments `<i>`, au demeurant déprécié selon les normes actuelles du web, par des éléments `<hi>` muni d'un attribut `@rend` avec pour valeur "italic"
  - les éléments `<p>` marquant les paragraphes ont été conservés mais sans leur attribut `@class` de valeur "txt"
  - les éléments `<h1>` marquant les titres ont été remplacé par des éléments `<head>`
  - les éléments `<h2>` marquant les titres de sous-sections (par exemple « L'intrigue », sous-section de « De quelques notions périmés ») par des éléments `<head>` avec un attribut `@type` ayant pour valeur "subsection\_head"
  - les éléments `<small>` par des éléments `<hi>` munis d'un attribut `@rend` avec pour valeur "small-caps"
  - les éléments `<sup>` par des éléments `<hi>` munis d'un attribut `@rend` avec pour valeur "exposant"
  - les éléments `<blockquote>` par des éléments `<cit>`
  - les éléments `<p>` et ses attributs marquant la mise en forme du nom de l'auteur de la citation mise en exergue, par un élément `<ref>`

Notons que le nettoyage a été effectué en conservant, grâce à des commentaires, les traces de balises supprimées que l'on pourrait vouloir restaurer (tels les éléments `<div>` utilisés dans les fichiers html d'origine pour insérer du blanc dans le corps du texte).

Nous ajoutons des balises ouvrantes `<quote>` à chaque fois que le script de nettoyage rencontre le caractère "«" et ferme après le caractère "»", afin d'effectuer un premier repérage automatique des citations ou emprunts, qui seront ensuite complété et corrigé à la main si besoin. Notons que le script de nettoyage ajoute également ces éléments en début et en fin de paragraphe du segment « Joë Bousquet le rêveur » où nécessaire (lorsque Alain Robbe-Grillet cite plus d'un paragraphe il n'insère pas de guillemets, rendant le balisage automatique plus laborieux).

Nous ajoutons également des éléments `<div>` marquant les sections du texte autour de chaque article du recueil ainsi qu'autour des passages identifiables à des sous-sections (telles les « notions périmées ») cette fois munis d'attributs `@type` ayant pour valeur "subsection".

Nous obtenons alors un fichier « .xml » valide qui n'est encore qu'une étape pour l'encodage complet.

### 3.2.2 Vers un encodage XML-TEI en vue d'une édition numérique

Afin d'intégrer les références transtextuelles de *Pour un nouveau roman* à notre édition pour permettre l'édition enrichie que nous nous proposons de réaliser nous employons un encodage de type « corpus ».

```

1 <?xml-version="1.0"-encoding="UTF-8"?>
2 <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml" schematypens="http://relaxng.org/ns/structure/1.0"?>
3 <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml"
4   schematypens="http://purl.oclc.org/dtd/schematron"?>
5 <teiCorpus version="3.3.0" xmlns="http://www.tei-c.org/ns/1.0">
6   <teiHeader>
19     <teiHeader>
20       <fileDesc>
33         <title>
38           <title>
40             <title>
53               <title>
58             <title>
59             <title>
60             <title>
61             <title>
62             <title>
63             <title>
64             <title>
65             <title>
66             <title>
67             <title>
68             <title>
69             <title>
70             <title>
71             <title>
72             <title>

```

FIGURE 4 – La structure du fichier d’encodage

Alors que la majorité des encodages TEI se contente d’un seul élément <TEI> contenant l’œuvre ou le manuscrit encodé nous employons un élément <corpus> qui contiendra plusieurs éléments <TEI> identifiés grâce à des attributs @xml:id. Une version vide du fichier xml a pour cela été produite. Cet xml vide converti au format texte brut est ensuite injecté par le script de fusion et de nettoyage des fichiers qui compose *Pour un nouveau roman*. Il contient :

- un élément <teiHeader> (sorte de carte d’identité du document ou du texte) pour l’ensemble du fichier, contenant des informations succinctes sur le projet d’édition auquel est rattaché le fichier
- et quelques éléments <TEI> accompagnés de <teiHeader>, vides pour les extraits des références transtextuelles
- un élément <TEI> et <teiHeader> contenant les informations relative à l’édition de *Pour un nouveau roman*. C’est cet élément <TEI> dans lequel sera injecté le texte nettoyé et pré-encodé par le script.

Si nous reproduisons l’intégralité de *Pour un nouveau roman* dans l’élément <text> qui lui correspond nous n’insérons dans les autres éléments <text> que les extraits qui nous intéressent : nous produisons bien une édition de *Pour un nouveau roman* inscrit dans un corpus plus vaste, pas l’édition d’un corpus dont *Pour un nouveau roman* ne serait qu’un élément. Extraits transtextuels et passages de *Pour un nouveau roman* correspondant sont ensuite liés via un jeu d’attributs @corresp et @xml:id.

The screenshot shows a text editor with two panes. The left pane displays XML code with attributes like @corresp and @xml:id. The right pane displays a French text passage from 'Pour un nouveau roman'.

FIGURE 5 – Exemple de correspondance

### 3.2.3 Encodage sémantique automatisé

On désigne généralement par « encodage sémantique » l’ajout d’éléments xml identifiant des noms de personnages, des toponymes, des dates, des références, etc. Afin d’accélérer cette étape fastidieuse nous optons pour un premier encodage automatisé. Pour cela, nous ajoutons à notre script de nettoyage des lignes servant à repérer les noms d’auteurs et à les baliser selon le modèle suivant :

- un élément <ref> (référence)
- muni d’un attribut @type avec pour valeur "author"

- et d'un attribut @cRef (référence canonique) dont la valeur est constituée sur le modèle « nom\_prenom ». Générée automatiquement par le script, cette valeur comprendra des majuscules et/ou des lettres accentuées (puisque Alain Robbe-Grillet écrit « Samuel Beckett » la valeur sera "Beckett\_Samuel"), elle fera l'objet d'une correction grâce à l'emploi d'une feuille xsl.

Afin d'éviter le double balisage, nous balisons la dénomination complète (« Samuel Beckett ») en ajoutant un espace insécable entre le prénom et le nom, avant de baliser « Beckett » seul, s'il est précédé d'un espace sécable.

Puis grâce à une transformation xsl (voir 4)sommaire, nous générons l'affichage des éléments <quote>, <ref> et <hi> munis d'un attribut @rend="italic" afin de les identifier, corriger et compléter si besoin. Cette vérification se fait livre en main : la transformation renvoie également pour chaque élément identifié la page à laquelle il apparaît. Cette étape est répétée plusieurs fois, ce qui nous permet de raffiner le balisage automatique à chaque passage.

Avant de passer à l'encodage manuel nous utilisons une autre transformation xsl afin de faciliter cette étape fastidieuse :

- Nous ajoutons des attributs aux éléments <quote> qui n'en ont pas, leurs valeurs doivent correspondre autant que possible aux éléments répertorié dans notre base de donnée (voir 2.2.2).
  - @type correspondant à ttNature
  - @corresp qui correspond à l'@xml :id des extraits cités (encodé en <text>) et aux valeurs de ttIdent définie dans la base de donnée préfixé de "tt"
  - @cert qui correspond à mReferenceStatu, soit le statut de la référence (est-elle mentionnée, citée, etc.)
  - @ana correspondant au mAxiologicStatus (s'agit-il d'un blâme, d'un éloge, ou d'une mention indifférente?).
- Nous modifions les valeurs des @xml :id des éléments <div> qui entourent chacun des articles du texte, afin de les faire correspondre aux valeurs de notre base de donnée.
  - Par exemple, le chapitre « À quoi servent les théories » encodé par un <div> n'aura plus comme valeur de l'@xml :id "page006" mais "1". L'utilisation des pages comme identifiants nous paraît superflu, vu la présence d'élément <pb> à chaque début de page.
- Nous retouchons les titres (de sections et de sous-sections) afin de corriger l'encodage d'origine qui encodait chacune des lignes d'un titre ou sous-titre ainsi que les dates dans deux éléments <head> différents. Les capitales sont également remplacés par des minuscules qui seront plus tard affichées en petites capitales.
- Les mentions des dates sont intégrées aux titres et encodées en tant que <date> avec un attribut @rend dont la valeur est "italic". Par ailleurs, nous en retirons les parenthèses qui provoquaient une erreur sans doute dû au moteur d'expression régulière d'Oxygen<sup>1</sup>, les parenthèses seront rétablies dans la version html de l'édition.
- Tous les autres éléments et attributs déjà présents sont reproduits à l'identique. Par là nous nous assurons de pouvoir réappliquer la transformation sur notre fichier d'encodage manuel autant de fois que nécessaire. Nous nous contenterons de modifier le nom du fichier de sortie afin de ne pas écraser les précédentes itérations. Ainsi nous pouvons appliquer des corrections « en cours de route » sans perdre les ajouts manuels.

<sup>1</sup>En expression régulière, les signes « ( ) » servent à garder en mémoire les caractères qu'ils contiennent. Or, notre transformation recourait au moteur d'expression régulière inclut dans Oxugen pour intégrer les dates au titre. Ceci provoquait une erreur : la transformation considérait que les caractères « ( ) » devait être interprété comme des délimiteurs d'expression régulière et non comme le contenu à supprimer, ainsi les dates étaient bien supprimées mais pas les parenthèses les entourants.

### 3.2.4 Encodage sémantique à la main

Certains contenus textuels ne peuvent être repérés automatiquement et nécessitent donc d'être encodés à la main.

En effet il convient d'attribuer les bonnes valeurs aux attributs voire, de rectifier l'encodage automatique qui ne peut, par exemple distinguer entre la mention à la page 164 d'un titre « L'Année dernière à Marienbad » et la mention d'un élément diégétique de cette œuvre « se sont-ils vraiment rencontrés, aimés, l'année dernière à Marienbad ? » un peu plus loin.

Notons par ailleurs les difficultés à limiter le balisage. En effet l'on pourrait considérer certains éléments diégétique, tels les allusions aux personnages des œuvres de Beckett ou les supposées réactions sus-citées que Alain Robbe-Grillet prête au public comme devant être encodées en tant que <quote> : citer un personnage d'une œuvre, n'est-ce pas citer l'œuvre ? ces réactions mêmes (re)constituées par Alain Robbe-Grillet ne sont-elles pas des éléments textuels mobilisés par Alain Robbe-Grillet à la manière de citation à réfuter ?

Pour régler ces difficultés nous nous appuyons sur les objectifs que nous nous sommes fixés au moment où nous avons élaboré ce projet d'édition numérique : nous souhaitons produire une édition qui expose les relations transtextuelles de *Pour un nouveau roman* pour le replacer dans son époque, ou plutôt pour donner les représentations que le ce texte produit de son époque. Dès lors, il nous paraît intéressant de baliser les réactions (surtout (re)constituées) du public pour permettre de les comparer à la réalité de ces réactions. Au contraire les paraphrases (surtout si elles sont exactes) d'œuvres longuement commentées par Alain Robbe-Grillet, ne nous intéresse que modérément. La source identifiée, ou plutôt, vérifiée, l'intérêt des passages cités n'ont que peu d'intérêt par rapport au commentaire dans son ensemble.

De manière plus anecdotique, la recherche de chaîne de caractères contenant des apostrophe telles « L'Étranger, L'Immortelle » pose un problème épineux à résoudre car l'apostrophe est le signe utilisé par le moteur de recherche pour délimiter la chaîne. On écrit

- 'L'étranger',
- 'L&apos;étranger',
- 'L'"étranger' ;

le moteur renvoie une erreur. Une solution existe sans doute mais pour moins d'une dizaine de mentions aisément identifiées à la main, la chercher trop longtemps ne présentait aucun intérêt.

L'encodage des sources citées se fait également à la main, chacun des textes du corpus sont ajoutés aux éléments <TEI> munis d'un attribut @xml:id servant à les identifier sur le modèle : initialdel'auteur\_numérodelaœuvre » ainsi *Mahu ou le matériau* sera désigné par « pr\_01 » et *Le Renard et la boussole* par « pr\_02 ». Chacun des extraits est ensuite inséré dans un élément <div> ayant aussi un attribut @xml:id construit sur le modèle : id-del'œuvre\_pagededébutdelacitation », ainsi le premier extrait de *Le Renard et la boussole* correspondra à « pr\_02\_09 ».

## 4 Vers l'édition numérique : transformation XSL

Une fois l'encodage terminé, l'encodeur conçoit une transformation XSL, soit un fichier contenant des informations de traitement afin de passer d'un fichier XML peu lisible pour le lecteur à une édition numérique au format HTML. En effet les feuilles de styles XSL permettent de conserver le fichier XML originel pour en créer d'autres de types divers, en l'occurrence nous nous contentons de produire un site internet, soit des pages au format HTML. Notons que le langage de balisage HTML est un dérivé de l'XML qui ne permet pas de structurer le contenu aussi finement que l'XML mais permet un affichage via navigateur web pour lecture.

## 4.1 Principes généraux

Les transformations XSL fonctionnent par *template*, patron, qui commande le traitement d'un ou de plusieurs éléments XML selon des restrictions diverses laissées au soin de l'auteur de la transformation. On peut par exemple transformer un élément `<quote>` muni d'un attribut `@corresp` en un lien hypertexte, qui, lié à des scripts (voir 5) permettra l'affichage de contenus supplémentaires.

## 4.2 Mise en œuvre

Pour tenter d'éviter une écriture redondante nous nous efforçons de produire des templates efficace et réemployable. Par exemple pour constituer les pages html de notre édition numérique il nous faut générer autant de « `<header/>` » (soit la section au sommet de la page) qu'il y a de page. Aussi, nous contentons-nous de n'écrire qu'une seule fois le « `<header/>` » (et tous les éléments identiques sur toutes les pages) au sein d'un template nommé qui est ensuite appelé à charge génération de page avec des paramètres permettant de modifier quelques éléments essentiels qui doivent bien être unique (tel le titre de la page).

```

xsl:stylesheet
93 <xsl:template name="footer">
94 <!--
95 <a href="https://www.univ-lille.fr/" target="blank"></a>
97 </xsl:template>
98 <!--
99 <xsl:template name="header">
100 <!--
101 <div class="header">
102 <div class="header_div">
103 <div class="header_div">
104 <div class="header_div">
105 <div class="header_div">
106 <div class="header_div">
107 <div class="header_div">
108 <div class="header_div">
109 <div class="header_div">
110 <div class="header_div">
111 <div class="header_div">
112 <div class="header_div">
113 <div class="header_div">
114 </div>

```

FIGURE 6 – Templates gérant le bas (footer) et le haut (header) des pages

Notons que la production des pages est générée par un template matchant la racine du document xml et appelant le template nommé « `body` » qui lui-même appelle les templates « `<header>` », « `<footer>` » etc., adaptant le contenu de la page selon des paramètres déclarés au moment de l'appel du template.

```

271 <xsl:template match="/">
272 <!-- Appel des pages, une par chapitre -->
273 <xsl:for-each
274   select="/div[ancestor::TEI/@xsl:id=punr][@type=pagechap]">
275   <xsl:call-template name="body">
276     <xsl:with-param name="doc">
277       <xsl:with-param name="title">
278         <xsl:with-param name="content">
279       </xsl:call-template>
280     </xsl:with-param>
281   </xsl:call-template>
282 </xsl:for-each>
283 <!-- Appel de HOME -->
284 <xsl:call-template name="body">
285   <xsl:with-param select="$home" name="doc">
286   <xsl:with-param name="title">
287     <xsl:with-param name="content">
288   </xsl:call-template>
289   <xsl:with-param select="$presentation" name="doc">
290   <xsl:with-param select="présentation de l'œuvre" name="title">
291     <xsl:with-param select="pres" name="content">
292   </xsl:call-template>
293   <xsl:with-param name="body">
294   <xsl:with-param select="$db" name="doc">
295   <xsl:with-param name="title">
296     <xsl:with-param name="content">
297   </xsl:call-template>
298   <xsl:with-param name="body">
299   <xsl:with-param select="$db" name="doc">
300   <xsl:with-param name="title">
301     <xsl:with-param name="content">
302   </xsl:call-template>

```

FIGURE 7 – Extrait du template gérant la génération des pages du site

## 5 Une expérience de lecture : ajouts de scripts

### 5.1 Script pour la lecture

#### 5.1.1 Javascript, présentation générale

Afin de permettre l'interactivité d'une page, nous employons un langage de programmation extrêmement courant : javascript. Ce langage de programmation permet la programmation de fonction qui, selon l'élément sur lequel clique l'utilisateur provoquera tel ou tel comportement au sein de la page.

Une « fonction » est une suite d'instructions parfois conditionnés par des « paramètres », des informations extérieures à la fonction qui y sont injectées.

#### 5.1.2 Afficher les citations

La première fonction que nous développons sert à afficher les extraits des autres œuvres cités ou mentionnés par Alain Robbe-Grillet. Nous choisissons de programmer un affichage que nous espérons élégant et pratique. En bas à droite de la page apparaît un encart contenant l'extrait cité, sa source et des informations quant à l'emploi de la citation (est-ce une « mention », est-ce un « blâme » ou un « éloge » ?). L'encart est censé permettre de continuer à lire *Pour un nouveau roman* sans avoir à le refermer. Moins élégant peut-être qu'une version qui obscurcirait le reste de l'écran, nous pensons que permettre tel usage correspond davantage à ce que souhaiterait un lecteur effectif.

longue aguilte... » Il faut préciser : « La naissance d'un objet, j'ai remarqué, n'a pas lieu aujourd'hui, il y a du mouvement tout autour qui l'empêche de montrer la tête et demain tu t'avises qu'il existe. Par conséquent la meilleure explication des origines serait de commencer par des bruits de bouche et de glisser progressivement vers des paroles articulées jusqu'au moment où l'auditeur sans se poser aucune question participe à tout histoire. » C'est justement ce qui se produit ici : peu à peu, au milieu des digressions et des faillites, une tache de couleur rousse s'impose dans le tableau (les héros font de la peinture dans *Le Renard*, comme ils écrivaient des romans dans *Malin*). Cette masse d'abord indistincte prend bientôt la forme d'un renard, qui se dédouble en plusieurs personnages dont l'un n'est autre que David, le Juif errant. Ce renard invente un voyage, un voyage en Israël. Suit une espèce de reportage sur la vie dans les kibboutzim, interrompu çà et là par des évocations bibliques ou autres, des apparitions des pharaons et de sultans, des rencontres plus imprévues encore – celles par exemple de Don Quichotte et de son désert castillan. Le documentaire traîne en longueur, dans la chaleur de l'été palestinien. Le narrateur s'inquiète de cette lassitude qu'il constate chez le lecteur et chez lui-même, comme chez ses voyageurs : « Ça ne va pas, dit-il, seraient-ils déjà revenus ? » Pourtant Renard et David font encore la connaissance de Marie-Madeleine, dite « Mama », qui les aidera à se rembarquer, puis celle plus importante du Scribe accroupi : « Méfiez-vous. Ne lui parlez pas, il écrit tout. » En effet il note aussitôt lui-même sa propre rencontre... « Un fait, un fait entre des milliards, on leur dénie toute valeur dans la pratique, on les note, on les note c'est tout. »

Et le plus naturellement du monde, Renard, qui s'est depuis longtemps confondu avec le narrateur (nommé John Tintouin Porridge), se retrouve à Fantoine, Malu et les autres lui reprochent de les avoir un instant quittés. Ne les aimait-il plus ? « Mais oui je vous aime, nous sommes liés pour la vie... me voici. » Au milieu des apéritifs et des déambulations valseuses du retour, surgit mademoiselle Lorpailleur, la romancière ; elle demande à John où il voulait en venir. Il se rappelle

Statut de la référence : mention ; éloges. [13] [...] Une vraie conversation est avant tout futile me plaît, il me ressemble, je suis et quant à la sonorité remplie un rōdeur, je grapple à la sauvette, l'hiérisations et de sous-entendu, mon langage est incohérent, mais un vrai livre doit être le contraire la vérité est tordue et que j'ai dit-on. Quand je pense au décidé de lui laisser ma peau ? Il Don Quichotte je sais que les faut me pardonner mes singeries, inflexions de vois n'y sont pas rappelle-toi que don Quichotte traduites, mais l'art est si grand avait un plat à barbe sur la tête, qu'aucune modulations n'échappe, c'est ma sauvegarde, aucune. Vive Cervantes. Nous laissons l'Espagne très loin derrière nous. Si j'avais su je serais retourné en Espagne. Mal comme. Jusqu'à Tolède seulement. Le Renard et la Juive, Roger Robert, Paris, Les Éditions de Minuit, [1933] 2000.

FIGURE 8 – Capture d'écran du résultat obtenu

Au moyen de notre transformation xsl, nous créons pour chacune des citations l'appel d'une fonction qui recevra en paramètre (selon la citation) l'identifiant du passage cité :

On remarque que l'on injecte en paramètre l'identifiant (@corresp) du passage cité, et les statuts référentiels et axiologiques de la référence (@cert et @ana, respectivement) ainsi qu'un dernier paramètre dont les valeurs possibles sont « 1 » ou « 0 » qui sert à orienter le comportement de la fonction nommée « displayExtract ».

```

58 // FERMETURE DES EXTRAITS DÉJÀ OUVERTS
59 fonction closeExtract(){
60     let toHide=document.getElementsByClassName('extractDisplay');
61     for (var i = toHide.length-1; i >= 0; i--) {
62         toHide[i].setAttribute('class','extractHide');
63     }
64     // Suppression des paragraphes d'annotation créés, pour éviter les doublons.
65     let toDelete=document.getElementsByClassName('deleteMe');
66     for (var i = toDelete.length-1; i >= 0; i--) {
67         toDelete[i].remove();
68     }
69 }
70 }
71 fonction displayExtract(corresp,mAxiologicStatus,mReferenceStatus,c){
72 // FERMETURE DES EXTRAITS DÉJÀ OUVERTS
73     if(c=='1'){closeExtract();}
74 // IDENTIFICATION DE L'EXTRAIT À AFFICHER
75     let toDisplay=document.getElementById(corresp);
76 // CONSTITUTION DES ÉLÉMENTS DE COMMENTAIRES, mention du statut axiologie et référentiel
77 // 1. Tester Les valeurs de L'attributs, le traduire en toutes lettres et lui donner de la couleur.
78     var mAxiologicStatusTreated='';
79     var mReferenceStatusTreated='';
80     var axiSpan=document.createElement('span');
81     var refSpan=document.createElement('span');
82     if (mAxiologicStatus=='0') {
83         mAxiologicStatusTreated='blâme';
84         axiSpan.setAttribute('class','Axi0');
85     }
86     if (mAxiologicStatus=='1') {
87         mAxiologicStatusTreated='indifférent';
88         axiSpan.setAttribute('class','Axi1');
89     }
90     if (mAxiologicStatus=='2') {
91         mAxiologicStatusTreated='éloge';
92         axiSpan.setAttribute('class','Axi2');
93     }
94     if (mAxiologicStatus=='3') {
95         mAxiologicStatusTreated='ambiguë';
96         axiSpan.setAttribute('class','Axi3');
97     }
98     if (mReferenceStatus=='0') {
99         mReferenceStatusTreated='citation explicite';
100         refSpan.setAttribute('class','Ref0');
101     }

```

FIGURE 9 – Extrait du script gérant l'affichage des extraits cités

```

110     if (mReferenceStatusTreated=='emprunt non déclaré fortement suggéré';
111         refSpan.setAttribute('class','Ref3');
112     }
113     if (mReferenceStatusTreated=='emprunt ou mention non déclaré(e) non suggéré(e) reconstitué(e)';
114         refSpan.setAttribute('class','Ref4');
115     }
116     let mReferenceStatusTreatedObj=document.createTextNode(mReferenceStatusTreated);
117     let mAxiologicStatusTreatedObj=document.createTextNode(mAxiologicStatusTreated);
118     let separator=document.createTextNode('<small> ');
119     let point=document.createTextNode('.');
120     refSpan.appendChild(mReferenceStatusTreatedObj);
121     axiSpan.appendChild(mAxiologicStatusTreatedObj);
122     // 2 Création des éléments html etc.
123     let mQuoteStatus=document.createElement('p');
124     mQuoteStatus.setAttribute('class','editor deleteMe');
125     let mQuoteStatusTxt=document.createTextNode('Statuts de la référence<small> ');
126     mQuoteStatus.appendChild(mQuoteStatusTxt);
127     mQuoteStatus.appendChild(refSpan);
128     mQuoteStatus.appendChild(separator);
129     mQuoteStatus.appendChild(axiSpan);
130     mQuoteStatus.appendChild(point);
131     // 3 Déduction de la position à laquelle on ajoute les info
132     // On insérera avant le frère du premier fils de la divtoDisplay,
133     // soit avant l'élément qui suit la croix pour fermer (premier fils de ToDisplay),
134     // ie le 1er paragraphe (de l'extrait ou de commentaire de l'éditeur).
135     let firstChild=toDisplay.firstChild;
136     let theOneEvenbefore=firstChild.nextSibling;
137     let theOnebefore=theOneEvenbefore.nextSibling;
138     // 4 Apendage
139     toDisplay.insertBefore(mQuoteStatus,theOnebefore);
140 // AFFICHAGE DE L'EXTRAIT CLIQUÉ
141     toDisplay.setAttribute('class','extractDisplay');
142 }

```

FIGURE 10 – Extrait du script gérant l'affichage des extraits cités

En effet, cette fonction n'a qu'à rendre visible le passage concerné lorsque l'utilisateur clique sur une citation. Cependant les choses se compliquent dès lors que nous souhaitons afficher plus d'un extrait pour une même citation. En effet, la même citation par exemple « Don Quichotte » dans l'article « Un roman qui s'invente lui-même » peut renvoyer à plusieurs passages de *Le Renard et la boussole*. Or l'encodage xml-tei nous permet précisément d'insérer plusieurs valeurs à l'attribut @corresp séparées par un espace.

On aura donc :

<quote corresp="pr\_02\_25 pr\_02\_50">Don Quichotte</quote>



Nous souhaitons arriver au résultat suivant :

, soit « lorsque l'utilisateur clique sur ce segment la fonction `displayExtract` est appelée deux fois sur deux extraits différents. Pour ce faire nous devons séparer les deux valeurs et répéter la valeur de l'attribut `@onclick` en ne changeant qu'un seul paramètre. Dans un premier template xsl nous testons la présence ou non d'un espace au sein de la valeur `@corresp`. S'il y a un espace une première partie de la valeur de l'attribut `@onclick` (soit un premier extrait) est gérée après quoi un autre template est appelé. Ce template nommé « `correspAffect` » reçoit en paramètre tout le contenu de l'attribut qui suit l'espace, il procède ensuite de même : gère le premier extrait en créant l'appel de la fonction sur ce qui précède l'espace (soit le deuxième extrait), puis via un appel récursif va gérer les uns après les autres toutes les valeurs de l'attribut `@corresp` après l'espace. S'appelant lui-même le template boucle sur une partie toujours plus réduite de l'attribut `@corresp` qui correspondront à autant de paramètres ensuite envoyé dans le Javascript jusqu'à ce qu'il n'y ait plus d'espace au sein du reste de `@corresp`, alors, le template produit un dernier appel à la fonction (s'il n'y a plus d'espace il reste l'identifiant d'un extrait) puis s'arrête.

FIGURE 11 – Le template correspAffect qui sépare les correspondances multiples

15



Si c'est bien Javascript qui produit l'affichage et lie plusieurs éléments html générés via xsl depuis l'xml, c'est ici encore l'xsl qui nous permet de produire très efficacement des liens complexes entre différents langage de programmation.

### 5.1.3 Afficher les variantes

## 5.2 Interactivité de la base de donnée

Afin de rendre la base de donnée lisible, plus aisément que dans un tableau nous employons l'outil GraphCommons<sup>2</sup>.

Cet outil permet la production de graphique depuis un fichier de type .csv, (soit un document texte, interprétable comme un tableau par les application de type tableur et assurant une interopérabilité très élevée notamment au sein de script perl ou javascript). Ainsi, nous exportant le contenu de la base de donnée au bon format puis retravaillons les données pour un meilleur affichage dans l'application. La base de données est d'abord exportée au format .csv, le tableau au fichier texte que l'on obtient a alors une colonne (les colonnes sont séparées au sein du texte par des virgules) par propriété des entités (voir : 2.2.2) de la base afin d'être utilisable par l'application GraphCommons, les colonnes nécessitent d'être renommée et quelques valeurs ajustées. En effet, une colonne « weight » qui détermine la taille d'un lien est interprétée de manière croissante, or nous avons choisis de représenter l'intensité des liens de manière décroissante (« 0 » étant la citation et « la mention reconstituée »), pour un affichage appropriée nous modifions donc les valeurs pour les inverser et renommons les colonnes notamment grâce à des scripts perl rudimentaires.

## 6 Conclusion

Tout au long de notre travail nous nous sommes efforcé d'employer la technique au service de questions scientifiques. Et si, bien souvent, la technique interroge le choix scientifique, nous nous sommes évertué à ne pas confondre questions et difficultés techniques. En effet trop souvent, sommes-nous tenter de remettre en question le

sans doute, tout n'y est pas, sans doute chercher à identifier et produire toutes les citations en contextes relève de l'impossible : Alain Robbe-Grillet peut s'être trompé, l'édition avoir disparu, une citation en appelle une autre, etc. Pareil projet cependant se mène non en ayant à cœur de couvrir l'ouvrage entier (du moins sans s'en faire l'illusion) et toute son époque, mais bien en sachant que l'on ne propose qu'une lecture, partielle partielle parmi d'autres, qui n'épuise pas l'œuvre.

<sup>2</sup>Voir : <https://graphcommons.com/graphs/2ffc7c8c-3d1b-4814-966b-593b1c206f3c>.