

LES SOURCES INÉDITES EN HISTOIRE

Travaux issus des journées d'étude de jeunes chercheurs organisées à l'École nationale des chartes, les 25 et 26 mai 2018, par le centre Jean-Mabillon et l'association Chroniques chartistes.

Études réunies par Léo Davy

École nationale des chartes

Date de mise en ligne : 13 décembre 2021.

*Contenu mis à disposition selon les termes de la licence
Creative Commons : attribution, pas d'utilisation
commerciale, pas de modification.*

POSSIBILITÉS ET VALORISATION D'UNE ÉTUDE SUR LE VOCABULAIRE DU ROI-PRÊTRE MÉDIÉVAL GRÂCE AUX HUMANITÉS NUMÉRIQUES

par MARIE-ASTRID HUGEL ◆

Possibilités et valorisation d'une étude sur le vocabulaire du roi-prêtre médiéval grâce aux humanités numériques

MARIE-ASTRID HUGEL ◆

Dans leurs réflexions respectives de 2004 et 2008, Joseph Morsel¹ et Jérôme Baschet² rappellent un écueil trop fréquent en histoire, celui de considérer la source par son contenu au détriment de la source comme objet. Cette dématérialisation a encore été renforcée par la numérisation des sources, nouvelle barrière entre l'historien et le produit « source », accompagnée d'un renouveau de l'utilisation de l'informatique grâce à la diffusion d'outils d'analyse et de visualisation. Cette croissance des méthodes informatisées de recherche a permis d'approfondir des sujets qui jusqu'alors ne pouvaient être étudiés qu'à partir d'un nombre limité de fonds. La question du *rex et sacerdos* médiéval, l'union du temporel et du spirituel qui traverse plusieurs types de documents, peut assurément profiter de ces nouvelles méthodes, car son étude était ardue en raison des difficultés à constituer un corpus cohérent puis à l'étudier sur plusieurs niveaux.

La présente contribution entend ainsi explorer ces nouvelles possibilités, partant de l'exemple de notre sujet de thèse sur l'expression *rex et sacerdos*. La réflexion se déploiera sur trois axes : le travail sur des sujets difficiles à appréhender sans ces réflexions et outils, la

1 Joseph Morsel, « Les sources sont-elles "le pain de l'historien" ? », dans *Hypothèses*, t. 7, 2004, p. 271-286, à la p. 280.

2 « Et c'est pourquoi on propose la notion d'*image-objet*, afin de souligner que l'image est inséparable de la matérialité de son support, mais aussi de son existence comme objet, agi et agissant. » Jérôme Baschet, *L'icongraphie médiévale*, Paris, 2008, p. 33-34.

méthodologie d'étude du produit « source » et de sa transformation en données, et enfin les usages futurs ouvrant la voie à des approches inédites sur les *corpora* ici constitués.

I. Le *rex et sacerdos*, entre absence de définition et oubli de la recherche

1. Origines et développement jusqu'au XIII^e siècle

Pourquoi parler d'expression ? L'idée suppose une architecture théorique claire, des contours définis, un usage peu ou pas ambigu. La thématique du *rex et sacerdos* échappe à toutes ces règles : elle est mal connue, se dissimule et se retrouve dans des sources diverses, utilisée dans des contextes inattendus.

1.1. Mentions bibliques : Melchisédech et le Christ

L'expression est associée à deux figures de la Bible : le Christ, véritable roi et véritable prêtre, mais aussi Melchisédech. C'est dans la titulature de ce deuxième personnage qu'apparaît la première mention de l'expression : au chapitre XIV de la Genèse, alors qu'Abraham revient de son combat pour libérer son neveu Loth, il rencontre Melchisédech (fig. 1). Ce dernier n'est l'objet que de trois versets³.

On retrouve ensuite le personnage surtout dans l'Épître aux Hébreux⁴. L'auteur anonyme voit dans ce mystérieux roi-prêtre de nombreuses similarités avec le Christ, et le décrit comme *assimilatus [...] Filio Dei*⁵. Cette description lie de manière unique le Christ et Melchisédech en faisant de ce dernier la première préfiguration explicite du Fils. Dans les premiers siècles du Moyen Âge,

³ Gn 14, 18-20 : « at vero Melchisedech rex Salem proferens panem et vinum / erat enim sacerdos Dei altissimi benedixit ei et ait benedictus Abraham Deo excelso qui creavit caelum et terra et / benedictus Deus excelsus quo protegente hostes in manibus tuis sunt et dedit ei decimas ex omnibus ».

⁴ Ps 110, 3 et Hebr 5,6.

⁵ Hebr 7,3.



Fig. 1 | Dieric Bouts l'aîné, *La rencontre d'Abraham et de Melchisédech*, huile sur bois, 1464-1467, Louvain, collégiale Saint-Pierre.

Melchisédech est avant tout connu et diffusé comme préfigurateur de l'Eucharistie par son don de ce pain et de ce vin et ainsi incarne le modèle du prêtre catholique par excellence. Avant d'être roi-prêtre, Melchisédech est prêtre.

1. 2. Le cas du Prêtre Jean

Outre Melchisédech, le Prêtre Jean est le seul roi-prêtre excepté le Christ (fig. 2). Sa première mention sous la plume d'Otton de Freising en 1146⁶ décrit un mystérieux roi appelé Prêtre Jean, qui pourrait soutenir les chrétiens face aux musulmans. Dans le contexte des croisades, cette aide bienvenue délie les imaginations des contemporains,

6 « Narrabat etiam quod ante non multos annos lohannes quidam qui ultra Persidem et Armeniam in extremo oriente habitans rex et sacerdos cum gente sua Christianus est, [...] bello petierit. » Otto de Freising, *De duabus civitatibus*, éd. Adolf Hofmeister, Hanovre, 1912 (MGH, SS. rer. Germ. 45), p. 363.



Fig. 2 | Diogo Homem, Le Prêtre Jean, *Queen Mary's Atlas*, ca. 1555-1559, Londres, British Library, ms. add. 5415 A, fol. 16, détail.

renforcées par l'apparition de lettres décrivant un royaume merveilleux. La légende se diffuse par la Lettre et par des romans, tel le *Jüngere Tituel* de la fin du XIII^e siècle dans lequel le héros rencontre le Prêtre Jean. Cet espoir dure jusqu'au XVI^e siècle puis tend à disparaître⁷.

En comparant les deux figures retenues, le principal élément commun repose dans une absence : celle d'une explication sur le fait qu'ils incarnent le *rex et sacerdos*, alors qu'en théorie seul le Christ peut légitimement porter ce double titre, car il est le vrai roi-prêtre⁸.

7 Keagan Brewer, *Prester John: The Legend and its Sources*, Farnham, 2015, qui a rassemblé des sources allant jusqu'au XVIII^e siècle.

8 Gélase I^{er}, *De anathematis vinculo* (PL 59) : « verum ventum est eundem regem atque pontificem ».

2. Contradiction médiévale : le problème de la dualité

1.1. Expression sans théorisation

L'absence d'argumentation s'explique par la contradiction inhérente qui entoure l'expression. Le fondement théologico-politique chrétien réside dans la séparation entre le souverain temporel, empereur ou roi, et le prêtre chrétien officiant réellement⁹. Seul le Christ dans sa perfection peut posséder les deux pouvoirs¹⁰. Mais l'existence d'autres figures incarnant le *rex et sacerdos* est avérée et pose problème.

Cet embarras autour de l'expression se retrouve dans les sources : il n'existe aucun texte évoquant explicitement la contradiction posée par Melchisédech après l'Antiquité tardive. Les textes patristiques portent principalement sur les Melchisédeciens, censés avoir placé Melchisédech au-dessus du Christ¹¹ : les Pères ont concentré leur attention sur la distinction entre ces deux derniers¹². Dans les sources médiévales, la figure est souvent mentionnée dans un texte, sans véritable développement. On trouve ainsi un corpus constitué essentiellement de textes théologiques.

1.2. Limites de la recherche traditionnelle sur la question

Les deux figures ont déjà fait l'objet de recherches historiques séparées. En ce qui concerne la bibliographie traitant du Prêtre Jean,

9 Constantin comme empereur romain portait le titre de *pontifex maximus*. Pouvoir temporel et pouvoir spirituel sont liés par essence dans le pouvoir impérial romain.

10 Le Christ possède de fait trois titres : il est prêtre, prophète et roi. La question de cette triade en lien avec la Trinité ne peut être abordée ici et un développement y sera consacré dans le travail de thèse.

11 Fred L. Horton, *The Melchizedek Tradition: A Critical Examination of the Sources to the Fifth Century A.D. and in the Epistle to the Hebrews*, Cambridge/Londres/ New York, 1976, p. 90-113.

12 Ambroise de Milan précise que Melchisédech n'est pas un ange mais un prêtre, un homme mortel : « Nam et illum Melchisedech, [...] non angelum utique secundum judaica ludibria intellegit ecclesia, sed virum sanctum et sacerdotem dei », dans *De fide libri V*, liv. III, chap. xi, l. 83 (PL 16, t. II, p. 607, 88). Consulté sur le Cross Database Searchtool de Brepols, le 13 décembre 2019.

il convient de remarquer que la plupart des études se concentrent sur le royaume de Jean¹³ et très peu sur ses pouvoirs, sujet de l'étude de Istvan Becjzy¹⁴.

La bibliographie portant sur Melchisédech est plus vaste et dispersée, du fait de sa présence dans la Bible et de sa place dans la théologie chrétienne. Dans la majorité des cas, la perspective reste similaire : la mention de Melchisédech dans une source suffit à l'ériger en étude de cas. En définitive, les seules synthèses, telle celle de Lucien-Jean Bord parue récemment¹⁵, émanent de théologiens souvent catholiques. Une exception existe cependant : Gilbert Dagron consacre en 1998 une étude complète sur l'idée de double pouvoir au prisme du césaropapisme dans l'Empire byzantin¹⁶.

3. Bornes et problématiques

Pour étudier le *rex et sacerdos*, la chronologie retenue par notre thèse va du XIII^e siècle, partant de l'avènement du pape considéré comme le plus théocratique, Innocent III (1198-1216), jusqu'au début du XVI^e siècle, en prenant en compte le règne de Maximilien I^{er} (1508-1519), mais en excluant la question du protestantisme¹⁷. Deux espaces ont été privilégiés : le royaume de France et le Saint-Empire romain germanique. Ce choix repose sur la différence de culture et de structure politique, distinguant un royaume de France porté par un pouvoir royal toujours plus grand, et un Saint-Empire davantage morcelé.

Dans ces limites temporelles et géographiques, la question la plus évidente mais aussi la plus difficile consiste à comprendre ce qu'est le *rex et sacerdos* et ce qu'il n'est pas. Cette approche doit en effet

¹³ Ulrich Knefelkamp, *Die Suche nach dem Reich des Priesterkönigs Johannes: dargestellt anhand von Reiseberichten und anderen ethnographischen Quellen des 12. bis 17. Jahrhunderts*, Gelsenkirchen, 1985.

¹⁴ István Pieter Becjzy, *La lettre du prêtre Jean : une utopie médiévale*, Paris, 2001, p. 61-62.

¹⁵ Lucien-Jean Bord, *Melchisedek : formation, histoire et symbolique d'une figure biblique*, Paris, 2013.

¹⁶ Gilbert Dagron, *Empereur et prêtre : étude sur le césaropapisme byzantin*, Paris, 1998.

¹⁷ Marie-Astrid Hugel, « Du roi-prêtre au roi », dans *Trajectoires*, t. 13, 2020, en ligne : <https://doi.org/10.4000/trajectoires.5294> (consulté le 22 mars 2021).

permettre de déterminer s'il existe un sens « strict » de *rex et sacerdos*, celui d'un dignitaire pleinement roi et pleinement prêtre, ou bien s'il existe d'autres acceptions et usages.

II. La méthodologie

1. La base théorique

1.1. Image médiévale : théorie et pratique

L'image occupe dans notre sujet une place équivalente au texte. Les textes parlant du *rex et sacerdos* sont surtout des courts extraits : ce corpus textuel est constitué de paragraphes, de phrases, voire de simples mentions. La minceur du matériau a de quoi surprendre : ainsi en est-il de la majorité des mentions qui ne signalent que *secundum ordinem Melchisedech*. Ces occurrences éparpillées ont pu décourager les historiens : c'est ici que l'image se révèle centrale pour eux.

Considérer l'image médiévale comme une source à part entière ne fait partie intégrante de la méthodologie de l'historien que depuis les années 1980, avec les travaux de Jacques Le Goff, de Michel Pastoureau ou de Jean-Claude Schmitt. L'image est un medium dont le sens n'a pas fait l'objet d'une théorisation poussée¹⁸. Parmi toutes les possibilités offertes par l'image, l'une d'entre elles a particulièrement retenu notre intérêt : elle abolit la distance entre le lointain ou l'invisible et la personne qui la regarde¹⁹. Cette absence constitue la caractéristique même du *rex et sacerdos*. Si le texte permet un recul critique sur l'expression, tout en formant une gêne à sa compréhension par les contemporains, l'image donne à voir, au sens plein, les multiples possibilités offertes par cette formule, c'est-à-dire les signes visibles d'un imaginaire qui ne l'a pas exclue. L'image médiévale ne s'offre pas seulement à l'intellect, mais anime aussi la sensibilité.

¹⁸ Jean-Claude Schmitt, « Introduction générale », dans *Les images dans l'Occident médiéval*, dir. Jérôme Baschet et Pierre-Olivier Dittmar, Turnhout, 2015 (L'atelier du médiéviste, 14), p. 10.

¹⁹ *Ibid.*, p. 11.

Prendre le chemin de l'analyse des images, c'est étudier les contours d'une expression dont les perceptions médiévales se sont emparées pour la façonner et en tirer des représentations mentales démultipliées : quand il faut bien mettre en image Melchisédech ou le Prêtre Jean, comment leur donner corps ?

1. 2. Constitution des *corpora* textuel et iconographique

Pour constituer le corpus textuel en langue médiolatine, en ancien français et en vieil-allemand, la sélection s'est faite à partir de la pertinence réelle pour le sujet : Melchisédech et le Prêtre Jean sont-ils bien l'objet d'une phrase, d'un paragraphe ou d'un chapitre ? Nous écartons ainsi les sources ne contenant que des occurrences simples énoncées plus haut, comme la seule mention de Melchisédech ou du Prêtre Jean sans autre développement. À ce premier critère s'ajoute une sélection des sources portant sur des figures étudiées pour compléter l'étude, tels Aaron ou Gengis Khan : dans ce cas, le texte doit mettre en scène ou analyser les particularités du pouvoir de ce personnage. Si le texte en question donne un titre similaire à celui de *rex et sacerdos*, comme c'est le cas pour les rois Hasmonéens, le texte est alors retenu.

Le corpus iconographique a été construit selon les caractères internes de notre démarche²⁰ : la présence d'une double couronne, de Melchisédech ou du Prêtre Jean, en plus de figures complémentaires, comme Gengis Khan et les rois Hasmonéens. Certaines images présentes dans un seul but comparatif sont également annexées et partiellement indexées. Les manuscrits ne contenant aucune représentation iconographique ont été écartés, afin de concentrer l'étude sur la relation entre le texte et l'image. Seules les sources éditées, notamment les écrits de nature politique, ont été retenues, afin d'étudier la question du double pouvoir dans les conflits entre princes temporels et papes. De même, quelques sources iconographiques ont été sélectionnées, comme le vitrail de Melchisédech de la cathédrale de Chartres.

²⁰ Les caractères internes sont les éléments de l'image en elle-même et sont propres au projet de recherche : une couleur, un geste, un vêtement, un symbole. Au contraire, les caractères externes sont les éléments situés en dehors de l'iconographie : la date, le lieu de production, le type de support. Voir Séverine Lepape, « Formalisation et analyse statistique d'un corpus d'images », dans *Les images dans l'Occident médiéval...*, p. 333-350, à la p. 334.

Nous avons donc constitué deux *corpora* : le premier, textuel, comporte des sources politiques, des sources théologiques au sens large parmi lesquelles figurent une lecture typologique de la Bible, des sources historiques telles les chroniques universelles, et des œuvres littéraires ; le second, de nature iconographique, comporte les mêmes sources si elles sont imagées, auxquelles s'ajoutent des vitraux et des sculptures.

2. L'analyse

1.1. Lecture séquentielle et non séquentielle

Deux approches sont possibles pour le chercheur étudiant textes et images : d'une part l'analyse de la source globale, d'autre part l'analyse d'un aspect spécifique sur un ensemble de sources. Cette méthodologie est similaire pour les deux types de sources, sans être toutefois identique.

L'étude d'une image médiévale doit se faire sur plusieurs niveaux : l'image comporte un thème iconographique, qui peut être représenté seul mais qu'il convient de saisir dans son ensemble, comme la rencontre entre Abraham et Melchisédech. Le motif, au contraire, est un élément qui ne peut être représenté seul, ici le double pouvoir, souvent inscrit dans des scènes bibliques²¹. Il ne s'agit pas de privilégier l'un ou l'autre aspect, mais davantage d'équilibrer ce que Jérôme Baschet a appelé l'analyse structurale, c'est-à-dire de l'œuvre en elle-même, et l'analyse sérielle, c'est-à-dire des thèmes recoupant les sources mais aussi des points précis²². Ces différences de niveaux de lecture se retrouvent dans les analyses textuelles sous les termes de *close* et de *distant reading*.

Le *distant reading* a été défini par Franco Moretti en 2011 : cette méthodologie consiste en une lecture non séquentielle qui permet de dégager des motifs offrant de comprendre un ou plusieurs textes par d'autres voies que la seule information linguistique dont est porteur

²¹ Jérôme Baschet, « Corpus d'images et analyse sérielle », dans *Les images dans l'Occident médiéval...*, p. 319-332, aux p. 323-324.

²² *Ibid.*

le texte²³. La visualisation de données sous forme de graphes ou de cartes est centrale ici. Le *distant reading* est l'opposé de ce que Franco Moretti nomme le *close reading*, soit la lecture séquentielle d'un passage ou d'une œuvre pour en comprendre les différents niveaux de signification²⁴, telle que nous la pratiquons au quotidien.

L'une des grandes critiques opposées au *distant reading* est de réduire la recherche à la mise en place d'un ensemble d'outils de visualisation sur des jeux de données, et surtout de considérer que le résultat serait exempt de toute subjectivité : le *distant reading* apporterait « la clarté de la confirmation empirique »²⁵ sur une matière inerte qu'est le texte. Shawna Ross rejoint ici Joseph Morsel face au problème de ne considérer la source historique que sous l'angle textuel, inerte, en dehors de son contexte de fabrication et, surtout, de conservation, « véritablement signifiant »²⁶ et trop souvent ignoré. Jérôme Baschet invite à bien tenir compte des risques de l'analyse sérielle soit non séquentielle, qui peut conduire à « un défaut de contextualisation de l'œuvre »²⁷. Ce rappel souligne l'importance de comprendre la source dans sa singularité et son contexte.

Cette présentation de deux systèmes d'analyse, respectivement pour le texte et l'image, développés chacun pour les besoins d'un support, ne doit jamais faire oublier à quel point l'articulation entre texte et image est essentielle, car cette étude s'inscrit dans la perspective définie par Jérôme Baschet : « au lieu de séparer, articule[r] texte et image, image et objet, corps sensible et corps signifiant, sens et effet »²⁸.

²³ Franco Moretti, *Graphs, Maps, Trees : Abstract Models for a Literary History*, Londres, 2005, p. 57.

²⁴ Stefan Jänicke, Greta Franzini, Gerik Scheuermann et al., « On close and distant reading in digital humanities: a survey and future challenges », 2015, en ligne, p. 2.

²⁵ « [Le chapitre « The Novel, History and Theory » de Franco Moretti révèle] one of the primary dangers of digital literary studies to be the adoption of an aggrandized, even hubristic attitude toward literature as so much inert stuff being pocked at. » Shawna Ross, « In praise of overstating the case: a review of Franco Moretti, *Distant Reading* (London: Verso, 2013) », dans *Digital Humanities Quarterly*, t. 8/1, 2014, en ligne : <http://www.digitalhumanities.org/dhq/vol/8/1/000171/000171.html>, paragraphe 5 (consulté le 25 mars 2021).

²⁶ J. Morsel, « Les sources... », p. 281.

²⁷ J. Baschet, « Corpus d'images et analyse sérielle », dans *Les images dans l'Occident médiéval...*, p. 323.

²⁸ J. Baschet, *L'iconographie médiévale...*, p. 348.

1. 2. Caractères internes des sources pour le projet

Les caractères internes des sources textuelles et iconographiques ainsi définis sont multiples et ne se situent pas au même niveau. Les caractères externes sont bien sûr indispensables et sont ajoutés dans la description de la source même.

Les caractères internes pour l'analyse textuelle reposent sur le vocabulaire associé au double pouvoir, à savoir les mots utilisés pour désigner le temporel et le spirituel, qu'il s'agisse de l'expression figée *rex et sacerdos* ou de variantes de vocabulaire, par exemple en latin comme *imperator et pontifex*, ou bien syntaxiques comme l'affirmation que *sacerdotium et regnum in unam convenere personam*²⁹.

L'analyse de la source iconographique dans le sujet a pour base la méthodologie de l'anthropologie historique sur les images. Afin de créer une « typologie élémentaire des modes de figuration »³⁰, il est opportun de préciser les caractères internes sur lesquels repose cette typologie, de façon à mieux en relever les variantes. Parmi les symboles, les plus essentiels et centraux sont la couronne, puis la tenue. Ces premiers critères mettent en lumière les variantes possibles : il existe des représentations de Melchisédech et du Prêtre Jean qui ne mettent en avant aucun des deux pouvoirs, mais ce sont là des exceptions.

L'exemple retenu ici est un folio issu d'un manuscrit conservé à la bibliothèque de Heidelberg : une *Biblia pauperum* composée en allemand, datant des années 1430 et produite dans le Sud de l'Allemagne. Sur le folio sélectionné, trois épisodes bibliques sont réunis pour mettre en évidence leur sens typologique : au centre, le dernier repas du Christ, à droite la manne tombant du Ciel lors de l'errance des Hébreux, et enfin à gauche la rencontre entre Abraham et Melchisédech. Le texte précise qu'il est un « roi [et un] prêtre du Dieu Très-Haut »³¹. Comparons maintenant ce texte à l'image

²⁹ « Orationes in Consistoria Habitaie », Acta Regni Karoli IV 1346 Nr. 100, MGH, Const. 8, 1910-1926, à la p. 155.

³⁰ J. Baschet, « Corpus d'images... », p. 325.

³¹ « Là il [Abraham, ndla] rencontra le roi Melchisédech qui avait été victime du pillage, et celui-ci apporta pain et vin en sacrifice à Abraham, car le roi était le prêtre du Dieu Très-Haut. » Texte transcrit et traduit par l'auteure.



Fig. 3 | Anonyme, Abraham et Melchisédech, *Biblia Pauperum*, enluminure, 1430, Heidelberg, Universitätsbibliothek, Cod. pal. germ. 148, fol. 65.

(fig. 3) : face à Abraham en armure, c'est un souverain barbu, en habit royal laïc, qui accueille le vainqueur. Aucun symbole ne permet de comprendre son rôle spirituel. Melchisédech demande sa part du butin et laisse le reste à Abraham³², comme s'il ne s'agissait que d'une rencontre entre deux chefs de guerre.

La différence d'interprétation entre le texte et l'image est d'autant plus frappante que le manuscrit portait la finalité d'aider les prêtres à mieux enseigner les rudiments de la foi à leurs fidèles. Il n'existe pas de continuité directe entre le texte et l'image : le texte parle bien du pouvoir temporel, et l'image représente bien un roi, mais la mention du pouvoir sacerdotal de Melchisédech ne trouve aucun écho dans l'image. Ce constat peut être replacé dans le contexte des années 1430, celui de la crise du pouvoir pontifical ouverte par les

³² J. Baschet, « Corpus d'images... », p. 331.

assemblées conciliaires. Remis dans une perspective sérielle, celle des représentations de la rencontre entre Abraham et Melchisédech dans les ouvrages présentant l'histoire typologique de la Bible, on constate que ce choix d'un roi-prêtre uniquement roi est rare³³.

III. La mise en œuvre numérique

1. Le choix d'un encodage en TEI

1.1. Souplesse et richesse de la TEI

La TEI, *Text Encoding Initiative* est un projet pour la communauté scientifique en sciences humaines et sociales, qui propose un standard de balisage en XML pour l'encodage des sources. Nombreux sont les travaux d'édition numérique qui utilisent ses recommandations. Ce standard a été adopté dans la recherche, et reste le meilleur choix pour encoder des textes. Certaines équipes ont utilisé le module propre des manuscrits, *msdescription*³⁴, pour encoder d'autres types de sources, du moment qu'elles contiennent un texte³⁵, par exemple des inscriptions antiques sur marbre³⁶.

La TEI est assez souple pour permettre de ne pas encoder un seul texte par document .xml, mais plusieurs. L'encodage du texte et de l'image tirés de la source permet ainsi de constituer une base de données où sont rassemblés les deux *corpora*, sans les fusionner pour autant. Chaque entrée a un identifiant, composé des initiales du lieu de conservation et de sa cote (par exemple « HD_Cod.pal.germ.148 » pour l'exemple étudié plus haut). À l'intérieur de chaque entrée, des identifiants uniques sont attribués à la partie textuelle (HD_Cod.

³³ On trouvera un exemple similaire dans la même bibliothèque, Universitätsbibliothek Heidelberg, Cod. Pal. germ 432, fol. 22, en ligne : <https://digi.ub.uni-heidelberg.de/diglit/cpg432/0051> (consulté le 22 mars 2021).

³⁴ <https://tei-c.org/release/doc/tei-p5-doc/en/html/MS.html> (consulté le 22 mars 2021).

³⁵ « Text-Bearing Object », en ligne : <http://www.stoa.org/epidoc/gl/latest/supp-history.html> (consulté le 22 mars 2021).

³⁶ Voir par exemple le projet Aphrodisias, en ligne : <http://insaph.kcl.ac.uk/iaph2007/index.html> (consulté le 22 mars 2021).

pal.germ.148__texte) et à la partie iconographique (HD__Cod.pal.germ.148__img), qui sont placés au même niveau dans l'arborescence XML-TEI.

1. 2. Le défi d'encoder l'image médiévale en TEI

L'encodage de l'image, peu répandu, pose le défi de la constitution de cette base de données. Ce choix d'un encodage précis en XML-TEI peut sembler étonnant. L'exemple proposé par Séverine Lepape³⁷ représente une étude sérielle d'images sur plusieurs caractères en formalisation mathématique des modalités³⁸ sous la forme d'un tableur Excel. Elle a retenu trois modalités tirées des caractères relevés et exprimés en booléen (1 pour vrai, 0 pour faux) : la présence de la Vierge à l'Enfant, le support manuscrit et enfin la présence du cœur de la tige surgie du corps de Jessé. Le tableur résultant de ces trois modalités, appliquées ensemble, sert de base aux analyses quantitatives présentées ensuite.

Notre choix de préférer un encodage en XML-TEI repose sur plusieurs considérations. La première est la possibilité de décrire en profondeur les symboles étudiés dans notre sujet de thèse, comme le préconise Séverine Lepape³⁹, ici sous une forme multi-scalaire. Nous allons du général (la position par exemple) au particulier (comme la position de chaque personnage). Ce choix implique de remplacer une liste de mots clefs, qui décrivent en fait différents niveaux d'une image, par une description structurée et complète en arborescence.

Un autre avantage, central pour notre démarche, consiste à pouvoir encoder côte à côte le contenu textuel et le contenu iconographique du manuscrit retenu. Ce double encodage permet de préserver, au sein de la base de données, une forme de lecture séquentielle entre le texte et l'image, tout en ouvrant la possibilité d'une lecture non séquentielle sur le long terme. Ce choix permet à terme de comparer la relation texte-image de manière quantitative : peut-on de la sorte

³⁷ S. Lepape, « Formalisation et analyse statistique... » p. 333-335.

³⁸ Les modalités désignent les « réponses observables » depuis les caractères retenus, en statistiques. *Ibid.*, p. 334.

³⁹ *Ibid.*, p. 333.

percevoir une dynamique de cooccurrences textuelle et iconographique sur le plan chronologique ?

Enfin, cette description fine de l'image peut être interrogée par le très puissant langage Xpath, qui permet la sélection précise de données, en descendant ou en remontant dans l'arborescence de nœuds. Des balises très différentes peuvent ainsi être recherchées ensemble, pour répondre à des questions appliquées aux sources. Par exemple, Abraham est-il davantage agenouillé devant un Melchisédech évêque ou un Melchisédech roi ?

```
<figure corresp="#Melchi">
  <head>
    <title xml:lang="deu" corresp="#Melchi">Melchisédech</title>
  </head>
  <figure type="position" subtype="standing">
    <figDesc>Melchisédech se tient debout, face à
    <persName ref="#Abr">Abraham</persName>
    </figDesc>
  </figure>
  <figure type="clothing">
    <figDesc>Vêtement rouge brodé d'or, couvert d'un manteau de fourure
    bleu bordé de blanc.</figDesc>
    <figure type="length" subtype="mid"/>
    <figure type="power">
      <figure type="temporal" subtype="royal"/>
    </figure>
    <figure type="headgear">
      <figure type="temporal" subtype="royalcrown"/>
    </figure>
  </figure>
</figure>
```

L'encodage de l'image montre bien cette souplesse et cette richesse : l'unique balise <figure> est utilisée pour proposer un encodage structuré des images grâce à une arborescence précise. Elle se situe sur deux nœuds de même niveau : le personnage et l'arrière-plan. Au sein de l'arborescence du personnage sont décrits par exemple ses habits, son couvre-chef, sa position et celle de ses mains. Si un groupe l'accompagne, il sera décrit à ce niveau. Dans chaque balise <figure>, il est possible d'ajouter des commentaires avec <figDesc>.

L'encodage des types de pouvoir s'est révélé plus difficile, car toutes les images ne présentent pas de couronne assez caractérisée pour distinguer le pouvoir signifié par ce symbole. Le compromis

a consisté à utiliser l'attribut @cert (*certainty*) quand le doute était présent. Ces pouvoirs sont distingués selon plusieurs catégories : noble, royal, impérial, seigneur de guerre d'un côté, et sacerdotal, épiscopal, papal et monacal de l'autre. Dans les deux cas, une valeur « non identifié » est ajoutée afin de ne pas surinterpréter un symbole en cas de doute.

Les inconvénients demeurent, mais peu nombreux. Le principal risque est de nature scientifique, à savoir de réduire l'analyse à des critères prédéfinis pour ne plus ensuite s'en éloigner. Mais il est aisé de modifier le schéma pour ajouter ou retrancher des éléments d'analyse sans pour autant compromettre l'intégralité de la base de données : nous avons ainsi pu faire évoluer cette base de données en même temps que notre réflexion sur le *rex et sacerdos* tout au long de la thèse.

Le travail informatique commence par un travail scientifique sur le sujet étudié pour éviter les écueils aboutissant à un mauvais traitement statistique.

2. L'exploitation future des données

1.1. Vers une analyse quantitative ?

Les méthodes quantitatives employées en histoire existaient dès les années 1970, mais la mauvaise compréhension des biais statistiques et les anachronismes de catégorisation ont fini par tarir cette démarche au profit d'un retour à la microhistoire⁴⁰.

Un autre écueil réside dans le manque de précision de certains caractères externes ou métadonnées, pour la période médiévale : la date exacte peut être inconnue au profit d'un siècle entier, tout comme ce peut être le cas pour le lieu d'origine. Ce défaut peut conduire à un manque d'homogénéité des *corpora* que la logique de classement des fonds ne permet pas de résoudre, alors que ce classement se révèle d'une grande aide pour les chartes conservées dans

⁴⁰ Claire Lemerrier et Claire Zalc, *Méthodes quantitatives pour l'historien*, Paris, 2008, p. 13.

des cartulaires⁴¹. Un découpage chronologique précis en « tranches » de trente à soixante ans, comme l'a opéré Séverine Lepape, serait difficile sans un travail préparatoire de datation des sources, en particulier pour les lettres du Prêtre Jean⁴².

Malgré ces possibilités, l'analyse factorielle de la relation texte et image a encore été peu exploitée pour les sources médiévales. L'un des meilleurs exemples d'une telle démarche reste le travail de Jean-Baptiste Camps sur la relation entre texte et image dans les chansons de troubadours, qui a mis en évidence la différenciation iconographique entre pauvres chevaliers, qui se poursuit dans une gradation visuelle des rangs de la noblesse⁴³. Ces études, trop peu nombreuses, requièrent une interdisciplinarité encore peu répandue.

1. 2. Possibilités de visualisation : Gephi

La visualisation des données a connu de grands changements au cours de ces dernières années, comme l'atteste Shinya Saito en 2010 : l'abondance d'informations a conduit à un plus grand besoin d'outils de visualisation pour mieux les étudier⁴⁴. Des outils tels que les camemberts ou les bâtons de Word ou LibreOffice limitent les modalités d'étude (fig. 4), mais ils permettent de visualiser des statistiques générales, par exemple pour cerner l'homogénéité d'un

⁴¹ Nicolas Perreaux, « L'écriture du monde (I). Les chartes et les édifices comme vecteurs de la dynamique sociale dans l'Europe médiévale (VII^e-milieu du XIV^e siècle) », dans *Bulletin du Centre d'études médiévales d'Auxerre*, t. 19/2, 2015 (consulté le 25 mars 2021).

⁴² Bettina Wagner, *Die « Epistola presbiteri Johannis » lateinisch und deutsch: Überlieferung, Textgeschichte, Rezeption und Übertragung im Mittelalter, mit bisher unedierten Texten*, Tübingen, 2000.

⁴³ Jean-Baptiste Camps, « Troubadours et analyses factorielles : approches statistiques à la représentation de l'auteur dans les chansonniers occitans A, I et K », dans *Nouvelles recherches en domaine occitan : approches interdisciplinaires. Colloque de l'Association internationale d'études occitanes, Albi, 11-12 juin 2009*, Turnhout, 2016.

⁴⁴ Shinya Saito, Shin Ohno, Mitsuyuki Inaba : A Platform for Cultural Information Visualization Using Schematic Expressions of Cube, dans *Proceedings of the Digital Humanities*, 2010, en ligne : <http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-796.html> (consulté le 25 mars 2021).

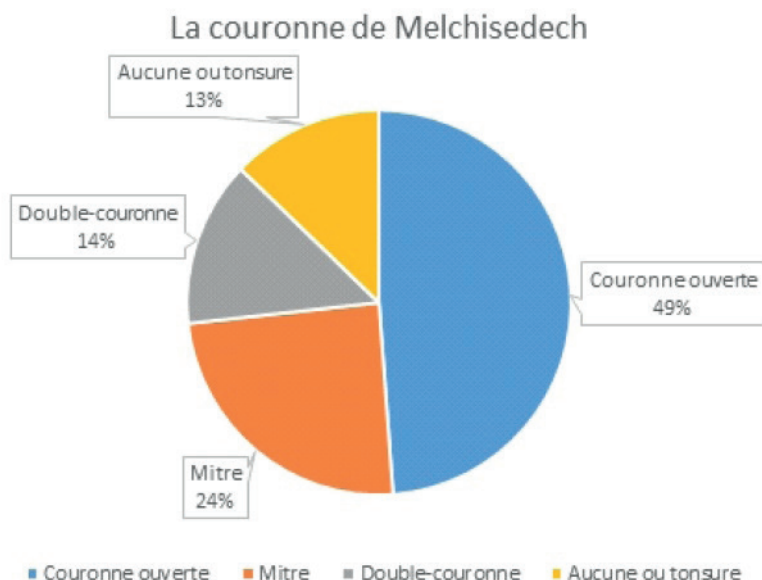


Fig. 4 | Marie-Astrid Hugel, « Le critère "double-couronne", représentation sous forme de camembert », dans *Pensée, représentation et imaginatio autour du double-pouvoir : la figure du roi-prêtre Melchisédech dans la culture médiévale du XIV^e et du XV^e siècle*, Paris, 2015, p. 383.

corpus⁴⁵. Séverine Lepape propose quant à elle une visualisation sous forme de nuage de points sur un plan factoriel – un point par objet étudié – permettant de croiser les modalités Christ et Vierge selon la modalité temps : en conséquence, plus les points sont rapprochés, plus la ressemblance entre les modalités étudiées est grande, ce qui permet de déterminer des tendances propres aux représentations de l'Arbre de Jesse⁴⁶.

Mais les outils de visualisation développés pour représenter des réseaux ou des nuages complexes ont été rendus accessibles, et l'un

⁴⁵ Jean-Baptiste Camps et Florian Cafiero, « Setting bounds in a homogeneous corpus: a methodological study applied to medieval literature », dans *Revue des Nouvelles Technologies de l'Information*, t. SHS-1, 2012, en ligne : <https://halshs.archives-ouvertes.fr/halshs-00765651/document> (consulté le 11 octobre 2021).

⁴⁶ Séverine Lepape, « Formalisation... », p. 346.

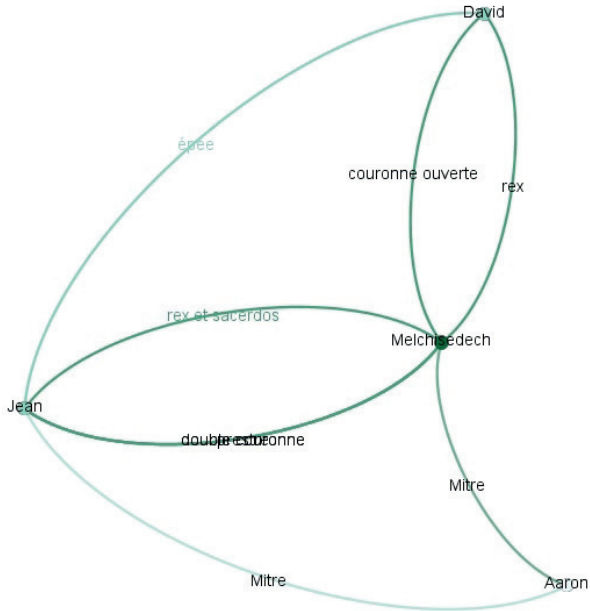


Fig. 5 | Visualisation des relations figures-critères avec le logiciel Gephi.

des logiciels les plus utilisés actuellement est Gephi, développé par l'université de Compiègne en 2008.

Ce logiciel ouvert⁴⁷ permet des visualisations de réseaux complexes et multicritères. Les nœuds et les arêtes qui lient ces nœuds entre eux sont les deux éléments visuels. Les nœuds sont ici des figures représentées dans le corpus. Les arêtes liant ces nœuds représentent le vocabulaire commun à ces figures : plus l'arête est foncée, plus l'indice de pondération de l'arête est élevé, et plus la ressemblance de vocabulaire utilisé pour décrire ces figures est élevée (fig. 5). Il est aussi possible de transformer ce réseau statique en un réseau dynamique⁴⁸ par l'entrée des dates dans le laboratoire de données du

⁴⁷ On trouvera l'excellent tutoriel de Martin Grandjean traduit pour Data Hist sur le lien suivant : <https://datahist.hypotheses.org/13> (consulté le 22 mars 2021).

⁴⁸ Un tutoriel sur github : <https://seinecle.github.io/gephi-tutorials/generated-html/convertng-a-network-with-dates-into-dynamic.html> (consulté le 25 mars 2021).

logiciel. Cette visualisation est aisée à déchiffrer, mais peut tromper : Florian Cafiero parle du « syndrome du Petit Prince » pour désigner la tendance à refaire une visualisation jusqu'à ce qu'elle ait la forme désirée⁴⁹.

IV. Conclusion

Ce sujet de thèse n'aurait pas pu voir le jour sans l'utilisation des outils informatiques et le standard de la TEI. L'accessibilité des méthodologies statistiques, adaptées aux sciences humaines pour infirmer ou confirmer des hypothèses de lecture non séquentielle sur le corpus, est un progrès certain. L'un des risques majeurs, visible tout au long de l'article, est de n'utiliser ces outils que de façon confirmatoire et non par exploratoire : il faut aussi savoir analyser ses données avec aussi peu d'*a priori* que possible, en travaillant des modalités que nous n'aurions pas sélectionnées dans un seul esprit de confirmation⁵⁰.

Mais face à ces outils de travail qui ont renforcé des méthodologies quantitatives d'analyse, le risque de perdre le contact avec l'objet « source » est encore plus grand, comme le rappellent Joseph Morsel et Jérôme Baschet. Le gouffre se creuse encore quand les sources rassemblées dans le corpus ou les *corpora* sont réduits à leur textualité ou médialité par extrait inerte, arrachées à leur contexte de conservation et ne devenant qu'une simple entrée parmi d'autres dans une base de données. Des institutions comme la Bibliothèque nationale de France compensent ce manque de matérialité et de contextualité en donnant de nombreuses précisions sur le fonds contenant le manuscrit et son aspect physique. Mais ce n'est pas le cas de tous les institutions patrimoniales. L'un des défis pour les chercheurs du XXI^e siècle est de savoir gérer intelligemment ces méthodes inédites de recherche,

⁴⁹ Florian Cafiero, « Analyse de réseaux (3) : s'il vous plaît... dessine-moi un réseau », dans *Sacré Graal : histoire, philologie, programmation et statistiques*, carnet de recherche, en ligne : <https://graal.hypotheses.org/716> (consulté le 26 mars 2021).

⁵⁰ Un exemple de méthode exploratoire sur un corpus médiéval : J.-B. Camps et F. Cafiero, « Setting bounds... ».

de s'ouvrir à de nouveaux sujets jusque-là peu adaptés aux méthodes traditionnelles, tout en sachant revenir à la source de son travail, le produit « source » ou l'image-objet. Le cœur de ce travail, la comparaison texte-image sur une question centrale du pouvoir médiéval, en combinant des lectures séquentielles et non séquentielles, promet d'offrir de nouvelles perspectives à même d'enrichir des études séquentielles de sources individuelles parlant de *rex et sacerdos*.

MARIE-ASTRID HUGEL

Diplômée de master à l'EHESS et l'université de Heidelberg
et du master TNAH de l'École nationale des chartes