

ÉCOLE NATIONALE DES CHARTES
UNIVERSITÉ PARIS, SCIENCES & LETTRES

Théo Burnel
docteur en histoire moderne

Exploration visuelle d'une base de données : apports de la datavisualisation à l'interprétation des analyses physico-chimiques des enluminures

Carte interactive, frise chronologique et facs-similés numériques

Mémoire pour le diplôme de master
« Technologies numériques appliquées à l'histoire »

2024

Résumé

Ce mémoire propose de revenir sur les trois visualisations produites autour des projets « La couleur : artefacts, matière et cognition » de Charlotte Denoël et « La fabrique matérielle du visuel : panneaux peints en Méditerranée » de Sigrid Mirabaud. Une cartographie interactive (avec Leaflet) et une frise chronologique (avec Vikus) proposent une mise en ordre du temps et de l'espace selon l'histoire des enluminures des manuscrits et des panneaux peints. Une édition de fac-similés numériques (avec des manifestes IIIF) contribue à la communication et à la médiation de la recherche.

Mots-clés : visualisation ; carte interactive ; frise chronologique ; fac-similé numérique

Informations bibliographiques : Théo Burnel, *Exploration visuelle d'une base de données : apports de la datavisualisation à l'interprétation des analyses physico-chimiques des enluminures. Carte interactive, frise chronologique et facs-similés numériques*, mémoire de master « Technologies numériques appliquées à l'histoire », dir. Chloé Pochon, École nationale des chartes, 2024.

Remerciements

Mes premiers remerciements vont à Madame Charlotte Denoël pour m'avoir accueilli au sein du département des manuscrits de la Bibliothèque nationale de France pendant la durée de ce stage. Je tiens à la remercier publiquement de m'avoir permis de réaliser ce travail dans des conditions si agréables et enrichissantes.

Je tiens également à remercier Madame Chloé Pochon d'avoir accepté de diriger ce mémoire et de m'avoir proposé des solutions techniques au cours de ces trois mois.

Enfin, je souhaite exprimer ma gratitude à Madame Emmanuelle Bermès pour m'avoir accepté dans son Master et pour son suivi tout au long de cette année.

Bibliographie

- AGAFONKIN (Volodymyr), *Leaflet*, en, URL : <https://leafletjs.com/> (visité le 26/07/2024).
- APPLIED SCIENCES POTSDAM (University of), *VIKUS Viewer : Explore cultural collections along time, texture and themes*, en, URL : <https://vikusviewer.fh-potsdam.de/> (visité le 17/06/2024).
- *Visualizing Cultural Collections*, URL : <https://uclab.fh-potsdam.de/vikus/> (visité le 26/07/2024).
- ArcGIS Online*, URL : <https://www.arcgis.com/index.html> (visité le 26/07/2024).
- BIBLISSIMA, *Qu'est-ce que IIIF?*, URL : <https://iiif.biblissima.fr/> (visité le 04/07/2024).
- CERTEAU (Michel de), *L'écriture de l'histoire*, Paris, 1975.
- CULTURE (Ministère chargé de la), *culturecommunication/ginco*, original-date : 2013-01-25T16:31:32Z, nov. 2022, URL : <https://github.com/culturecommunication/ginco> (visité le 04/07/2024).
- D3 / The JavaScript library for bespoke data visualization*, URL : <https://d3js.org/> (visité le 26/07/2024).
- Charlotte Denoël (éd.), *Le Beatus de Saint-Sever*, Édition originale, tirage limité et numéroté, Paris, 2022 (BnF éditions).
- GAMESON (Richard), *The pigments of British medieval illuminators : a scientific and cultural study*, London, 2023.
- Git Graph - Visual Studio Marketplace*, en-us, URL : <https://marketplace.visualstudio.com/items?itemName=mhutchie.git-graph> (visité le 26/07/2024).
- GitLens — Git supercharged - Visual Studio Marketplace*, en-us, URL : <https://marketplace.visualstudio.com/items?itemName=eamodio.gitlens> (visité le 26/07/2024).
- Google My Maps*, URL : <https://www.google.com/maps/d/> (visité le 26/07/2024).
- INHA, *Thésaurus et vocabulaires contrôlés de l'Institut national d'histoire de l'art*, URL : <https://thesaurus.inha.fr/thesaurus/> (visité le 04/07/2024).
- International Image Interoperability Framework*, en, juin 2024, URL : <https://iiif.io/> (visité le 26/07/2024).
- LÉNART (Michèle), « SKOS, un langage de représentation de schémas de concepts », *Documentaliste-Sciences de l'Information*, 44–1 (2007), Place : Paris Publisher : A.D.B.S., p. 75-75, DOI : 10.3917/docsi.441.0075.

- Live Server - Visual Studio Marketplace*, en-us, URL : <https://marketplace.visualstudio.com/items?itemName=ritwickdey.LiveServer> (visité le 26/07/2024).
- Mapbox / Maps, Navigation, Search, and Data*, URL : <https://www.mapbox.com/> (visité le 26/07/2024).
- Omeka S*, URL : <https://omeka.org/s/> (visité le 26/07/2024).
- ONLINE (Conservation & Art Materials Encyclopedia), *Pigment*, en, URL : <https://cameo.mfa.org/wiki/Pigment> (visité le 17/06/2024).
- OpenLayers*, URL : <https://openlayers.org/> (visité le 26/07/2024).
- PASTOUREAU (Michel) et SIMONNET (Dominique), *Le petit livre des couleurs*, Paris, 2014 (Points, H377).
- POCHON (Chloé), *AGORHA : Plateforme de données de la recherche de l'Institut national d'histoire de l'art*.
- ROBSON (Glen), *IIIF Online Workshop*, URL : <https://training.iiif.io/iiif-online-workshop/> (visité le 04/07/2024).
- ROGER (Patricia), « Étude des couleurs et de la pratique picturale », *Art de l'enluminure L'Évangéliaire de Charlemagne*-20 (2007), p. 46-65, URL : <https://www.faton.fr/art-de-lenluminure/numero-20/l-evangeliaire-charlemagne.2451.php> (visité le 02/04/2024).
- THACH (Thiso), *Qu'est-ce qu'un Iframe et comment l'intégrer dans une page ?*, fr, fr, URL : <https://blog.hubspot.fr/website/iframe> (visité le 02/07/2024).
- TOUBKIS (Dimitri), « L'ordre de la chronologie », *Hypothèses*, 7-1 (2004), Place : Paris Publisher : Éditions de la Sorbonne, p. 133-145, DOI : 10.3917/hyp.031.0133.
- Visual Studio Code - Code Editing. Redefined*, en, URL : <https://code.visualstudio.com/> (visité le 26/07/2024).
- Web Annotation Data Model*, en, févr. 2017, URL : <https://www.w3.org/TR/annotation-model/> (visité le 26/07/2024).

Introduction

I. L'exemple du bleu égyptien, ou ce que l'investigation physico-chimique apporte à l'histoire de l'art

Les couleurs connaissent des trajectoires, changeantes avec les époques, les lieux et les découvertes. Elles témoignent des connaissances d'une société, des modes d'un milieu, des besoins, aussi, des artistes et des commanditaires. Les couleurs sont un reflet des savoirs techniques de leur création, des échanges commerciaux pour leur acheminement et d'une interprétation sémantique de leur usage, de la force symbolique de leur choix. La « fabrique visuelle » d'une enluminure n'est pas qu'une création artistique, à moins de considérer cette dernière comme le miroir de ce qui précède¹.

L'histoire de l'utilisation du bleu d'Alexandrie, communément appelé « bleu égyptien », semble être un parfait exemple de ce que la présence – ou l'absence – d'une couleur peut nous apprendre sur la localisation et la datation des œuvres, sur les pratiques d'un artiste ou d'un atelier, sur les différentes interventions au sein d'une même œuvre, sur les circuits commerciaux et la circulation des pigments ou encore sur les transferts techniques entre les cultures.

Le bleu ne se résume pas à une couleur, mais à une multitude de teintes et de nuances, tantôt lumineuses, tantôt sombres. Parmi les variétés de bleu, l'égyptien se démarque par son origine : il est le plus ancien pigment de synthèse connu. Charlotte Denoël, dans l'édition critique du *Beatus de Saint-Sever*, retrace son histoire :

« Apparu en Égypte sous la IV^e dynastie, vers 2620 av. J.-C., et très employé en peinture, notamment sur les portraits du Fayoum, ce composé de silicate double de cuivre et de calcium était obtenu artificiellement par cuisson à 850°C dans des conditions très contrôlées, avec un savoir-faire élaboré, le plus souvent au contact de métallurgistes. C'est pourquoi ce pigment est longtemps passé pour disparu avec la fin du monde antique, sa toute dernière occurrence apparaissant dans le célèbre Pentateuque d'Ashburnham exécuté aux environs de 600, probablement à Rome. De récentes analyses ont cepen-

1. Pour une introduction à la trajectoire des couleurs travers le temps, voir Michel Pastoureau et Dominique Simonnet, *Le petit livre des couleurs*, Paris, 2014 (Points, H377).

dant permis d'attester la présence du bleu égyptien dans quelques peintures murales et manuscrits enluminés du premier Moyen Âge, dont l'Évangéliaire de Charlemagne, daté de 781-783. Son identification dans le Beatus vient documenter de façon remarquable la permanence de l'utilisation de ce pigment au XIe siècle, en même temps qu'elle souligne l'extraordinaire niveau de technicité des artistes de Saint-Sever.² »

Des investigations physico-chimiques permettent ainsi de retracer et de comprendre des conditions matérielles et un contexte iconographique d'utilisation de ce pigment. Le bleu égyptien s'inscrit dans une chronologie et une géographie qui, par l'analyse d'un corpus de manuscrits, s'éclaircissent progressivement. Il en est du bleu égyptien comme d'autres couleurs et le projet de recherche sur la « La couleur : artefacts, matière et cognition », qui a obtenu un financement dans le cadre de l'appel 2020-23, permet de proposer aux chercheurs en sciences des matériaux du patrimoine, aux historiens ou aux historiens de l'art des informations précieuses sur les enluminures des manuscrits.

II. Une recherche scientifique et culturelle

Le cas du bleu égyptien semble représentatif des nouvelles recherches qui sont menées sur les enluminures médiévales depuis une dizaine d'années maintenant. Plus que d'autres recherches, les travaux sur les manuscrits anciens lient, avec bénéfice, les différents apports de l'interdisciplinarité – qui dépassent ceux des sciences humaines. La recherche sur la couleur, la matière et la technique utilisées dans la composition des enluminures permet d'enrichir scientifiquement et culturellement nos connaissances sur ce monde passé et encore préservé, où chefs-d'œuvre textuels et picturaux cohabitent.³

La découverte est dans un premier temps scientifique. Elle s'appuie sur les dernières techniques mises au point dans l'investigation des manuscrits – des techniques qui, à la différence d'un passé récent encore, sont désormais non invasives et permettent d'explorer les manuscrits les plus précieux et en plus grand nombre. Il n'est pas lieu ici de présenter les différentes techniques qui permettent la réalisation des analyses physico-chimiques ; les personnes intéressées pourront se référer au chapitre dédié dans l'ouvrage *The pigments of british medieval illuminators*⁴. À titre d'exemple, dans le cas des bleus rencontrés dans

2. Charlotte Denoël (éd.), *Le Beatus de Saint-Sever*, Édition originale, tirage limité et numéroté, Paris, 2022 (BnF éditions). Pour son utilisation dans l'évangéliaire de Charlemagne, voir : Patricia Roger, « Étude des couleurs et de la pratique picturale », *Art de l'enluminure L'Évangéliaire de Charlemagne-20* (2007), p. 46-65, URL : <https://www.faton.fr/art-de-lenluminure/numero-20/1-evangeliaire-charlemagne.2451.php> (visité le 02/04/2024). Pour une présentation des l'analyse physico-chimique du bleu : Richard Gameson, *The pigments of British medieval illuminators : a scientific and cultural study*, London, 2023.

3. La création en 2022 d'un musée à la Bibliothèque nationale de France sur le site de Richelieu participe à la connaissance du grand public des trésors de la bibliothèque, dont les manuscrits occupent la première place.

4. Pp. 1-42. *Ibid.*

le *Beatus de Saint-Server*, trois technologies sont combinées afin de produire une analyse ponctuelle, une cartographie de la surface des feuillets et une de ses différentes couches :

« [...] la tomographie en cohérence optique, pour rendre compte de la structure des couches colorées et de leurs épaisseurs ; l'imagerie hyperspectrale dans le visible et dans l'infrarouge, pour accéder à la réflectance des matériaux et identifier certaines espèces colorantes ; la cartographie de fluorescence X et UV pour rendre compte de l'analyse élémentaire composant les matériaux et renseigner certains composés de notre organique ou inorganique⁵ ».

Ces analyses permettent de mettre en avant différentes teintes, nuances et techniques autour de la couleur bleu. Dans le seul cas de la représentation du paon et du serpent dans le *Beatus de Saint-Server*, quatre bleus sont ainsi identifiés : l'indigo, l'azurite, l'outremer et, enfin, le bleu égyptien⁶. La découverte devient aussi culturelle, tant l'analyse apporte des éléments sur nos connaissances des colorants, des matériaux, des enlumineurs et des manuscrits d'une époque. L'exploration d'un corpus aussi important que celui réalisé dans le cadre du projet « La couleur : artefacts, matière et cognition » permet d'interroger à nouveau frais la circulation des matières premières essentielles à l'enlumination, la sémantique des couleurs ou encore la localisation et la temporalité des productions artistiques.

Les analogies, les rapprochements sont essentiels à l'heuristique dans le cadre d'une histoire transculturelle sur plus d'un millénaire et marquée par une pauvreté des sources primaires comme il en est ici. Mais, sans rien ôter à la singularité de chaque manuscrit, en témoigne le regard porté sur les artistes eux-mêmes, la comparaison des résultats met en avant des transferts ou des influences culturelles qui peuvent représenter de nouvelles interprétations et analyses à la réalisation des enluminures sur un temps long. Les questions sont nombreuses et quelques exemples peuvent illustrer de la richesse de la recherche : comment apparaît et se diffuse l'utilisation du vert de cuivre ? le sens d'un rouge ou d'un pourpre dans une enluminure change-t-il au cours du temps ? le jaune est-il nécessairement doré pour le sacré ? Les réponses à ces différentes interrogations trouvent leur élucidations dans la comparaison des occurrences du thésaurus constitué et dans son architecture⁷. Un répertoire structuré de termes sur des critères comme les matières colorantes, les objets, les supports, les centre de production, les époques et les aires géographiques est essentiel au comparatisme. Les recherches conduites dans le cadre du projet, sur près de cinq milles couches matérielles, permettent dès lors l'élaboration d'un corpus quantitatif, qui vient s'ajouter à l'existence des données qualitatives obtenues par les analyses. Ces données, complexes et disséminées dans l'immensité de la recherche effectuée, peuvent être utilisées, avec bénéfice, dans le cadre d'une projet de visualisation.

5. Pp. 78., *Le Beatus de Saint-Sever...*, Ces analyses ont été réalisées par le centre de recherche et de restauration des musées de France et le Centre de recherche sur la conservation.

6. Cf. annexe 1.

7. Je reviendrai plus longuement sur l'importance du thésaurus dans la première partie de ce mémoire.

III. La visualisation : la représentation au service de l'interprétation

Les apports de la conduite d'un projet de visualisation pour accompagner la recherche scientifique sont aujourd'hui suffisamment reconnus pour qu'il ne faille pas à nouveau les expliciter et les justifier ici. Mais plus que des apports, ces projets sont en passe de devenir des conditions *sine qua non* à l'obtention de contrats de recherche. La visualisation, définie comme une représentation de données, est conçue pour aider à comprendre, interpréter et communiquer des informations et des résultats de manière efficace. La visualisation n'est pas seulement la dernière séquence d'un projet scientifique qui, d'une certaine façon, matérialiserait la recherche menée – un faire-valoir des résultats ; elle est essentielle à tous les niveaux de la recherche, depuis la collecte et l'analyse des données jusqu'à la communication des résultats. Dans le cadre d'un projet comme celui qui m'intéresse dans le cas présent, il me semble que la conduite d'un projet de visualisation comporte au moins trois avantages.

1. Elle participe à la compréhension des données. Les visualisations permettent aux chercheurs d'interpréter rapidement et intuitivement des données initialement complexes, parfois illisibles – illisibles dans le sens où elles nécessitent de parcourir et de retenir près de cinq mille lignes pour en comprendre le sens. En représentant des relations entre des variables, comme par exemple des années, des couleurs ou des matériaux, des tendances et des modèles peuvent être identifiés plus facilement. Les analyses peuvent ainsi être multipliées en des temps réduits. Les visualisations permettent aux chercheurs de repérer d'autant plus facilement des tendances lorsqu'une option d'éditorialisation des jeux de données est mise en place⁸. Les chercheurs peuvent alors explorer les données sous différents angles et identifier des associations qui pourraient ne pas être évidentes autrement.
2. La communication et la méditation des résultats sont facilitées. Les représentations rendent les résultats de la recherche plus accessibles et attrayants pour un public plus large, y compris les chercheurs dans d'autres domaines. L'aridité de la recherche scientifique, l'initiation préliminaire qu'elle demande, sont compensées par une clarté et une simplicité inhérente à la visualisation. L'exemple de la carte interactive qui sera présenté plus avant me semble parfaitement éclairant sur ce point : la concentration ou, au contraire, la dissémination des marqueurs, qui représentent les résultats du filtre dans les jeux de données, est parfaitement interprétable par un public qui ne serait pas familiarisé avec les manuscrits et les enluminures.

8. Les chercheurs ont la possibilité d'éditorialiser leur recherche sur la carte interactive et la frise chronologique.

3. Elle favorise la collaboration interdisciplinaire. Les visualisations peuvent servir de point de départ pour des discussions collaboratives entre chercheurs de différents domaines. En représentant visuellement les données, les chercheurs peuvent mieux comprendre et apprécier les contributions des autres membres de l'équipe, ce qui favorise la collaboration interdisciplinaire. Dans le cadre d'annotations sémantiques d'images à partir des technologies IIIF, il me semble ainsi que le dialogue entre l'historien de l'art, le conservateur et le scientifique réalisant l'analyse physico-chimique est renforcé et simplifié par la désignation précise de l'attendu et sa représentation.

En somme, la conduite d'un projet de visualisation me semble tenir un rôle primordial lors d'un projet de recherche. Elle permet d'interroger à nouveau frais les données récoltées et les résultats obtenus, elle facilite la communication et la médiation de la recherche, et elle peut instaurer un langage commun entre les cultures scientifiques diverses. Dès lors, la représentation des données ne saurait être reléguée au second plan, marginalisée ou conditionnée à un reliquat budgétaire. Elle est et doit être présente dans la réflexion du chercheur, au même titre que celle sur la communication écrite, qu'elle soit d'articles scientifiques ou de vulgarisation.

Afin d'appuyer cette hypothèse, ce mémoire est composé en quatre temps. Le premier chapitre propose un retour sur les préliminaires de la conduite d'un projet de visualisation. Il s'agit d'un chapitre introductif qui revient sur l'antécédent de la base de données qui sert de socle à la visualisation. Le deuxième et le troisième chapitres retracent les choix et la réalisation de l'éditorialisation des données de la base. Ils présentent les réflexions et les étapes de la construction d'une carte interactive et d'une frise chronologique, les deux représentations qui permettent d'ordonner les données selon un espace et un temps – en d'autres termes, de rendre intelligible un fichier CSV à plusieurs milliers de lignes. Ce mémoire se clôture par la présentation de la réalisation de fac-similés numériques. L'édition enrichie des manuscrits selon ce modèle permet une nouvelle communication de la recherche auprès du public, plus interactive et attractive.

Chapitre 1

Présentation du projet de recherche

Dans le cadre de la réponse à l'appel à projet pour le plan triennal de la recherche de la Bibliothèque nationale de France, le programme « La couleur : artefacts, matière et cognition » prévoit tout un volet d'éditorialisation à partir des données récoltées et analysées. L'éditorialisation doit permettre de « répondre aux différents besoins des communautés scientifiques travaillant sur la couleur et ses matériaux¹ ». Les données sont, pour leur part, renseignées dans une base qui est développée et hébergée par l'Institut national d'histoire de l'art (INHA). La base repose sur un modèle structuré : les données sont hiérarchisées et suivent des référentiels spécialisés. Le programme de recherche « La couleur : artefacts, matière et cognition » est également lié au programme de recherche de Sigrid Mirabaud sur « La fabrique matérielle du visuel : panneaux peints en Méditerranée, XIIIe-XVIe siècle ». Les visualisations envisagées et évoquées au cours de ce mémoire concernent les deux projets.

Ce chapitre revient sur les étapes qui précèdent et qui permettent l'éditorialisation des données. Il s'agit, d'une certaine façon, d'un chapitre préliminaire à la présentation du travail réalisé au cours de ce stage. Cependant, les pages qui suivent permettent de comprendre les chemins que pouvaient prendre le travail de visualisation à partir des données présentes.

1. Appel à projet « La couleur : artefacts, matière et cognition »

I. Les bases de donnée

Les programmes de recherche sur « La couleur : artefacts, matière et cognition » et sur « La fabrique matérielle du visuel : panneaux peints en Méditerranée » débutent en janvier 2021. Les premiers temps du projet sont consacrés à la constitution d'un corpus d'œuvres à soumettre aux analyses. Il s'agit principalement de collections du département des manuscrits pour le premier projet. Un peu plus de trois années après le lancement, en juin 2024, trois cent trente-six manuscrits et panneaux peints en ont fait l'objet. Le résultat est un ensemble de données, issues de rapports d'analyses physico-chimiques. Elles proviennent de différents laboratoires avec lesquels la Bibliothèque nationale de France a conclu des accords : le Centre de Recherche et de Restauration des Musées de France (C2RMF), l'Institut de recherche sur les archéomatériaux Centre Ernest-Babelon (IRAMAT-CEB), le laboratoire de recherche des monuments historiques (LRMH) ou encore le Centre de Recherche sur la Conservation (CRC). Depuis le début du projet, un carnet de recherche accompagne les travaux menés et leur assure une publicité et une diffusion auprès du public (<https://couleur.hypotheses.org>). Ce carnet permet également de refléter les croisements interdisciplinaires qui entourent le projet : histoire de l'art, histoire, archéologie, linguistique, anthropologie, mais aussi, nous l'avons vu, des disciplines issues des sciences dites « dures ». La réunion de chercheurs de traditions scientifiques différentes a produit de nouveaux thésaurus qui viennent enrichir la description des enluminures, que ce soit pour les matériaux ou les techniques. Aujourd'hui, le projet est à sa phase finale : les dernières données collectées sont intégrées dans Agorha et la migration vers deux sites Omeka S est prévue pour l'automne. Les visualisations réalisées dans le cadre de ce stage trouveront place sur ces derniers.

La base produite à l'issue des programmes de recherche contient des données de différentes natures sur les manuscrits et les panneaux peints.

1. Premièrement, elle contient des informations propres à chaque objet du corpus : l'identifiant, la cote ou l'inventaire, l'auteur, la date, l'origine ou encore les dimensions. Ces informations permettent d'identifier l'objet dont il est question, de le recenser au sein du corpus et de le situer pour le public. Ces métadonnées sont le point de départ des interprétations qui sont réalisées avec le résultat des analyses. Elles attachent ces dernières à un contexte et à un cadre de production auxquels il faut toujours se référer.
2. Ensuite, chaque entrée de la base de données est reliée à une image numérique. Cette image est une reproduction de l'objet analysé. Elle a une double utilité. Elle permet, dans un premier temps, à l'utilisateur de prendre connaissance de l'enluminure avec les métadonnées précédemment décrites. Ensuite, l'image numérique change de sens pour devenir un outil, pour un regard plus précis sur les

zones analysées et renseignées dans le point suivant.

3. En troisième lieu, la base contient toutes les analyses physico-chimiques et techniques réalisées au cours des deux projets de recherche. Ces données renseignent sur la localisation des analyses, avant de passer à la présentation des matériaux pour les couleurs observées : pigments, colorants, liants, vernis. Les données sont interprétées, datées et sourcées. Elles sont issues des thésaurus réalisés et sont ainsi normalisées.
4. Enfin, la base fournit une documentation complémentaire relative à la conservation et à la restauration, une bibliographie spécialisée sur les matériaux, des liens vers les manuscrits numérisés (Gallica) ou encore vers des corpus de sources textuelles sur les pigments (recettes médiévales, traités théoriques...).

Le travail d'éditorialisation des données des bases est réalisé à partir de fichiers CSV. Cependant, les données sont visibles pour le public sur le site Agorha pour le moment, puis sur les sites Omeka S à l'automne 2024. Cette exploration des données par une navigation internet fait partie prenante des projets et conditionne, aussi, les visualisations qui seront produites.

II. La plateforme Agorha

Mise en ligne en 2011, refondée en 2021, Agorha a aujourd’hui plus de 250 000 notices sur son site. Il donne un accès aux bases de données produites par l’INHA et ses partenaires, comme le département des manuscrits dans le cas présent. Ces données sont issues de programmes de recherche scientifique en histoire de l’art et archéologie. Agorha n’a pas donc pas vocation à l’exhaustivité, mais à proposer une sorte de bibliothèque numérique choisie à ses utilisateurs. Une bibliothèque de référence, qui témoigne d’une recherche numérique ouverte, d’une science partagée.

Les données produites sont organisées selon des notices décrivant des œuvres, des personnes et organismes, des références bibliographiques et d’archives, des collections et des événements. Ces notices sont enrichies collaborativement. La plateforme est pensée pour que les notices puissent être liées entre elles, indépendamment de leur type ou des programmes de recherche auxquels elles sont rattachées ; une même notice peut être ainsi reliée à différents programme de recherche. Les données sont également constituées de médias. Elles enrichissent les notices avec des fichiers, principalement d’images. Outre les liens faits entre les notices, celles-ci peuvent aussi être liées à des thésaurus et ainsi renvoyer à des référentiels de vocabulaire interne ou externe de l’INHA. Je reviendrai sur la question des termes contrôlés plus en avant.

Le 1er juillet 2024, Agorha compte 1136 notices pour le projet sur la couleur du département des manuscrits. Toutes ne sont pas encore publiées et disponibles en ligne à cette date, mais elles le seront d’ici l’automne. Ces notices sont majoritairement des œuvres (1013), mais elles portent aussi sur des références (93), des personnes (28) et des collections (2). Il n’y a pas de notice sur des événements. Le projet sur les panneaux a, à la même date, 762 notices : 227 œuvres, 5 personnes et 530 références.

La structuration d’une notice Agorha tient en deux temps pour l’utilisateur. Le premier est celui où il la découvre par son média – ici son image – et les liens qu’elle peut entretenir avec d’autres notices. Il s’agit d’un premier carrousel d’images. En bas de la page, une second carrousel permet d’afficher toutes les œuvres, toutes les personnes ou encore toutes les références liées à la notice en cours. *Voir la figure 1.*

Ensuite, la notice repose sur des champs de données structurées. Les informations sur les feuillets des manuscrits sont divisées entre plusieurs catégories, selon le besoin des projets. Pour le projet sur la couleur, huit champs de données sont retenus : l’identification, la localisation, la description (avec les analyses), la création, la qualification entre l’imprimé et le manuscrit, les liens entre les œuvres, la documentation et, pour finir, la gestion de la notice. Ces champs sont à leur tour divisés en de nouveaux qui contiennent les résultats des projets de recherche, soit sous la forme d’une entrée de thésaurus, soit d’un texte libre entré par la personne qui saisit. *Voir la figure 2.*

Présentation du projet de recherche



FIGURE 1.1 – Le média d'une notice Aghora

Localisation

Localisations

Lieu de conservation : **BnF - département des Manuscrits**

Type de Cote / numéro : **cote**

Cote / numéro : **Français 8**



Description

Restaurations



Matérialité

Type de description : **Couche colorée**
matérielle :

Commentaire Type identifiant : **I8F11RxrfP01 ; I8F11RrubyP01 - localisation : vêtement ; de description robe ; Caïn - couleur : rose**
matérielle :

Caractéristique de description : **Couche picturale** **Rose**
matérielle :

Matériau : **blanc de plomb** **colorant rouge**

Commentaire Matériau : **[blanc de plomb] ; certitude : observé ; technique : vidéomicroscope numérique ; date : 2021 - matériau : [blanc de plomb] ; certitude : caractérisé ; technique : XRF ; dimension : 1 mm2 ; date : 2021 - matériau : [indéterminé] ; certitude : indéterminé ; technique : spectrophotocolorimétrie ; dimension : 2 mm ; date : 2021**

Technique : **gouache** **enluminé** **mélange**

Commentaire Spectre de réflectance ne permet pas une identification.

Matérialité : **Blanc de plomb + colorant rouge**



FIGURE 1.2 – Les champs structurés sur Agorha

Les notices ont des identifiants uniques attribués automatiquement par le système informatique d'Agorha. L'identifiant est présent dans l'URL, c'est un UUID (Universally unique identifier). Pour retrouver une notice, il suffit ainsi de le renseigner à la fin de l'adresse d'Agorha :

<https://agorha.inha.fr/ark:/54721/> + UUID

En anticipant un peu sur la suite de ce mémoire, il convient ici de souligner que les notices œuvres concernent des feuillets de manuscrit. À titre d'exemple, le verso du feuillet 91 du manuscrit syriaque 47 a pour adresse <https://agorha.inha.fr/ark:/54721/4308c58d-6569-489e-8d7d-a7436872b5cd?database=89>. Son identifiant est donc : 4308c58d-6569-489e-8d7d-a7436872b5cd. Ce système d'identification unique, avec le même nombre de caractères, facilite la récupération d'informations et la structuration des différentes notices entre elles, nous le verrons plus en avant.

Pour revenir à Agorha, ces identifiants uniformes permettent à la plateforme de s'insérer pleinement dans le web de données liées². Ce dernier repose sur un ensemble de bonnes pratiques pour publier et relier des données structurées sur le web. Il vise principalement à faciliter l'intégration et l'interconnexion de données provenant de diverses sources. Agorha suit le modèle RDF (Resource Description Framework) qui représente les informations sous forme de triplet : un sujet, un prédicat, un objet. Par exemple, un feuillet (sujet) est réalisé à (prédicat) tel endroit (objet). La multiplication de triplets RDF permet de constituer un ensemble qui structure et interconnecte les informations. Les données liées peuvent ensuite être sérialisées avec du JSON-LD. Cette syntaxe permet de représenter des graphes RDF de manière compacte et facile à intégrer dans des applications web, tout en conservant les capacités de lien et de description du RDF. Dans Agorha, toutes les notices reposent sur du JSON-LD. Pour faire apparaître la méthode, il suffit de compléter les adresses de cette façon :

<https://agorha.inha.fr/ark:/54721/> + UUID + .json.

Si nous nous penchons sur le cas du manuscrit syriaque précédent, sa localisation apparaît ainsi en JSON-LD :

```

1 "localizationInformation" : {
2   "localization" : [ {
3     "place" : {
4       "thesaurus" : {
5         "ref" : "https://thesaurus.inha.fr/thesaurus/resource/
6           ark:/54721/bd1b5af3-9b44-49e9-98b8-b97ca10dd7f2",
7         "prefLabels" : [ {
8           "label" : "Manuscrit syriaque 47a"
9         }

```

2. Chloé Pochon, *AGORHA : Plateforme de données de la recherche de l'Institut national d'histoire de l'art.*

```
7         "value" : "\indexmot{Bibliothèque nationale de France
8             } (Paris)",
9         "language" : "fre"
10        } ],
11        "conceptPath" : "/Europe/France/Ile-de-France/Paris/\\
12            indexmot{Bibliothèque nationale de France} (Paris)",
13        "geoPoint" : "48.834355,2.373378"
14    }
15 },
16 "institutionIdentifierType" : {
17     "thesaurus" : {
18         "ref" : "https://thesaurus.inha.fr/thesaurus/resource/
19             ark:/54721/fecfc54-f2da-458a-a42c-530037a5132b",
20         "prefLabels" : [ {
21             "value" : "cote",
22             "language" : "fre"
23         }],
24         "conceptPath" : "/cote"
25     }
26 },
27 "institutionIdentifier" : {
28     "value" : "Syriaque 47",
29     "language" : "fre"
30 },
31 "sourcing" : [ {
32     "database" : {
33         "ref" : "89",
34         "value" : "La fabrique de l'art. Couleurs et matériaux
35             de l'enluminure"
36     }
37 } ]
38 } ]
```

Agorha permet donc une intégration structurée et interconnectée des résultats des deux programmes de recherche sur le web. La plateforme de l'INHA contribue à la science ouverte et la communication des résultats des analyses sur les enluminures et les panneaux peints selon les principes du web de données liées.

III. Deux sites Omeka S

Les deux projets ne seront pas seulement présents sur la plateforme Agorha. L'INHA développe en ce moment deux sites Omeka S qui permettront d'héberger, à partir de l'automne, les bases de données ainsi que les visualisations réalisées. Cette plateforme open source est spécifiquement conçue pour gérer des collections numériques dans l'optique d'une publication en ligne³. Elle semble parfaitement correspondre aux attentes d'un projet de recherche scientifique comme ceux sur les manuscrits ou les panneaux peints.

En premier lieu, le site Omeka S permet de créer des collections numériques. Ces dernières peuvent contenir les différentes notices présentes sur Agorha, mais avec une organisation et un archivage qui sont facilités. De plus, l'utilisation du schéma de métadonnées Dublin Core permet de reprendre les éléments des descriptions détaillées des notices. Il n'y a pas de perte de l'information et la normalisation reste. Ensuite, Omeka S est conçu pour intégrer les données RDF que j'ai présentées précédemment. L'interconnexion entre les différentes données des résultats existera toujours. Il en est de même avec le JSON-LD. Omeka S permet aussi d'accéder et de manipuler les données via des API, ce qui correspond aussi au critère d'interopérabilité voulu. Le principal avantage, au regard de la plateforme Agorha, réside dans la possibilité de la personnaliser pour les besoins spécifiques du projet. Le site public peut être structuré pour correspondre le mieux possible aux deux bases : navigation entre les onglets, liens entre les pages, visibilité des résultats ou encore insertion des visualisations. Le module de recherche lui-même peut être amélioré, afin d'orienter l'utilisateur vers des filtrages, option indispensable pour les projets scientifiques aussi complexes. Avec le site Omeka S, les utilisateurs pourront sélectionner des paramètres parfois très précis pour isoler certains résultats ou croiser plusieurs filtres afin de faire apparaître des correspondances. En dernier lieu, un site conçu sur Omeka S est traditionnellement conforme aux standards archivistiques et de gestion des collections, garantissant la pérennité des données. En outre, sa conception est conforme aux normes d'accessibilité, rendant le site utilisable par toute sorte de public.

En ce qui concerne les trois visualisations proposées pour les projets de recherche, leur intégration dans Omeka S varie. L'intégration par iframe semble être la meilleure solution pour ce qui est de la carte interactive et de la frise chronologique. Un iframe, c'est :

« un élément HTML qui permet d'intégrer un contenu externe dans une page web. L'Iframe peut contenir une vidéo, une carte, une publication de réseau social ou encore une page web tierce. Le contenu s'affiche dans un cadre dont les propriétés peuvent être paramétrées à l'aide d'attributs.⁴ »

3. *Omeka S*, URL : <https://omeka.org/s/> (visité le 26/07/2024).

4. Thiso Thach, *Qu'est-ce qu'un Iframe et comment l'intégrer dans une page ?*, fr, fr, URL : <https://blog.hubspot.fr/website/iframe> (visité le 02/07/2024).

En récupérant le contenu des deux visualisations et en personnalisant leur affichage, le site Omeka S devrait permettre une navigation confortable pour l'utilisateur. Pour ce qui est des manifestes IIIF et de leur exploration dans une visionneuse, il ne pouvait être raisonnablement proposé de passer par la solution d'un Iframe. En effet, Omeka S permet d'implémenter directement une visionneuse, comme Mirador, dans le site réalisé. Cependant, la visionneuse étant insérée par un module propre à Omeka, elle ne possède pas les mêmes propriétés que celle en libre accès sur internet. Un problème de compatibilité des manifestes IIIF est apparu alors, j'y reviendrai plus un avant.

IV. L'enjeu du thésaurus

Les visualisations sont, dans le fond, une certaine forme d'éditorialisation de la donnée. Elles reposent sur une interrogation des jeux et des entrées présents dans la base. Ainsi, plus les termes et les catégories utilisés lors de sa conception sont précis et cohérents, plus la visualisation sera pertinente d'un point de vue scientifique au moment de sa réalisation. Il me semble que la constitution des thésaurus a joué un rôle déterminant dans la réussite des deux projets et a permis d'envisager les différentes visualisations. Je souhaite donc m'y attarder un moment pour finir ce chapitre introductif.

Un thésaurus est, en quelque sorte, une standardisation du vocabulaire. Son rôle est de garantir que les termes sont utilisés de manière uniforme tout au long du travail de recherche. C'est un vocabulaire contrôlé, qui doit prévenir les éventuelles ambiguïtés et incohérences dans le remplissage des données dans la base. Mais au-delà de cette normalisation, l'instauration d'un thésaurus est fondamentale en ce qu'il constraint les chercheurs à une réflexion sur les termes utilisés et leur articulation les uns avec les autres. Ainsi, le thésaurus aide à structurer la description. L'attribution d'un terme à une donnée, à une idée, à une réalité – ce qui est donc une indexation – peut induire une hiérarchie en fonction du vocabulaire retenu. Les données sont alors regroupées selon leur similarité et classées par une logique définie en amont. Ce vocabulaire nécessite un effort théorique de la part des coordinateurs du projet : les termes doivent être justifiés et explicités et, bien souvent, une documentation est produite pour accompagner un thésaurus. Elle permet de fournir des définitions claires et précises des termes, des orientations aussi quant à l'indexation à choisir.

Les porteurs du projet ne cachent pas que l'un des principaux défis a bien été la constitution d'un référentiel d'autorité sur les matériaux et les couleurs qui puisse être applicable à des techniques, des époques et des aires géographiques très diverses⁵. S'il existait déjà un thésaurus de référence pour les sciences de la conservation⁶, ce n'était pas le cas pour la couleur et ses composants. Quelques précédents étrangers peuvent servir d'exemples cependant, comme la base américaine CAMEO⁷. L'élaboration du référentiel des matières colorantes est un parfait exemple de la synergie entre les différentes traditions scientifiques qui entourent le projet, puisqu'elle est le fruit d'ateliers réunissant des spécialistes de diverses disciplines (historiens de l'art, chimistes, spécialistes des textes anciens, naturalistes, anthropologues, etc.) et des professionnels issus des institutions partenaires. Une fois créés, les thésaurus « matériaux » et « techniques » viennent enrichir GINCO (Gestion d'Informations Communes). GINCO donne accès à l'ensemble des vocabulaires

5. Appel à projet « La couleur : artefacts, matière et cognition »

6. Vocabulaire contrôlé réalisé pour les projet PARCOURS, puis SOCORÉ.

7. Conservation & Art Materials Encyclopedia Online, *Pigment*, en, URL : <https://cameo.mfa.org/wiki/Pigment> (visité le 17/06/2024).

scientifiques et techniques produits par le Ministère de la Culture et ses partenaires⁸. L'INHA, pour ses projets, se fournit et alimente les vocabulaires établis, au moyen du langage de requête SPARQL⁹. De façon concrète, dans le cadre des deux projets qui nous intéressent, le chercheur qui saisit des données dans la base le fait par l'intermédiaire de champs de recherche. À ce moment, un petit moteur de recherche suggère par auto-compléTION les termes qui pourraient convenir et qui sont présents dans GINCO. Si aucun terme ne convient, il faut que le chercheur demande son ajout dans le thésaurus, le vocabulaire serait alors renseigné au format SKOS¹⁰ pour pouvoir intégrer GINCO. Cette méthode de travail et de compléTION de la base permet de s'assurer de la cohérence du vocabulaire utilisé et de sa normalisation.

Concernant les visualisations des jeux de données, la constitution d'un thésaurus commun aux deux projets permet de les réunir au sein d'un même travail d'édition. L'utilisation d'un vocabulaire contrôlé identique, pour la recherche sur « La couleur : artefacts, matière et cognition » et sur « La fabrique matérielle du visuel : panneaux peints en Méditerranée », autorise leur interrogation selon les mêmes critères et pour l'ensemble des résultats. Cette démarche simplifie aussi le travail technique qui poursuit les projets. La définition des filtres, qui sont les pivots de l'édition, ne représente dès lors plus un enjeu puisqu'ils sont communs et choisis par les chercheurs. Il n'y aura ni ambiguïté, ni incohérence.

8. Ministère chargé de la culture, *culturecommunication/ginco*, original-date : 2013-01-25T16:31:32Z, nov. 2022, URL : <https://github.com/culturecommunication/ginco> (visité le 04/07/2024), En 2024, le logiciel n'est plus maintenu à jour et semble voué à disparaître.

9. INHA, *Thésaurus et vocabulaires contrôlés de l'Institut national d'histoire de l'art*, URL : <https://thesaurus.inha.fr/thesaurus/> (visité le 04/07/2024).

10. « SKOS (Simple Knowledge Organisation System ou Système simple d'organisation des connaissances) est un langage de représentation de schémas de concepts, ce qui recouvre les langages documentaires tels que les thésaurus, classifications, listes de vedettes matière, taxonomies, folksonomies, etc. Son nom a été choisi pour mettre en évidence l'objectif même visé par ce langage : proposer un système permettant d'exprimer et de gérer des modèles interprétables par les machines dans la perspective du web sémantique », Michèle Lénart, « SKOS, un langage de représentation de schémas de concepts », *Documentaliste-Sciences de l'Information*, 44-1 (2007), Place : Paris Publisher : A.D.B.S., p. 75-75, DOI : 10.3917/docs.441.0075.

*

Depuis le début du projet, en janvier 2021, les différents porteurs ont œuvré pour que les données constituées dans la base s'inscrivent dans les nouvelles recommandations de l'interopérabilité et dans une politique de science ouverte. La communication des résultats avec le public des chercheurs est pensée en deux temps : une première sur la plateforme Agorha de l'INHA, où les données rejoignent – et sont valorisées avec – d'autres programmes de recherche en histoire de l'art et archéologie ; et une seconde, avec une migration des données vers deux sites Omeka S, afin d'individualiser et de personnaliser les résultats du projet. La constitution de thésaurus identiques aux deux recherches permet de proposer des visualisations communes et ainsi de poursuivre leur collaboration.

Chapitre 2

Ordonner les données dans l'espace, la carte interactive

Le projet « La couleur : artefacts, matière et cognition » repose sur plus de 8 000 analyses. Ces dernières se déclinent en une multitude d'informations, de définitions et de commentaires qui sont, en les situant dans un contexte informatique, autant de variables pour un traitement personnalisé de la recherche. Dans le cadre d'une politique d'éditionnalisation des données par les utilisateurs, il me semble que proposer la mise en ordre des données dans l'espace est l'une des plus importantes. Les chercheurs ont réuni, au cours des années du projet, des informations précieuses sur les lieux de création des œuvres. Illisibles dans l'immensité d'un fichier CSV, la représentation de ces données sur une carte permet de produire une distribution dans l'espace de la production artistique et, ainsi, de faciliter l'interprétation avec une approche géographique.

I. Leaflet : une bibliothèque JavaScript open-source et légère

La cartographie interactive est un outil précieux pour visualiser et interagir avec des données géographiques de manière dynamique et intuitive.

Les données de la base des projets de recherche sont une véritable mine d'informations. Elles n'auraient pu que s'appauvrir en les représentant avec une carte statique classique. Premièrement, cette dernière est contrainte par une image numérique qui prédéfinit une dimension et une résolution, là où la carte interactive permet des manipulations d'échelles qui évitent les excessives accumulations de données. Ensuite, elle ne permet pas de représenter les variables temporelles, à moins d'être multipliées en autant d'exemplaires que de dates voulues. Enfin, la carte classique ne représente finalement qu'une visualisation des informations conçue et produite par un cartographe, sans personnalisation possible par l'utilisateur. Sans lister exhaustivement toutes les contraintes – et les limites – d'une telle représentation, ces petits exemples permettent de mesurer la nécessité de réaliser dans le cas présent une cartographie interactive, plus complexe, mais plus fidèle à la richesse des données de la base.

Aujourd’hui une carte interactive peut être réalisée à l'aide d'outils et de plateformes en ligne, comme Google My Maps¹, Mapbox² ou encore ArcGIS Online³. Ces plateformes sont gratuites et accessibles à toute personne possédant des rudiments d'informatique, mais elles ne parviennent pas à offrir autant d'avantages et de fonctionnalités que les outils qui nécessitent des connaissances en codage. Ces derniers permettent de personnaliser presque toutes les options de la carte, y compris l'apparence, le comportement et les fonctionnalités. Ils sont réputés avoir une meilleure performance si le nombre de données est amené à être important et, surtout, les outils reposant sur du codage peuvent être plus facilement intégrés dans des applications ou sites Web – comme il en est dans le cas présent. Les outils de cartographie sont, pour l'essentiel, des bibliothèques JavaScript, en raison de la simplicité, de l'interactivité et de la performance de ce langage. Il possède également une grande communauté de développeurs qui contribuent à le maintenir à jour et à proposer toujours de nouvelles solutions aux utilisateurs. La communauté JavaScript a développé quelques outils populaires, comme Leaflet.js⁴, OpenLayers⁵, D3.js⁶ ou encore Mapbox GL JS. Dans le cadre de ce stage, le choix de la bibliothèque Leaflet.js,

1. *Google My Maps*, URL : <https://www.google.com/maps/d/> (visité le 26/07/2024).

2. *Mapbox / Maps, Navigation, Search, and Data*, URL : <https://www.mapbox.com/> (visité le 26/07/2024).

3. *ArcGIS Online*, URL : <https://www.arcgis.com/index.html> (visité le 26/07/2024).

4. Volodymyr Agafonkin, *Leaflet*, en, URL : <https://leafletjs.com/> (visité le 26/07/2024).

5. *OpenLayers*, URL : <https://openlayers.org/> (visité le 26/07/2024).

6. *D3 / The JavaScript library for bespoke data visualization*, URL : <https://d3js.org/> (visité le 26/07/2024).

bien qu'elle soit réputée comme étant la plus légère, est apparu comme une évidence. Elle propose, malgré sa simplicité, une large gamme de fonctionnalités de cartographie, notamment en ce qui concerne les marqueurs, les pop-ups, la possibilité de cumuler les filtres ou encore de faire apparaître une barre latérale. Elle propose déjà, en somme, toutes les options qui ont été retenues pour la visualisation interactive⁷. De plus, elle est utilisée couramment par le service informatique de l'INHA, service qui hébergera la base Omeka S du projet⁸. Dans le cadre d'un stage d'une durée de trois mois et avec les conditions énoncées préalablement, il ne m'a pas semblé opportun de continuer l'investigation vers les autres bibliothèques, mais au contraire plus pertinent de chercher à déployer pleinement toutes les potentialités de la bibliothèque Leaflet. Ainsi, les fonctionnalités plus avancées d'OpenLayers, notamment quant aux tuiles ou aux images raster n'étaient daucun intérêt dans le cas présent, D3.js n'est pas nativement conçu pour de la cartographie et aurait complexifié inutilement le travail, tandis que la dernière bibliothèque, Mapbox GL JS, semble plus convenir à un autre public que celui de chercheurs au regard de l'éditorialisation à venir.

À la date du 18 mai 2023, Leaflet est à sa version 1.9.4⁹. La bibliothèque est déposée et peut-être librement téléchargée sur la plateforme GitHub – elle pèse seulement 42 ko de JS. Son premier développeur, l'Ukrainien Vladimir Agafonkin, s'est entouré depuis une dizaine d'années d'une petite équipe qui veille au maintien de Leaflet. La bibliothèque répond ainsi aux innovations du web et aux nouveaux besoins des utilisateurs. Pour ces derniers, les développeurs déposent à l'adresse <https://leafletjs.com> des tutoriels et une riche documentation qui apportent une réponse à la plupart des questions et des difficultés rencontrées.

Une fois téléchargé, le code se divise en trois fichiers et un dossier¹⁰ :

- `leaflet.js` : il s'agit du code minifié¹¹
- `leaflet-src.js` : le code non-minifié cette fois-ci, ce qui peut permettre un débogage
- `leaflet.css` : la feuille de style
- `images` : un dossier contenant les images référencées par le fichier précédent

Cet ensemble doit être décompressé et placé dans le répertoire du site web. Sur la page HTML, il suffit d'ajouter ces deux lignes pour créer le lien avec la bibliothèque Leaflet :

7. Cf. infra.

8. À titre personnel, la bibliothèque Leaflet est aussi celle qui me fut enseignée au cours de cette année de master.

9. Pour notre part, nous avons réalisé le travail à partir de la version 1.8.0.

10. Id., *Leaflet...*

11. En programmation, minifier signifie réduire la taille du code.

```
<link rel="stylesheet" href="/chemin/vers/le/fichier/leaflet.css" />
<script src="/chemin/vers/le/fichier/leaflet.js"></script>
```

Les fichiers JavaScript sont alors correctement liés à la page HTML qui permet de visualiser et déployer la carte sur le site web. Avant de débuter le traitement de la donnée, il convient de poursuivre le travail préparatoire sur les objectifs de la visualisation et de définir, avec une plus grande précision, les attendus de l'éditorialisation.

II. Éditorialiser une base de données avec une cartographie interactive

La constitution de la base de données par les chercheurs du projet, notamment à partir des thésaurus, permet d'avoir à disposition des variables qui seront à l'origine de l'éditorialisation. Concrètement, il s'agit dans le cadre d'une cartographie interactive de réaliser le choix d'un ou de plusieurs filtres qui s'appuient sur des clefs, représentées par des colonnes dans les fichiers CSV, et de porter la sélection sur les valeurs associées, limitées raisonnablement par le thésaurus. À titre d'exemple, que ce soit pour la base materiality⁸⁸ ou celle ⁸⁹, au sein de la colonne « schema :color » qui représente les couleurs identifiées au moment de l'analyse, le filtre porte sur une sélection entre quinze valeurs : argenté, beige, blanc, bleu, brun, cuivré, doré, gris, jaune, noir, orange, rose, rouge, vert et violet. Une attribution définie au moment de l'analyse scientifique de l'objet devient ainsi ici une variable qui permet d'éditorialiser – de filtrer – le résultat pour l'utilisateur.

La cartographie interactive est une visualisation parfaitement adaptée à ce processus de sélection, de désélection et de combinaison de filtres. À la manière des cartes statiques qui laissent une place aux légendes, celles interactives permettent de faire apparaître au plus près les besoins en information et en usage pour l'utilisateur. Elles prennent la forme de quadrilatères équiangles pouvant être disposés selon les souhaits du cartographe. Dans notre cas présent, l'éditorialisation se réalise à partir de quatre filtres, qui peuvent tous s'additionner afin d'affiner la recherche.

La visualisation a pour objectif de traduire la circulation et l'utilisation des matériaux de couleur dans la fabrique des enluminures et des panneaux peints sur plus d'un millénaire. C'est une visualisation qui est non seulement géographique, mais également chronologique. Elle repose ainsi sur quatre variables présentes dans les bases de données : le matériau (schema :material), la couleur obtenue (schema :color), la localisation (schema :geo) et la date de création (crm :P4_has_time-span). Toutes les autres informations présentes dans les bases servent à enrichir les informations proposées par les résultats du filtre, mais elles ne constituent pas les pivots de l'éditorialisation.

La première option à la disposition de l'utilisateur émane d'une demande des responsables du projet. Madame Charlotte Denoël, responsable du projet « La couleur : artefacts, matière et cognition », et madame Sigrid Mirabaud, responsable du programme « La fabrique matérielle du visuel », souhaitent laisser le choix entre une navigation qui réunirait les deux projets et une autre, qui porterait exclusivement sur l'un d'entre eux. Si les deux projets sont réunis sur la plateforme Omeka S en raison de leur synergie, ils doivent pouvoir être individualisés à la demande. Ainsi, il convient d'insérer un filtre qui permet la bascule d'un projet à l'autre en première option, tout en proposant leur réunion – choix initialisé par défaut.



FIGURE 2.1 – Filtrer par projet sur la carte

Afin de visualiser l'évolution dans le temps des données, le filtre sur les dates de création est sous la forme d'une jauge. Cette dernière, plus intuitive, est plus à même d'offrir une interaction fluide à l'utilisateur, qui n'a qu'à faire glisser le curseur pour sélectionner une année spécifique. Quant à la granularité de la jauge, les deux bases de données couvrant plus d'un millénaire, j'ai décidé de réaliser une sélection par décennie plutôt que par années. Ce choix doit rendre plus aisés les tendances sur la carte, par une apparition plus fréquente des données. De plus, ce filtre n'est pas, en quelque sorte, « excluant », mais au contraire inclut et ajoute progressivement de la donnée à la carte : le filtre ne représente pas seulement une décennie, mais additionne les créations de ces dernières à celles préexistantes. Ainsi, si nous déplaçons le curseur sur l'année 1060, les manuscrits créés au cours de la décennie s'ajouteront à ceux antérieurs, sans les remplacer. Il me semble que c'est ici proposer une vision plus globale d'un contexte de production, en ne l'isolant pas des dynamiques passées.



FIGURE 2.2 – Filtrer par année sur la carte

La troisième variable est la plus importante du projet. Elle repose sur l'analyse des matériaux et sur leur ontologie. Que ce soit pour la base materiality88 ou celle 89, ces données sont entrées au sein de l'entité « schema :material ». Fruit d'un long travail, elles reflètent aussi la difficulté de la coordination et de la description manuelle sur un tel corpus. Un script Python révèle qu'à la date du 15 mai 2024, lors d'un travail de normalisation encore en cours, l'entité « schema :material » possédait 258 valeurs différentes, loin du thésaurus constitué. Des erreurs d'écriture, plus que de choix, expliquent le résultat – nous le remarquons notamment par la présence d'apostrophes qui déprécie l'uniformité¹². De ce fait, un filtre reposant sur une simple sélection n'est pas envisageable. Je propose, pour le choix du matériau, un champ de recherche avec de l'auto-complétion afin de guider l'utilisateur parmi toutes les valeurs – l'ordre de la valeur recherchée n'a pas d'importance et le filtre peut être nettoyé à tout moment.

12. On rencontre ainsi 1050 occurrences de « blanc de plomb », 12 de « blanc de plomb' » et 10 de « 'blanc de plomb ».



FIGURE 2.3 – Filtrer par matériau sur la carte

Le dernier et quatrième filtre visible sur la carte porte sur les couleurs. Il semble utile de faire apparaître la concordance ou, au contraire, la discordance de cette variable avec la précédente. Le cumul de ces deux filtres fait ainsi apparaître un maintien ou une disparition de marqueurs que les chercheurs pourront interpréter. Cette entité ne reposant que sur quinze valeurs, j'ai opté pour un sélecteur à choix multiples.



FIGURE 2.4 – Filtrer par couleur sur la carte

L'ensemble de la carte interactive repose la présence de marqueurs, qui représentent la ou les localisations des feuillets analysés. Ces informations sont entrées au sein de l'entité « schema :geo ». Cette dernière contient des données de coordonnées géographiques, au format « latitude, longitude ». Un feuillet peut avoir plusieurs localisations possibles, dans le cadre de contexte de production pluriel ou indéterminé. Au lieu de choisir arbitrairement une seule coordonnée, un même feuillet peut apparaître sur la carte à plusieurs reprises – à la condition que les localisations soient éloignées d'au moins trois points de coordonnées pour ne pas surcharger inutilement la visualisation. Ainsi, le *Pentateuque dit d'Ashburnham ou de Tours* apparaît aussi bien en Afrique du nord, en Espagne, en Italie

qu'en France¹³. Lorsqu'ils sont en nombre, les marqueurs sont réunis en clusters – ils sont groupés. Ces derniers indiquent par un chiffre ou un nombre le total des feuillets qu'ils contiennent. Par un agrandissement de la carte ou un clic sur le repère, les marqueurs des feuillets apparaissent. Dans l'exemple ci-dessous, un clic sur le cluster vert situé au niveau de la ville de Milan fait apparaître quatre feuillets analysés du *Liber comitis dit Lectionnaire pourpré*¹⁴.

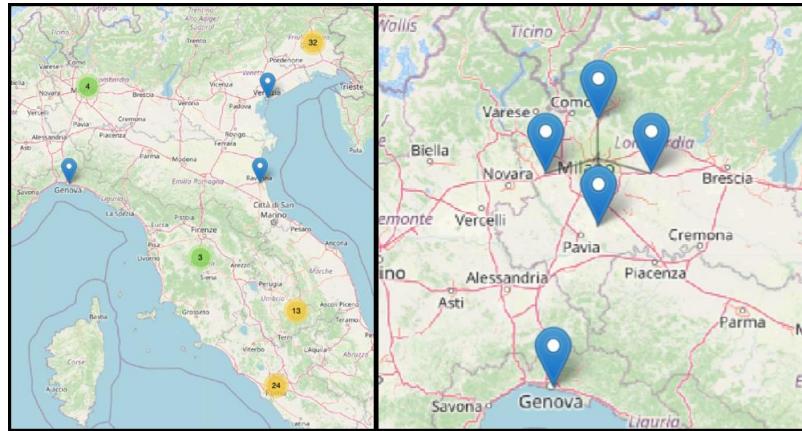


FIGURE 2.5 – Marqueurs et clusters

Il est apparu, dans le cadre du projet, qu'il est parfois impossible de proposer une localisation, même indécise, à un manuscrit. Dans ce cas de figure, je prends le parti de définir une localisation par l'absurde afin de faire figurer le marqueur. Situé au large des côtes bretonnes, dans l'Atlantique, le marqueur ne laisse pas de place à une ambiguïté. Une fois le carrousel déployé, il apparaît comme titre du feuillet l'indication « Lieu de création inconnu ».

À son lancement, la carte a, par défaut, tous les filtres ouverts et donc tous les marqueurs de représentés. Le filtre « projet » affiche les résultats sur les deux bases materiality, la date est dans sa dernière décennie, aucun matériau n'est sélectionné, ni aucune couleur définie. L'affichage est centré sur l'Europe et le bassin méditerranéen, où se situent les principaux manuscrits analysés – je crois en l'interactivité de la carte pour que l'utilisateur ait la curiosité de la décentrer et de découvrir des analyses sur des manuscrits produits en d'autres continents.

13. *Pentateuque dit d'Ashburnham ou de Tours*, Paris, BnF, NAL 2334

14. *Liber comitis dit Lectionnaire pourpré*, Paris, BnF, Latin 9451

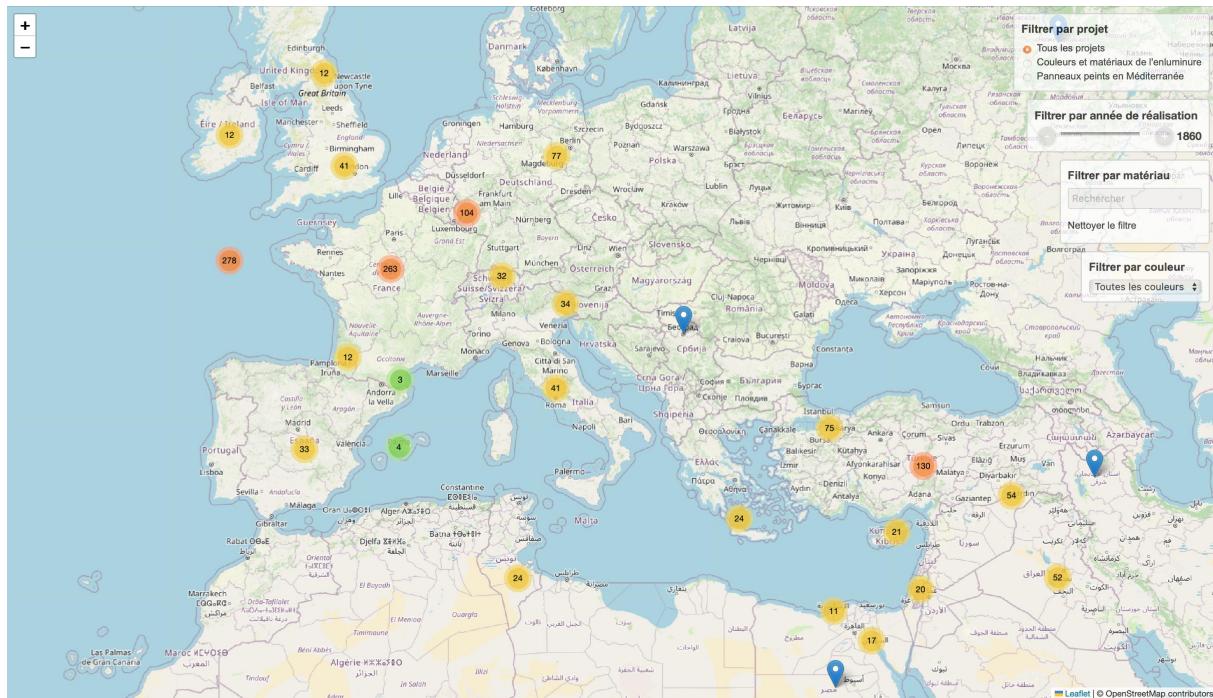


FIGURE 2.6 – Initialisation de la carte

Il a été loué précédemment la vertu heuristique du cumul des différents filtres — du croisement de variables en interrogeant la base — afin de faire apparaître certains résultats dans un contexte géographique et chronologique. La visualisation ci-dessous, en quatre temps, est insérée à titre d'exemple afin d'apprécier l'efficacité de l'entonnoir produit par les filtres. En partant de la carte initialisée comme précédemment, je définis successivement une base de recherche, les « Couleurs et matériaux de l'enluminure », un cadre chronologique avec l'année 1070, la cochenille pour matériau ainsi que la couleur rose. Il apparaît alors vingt-quatre résultats. *Voir la figure 2.7.*

Ordonner les données dans l'espace, la carte interactive

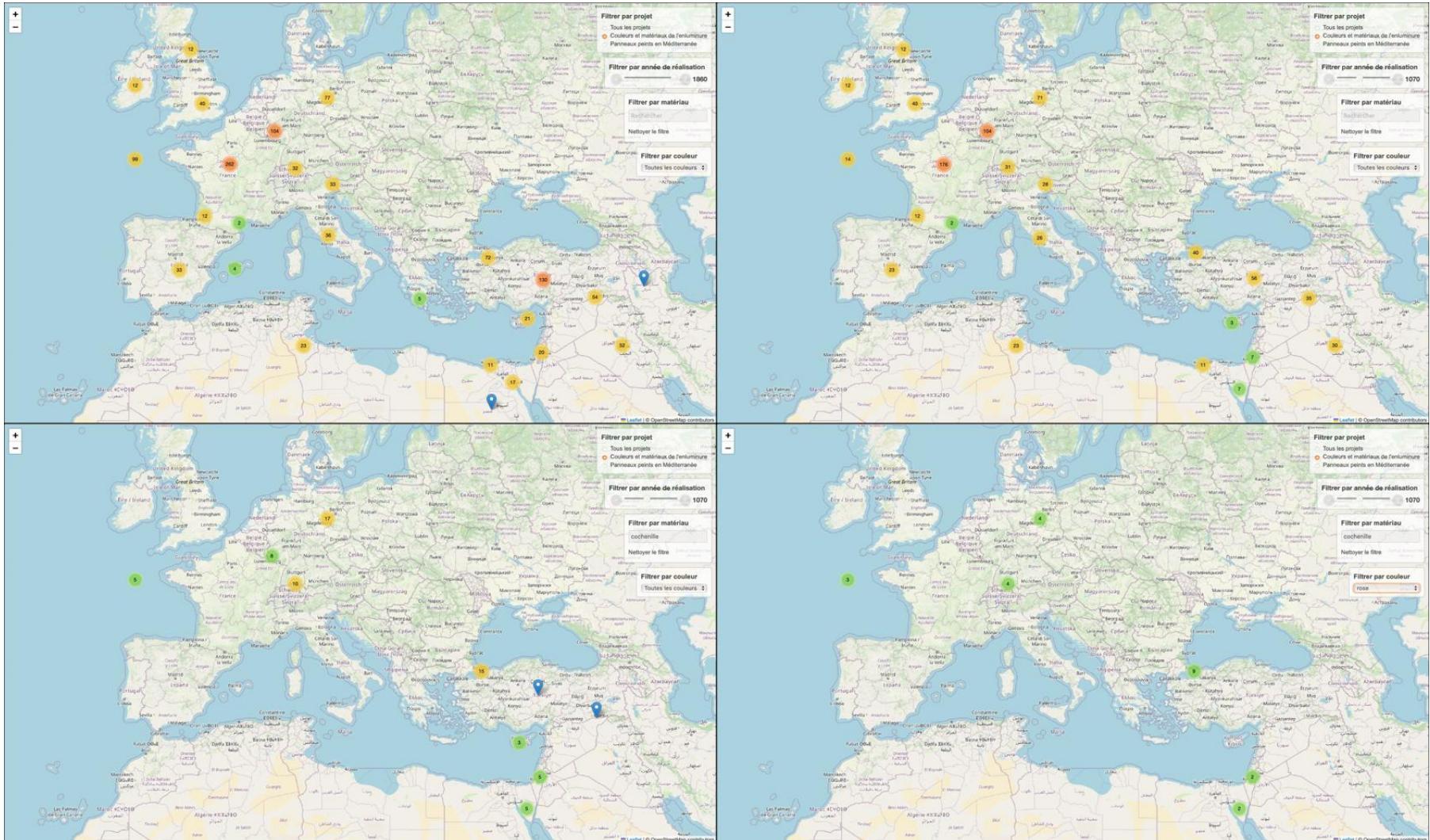


FIGURE 2.7 – Exemple d'éditorialisation

La lecture des résultats est proposée en trois temps, qui représentent autant de niveaux de détails. Une première info-bulle apparaît lorsqu'un marqueur est survolé par le curseur. Elle reprend l'information entrée dans la colonne « dcterms :title » et informe sur le titre du manuscrit, sa cote, le nom du feuillet et son numéro. Ce survole rapide accélère le temps de recherche lorsque l'utilisateur est à la recherche d'un feuillet ou d'un manuscrit précis. *Voir la figure 2.8.*

Ensuite, un clic permet de déployer un carrousel qui contient les données essentielles sur le feuillet : une reprise de la colonne « dcterms :title » comme titre, suivie de l'appel à une image IIIF – image qui peut être agrandie en haute résolution par un nouveau clic. Le carrousel permet une navigation entre les différentes analyses réalisées sur le feuillet. On y trouve ainsi la caractéristique, le motif, la technique, la couleur ainsi que le matériau. *Voir les figures 2.9 et 2.10.*



FIGURE 2.8 – Une info-bulle avec les premières informations

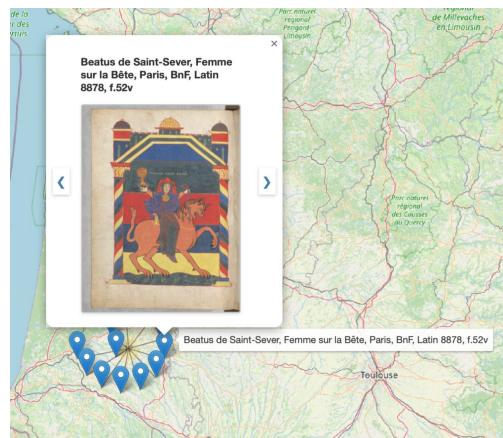


FIGURE 2.9 – Une image du carrousel

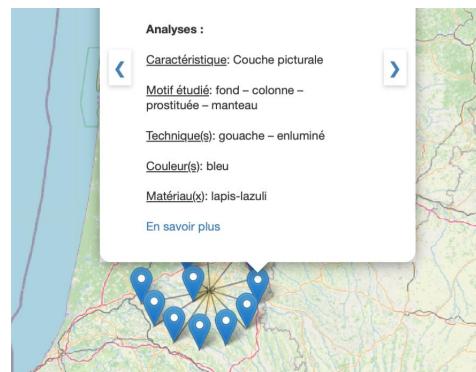


FIGURE 2.10 – Un descriptif du carrousel



FIGURE 2.11 – Une barre latérale avec la totalité des informations

Beatus de Saint-Sever, Femme sur la Bête, Paris, BnF, Latin 8878, f.52v

Identification

Identifiant: 72d9a7bf-87e4-469d-97e1-26408db436cb_001

Type d'œuvre: manuscrit - enluminure

Attribution: Placidus

Lieu de création: Abbatiale Saint-Sever - Sud-Ouest de la France

Date de réalisation: 1051 - 1072

Description

Type: Couche colorée

Caractéristique: Couche picturale

Motif étudié: fond - colonne - prostituée - manteau

Technologie(s): VNIR - scanner XRF

Couleur(s): bleu

Matériaux(x): lapis-lazuli

Certitude: caractérisé -- caractérisé

Date: 2022 - 2022

Rapport de l'analyse: Bleu: bleu outremer scanner XRF: Si, K, Ti, Fe / caméra VNIR: absorption à 600 nm + Ca

Source du rapport: Thomas CALLIGARO, Rapport d'analyses Beatus de Saint-Sever campagne janvier-mars 2022, - MICHELIN, Anne, Analyses Beatus de Saint-Sever campagne janvier 2022

Lieu de conservation

Localisation: BnF - département des Manuscrits

Cote / numéro: Latin 8878

Lorsque l'utilisateur trouve le feuillet qui l'intéresse, il peut faire apparaître une barre latérale qui reprend et complète les informations présentes. Il apparaît alors le lieu de création, une éventuelle attribution, d'avantage de détails sur l'analyse, avec notamment les références nécessaires pour la citer, ainsi que des informations sur la conservation actuelle. *Voir la figure 2.11.*

La carte interactive est ainsi le reflet de la richesse du travail réalisé sur les deux bases de données. Elle reprend, d'une façon ordonnée et hiérarchique, toutes les informations essentielles qu'elles contiennent. Si la visualisation repose sur plus de 8 000 lignes de CSV, elle permet aussi d'individualiser chacune d'entre elles : le feuillet est représenté dans le travail global du projet de recherche, mais les informations détaillées qui le concernent sont présentes également. La carte interactive est tout autant une synthèse qu'une analyse. Une synthèse en ce qu'elle propose une vue d'ensemble et cohérente du travail des chercheurs – elle est un résumé visuel ; elle est une analyse lorsqu'elle décompose les résultats en fournissant toutes les informations nécessaires pour comprendre les matériaux et les couleurs et que l'utilisateur peut les filtrer selon ses préférences. Pour ce faire, le code doit permettre des allers-retours, des personnalisations, et une exploration plus ou moins profonde des fichiers CSV qui révèlent progressivement les différents niveaux de détail de la recherche.

III. Préparer les données

L'écriture du code se fait sur Visual Studio Code, un éditeur développé par Microsoft pour Windows, Linux et macOS¹⁵. Actuellement dans sa version 1.90.0, cet éditeur inclue une prise en charge du débogage, une mise en évidence de la syntaxe, parfois une complé-
tion intelligente du code et, point important dans le cas présent, une intégration de Git . Ce dernier aspect est essentiel lorsque le code écrit repose sur des essais qui sont validés, corrigés ou invalidés. Le travail sur une bibliothèque de cartographie interactive est dans la pratique essentiellement empirique et ne nécessite que ponctuellement de s'aventurer dans les pages de la documentation. Des lignes de code sont écrites, on observe un résultat, on les corrige par petites touches ou retranchements jusqu'à ce que le résultat finalement soit satisfaisant et conforme à l'attendu. La possibilité de réaliser des allers-retours est ainsi essentiel, tout comme la celle de pouvoir comparer différentes étapes et les réunir le cas échéant. De ce fait, l'implémentation de Git Lens¹⁶ et de Git Graph¹⁷ à VSCode, deux extensions qui facilitent l'expérience du codage, notamment par des visualisations claires de l'historique des commits, est un plus. Live Server¹⁸ est la troisième extension nécessaire, elle permet de lancer un serveur web local pour prévisualiser les pages web en temps réel. Chaque modification apportée dans le code est alors aussitôt visible sur la carte interactive.

Une fois ces paramétrages effectués, le travail avec la bibliothèque Leaflet peut commencer. Je ne présenterai ici que les lignes directrices qui conduisent à la visualisation obtenue, pas une reprise détaillée de chaque ligne du code. Ce dernier est commenté et est disponible à l'adresse <https://github.com/TheoBurnel/Stage.git> le cas échéant.

Dans un premier temps, le travail repose sur quatre fichiers CSV fournis par l'INHA qui reprennent toutes les informations entrées dans Agorha par les porteurs du projet : artwork88, artwork89, materiality88 et materiality89. La visualisation est souhaitée à partir des feuillets analysés, et non pas des manuscrits. De ce fait, l'essentiel des données repose sur les deux derniers fichiers, les premiers n'étant présents dans mon code que pour la récupération des images des feuillets en question – la colonne « dcterms :image » qui contient les liens des images n'existe pas dans les fichiers materiality88 et 89. Un code Python permet de n'extraire de ces CSV que l'identifiant du feuillet ainsi que le lien vers son image, une liaison par l'identifiant peut alors se faire avec les deux autres bases CSV¹⁹.

15. *Visual Studio Code - Code Editing. Redefined*, en, URL : <https://code.visualstudio.com/> (visité le 26/07/2024).

16. *GitLens — Git supercharged - Visual Studio Marketplace*, en-us, URL : <https://marketplace.visualstudio.com/items?itemName=eamodio.gitlens> (visité le 26/07/2024).

17. *Git Graph - Visual Studio Marketplace*, en-us, URL : <https://marketplace.visualstudio.com/items?itemName=mhutchie.git-graph> (visité le 26/07/2024).

18. *Live Server - Visual Studio Marketplace*, en-us, URL : <https://marketplace.visualstudio.com/items?itemName=ritwickdey.LiveServer> (visité le 26/07/2024).

19. Voir le code « images.py », puis « lien.py » à l'adresse <https://github.com/TheoBurnel/>

Concernant les deux CSV principaux, le code Python « nettoyage » me permet d'isoler les identifiants des feuillets, de sélectionner les coordonnées voulues et au bon format, de supprimer d'éventuels doublons et enfin de les réunir en un seul fichier²⁰. Une fois la concaténation faite avec les images présentes dans les fichiers artwork, j'obtiens un tableau de 10701 lignes pour 37 colonnes, soit 395 937 entrées pour l'analyse. La cartographie est bien le reflet d'une grande étude, non seulement qualitative, mais également quantitative.

Après la phase préparatoire et avant de passer au code JavaScript, il est nécessaire de transformer le fichier CSV obtenu en données JSON. À ce moment, le langage Python est encore nécessaire. Il l'est dans un premier temps pour donner une structure appropriée au fichier de sortie – un fichier .json, mais il l'est également pour la définition d'une entité « parent » et d'une « enfant » qui serviront à la réalisation du carrousel de la carte. Puisque la visualisation repose sur les feuillets, il convient de définir les valeurs de la colonne « crm :P106i_forms_part_of » comme des parents des valeurs présentes dans celle « dcterms :title » où se trouvent les analyses réalisées. En d'autres termes, un feuillet qui a pour valeur – identifiant – « 02389266-ee27-48ea-be1e-0b40e5e0144e » sert de pivot – de socle à la rotation du carrousel – aux analyses « 02389266-ee27-48ea-be1e-0b40e5e0144e_001 », « 02389266-ee27-48ea-be1e-0b40e5e0144e_002 », « 02389266-ee27-48ea-be1e-0b40e5e0144e_003 », etc. Cette attribution faite, il possible d'encoder le tableau au format idoine pour débuter le travail de cartographie²¹.

Stage/tree/9c0841f622f3f39d51615e129507a4bf9a2abaf9/leaflet/python

20. Voir le code à l'adresse <https://github.com/TheoBurnel/Stage/tree/9c0841f622f3f39d51615e129507a4bf9a2abaf9/leaflet/python>

21. Un exemple de résultat pour une analyse est également proposée à l'annexe 2.

IV. Coder la bibliothèque JavaScript Leaflet

Le code repose sur un fichier JSON d'un peu plus de sept cents lignes. Ne pouvant le détailler ici, je me propose de me concentrer sur le traitement de l'entité « schema :material », qui concerne les entrées sur les matériaux. Sa présentation permet de montrer un travail réalisé sur une entité de première importance, mais également un modèle de création de filtre avec la bibliothèque Leaflet.

Après les premières lignes qui initialisent la carte, chargent les tuiles et instaurent les principes des clusters pour les marqueurs, il convient de reprendre le travail sur la définition des entités parent ou enfant. Le code est le suivant :

```

1 // Fonction pour vérifier si c'est un parent (adulte)
2 function isParent(titre) {
3     return titre && titre.trim() !== ''; // Considérer comme parent si le
4         titre n'est pas vide
5
6 // Fonction pour récupérer les enfants par parent
7 function getChildrenByParent(parent) {
8     var children = [];
9     // Logique pour récupérer les enfants
10    geojson_RAMA.features.forEach(function (feature) {
11        if (feature.properties.Parent === parent) {
12            children.push({
13                // Exemple pour récupérer la valeur des matériaux
14                // Il faut fonctionner de façon similaire pour toutes les autres
15                // entrées
16                materiaux: feature.properties.Materiaux || '?',
17                // etc etc.
18            });
19        }
20    });
21
22    return children;
23 }
```

À présent, le socle du carrousel peut être définie sur les entrées du titre, à savoir la colonne « dcterms :title », tandis que le défilement se fera sur l'ensemble des enfants récupérés, à savoir les autres colonnes de la même ligne.

Le contenu du carrousel tient en quelques lignes :

```

1 // Fonction pour créer le contenu du carrousel pour les enfants
2 function createCarousel(parent, identifiant) {
3     var carouselContent = "<div class='carousel'><div class='carousel-
4         content'>";
5
5     var children = getChildrenByParent(parent);
```

```

6     children.forEach(function (child, index) {
7         var displayStyle = index === 0 ? 'block' : 'none';
8
9         carouselContent += "<div class='slide' style='display: " +
10            displayStyle + "; padding-left: 30px;'>";
11
12        if (child.coordinates && child.coordinates.length === 2 && child.
13            coordinates[0] === -8 && child.coordinates[1] === 48) {
14            carouselContent += "<p><h4><u>Lieu de création inconnu</u></h4></p>
15            ";
16        }
17
18        carouselContent += "<h3>" + child.titre + "</h3>";
19
20        if (child.image && child.image !== '?') {
21            carouselContent += "<img id='carousel-image-" + index + "' src='"
22            child.image + "' onclick='expandImage(this)' />";
23        }
24
25        carouselContent += "<p class='carousel-legend'>Identifiant de la
26          photo :<br>" + child.name + "</p>";
27
28        carouselContent += "<p><h4>Analyses :</h4></p>";
29        // Jusqu'à présent, le carrousel a un titre, une valeur par défaut
          pour les lieux de création inconnus, une image téléchargée
          automatiquement avec une légende et également un paragraphe
          introduisant les résultats des analyses. Il convient à présent 'd'y
          mettre les informations souhaitées. Pour les matériaux :

```

Il est en de même pour les autres entrées. Le carrousel se termine par un lien vers la barre latérale, la création des flèches pour le défilement ainsi que la possibilité d'agrandir l'image téléchargée. Les lignes pour la définition du contenu de la barre latérale sont très proches de celles pour la carrousel, il convient simplement de bien veiller à la correspondance entre les éléments parents, autrement dit de faire correspondre les feuillets et les analyses. À la fin du carrousel, le lien est défini par cette ligne :

```

1     carouselContent += "<p class='slidebar-link' onclick='showSidebar(" +
2       JSON.stringify(child) + ")'>En savoir plus</p>";

```

Ce qui permet de récupérer les informations dans la barre latérale :

```

1 // Fonction pour afficher la sidebar avec les détails de l'élément

```

```

    sélectionné
2  function showSidebar(child) {
3      var sidebar = document.createElement('div');
4      sidebar.className = 'sidebar';
5
6      // Création du contenu HTML de la sidebar en vérifiant les valeurs
7      var sidebarContent = `
8          <div class="sidebar-content">
9              <button class="close-btn" onclick="hideSidebar()">&raquo;</button>
10             <div class="sidebar-header">`;
```

Puis d'appeler le contenu, avec l'exemple des matériaux :

```

1  if (child.materiaux && child.materiaux !== '?') {
2      sidebarContent += `<p><u>Matériaux :</u> ${child.materiaux}</p>`;
3 }
```

L'entité « schema :material », nous l'avons vu, est également à la source d'un filtre. Ses deux cent cinquante-huit valeurs différentes doivent constituer des variables sur les-quelles il est possible de réaliser une sélection. Afin de pouvoir permettre une éventuelle mise à jour de l'ontologie sur les matériaux, notamment sur les corrections que nous avons évoquées précédemment, les valeurs ne sont pas écrites en dur dans le code. Elles sont appelées pour être directement retranscrites dans un ensemble qui les stocke. À cette occasion, les données sont aussi partiellement nettoyées de leurs défauts. Ensuite, un tableau est créé, dans lequel les résultats individualisés sont à la fois clef et valeur pour le filtrage. Une fois ces données prêtes, l'ajout du panneau de contrôle, l'auto-complétion, ainsi que nettoyage du filtre se codent sans difficulté. Plus complexes, en revanche, sont le paramétrage du filtre et sa liaison aux trois autres.

Pour le premier, la solution est de réaliser des méthodes « querySelector » sur les valeurs rassemblées. « querySelectorAll » renvoie à une liste statique, ici l'ensemble des entrées du tableau ; elle permet ainsi d'initialiser le filtre au lancement de la carte, tout comme de nettoyer le filtre. De son côté, « querySelector » renvoie l'élément qui correspond à l'entrée voulue.

```

1  // Appeler la fonction d'initialisation une fois que le DOM est chargé
2  document.addEventListener('DOMContentLoaded', function () {
3      initMaterialItems();
4  });
5
6  // Fonction pour réinitialiser le filtre de recherche de matériau
7  function resetMaterialFilter() {
8      var materialElements = document.querySelectorAll('.material-item');
9
10     // Cacher tous les éléments matériau
11     materialElements.forEach(function (element) {
12         element.style.display = 'none';
```

```
13     });
14 }
15
16 // Fonction appelée lors de la sélection d'un matériau dans la liste
17 function selectMaterial(material) {
18     updateMateriauFilter(material);
19     // Réinitialiser le filtre après la sélection d'un matériau
20     resetMaterialFilter();
21
22     // Sélectionner le champ de recherche
23     var searchInput = document.querySelector('input[type="text"] [
24         placeholder="Rechercher"]');
25     // Mettre à jour la valeur du champ de recherche avec le matériau
26     // sélectionné
27     if (searchInput) {
28         searchInput.value = matériaux[material] || ""; // Utiliser le libellé
29             correspondant à la clé sélectionnée
30     }
31 }
32
33 // Fonction pour nettoyer le filtre de recherche de matériau
34 function clearMaterialFilter() {
35     // Réinitialiser le champ de recherche
36     var searchInput = document.querySelector('input[type="text"] [
37         placeholder="Rechercher"]');
38     if (searchInput) {
39         searchInput.value = ''; // Effacer le texte du champ de recherche
40     }
41
42     // Réinitialiser et afficher tous les éléments matériau
43     var materialElements = document.querySelectorAll('.material-item');
44     materialElements.forEach(function (element) {
45         element.style.display = 'block'; // Afficher tous les éléments
46             matériau
47     });
48 }
49
50 // Ajouter le contrôle de recherche à la carte
51 controlSelect.addTo(map);
```

Mais pour que ce paramètre fonctionne de manière synchronisée avec les autres, il faut que les filtres sur les projets, la décennie et la couleur soient mis à jour au moment de son lancement. Pour le filtre matière, la fonction tient en trois lignes – il convient de faire de même pour les autres filtres :

```
1 // Fonction pour mettre à jour le filtre par matériau
2 function updateMateriauFilter(value) {
3     currentMateriauFilter = value;
```

```

4     filterMarkersByDateTypeAndColorAndMateriau(currentYearFilter,
5         currentMateriauFilter, currentColorFilter, currentBaseFilter);
5 }
```

Dernière partie du code, les marqueurs sont la finalisation de tout ce qui précède. Ils font apparaître sur la carte interactive les résultats et ils communiquent les informations et les données correspondantes. Ce sont les marqueurs, d'abord réunis en cluster, qui contiennent dans un premier temps l'info-bulle, puis le carrousel et enfin la barre latérale. À chaque étape de la visualisation, ils délivrent ainsi les renseignements recherchés par l'utilisateur.

```

1 // FONCTION POUR RÉCUPÉRER LES DONNÉES
2 // Fonction pour filtrer les marqueurs par année, matériau, couleur et
   filtre de base
3 function filterMarkersByDateTypeAndColorAndMateriau(yearFilter,
4     materiauFilter, colorFilter, baseFilter) {
5     // Nettoie les anciens marqueurs de la carte
6     markers.clearLayers();
7
8     // Utilisation d'un objet pour suivre les marqueurs parent déjà ajoutés
9     var parentMarkers = {};
10
11    // Parcours de chaque entité (feature) dans les données géographiques (
12        geojson)
13    geojson_RAMA.features.forEach(function (feature) {
14        // Extraction de l'année à partir de la propriété 'Date_filtre' de l'
           entité
15        var dateYear = parseInt(feature.properties.Date_filtre.split("-")[0])
16            ;
17        // Récupération des autres propriétés de l'entité
18        var type = feature.properties.Type;
19        var couleur = feature.properties.Couleur;
20        var parent = feature.properties.Parent;
21        var projet = feature.properties.Projet;
22        var titre = feature.properties.Titre;
23
24        // Nettoyer la valeur du filtre de base en mettant en minuscule et en
           supprimant les apostrophes
25        var cleanedBaseFilter = baseFilter.toLowerCase().replace(/\'/g, '');
26
27        // Vérifier si l'entité est un parent en fonction de son identifiant
28        if (isParent(feature.properties.Identifiant)) {
29            // Appliquer les filtres : année, matériau, couleur et filtre de
               base
30            if (
31                dateYear <= yearFilter &&
32                (materiauFilter === '' || feature.properties.Materiaux.toLowerCase()
33
```

```
30     (.includes(materiauFilter.toLowerCase()) &&
31     (colorFilter === '' || couleur.toLowerCase().includes(colorFilter.
32        toLowerCase())) &&
33     (cleanedBaseFilter === '' || projet.toLowerCase().includes(
34         cleanedBaseFilter))
35   ) {
36     // Vérifier si le parent n'a pas encore été ajouté comme marqueur
37     if (!(parent in parentMarkers)) {
38       // Récupérer les coordonnées de l'entité pour créer un marqueur
39         \indexmot{Leaflet}
40       var coordinates = feature.geometry.coordinates;
41       var marker = L.marker([coordinates[1], coordinates[0]]);
42       var identifiant = feature.properties.Identifiant;
43
44       // Créer le contenu de la fenêtre contextuelle (popup) à
45       // afficher
46       var popupContent = createCarousel(parent, identifiant);
47
48       // Créer et attacher une infobulle (tooltip) au marqueur
49       var tooltip = L.tooltip().setContent(titre);
50       marker.bindTooltip(tooltip);
51
52       // Attacher le contenu de la fenêtre contextuelle (popup) au
53       // marqueur
54       marker.bindPopup(popupContent, {
55         maxHeight: 400
56       });
57
58       marker.on('popupopen', function (e) {
59         updatePopupMaxHeight(e.popup); // Mettre à jour la hauteur
60         // maximale de la fenêtre contextuelle lors de l'ouverture
61       });
62
63       marker.on('popupcontentupdate', function (e) {
64         updatePopupMaxHeight(e.popup); // Mettre à jour la hauteur
65         // maximale de la fenêtre contextuelle lors de la mise à jour
66         // du contenu
67       });
68
69       // Cacher la barre latérale lorsque la fenêtre contextuelle se
70       // ferme
71       marker.on('popupclose', function () {
72         hideSidebar();
73       });
74
75       // Ajouter le marqueur à la couche de marqueurs (markers)
76       markers.addLayer(marker);
77     }
78   }
79 }
```

```
67
68      // Marquer le parent comme ayant été ajouté
69      parentMarkers[parent] = true;
70  }
71  }
72  }
73 );
74
75 // Ajouter la couche de marqueurs à la carte
76 map.addLayer(markers);
```

V. La cartographie interactive et ses biais

Le chapitre précédent ne propose qu'un aperçu du code pour la réalisation de la carte interactive. Je me suis restreint à la seule entité des matériaux, soit une seule colonne parmi les trente-sept du tableau obtenues après le prétraitement avec le langage Python. Cependant, le cas des matériaux permet d'avoir un aperçu non seulement sur la récupération des valeurs des entités, mais aussi sur l'encodage d'un filtre. Si l'exercice de la présentation d'un code peut sembler relever d'un aspect purement technique, il me semble que sa compréhension est essentielle en ce qu'il conditionne lui-même l'interprétation.

La conception et l'encodage de la cartographie interactive façonnent la manière dont les utilisateurs interagissent avec les données et, par conséquent, comment ils les interprètent. Dans un premier temps, il convient de rappeler que les analyses présentées le sont par feuillet et non par manuscrit. Il existe ainsi, à la lecture de la carte, une surreprésentation des matériaux des manuscrits qui ont le plus de feuillets analysés. En outre, plus de soixante pour cent des analyses proviennent de fonds conservés à la Bibliothèque nationale de France.

Le choix et la présentation des filtres conduisent les utilisateurs à s'intéresser prioritairement aux matériaux, soit en lien avec une chronologie, soit avec une couleur. Les récurrences de motifs ou de techniques sont partiellement délaissées, du moins renvoyées à des précisions lors du déploiement de la barre latérale. Cependant, je ne pense pas que l'on puisse évoquer un possible biais ici. Il s'agit simplement de faire droit à ce qui est la raison même du projet sur « La couleur : artefacts, matière et cognition », bien que ce dernier ait pu apporter de nouvelles connaissances par la suite, par exemple sur les attributions des œuvres. En revanche, les filtres sur les matériaux et les couleurs ne proposent des visualisations que sur une seule sélection au sein de leur ontologie ; là où certains chercheurs auraient pu souhaiter questionner la coexistence de certains matériaux au sein d'un même feuillet²².

Il n'y a pas de véritable sélection dans le choix de données à inclure ou à exclure. L'affichage en deux temps que permet la carte interactive, d'abord avec le carrousel, puis avec la barre latérale, permet de retranscrire presque l'intégralité des informations présentes dans les bases materiality⁸⁸ et ⁸⁹. Cependant, l'affichage du carrousel conduit bien à une lecture sur les analyses mêmes, tandis que la barre latérale apporte les informations complémentaires sur le manuscrit. En dernier lieu, les marqueurs et les clusters sont neutres dans leur visualisation. La taille et l'icône ne disent rien, tandis que la couleur vive est réservée logiquement aux clusters avec le plus d'occurrences.

22. Il s'agit d'une motivation supplémentaire pour proposer la visualisation Vikus. Voir le chapitre 3.

*

La carte interactive ne se limite pas à éditorialiser les données ; elle les organise aussi selon des critères géographiques, rendant ainsi intelligible une base de données complexe. Elle offre une visualisation claire de la distribution géographique des enluminures et des panneaux peints analysés. Pour les chercheurs, elle facilite une interprétation principalement spatiale des données, tout en servant de moyen de communication auprès d'un public plus large. En permettant une navigation aisée à travers les données, elle permet de découvrir les œuvres dans leur contexte géographique et historique.

Chapitre 3

Ordonner les données dans le temps, la chronologie interactive

La cartographie interactive permet une éditorialisation des données selon des paramètres essentiellement géographiques, nous l'avons vu, même si d'autres filtres viennent la compléter et l'enrichir. Dans le cadre de mon stage, il m'a semblé ne pas pouvoir me restreindre à cette seule visualisation. L'exploration des résultats des analyses m'a convaincu que, dans le cadre d'une recherche sur les matériaux et les couleurs des enluminures, la chronologie est tout aussi importante que la géographie. Ainsi, je crois que deux œuvres créées à des époques très différentes dans la même région peuvent avoir peu de points communs ; là où des enluminures réalisées à la même époque, mais dans des régions différentes, peuvent parfois révéler des similitudes significatives dans leurs analyses. Il s'agit donc de proposer une interface interactive qui repose non plus sur une interaction entre une géographie et des mots-clés, mais entre ces derniers et une chronologie.

Le programme de recherche « La couleur : artefacts, matière et cognition » présente deux atouts majeurs pour une valorisation scientifique : la richesse tant qualitative que quantitative des analyses menées et la numérisation complète des manuscrits étudiés. Je crois que cette dernière doit être davantage mise en avant dans la conception d'une nouvelle visualisation. En d'autres termes, l'exploration des données doit permettre une plus grande appréciation des enluminures par l'utilisateur, plutôt que de se focaliser uniquement sur les données, parfois un peu arides. La visualisation doit être à la fois chronologique et visuelle, et ce sont ces raisons qui m'ont conduit à m'intéresser au projet *VIKUS Viewer*.

I. VIKUS Viewer : un système de visualisation web avancé

Christopher Pietsch, le principal développeur du projet, présente VIKUS Viewer comme un système de visualisation web avancé qui permet d'explorer des collections culturelles à travers des aspects thématiques et temporels¹. Le logiciel peut organiser des milliers d'artefacts culturels sur une toile dynamique. Il offre une exploration, simple d'utilisation, des motifs thématiques et temporels de grandes collections, tout en offrant un accès rapide à des images en haute résolution. Il appartient au projet *Visualisierung kultureller Sammlungen*, qui a vu le jour à l'Université des Sciences Appliquées de Potsdam dans le cadre d'une recherche interdisciplinaire financée de 2014 à 2017. Son objectif était d'étudier de nouvelles formes d'interfaces utilisateur graphiques pour soutenir l'exploration du patrimoine culturel numérique. Ce projet a abouti à plusieurs propositions de visualisations qui sont visibles sur sa page internet².

Actuellement, VIKUS Viewer est utilisé pour visualiser quelques collections, notamment les dessins de Frédéric-Guillaume IV de Prusse³, la bibliothèque de Johann Wolfgang von Goethe⁴, les pamphlets de la guerre de Sept Ans⁵, les dessins et peintures de Vincent van Gogh⁶ ou encore les pancartes de la Marche des femmes de Boston en 2017⁷. À ma connaissance, il n'y a pas encore d'utilisation connue de VIKUS Viewer en France. Avec le dépôt Github que j'ai réalisé au cours de ce stage, la Bibliothèque nationale de France est probablement l'une des premières à le faire⁸.

Le code derrière VIKUS est librement accessible sur un dépôt Github⁹ – sous une licence MIT – qu'il suffit de cloner et de faire tourner sur un serveur web¹⁰. Il est constitué de HTML, CSS et JS¹¹. Si la documentation est sommaire, les développeurs ont déposé des modèles pour l'élaboration des métadonnées qui peuvent être repris. Concernant la gestion des images, qu'il faut transformer, un lien vers un script est présent¹². Ainsi, si le futur fichier « data », qui contient les fichiers config.json, data.csv, timeline.csv et info.md, est réalisé selon les modèles préexistants en utilisant des transformations Python

1. University of Applied Sciences Potsdam, *VIKUS Viewer : Explore cultural collections along time, texture and themes*, en, URL : <https://vikusviewer.fh-potsdam.de/> (visité le 17/06/2024).

2. Id., *Visualizing Cultural Collections*, URL : <https://uclab.fh-potsdam.de/vikus/> (visité le 26/07/2024).

3. <https://vikusviewer.fh-potsdam.de/fw4/>

4. <https://vikusviewer.fh-potsdam.de/goethe/>

5. <https://vikusviewer.fh-potsdam.de/recs/>

6. <https://vikusviewer.fh-potsdam.de/vangogh/>

7. <https://vikusviewer.fh-potsdam.de/artofthemarch/>

8. https://github.com/TheoBurnel/vikus_bnf.git

9. <https://github.com/cpietsch/vikus-viewer.git>

10. Comme pour la bibliothèque Leaflet, j'ai travaillé sur VSCode avec Live Server.

11. Il faut ajouter les bibliothèques pixi.js et d3.js.

12. <https://github.com/cpietsch/vikus-viewer-script.git>

pour s'en approcher le plus possible ; les images, elles, sont traitées par un script fournit par les développeurs de VIKUS. Elles sont préparées en texture et en feuilles de « sprites », c'est-à-dire en une image qui contient plusieurs plus petites images.

Christopher Pietsch propose en complément à son code une vue alternative. Elle repose sur une mise en page t-SNE, basée sur la similitude de l'image. Un script crée un fichier tsne.csv qui contient un tri des images selon leur ressemblance. Cependant, l'essai que j'ai réalisé à partir d'un petit échantillon ne me semble pas suffisamment convaincant pour le proposer dans le cadre de ce stage. Je propose donc ici deux fonctionnalités de la visionneuse VIKUS : une articulation entre une chronologie et des mots-clefs, ainsi qu'une visualisation en haute résolution des feuillets avec un contenu explicatif.

II. Éditorialiser avec Vikus

VIKUS Viewer rend accessible de grandes collections de documents historiques avec une interface visuelle dynamique et intuitive. Le nombre de feuillets à faire apparaître au sein de la visualisation ne représente aucunement un obstacle, mais, bien au contraire, donne un sens supplémentaire au choix de ce logiciel. Plus le corpus est important, plus la visualisation dynamique gagne en pertinence. Elle permet de quantifier et d'observer les déplacements des feuillets selon les variables retenues, qu'elles soient chronologiques ou par mots-clés.

La capture d'écran ci-dessous représente l'évolution de la visualisation, depuis son lancement par l'utilisateur, en l'absence de filtre, puis avec la sélection du matériau « oxyde de plomb ». Le déplacement des icônes, qui sont des images des feuillets, représente le résultat alors obtenu. L'évolution de la hauteur des colonnes traduit la sélectivité effectuée par le filtre.

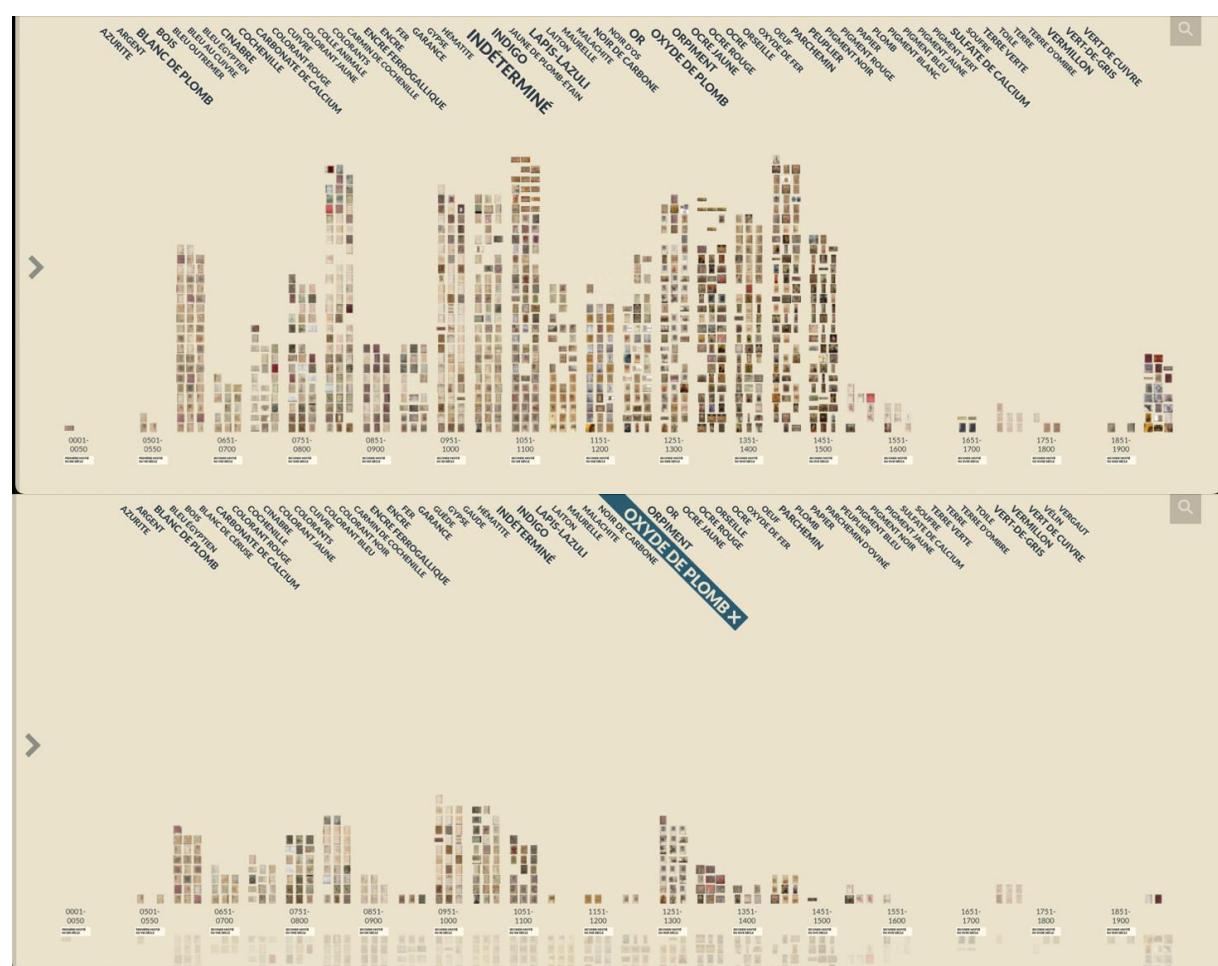


FIGURE 3.1 – Déplacement d’icônes dans Vikus

Le filtrage repose sur deux paramètres, des séquences chronologiques et des thèmes spécifiques. Les données suivent une distribution temporelle dans VIKUS. Cette chrono-

logie doit faciliter la compréhension d'un contexte historique et de tendances au fil du temps. La difficulté, pour un corpus comme le nôtre, réside dans la fragmentation et l'incertitude des datations. Les enluminures des manuscrits ne peuvent être que difficilement datées. Elles font bien souvent l'objet d'une estimation, d'une approximation chronologique afin de proposer un contexte historique vraisemblable. Parfois, l'ordre de grandeur est seulement de quelques années, parfois il est de plusieurs décennies. Dans une frise chronologique classique, ces incertitudes sont contournées par la figuration de jauges, de segment qui indiquent les bornes de la datation. Une telle approche n'est pas possible sur VIKUS. Les datations sont elles-mêmes des entités et pas seulement des valeurs. Elles ne peuvent pas se chevaucher ou se croiser. Chaque datation doit avoir une valeur unique et chaque feuillet doit avoir une seule datation. Dès lors, un certain arbitraire s'impose dans l'assignation d'un contexte de création à une enluminure.

L'autre difficulté réside dans le découpage de la chronologie. Les feuillets apparaissent par colonne au-dessus de leur datation : plus ces dernières seront nombreuses, plus les colonnes seront petites et peu lisibles ; moins les datations seront nombreuses, plus les résultats perdent en sens, diluant les mouvements artistiques dans de grandes ères historiques. Il me semble qu'une répartition par demi-siècle est, au regard du corpus, la meilleure division possible. Elle permet une visualisation lisible et fluide, tout en n'élargissant pas exagérément les séquences chronologiques. Elle a, néanmoins, l'inconvénient de rapprocher artificiellement des productions séparées parfois de près de quatre décennies et d'en séparer d'autres proches de seulement quelques années. Ainsi, sur une même colonne, pourraient figurer des enluminures des années 1350 et 1390, mais elles seraient séparées de celles du début de la décennie 1400. Il y a là un biais propre à VIKUS que je ne peux contourner. Néanmoins, la chronologie de plus d'un millénaire sur laquelle reposent les analyses effectuées permet de lisser ce défaut, lorsque le regard est porté sur un temps long. *Voir la figure 3.2.*

La seconde variable repose sur l'ontologie des matériaux. Comme pour la carte interactive, le matériau est à nouveau le pivot des résultats obtenus. La navigation se fait à travers les mots-clefs, isolés et rangés alphabétiquement, présents dans la partie supérieure de la fenêtre. La taille de ces valeurs varie en fonction de leur occurrence : si un matériau apparaît fréquemment dans le corpus, il aura une police supérieure à celle d'un autre plus rare. Le filtre par matériau proposé dans cette visualisation permet de combiner les valeurs, ce qui n'est pas le cas sur la carte interactive. La recherche s'affine et il est possible d'interroger la co-présence de matériaux au sein d'un même feuillet. Dans le cas ci-dessous, la première recherche est faite sur la seule présence de cochenille, puis elle se précise dans un second temps avec celle d'orseille. Le résultat obtenu est une visualisation très précise sur leur usage commun à partir de la seconde moitié du Xe siècle et qui demeure rare dans le temps. *Voir la figure 3.3.*

Les *artefacts*, comme les appelle Christopher Pietsch, qui sont traités par le logiciel,



FIGURE 3.2 – Les séquences chronologiques dans la frise

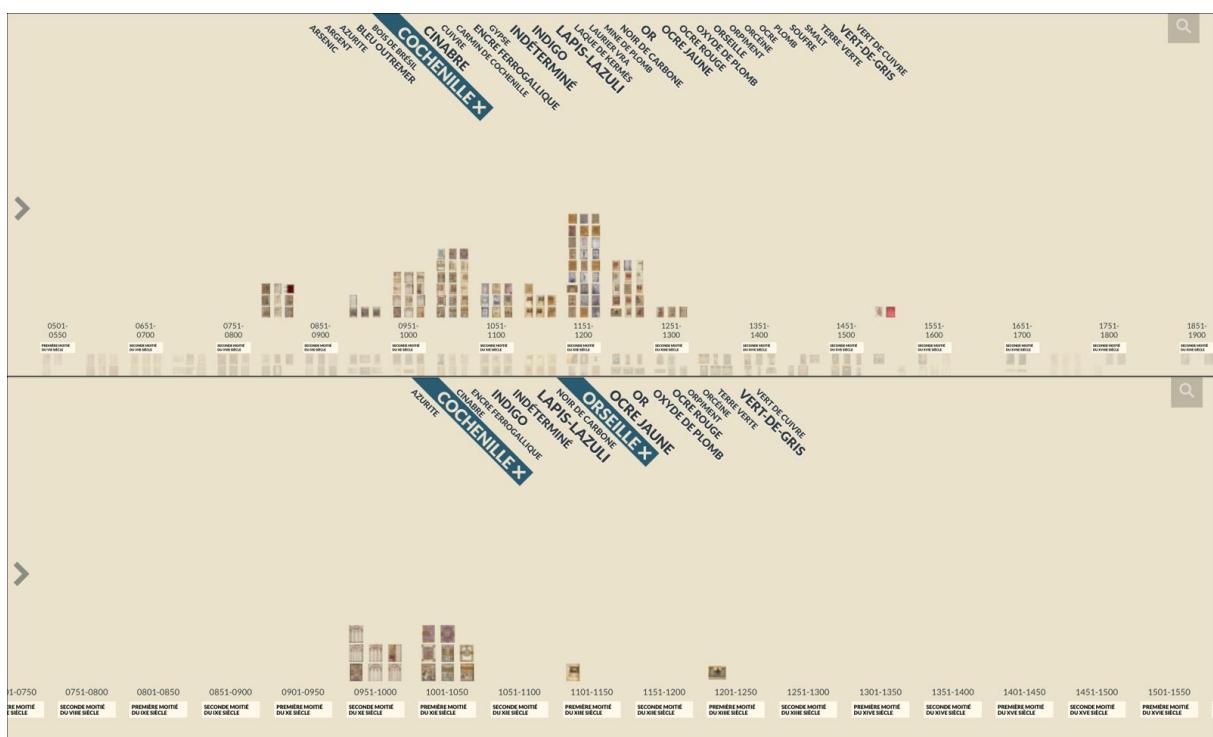


FIGURE 3.3 – Co-présence de matériaux

prennent l'apparence de l'image du feuillet. Elles se déplacent sur l'ensemble de l'interface selon le paramétrage du filtre. Avec un niveau de zoom suffisant ou avec un clic, l'interface fait apparaître une barre latérale à côté de l'image. Elle reprend les informations essentielles du feuillet. J'y fais figurer le titre de l'œuvre, les matériaux qui sont utilisés pour sa réalisation, les résultats des différentes analyses, son lieu de création ainsi que sa localisation actuelle. L'utilisateur a à sa disposition toutes les informations essentielles pour appréhender le manuscrit.

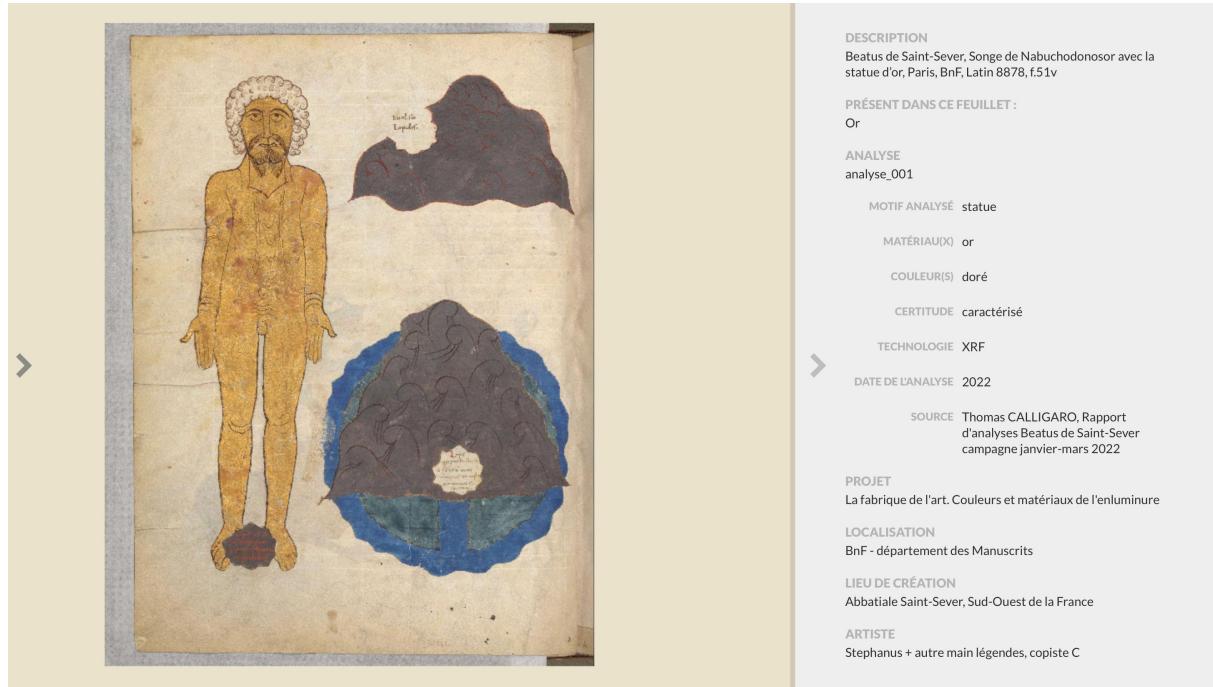


FIGURE 3.4 – Présentation des données sous Vikus

VIKUS Viewer permet de travailler à partir de toutes les données entrées dans les deux bases du projet, materiality88 et 89. Celles-ci sont accessibles dans le champ supérieur de la fenêtre, qui permet la recherche en texte libre, que ce soit par le nom du manuscrit, de la couleur, du lieu de création ou de toute autre information. Si la visualisation nécessite un certain appauvrissement de la chronologie, toutes les autres données peuvent être encodées et reportées sans difficulté. Comme pour la cartographie interactive, l'essentiel de la transformation repose sur le langage Python.

III. Traitement des données

La visualisation repose sur un fichier HTML qui est relié à des dossiers contenant du code JavaScript et du CSS¹³. À cette architecture, il faut produire et ajouter un dossier « data », qui est lu par les codes JavaScript et qui contient toutes les données qui sont traitées. La réalisation du dossier « data », nous l'avons vu, se fait essentiellement à partir d'une imitation des exemples proposés. Dans le cas présent, je me suis inspiré du modèle réalisé pour les peintures de Van Gogh.

Le travail repose sur la création de quatre fichiers de métadonnées qui décrivent la collection et les objets, et qui configurent la visualisation. Il convient d'ajouter à ces fichiers une collection d'images, qui représentent les artefacts. Le premier fichier se nomme config.json. Il s'agit du fichier de configuration qui définit le nom du projet, les lien de données, les colonnes des artefacts, les styles et le contenu de la barre latérale qui détaille la collection.

Le data.csv contient toutes les informations de métadonnées pour chaque objet de la collection. Il a des champs obligatoires qui permettent de faire le lien avec les autres données. Les lignes doivent contenir au moins un identifiant, un mot-clef et une année. L'identifiant doit correspondre au nom de l'image qui sera produite par la suite. Par exemple, dans notre cas, l'identifiant fd106807-3864-4227-87ef-1ff82988616d correspond à l'image des *Épîtres de saint Paul* (lettre ornée, Paris, BnF, Latin 10440, f.20r, f.20v). La complexité de l'identifiant peut surprendre, mais elle correspond en réalité à une simplification des données. Pour la préparation du fichier data.csv, je m'appuie sur le fichier préparé pour la cartographie. Afin de faire apparaître les résultats par feuillets, et non par analyses, je réalise un pivot de ces dernières sur chaque première ligne des feuillets, tout en les ordonnant selon une nouvelle numérotation. Ensuite, il suffit de procéder à un téléchargement des images en leur assignant le numéro de l'identifiant comme nom. En ce qui concerne les mots-clés, qui portent sur les matériaux, une concaténation des différentes analyses des feuillets, toutes présentes maintenant sur les mêmes lignes, permet de constituer la colonne. Il convient de réaliser un nettoyage des entrées ensuite, afin qu'elles puissent être correctement reconnues par le code JavaScript. Ensuite, la colonne 'year', elle aussi obligatoire, correspond à l'approximation par demi-siècle que je dois réaliser afin de proposer une visualisation intelligible. Les autres colonnes du fichier data.csv correspondent à différents champs de métadonnées qui peuvent être personnalisés. Ici, nous retrouvons les informations qui correspondent au titre du feuillet, aux différentes analyses, aux lieu de création et de conservation, ainsi qu'au projet de recherche auquel il se rattache.

13. Le clonage du dépôt GitHub inclut également un dossier « font » pour les polices et un « image » pour les icônes utiles pour la mise en page.

Le timeline.csv contient les informations pour la chronologie affichée sous les années. Si l'année peut être un nombre ou une chaîne de caractère, elle correspond ici aux demi-siècles de la colonne ‘year’. Ensuite, le fichier doit contenir d’autres renseignements : un titre de présentation sous les séquences chronologiques, un texte détaillé lorsqu'il y a un premier niveau de zoom, puis un texte supplémentaire lorsqu'il est zoomé au maximum. Ces éléments doivent contribuer à fournir un contexte historique à l’utilisateur lors de son exploration de la base. Cependant, avec une chronologie *artificielle* comme la nôtre, il n'est pas évident d'apporter des compléments d'informations sous les séquences chronologiques. J'ai simplement opté pour le fait de mettre en toutes lettres la période considérée ; ainsi la colonne *0951-1000* est accompagnée de la précision *Seconde moitié du Xe siècle*.

Enfin, le dernier fichier de métadonnées, info.md concerne les informations affichées sur le côté gauche lors de l’ouverture de la visualisation. D'une écriture simple en Markdown, j'y fais apparaître une présentation simple et concise du projet de recherche et de la prise en main de l'interface.

À côté de ces quatre fichiers, la visualisation repose sur la préparation et la configuration d'une collection d'images. Ces dernières doivent être téléchargées dans un seul dossier et nommées selon l'identifiant qui leur est assigné dans le fichier data.csv. Le traitement consiste en l'application du script vikus-viewer qui génère des sprites et des textures pour les différents niveaux de zoom.

En premier lieu, il faut cloner le dépôt et installer les dépendances :

```
git clone https://github.com/cpietsch/vikus-viewer-script cd vikus-viewer-script  
npm install
```

Ensuite, le script est exécuté pour générer les textures et les sprites :

```
vikus-viewer-script « /chemin/vers/les/images/*.+(jpg|jpeg|png) »
```

Le script crée un dossier de données pour les textures (1024 et 4096) ainsi qu'un dossier de sprites pour les spritesheets à l'intérieur du dossier actuel.

Les créations doivent être copiées à l'intérieur du dossier data, à côté des autres fichiers présentés précédemment.

Une fois que tous les liens sont bien vérifiés dans le config.json, il est possible de débuter la visualisation avec un serveur web sur le fichier index.html.

IV. Chronologie et interprétation

L'un des principaux avantages de l'utilisation de la chronologie est qu'elle permet aux utilisateurs de comprendre les objets dans leur contexte historique. Les manuscrits sont, dans le cas présent, ordonnancés selon l'époque à laquelle ils ont été créés. La géographie, les sujets iconographiques, les artistes et même les couleurs ne jouent ici aucun rôle dans le classement. C'est une distribution temporelle stricte. Il s'agit, avec VIKUS, de donner la primauté à une analyse historique du corpus. Nous pouvons en proposer quelques exemples.

Si nous revenons au cas de notre bleu égyptien déjà évoqué au cours de ce mémoire, la frise chronologique fait apparaître trois séquences à son utilisation dans les enluminures. La première est datée de la seconde moitié du VI^e siècle, avec le *Pentateuque dit d'Ashburnham*. Une deuxième séquence est très circonstanciée à la fin du VIII^e siècle, dans l'*Évangéliaire de Charlemagne*. Sa dernière apparition est au XI^e siècle, principalement dans le *Beatus de Saint-Sever*. Ces trois temps apparaissent sur la frise, sans qu'ils soient rapportés à d'autres variables que la date de réalisation et le matériau. La chronologie, plus que toute autre forme de visualisation, permet de représenter les discontinuités temporelles de l'emploi du bleu égyptien. Les interruptions et les reprises successives dans l'utilisation du matériau apparaissent et invitent le chercheur à s'interroger sur les raisons artistiques, culturelles et matérielles de ces retours et disparitions au cours du temps.

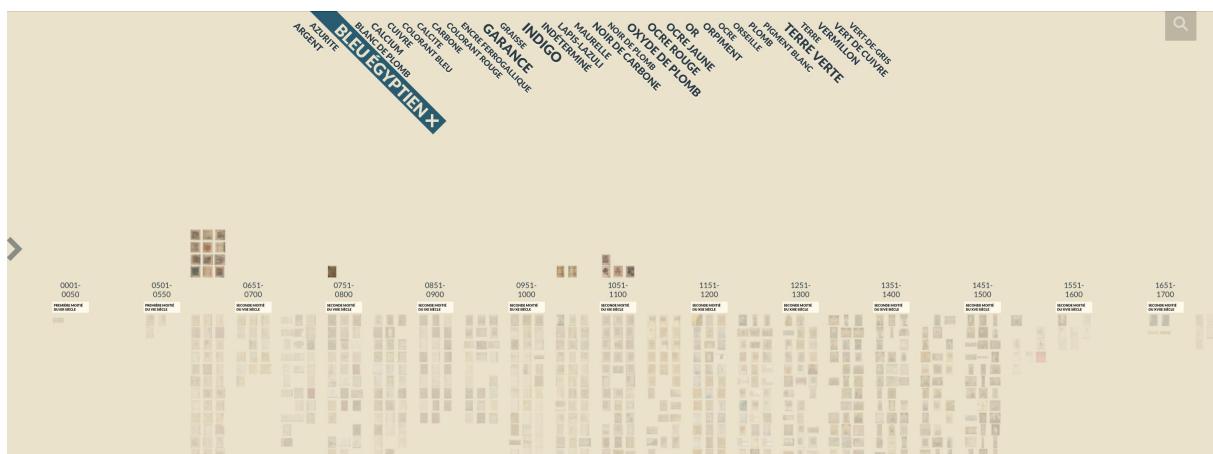


FIGURE 3.5 – Distribution temporelle du bleu égyptien

La chronologie permet d'observer les évolutions et les tendances au fil du temps. Le deuxième exemple se sert de la recherche en texte libre que propose VIKUS viewer pour porter un regard sur les contextes de production des enluminures. Ce champ permet d'interroger l'ensemble des données du fichier data.csv construit précédemment. En visualisant les occurrences du mot *abbaye*, le résultat fait apparaître une longue séquence chronologique allant du VII^e siècle au XI^e, avec une petite reprise aux alentours de l'année 1400. Une première hypothèque qui pourrait être formulée est celle de la mise en évidence,

par la chronologie, d'un déplacement dans le contexte de production du corpus analysé. Je n'évoque qu'une hypothèse, la base de données n'étant pas suffisamment fournie pour parvenir à une conclusion, mais la visualisation invite à porter notre regard vers d'autres collections, afin de constater une même délocalisation de la production artistique, une même sortie des abbayes vers le monde.

La visualisation reproduite ci-dessous rappelle que la chronologie est aussi un discours sur le temps¹⁴. Les séquences de production des enluminures que nous avons définies, par demi-siècle, renvoie peut-être trop la chronologie à un instrument taxinomique¹⁵. Avec une étude plus poussée des contextes de production de ces enluminures, il serait alors possible de proposer des chronologies-typologies sur VIKUS, comme des emboîtements thématisés. Serait-il dès lors possible de recomposer un récit temporel et de concevoir comme séquence chronologique un *temps de production abbatiale* par exemple ?

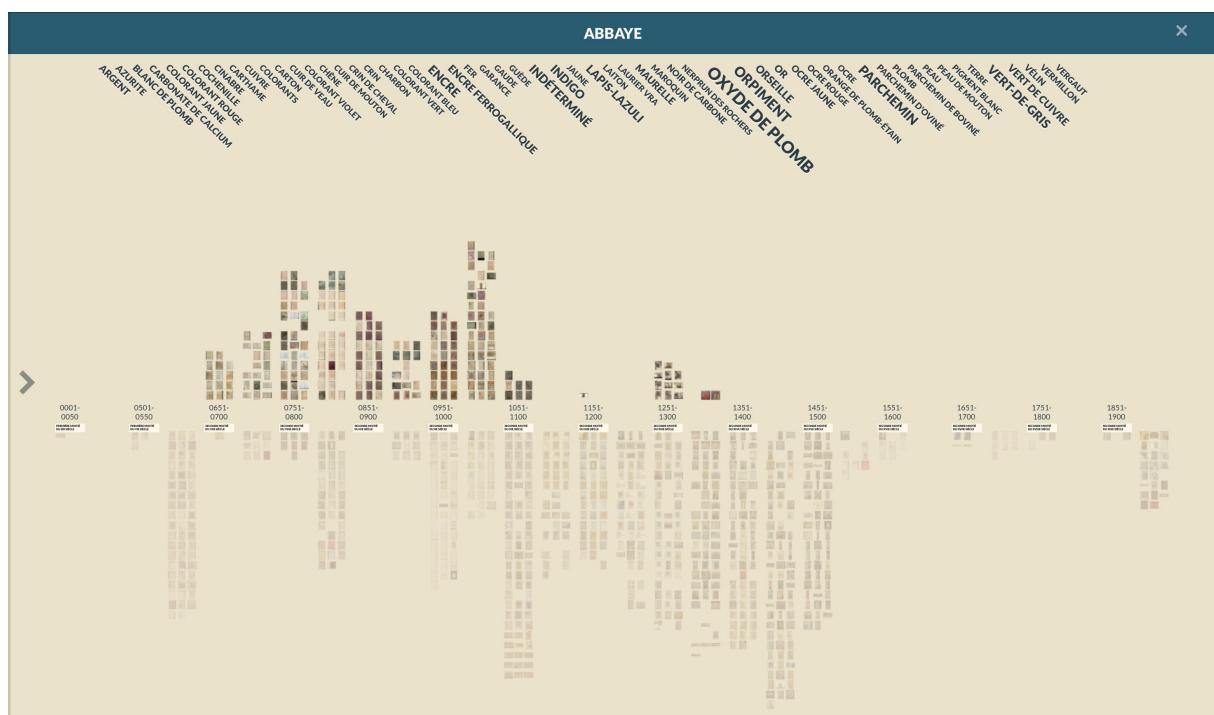


FIGURE 3.6 – Chronologie de la production abbatiale

14. Dimitri Toubkis, « L'ordre de la chronologie », *Hypothèses*, 7–1 (2004), Place : Paris Publisher : Éditions de la Sorbonne, p. 133–145, DOI : 10.3917/hyp.031.0133.

15. Michel de Certeau, *L'écriture de l'histoire*, Paris, 1975.

*

Ces exemples sont des parfaites illustrations des vertus heuristiques de certaines formes de visualisation évoquées au cours de ce mémoire. Le brassage des données, selon des critères tantôt géographiques, tantôt chronologiques, permet l'observation de disjonctions et de continuités. Ces observations sont à la source de nouvelles analyses pour les chercheurs.

En somme, l'utilisation d'une chronologie enrichie par des mots-clés, à l'image de celle proposée par VIKUS Viewer, permet une exploration plus approfondie et contextuelle des données de la recherche. Elle offre une compréhension temporelle du corpus, elle met en lumière les évolutions et les tendances historiques, et elle facilite la comparaison selon des thèmes au fil du temps, ce qui peut être particulièrement précieux pour les chercheurs, les historiens et le grand public.

Chapitre 4

Annotations sémantiques d'images

Les deux projets de recherche, « La couleur : artefacts, matière et cognition » et « La fabrique matérielle du visuel : panneaux peints en Méditerranée », ont chacune des entrées de leur base de données reliée à une image numérique. Il s'agit d'une photographie, en haute résolution, du feuillet de manuscrit ou du panneau qui est l'objet des analyses. Il y en a plus de mille cent. C'est un atout majeur du projet : elles permettent non seulement de faire connaître au grand public des chefs-d'œuvre conservés dans les manuscrits, mais également de servir de supports aux analyses pour les chercheurs.

Le musée d'art et d'antiquités de l'université de Cambridge, le Fitzwilliam Museum, a proposé une valorisation de ses manuscrits enluminés à l'adresse suivante : <https://www.fitzmuseum.cam.ac.uk/illuminated/>. À partir d'une petite sélection, les équipes du musée présentent les enluminures en haute résolution, avec des métadonnées descriptives et des points d'analyse sous forme d'annotations. Il me semble que le travail réalisé ici peut servir de modèle pour d'autres projets de recherche, comme ceux qui nous concernent aujourd'hui. Il est, d'une certaine façon, question de faire au cours de ce stage une démarche similaire avec des manuscrits sélectionnés par le département des manuscrits.

I. Un fac-similé numérique

L'idée est de travailler à partir des images numérisées des feuillets et de proposer des annotations qui pourront servir aux chercheurs. Une annotation est une information supplémentaire ajoutée à l'image, soit pour la décrire, soit pour y apporter des informations. Dans notre cas, il s'agit d'y renseigner les données de la base sur les analyses réalisées. Ces annotations se superposent à des images qui s'inscrivent dans un ensemble de protocoles, appelé IIIF.

L'International Image Interoperability Framework (IIIF) a été développé pour répondre à la nécessité d'un accès standardisé et interopérable aux images numériques dans le domaine des bibliothèques, archives et musées¹. L'idée de ce consortium², fondé au cours de la décennie précédente, est de changer la manière dont les images numériques sont partagées et utilisées. Les images seront numérisées une seule fois en haute résolution par une institution, et elles seront déposées sur une plateforme, d'où elles pourront être appelées par n'importe quel utilisateur, selon son besoin spécifique. Ainsi, pour ce qui est des images des manuscrits que nous souhaitons valoriser, il suffit d'appeler les images numériques présentes sur la plateforme Gallica. L'interopérabilité permet de ne plus se soucier du processus de numérisation, déjà assuré par d'autres services.

Avec les protocoles qui définissent IIIF, les images sont non seulement consultables par n'importe quelle application ou logiciel compatible, mais elles sont également manipulables et annotables³. Pour ce faire, il faut passer par des interfaces de programmation d'application (API). Dans le cadre de IIIF, ces dernières sont doubles. Elles concernent aussi bien l'image en elle-même, appelée l'API image, que sa présentation, c'est-à-dire avec ses métadonnées, appelée l'API présentation.

L'API image permet d'appeler les pixels d'une image et de manipuler cette dernière à distance à travers une syntaxe d'URL standardisée. Le modèle de toute API image suit ce schéma :

schéme ://server/prefix/identifier/region/size/rotation/quality.format

Pour les images que j'appelle sur Gallica pour les manifestes du projet, elles deviennent par exemple :

<https://gallica.bnf.fr/iiif/ark:/1214/btv1b6000718s/f1/full/full/0/native.jpg>

1. *International Image Interoperability Framework*, en, juin 2024, URL : <https://iiif.io/> (visité le 26/07/2024).

2. Parmi les premières institutions à travailler sur les protocoles IIIF, nous pouvons citer la Bibliothèque Bodleian de l'Université d'Oxford, la Bibliothèque de l'Université de Stanford ou encore la British Library.

3. La présentation de IIIF qui suit doit beaucoup à deux sites internet de vulgarisation. Le premier est le workbench de Glen Robson, le second est la documentation proposée Régis Robineau pour Biblissima Glen Robson, *IIIF Online Workshop*, URL : <https://training.iiif.io/iiif-online-workshop/> (visité le 04/07/2024) ; Biblissima, *Qu'est-ce que IIIF ?*, URL : <https://iiif.biblissima.fr/> (visité le 04/07/2024).

Il est dès lors possible de changer de région, de taille, de rotation, de qualité et de format selon les besoins spécifiques.

De son côté l'API présentation est décrite par Régis Robineau comme celle qui « spécifie les métadonnées (descriptives, structurelles, techniques) nécessaires à la présentation d'un objet numérique dans une interface. Toutes ces informations sont contenues dans un fichier appelé "manifeste", une sorte d'enveloppe virtuelle formant l'unité de distribution élémentaire dans l'univers IIIF⁴. C'est en général ce fichier que vont manipuler les logiciels pour interagir avec une ressource, la visualiser, ou la transférer vers un autre outil. »

Un manifeste IIIF peut être un fac-similé numérique d'un objet physique (livre, manuscrit, périodique, carte, peinture, photographie, partition, monnaie, objet archéologique, archive sonore, captation vidéo, etc.), mais aussi un objet virtuel et composite constitué d'une série d'images ou d'autres médias rassemblés à des fins scientifiques ou pédagogiques, et pouvant provenir de différentes collections⁵. Dans le cas présent, les manifestes produits pour le stage s'apparentent à la première catégorie. Il s'agit d'un fac-similé numérique de manuscrit avec quelques annotations supplémentaires.

Le schéma ci-dessous, qui est proposé dans la documentation Biblissima, reprend les principales composantes d'un manifeste IIIF et permet de s'en faire une idée plus juste.

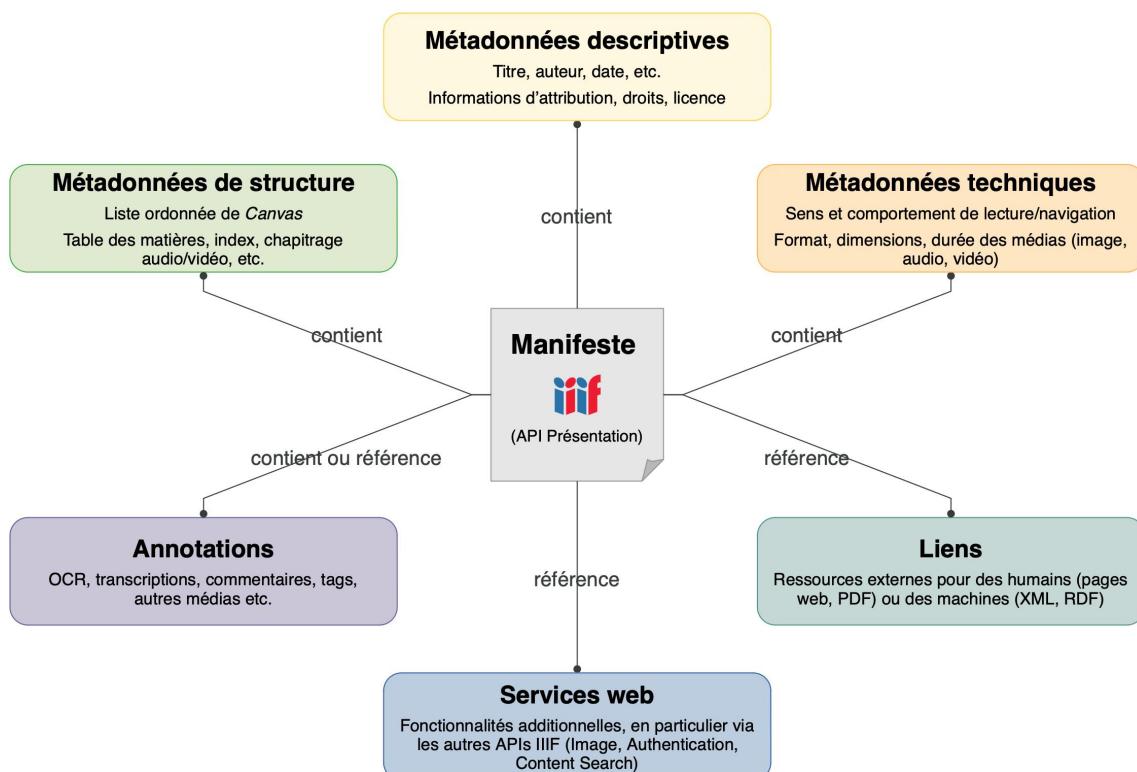


FIGURE 4.1 – Structure d'un manifeste IIIF

4. <https://doc.biblissima.fr/iiif/introduction-iiif/#principe>
5. <https://doc.biblissima.fr/iiif/introduction-iiif/#principe>

Les manifestes peuvent être affichés avec des visualiseurs IIIF. Ils proposent à l'utilisateur une expérience de visualisation riche, s'appuyant sur l'ensemble des informations contenues dans les manifestes, au format JSON-LD. Les deux images ci-dessous permettent d'illustrer où sont mobilisées les deux API image et présentation dans un visualiseur comme Mirador 2.

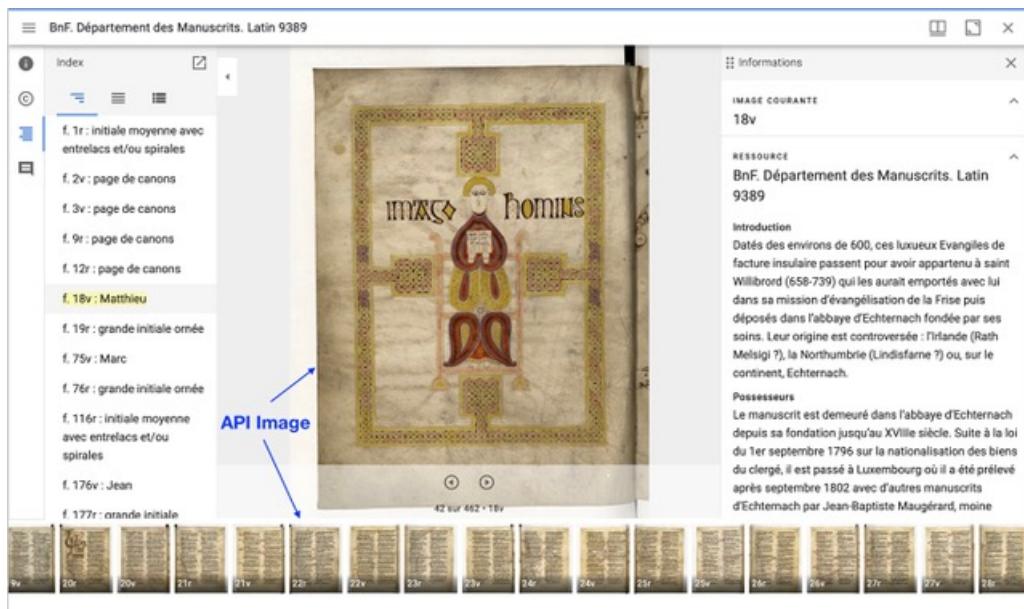


FIGURE 4.2 – API Image

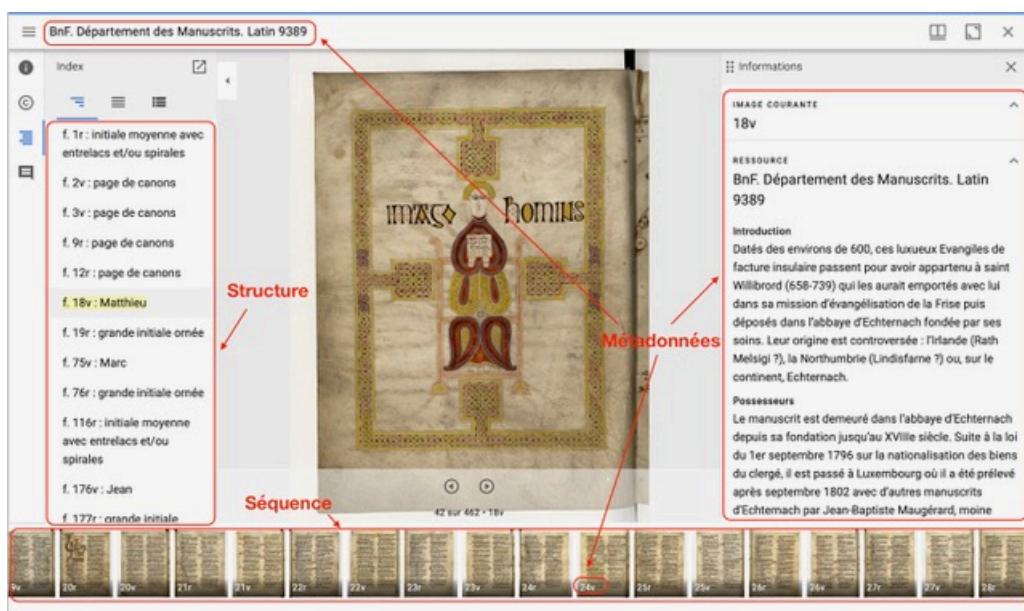


FIGURE 4.3 – API Présentation

II. Enrichir une image numérique

Les manifestes édités sont, je l'ai écrit précédemment, en quelque sorte un fac-similé numérique de manuscrit, avec quelques annotations supplémentaires. L'édition doit se faire au plus près de la version d'origine du manuscrit. Pour ce faire, et pour nos manuscrits, le point de départ est la bibliothèque numérique Gallica. La plateforme Gallica contient non seulement la numérisation de millions de documents de la Bibliothèque nationale de France, mais elle donne également accès à une API qui permet de récupérer des manuscrits numérisés avec leur manifeste IIIF. Ces derniers permettent de récupérer tous les feuillets du document en haute résolution, ainsi que ses métadonnées descriptives et de séquences. Si je prends l'exemple du manuscrit Arabe 5847, Gallica propose l'intégralité de sa numérisation (avec les pages de couverture, pages de garde, contreplats et même le dos) dans le manifeste <https://gallica.bnf.fr/iiif/ark:/12148/btv1b8422965p/manifest.json>. Ce dernier contient également toutes les séquences du manuscrit, un titre pour tous les feuillets renseignés, qu'ils soient recto ou verso. Cette récupération auprès de Gallica permet de disposer du manuscrit Arabe 5847 en haute résolution et de travailler à la personnalisation du manifeste à partir de 9 292 lignes de JSON-LD déjà écrite par la plateforme.

Pour personnaliser le fac-similé numérique, il faut lui attribuer un nouvel identifiant unique. Chaque manifeste ne peut avoir qu'un seul identifiant. Pour différencier le travail qui sera réalisé sur le manuscrit de ce que propose Gallica, il convient donc de le rééditer sous un autre nom. La Bibliothèque Bodleian de l'Université d'Oxford met à la disposition du public un outil en ligne pour aider à la réédition. Il se trouve à l'adresse : https://digital.bodleian.ox.ac.uk/manifest-editor/#/?_k=xjzu5x. L'éditeur permet de travailler soit à partir d'un nouveau manifeste, soit de reprendre un préexistant en le modifiant et en lui attribuant un nouvel identifiant. Dans notre cas, il suffit d'importer le canevas du manuscrit qui est renseigné sur Gallica. Ensuite, il est possible de personnaliser les métadonnées directement sur le site internet. Si les métadonnées descriptives de notre fac-similé numérique sont proches de celles de la plateforme de la BNF quant au titre, son introduction, son historique ou encore sa description technique, elles diffèrent au moment de les relier aux deux projets de recherche. Un paragraphe sur les matériaux utilisés dans le manuscrit est ainsi ajouté. En conséquence, lorsqu'une visionneuse IIIF interprétera le manifeste, il y aura, parmi les informations personnalisées, un premier résumé des résultats de la recherche.

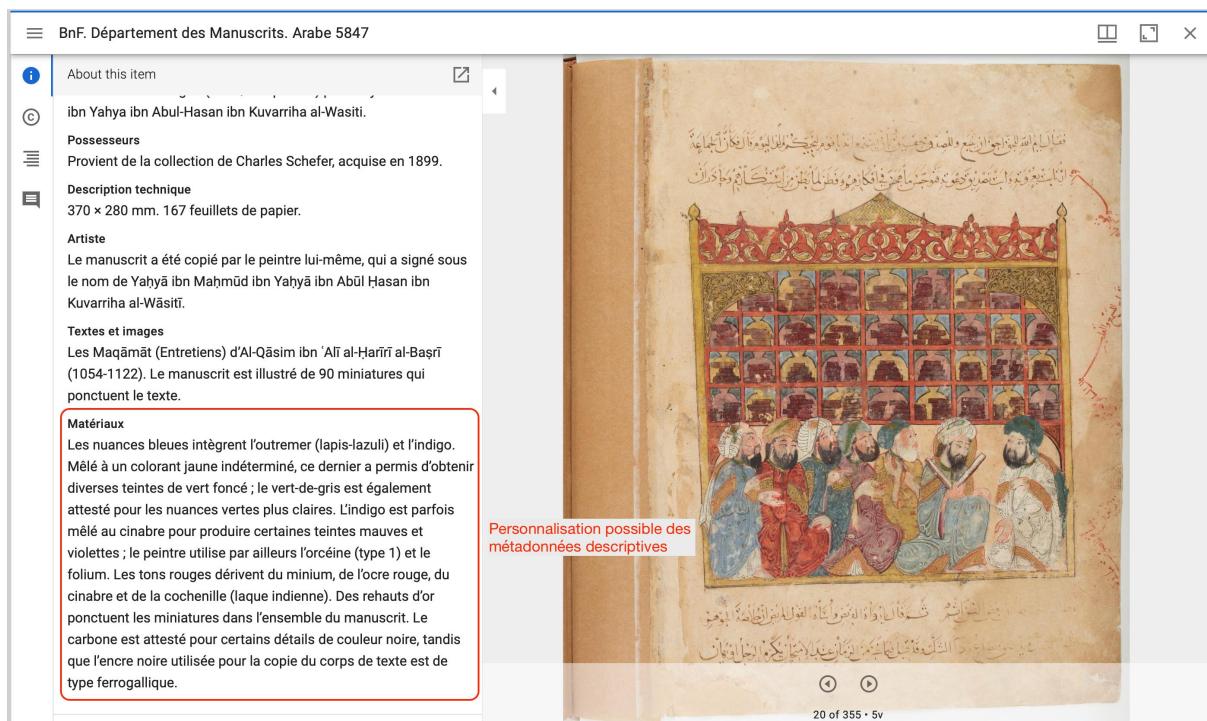


FIGURE 4.4 – Personnalisation des métadonnées descriptives

Le fac-similé numérique peut être enrichi d'annotations⁶. Ces dernières peuvent être invisibles, n'altérant pas la première lecture du manuscrit, ou visibles, permettant d'apporter des compléments d'informations. Les enrichissements possibles sur une image numérique sont nombreux. Il peut aussi bien s'agir de la transcription d'un document, d'un commentaire ou d'une analyse de contenu, ou encore de la mise en évidence d'une zone. Dans notre cas, il s'agit d'y renseigner les données de la base sur les analyses réalisées.

Les annotations IIIF suivent le Web Annotation Data Model⁷. Elles sont composées de deux parties. L'une renseigne le contenu de l'annotation, ce qui sera reporté sur l'image numérique ; l'autre indique la cible de l'annotation, le manifeste dont il est question et sa partie concernée par le commentaire. Une annotation simple a une structure standardisée composée de six points.

Structure d'une annotation :

1. @context : il spécifie le contexte JSON-LD utilisé pour interpréter les données d'annotation.
2. id : un identifiant unique pour l'annotation.
3. type : le type de l'annotation, généralement il suffit d'indiquer "Annotation".

6. Il existe actuellement une dizaine de serveurs d'annotations IIIF. Les principales possibilités sont listées à l'adresse suivante : <https://github.com/IIIF/awesome-iiif?tab=readme-overview#annotation-servers>. Dans le cadre de ce stage, les annotations sont réalisées avec SimpleAnnotationServer (SAS), un serveur écrit en Java et développé par Glen Robson (<https://github.com/glenrobson/SimpleAnnotationServer>).

7. *Web Annotation Data Model*, en, févr. 2017, URL : <https://www.w3.org/TR/annotation-model/> (visité le 26/07/2024).

4. motivation : la raison derrière l'annotation (par exemple, « commentaire »).
5. target : la ressource ou la partie de la ressource qui est annotée.
6. body : le contenu de l'annotation.

De façon simplifiée, parce que celles réalisées sont finalement un peu plus complexes, une annotation sur le manuscrit Arabe 5847 peut répondre au schéma suivant :

```
1  {
2      "@context": "http://iiif.io/api/presentation/2/context.json
3          ",
4      "@id": "http://localhost:8888/annotation/1719318035104",
5      "@type": "oa:Annotation",
6      "motivation": ["oa:commenting"],
7      "on": {
8          "full": "https://gallica.bnf.fr/iiif/ark:/12148/
9              btv1b8422965p/canvas/f34",
10         "@id": "http://744902fa-f92b-4edf-b593-f9260051c14d"
11     },
12     "resource": {
13         "chars": "<p>Motif : v&ecirc;tement, chausse dextre</p>
14             <p>Couleur : rouge</p>
15             <p>Mat&eacute;riaux : ocre
16             rouge</p>"
```

Ces annotations permettent des reports d'informations. Elles enrichissent les fac-similés numériques avec le travail réalisé lors des projets de recherche. Les données de la base, qui concernent les résultats sur les couleurs et les matériaux, peuvent être reportées et figurer directement sur les feuillets. Ainsi, le chercheur a sous les yeux les correspondances, sans avoir besoin de passer d'un document à l'autre. Mais un manifeste IIIF a encore d'autres possibilités d'enrichissement qui, combinées avec les annotations, permettent de croître encore l'intérêt d'un fac-similé numérique. L'appel d'un feuillet de manuscrit pour son affichage peut être inclus dans un système de calques. Autrement dit, d'autres images numériques sont superposables à la première appelée.

La Bibliothèque du Congrès propose un bel exemple de l'intérêt d'une combinaison de calques avec des images IIIF. Lorsqu'elle a récemment numérisé les papiers d'Alexander Hamilton, une lettre à Elizabeth Schuyler (datée du 6 septembre 1780, deux mois avant leur mariage), non éditée, avec quatorze lignes effacées par l'auteur, est devenue visible pour la première fois. Pour savoir ce qui se trouvait sous les rayures, la Bibliothèque du Congrès a utilisé l'imagerie hyperspectrale. Une analyse non invasive qui utilise la lumière à différentes longueurs d'onde pour capturer des informations non visibles à l'œil. Cette

technologie a permis de révéler le contenu. La bibliothèque a réalisé un manifeste IIIF de la lettre, pour permettre aux utilisateurs d'en découvrir la teneur par eux-mêmes en jouant sur le niveau d'opacité des calques⁸.

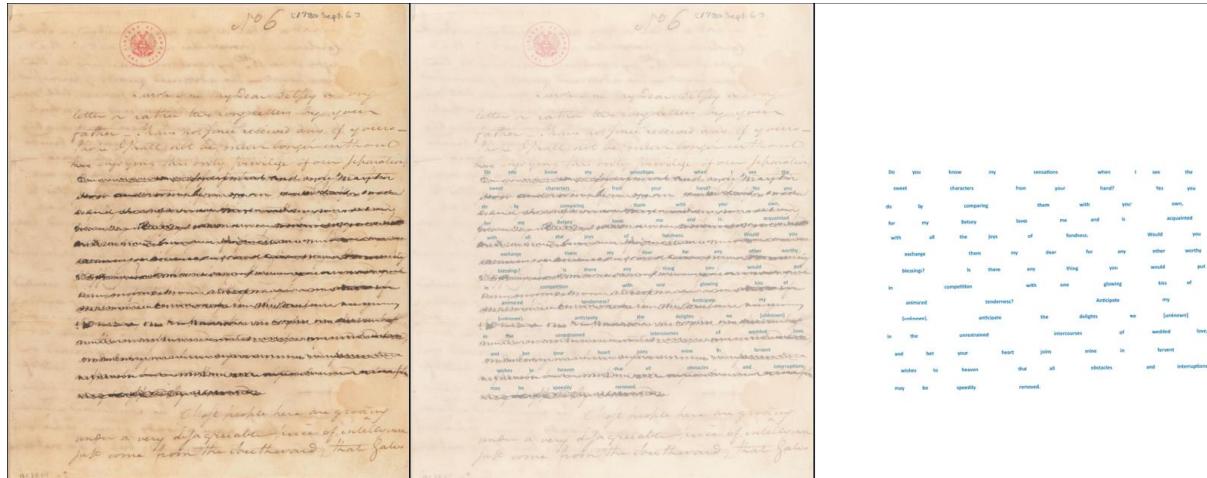


FIGURE 4.5 – Lettre d'Alexander Hamilton

Le projet de recherche « La couleur : artefacts, matière et cognition » possède également des images numériques issues d'analyses hyperspectrales. Les laboratoires en charge des rapports d'analyses physico-chimiques ont fourni, au moment de leurs résultats, des doubles numériques au commanditaire. Il est ainsi possible d'envisager, à l'image de ce qu'a proposé la Bibliothèque du Congrès, de réaliser des claques sur les feuillets de manuscrit, permettant de découvrir les images hyperspectrales correspondantes et donc les résultats.

8. <https://dvp.prt.d.app/hamilton/manifest.json>

III. La sélection du département des manuscrits

Le fac-similé numérique produit est considérablement enrichi au regard de simples reproductions numériques d'images habituellement disponibles. Douze manuscrits ont fait l'objet d'une transformation en manifeste IIIF⁹. Ils sont retenus en raison de la qualité exceptionnelle de leurs enluminures. Une fois les sites Omeka S accessibles au public, ils constitueront une petite sélection pour les utilisateurs, un modèle du travail effectué au cours des années du projet de recherche.

1. Le Psautier de Paris (BnF. Département des Manuscrits. Grec 139)
2. L'Evangéliaire dit de Charlemagne ou de Godescalc (BnF. Département des Manuscrits. NAL 1203)
3. Le Beatus a Liebana, Commentarius in Apocalypsin (BnF. Département des Manuscrits. Latin 8878)
4. Les Evangiles dits d'Echternach ou de Saint Willibrord (BnF. Département des Manuscrits. Latin 9389)
5. L'Ancien Testament et quelques feuillets du Nouveau Testament en syriaque (BnF. Département des Manuscrits. Syriaque 341)
6. Le Tétravangile (BnF. Département des Manuscrits. Copte 13)
7. Les Makamat de Hariri (BnF. Département des Manuscrits. Arabe 5847)
8. Le Codex Xolotl (BnF. Département des Manuscrits. Mexicain 1-10)
9. L'Évangéliaire dit de Poussay (BnF. Département des Manuscrits. Latin 10514)
10. Le Codex Sinopensis (BnF. Département des Manuscrits. Supplément Grec 1296)
11. Le Psautier dit de saint Louis (BnF. Département des Manuscrits. Latin 10525)
12. Un recueil d'œuvres en prose et en vers, probablement toutes dues à Nasīr al-Dīn Muḥ. b. Ibrāhīm b. ḤAbd-ullāh al-Rammāl al-Muazzim al-Sācatī al-Haykalī (BnF. Département des Manuscrits. Persan 174)

Les manuscrits sont tous enrichis de métadonnées descriptives. Ils offrent une première approche avec le manuscrit dans une courte introduction, puis quelques phrases qui retracent son historique, suit ensuite une petite fiche technique, puis un commentaire sur le texte et les images, une note sur le ou les artistes et, enfin, quelques grandes lignes sur les matériaux découverts lors des analyses.

Ci-dessous, l'exemple de la présentation du manuscrit Latin 10514 telle que visible

9. Les liens des manifestes IIIF sont reportés en annexe.

dans Mirador :

LIBELLÉ :

BnF. Département des Manuscrits. Latin 10514

DESCRIPTION :

Évangéliaire dit de Poussay

INTRODUCTION :

Réalisé dans l'abbaye de Reichenau à la fin du Xe siècle, vers 980, l'Évangéliaire de Poussay fait partie du groupe des manuscrits dit « de Ruodprecht », d'après le copiste du Psautier d'Egbert, également originaire de Reichenau. Il s'agit d'un évangéliaire festif contenant des enluminures sur fond pourpre. Sa destination est incertaine, mais il a été offert vers 1036 à l'abbaye de Poussay par Brunon d'Eguisheim-Dagsbourg, alors évêque de Toul et devenu pape entre 1049-1054 sous le nom de Léon IX.

POSSESSEURS :

Trésor de l'abbaye de Poussay jusqu'à la Révolution française (?) ; bibliothèque de la Ville de Mirecourt ; acquis en 1844 par la Bibliothèque royale contre des livres imprimés.

DESCRIPTION TECHNIQUE :

285 x 205 mm 133 feuillets de parchemin

TEXTES ET IMAGES :

13 peintures en pleine page : f. 3v Ecclésiastique offrant son livre, conduit par 2 anges ; f. 4r Christ trônant, main étendue vers le donateur ; f. 5v saint Luc inspiré par le bœuf ; f. 6r saint Marc inspiré par le lion ; f. 7v saint Jean examinant sa plume ; f. 8r saint Matthieu écrivant ; f. 9v Nativité ; f. 18v Adoration des mages ; f. 35v crucifixion ; f. 46v le lavement des pieds ; f. 50v deux saintes femmes au Tombeau ; f. 66v Ascension ; f. 69v Pentecôte ; 16 pages d'initia ornées : f. 10r ; f. 13r ; f. 19r ; f. 25r ; f. 36r ; f. 49r ; f. 51r ; f. 67r ; f. 70r ; f. 81r ; f. 84r ; f. 95r ; f. 97r ; f. 102r ; f. 119r ; 7 pages d'incipit ornées : f. 12v ; f. 48v ; f. 80v ; f. 94v ; f. 96v ; f. 101v ; f. 118v.

ARTISTES :

Le décor est de la main d'un seul artiste, dont le style se caractérise par la monumentalité des personnages représentés et la densité de certaines compositions. Son style a été rapproché de celui du Psautier d'Egbert (Cividale, Museo Archaeologico Nazionale, cod. 136) qui fut évêque de Trèves à partir de 977.

MATÉRIAUX :

Les fonds des enluminures à pleines pages sont pourprés. La palette utilisée est large mais délimitée : rouge, vert, jaune, doré, violet, des touches de bleus et des nuances de bruns et de gris. Le bleu est réalisé à partir d'outremer (lapis-lazuli) ; les

tons de rouge, de violet et parfois de brun contiennent de l'orseille, pigment obtenu à partir de lichen, mélangé parfois avec de l'oxyde de plomb pour le marron. Le jaune contient de l'ocre jaune, qui est associé au vert-de-cuivre pour obtenir les teintes de couleur verte.

En deuxième lieu, les manifestes sont annotés selon les points d'analyses qui sont renseignés dans les données de la base. Il faut faire un choix pour le commentaire des annotations, trois informations sont retenues : le motif analysé, la couleur et les matériaux découverts. Ces trois données sont celles qui correspondent le mieux au principe de l'annotation, puisqu'elles concernent des points d'analyses précis, là où d'autres concernent le feuillet ou le manuscrit. Si je reprends le cas du manuscrit Latin 9389, les informations de la base sont reportées selon le modèle ci-dessous, à gauche. Seuls les commentaires du Latin 8878 ont une autre présentation, comme visible à droite.

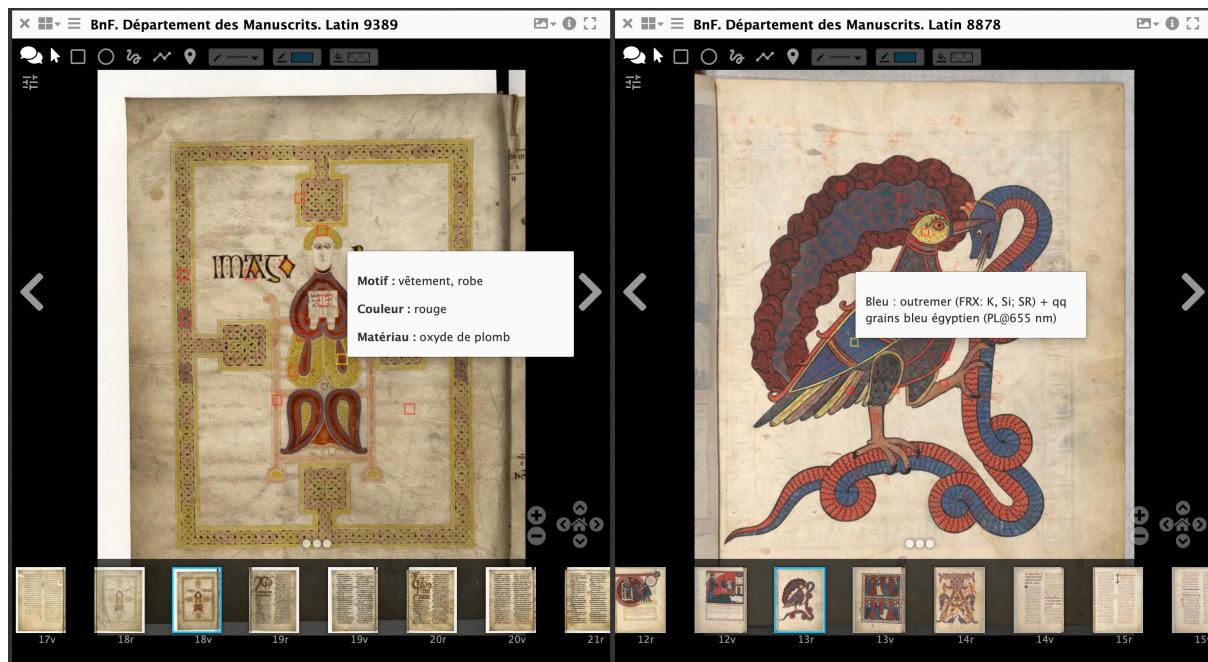


FIGURE 4.6 – Annotations dans Mirador

Enfin, un système de calque vient mettre en complément de l'image numérique du feuillet les résultats obtenus par les analyses hyperspectrales. Ils sont invisibles au chargement de la page, mais ils apparaissent une fois sélectionnés dans la barre latérale. L'affichage sous Mirador permet de cumuler l'affichage des calques à celles des annotations. Ainsi, il est permis d'additionner les résultats de différentes analyses pour avoir toutes les informations nécessaires sur le feuillet en question.

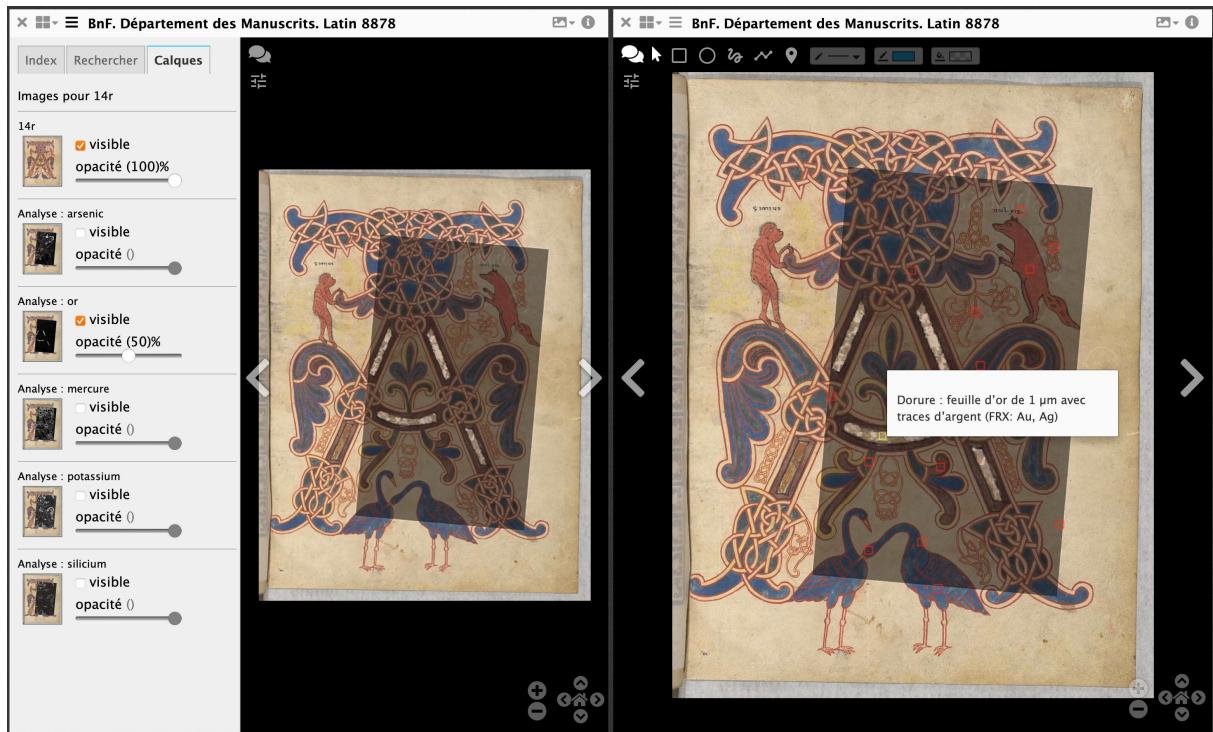


FIGURE 4.7 – Addition d'un calque et d'une annotation

Un fac-similé de cette nature facilite la communication et la médiation de la recherche. Les chercheurs ont à leur disposition les images numériques en haute résolution des feuillets, la présentation du manuscrit, ainsi que le résultat des analyses menées dans le cadre du projet de recherche. Les données de la base sont à la disposition du public, mais elles sont également intelligibles parce que contextualisées sur la source même du travail. Elles ne sont plus inscrites dans des CSV à plusieurs milliers de lignes, mais sur l'objet même de la recherche.

Le fac-similé numérique permet aussi d'instaurer un langage commun entre des cultures scientifiques diverses. Ainsi, le scientifique et les historiens, historiens de l'art, spécialistes des textes anciens ont sous les yeux le document qui est l'objet de leur discussion. Ils peuvent y rapporter directement leurs commentaires, leurs analyses et leurs résultats. Des étapes de traductions scientifiques, qui peuvent être parfois à la source d'erreurs, non plus lieu d'être et les processus de communication sont simplifiés.

*

Le projet détaillé ici n'a pu trouver son aboutissement dans le cadre de ce stage. Les deux projets scientifiques sont hébergés à partir de l'automne sur un site Omeka S. En conséquence, et en anticipation de la migration, un essai d'implémentation d'une visionneuse dans Omeka S¹⁰ a été réalisé. Cet essai a également pour objectif de s'assurer de la bonne conformité des manifestes IIIF précédemment cités. Cependant, il s'est avéré que le module Mirador proposé pour Omeka est appauvri par rapport aux versions de la visionneuse qui peuvent être trouvées sur le web. Si une solution a été trouvée pour la gestion des annotations, il n'en est pas de même pour les calques d'images¹¹. Les manifestes qui seront proposés à l'automne sur les sites Omeka S ne proposeront donc pas ce dernier enrichissement pour les fac-similés. Néanmoins, les facs-similés numériques produits demeurent de nouveaux outils de recherche mis à la disposition du public et utiles pour la communication de la recherche.

10. Pour l'implémentation de Mirador dans Omeka S, voir annexe 3.

11. Cf annexe 4.

Conclusion

Les programmes de recherche sur « La couleur : artefacts, matière et cognition » et sur « La fabrique matérielle du visuel : panneaux peints en Méditerranée » ont abouti à trois projets de visualisation : une carte interactive, une frise chronologique et une édition de fac-similés numériques. Ces visualisations prolongent et enrichissent les recherches des années précédentes. Elles permettent de réexaminer les données collectées et les résultats obtenus, elles facilitent la communication et la médiation de la recherche, et elles instaurent un langage commun entre diverses cultures scientifiques. Ainsi, une politique de visualisation des données doit être intégrée dans la réflexion du chercheur dès le lancement d'un programme de recherche.

Ce mémoire a démontré que les visualisations ne sont pas simplement des outils à disposition du chercheur, mais qu'elles produisent et reflètent également des connaissances historiques. La cartographie interactive et la frise chronologique proposent ainsi une mise en ordre du temps et de l'espace selon l'histoire des enluminures des manuscrits et des panneaux peints. Chaque réalisation est située par rapport aux autres, en fonction d'un système de repérage constitué de coordonnées ou de dates. Il s'agit de représenter, de distinguer ou de regrouper une réalisation par rapport à une autre, selon leur contexte géographique ou temporel.

Ces visualisations sont à l'origine d'un discours sur l'espace et sur le temps pour le chercheur. Elles permettent d'identifier des tendances historiques et des contextes de production, autrement dit des alignements de données, qui doivent être explorés. Le chercheur, parfois submergé par les milliers de lignes d'un fichier CSV, peut perdre de vue l'ensemble des données. Les visualisations redonnent un sens aux analyses produites, les orientant selon des préoccupations chronologiques et géographiques. Elles permettent également de faire ressortir les données essentielles et de reléguer les autres à leur place accessoire ou complémentaire, grâce à des politiques d'éditorialisation de la base, qui, comme nous l'avons vu, est avant tout un encodage de filtres sur les thésaurus constitués.

Les fac-similés numériques, troisième et dernier volet des visualisations, facilitent la communication et la médiation de la recherche. Les données de la base sont reportées sur des images numériques en haute résolution avec des métadonnées de présentation. Les analyses menées, qui constituent le travail scientifique des chercheurs lors du projet

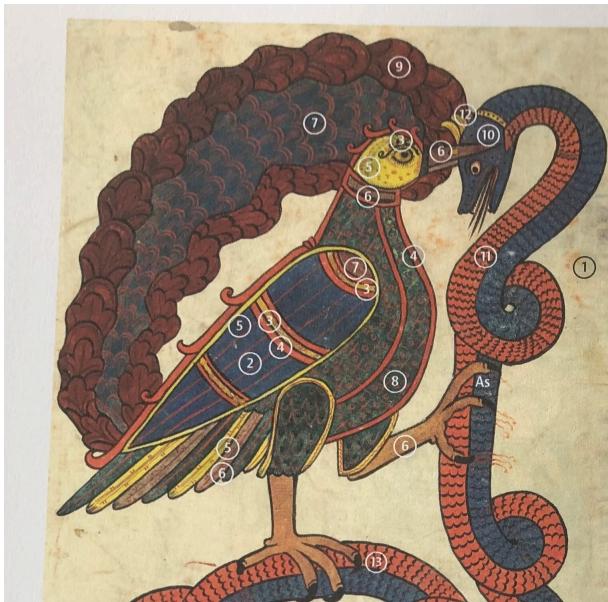
de recherche, deviennent intelligibles au public parce que contextualisées sur la source même du travail. Le fac-similé numérique permet aussi d'instaurer un langage commun entre des cultures scientifiques diverses. Ainsi, le scientifique et les historiens, historiens de l'art, spécialistes des textes anciens ont sous les yeux le document qui est l'objet de leur discussion. Ils peuvent y rapporter directement leurs commentaires, leurs analyses et leurs résultats. Des étapes de traduction scientifique, qui peuvent parfois être à la source d'erreurs, n'ont plus lieu d'être et les processus de communication sont simplifiés.

Ce mémoire fait apparaître que la visualisation n'est plus seulement la représentation de l'analyse d'un chercheur avec des outils informatiques. Dans un certain sens, une logique s'inverse : la visualisation n'est plus une fin, mais tend à devenir un moyen, une étape de la recherche. Avec cette redéfinition, le public de la visualisation change aussi. Elle ne se réduit plus au grand public, à des fins de communication de la recherche, mais elle devient un service de la recherche, désormais en soutien aux chercheurs.

Cette évolution conduit à repenser les différentes séquences des projets de recherche. D'un schéma en trois temps, où se succèdent « recherche », « analyse » et « publication (dont visualisation) » ; il conviendrait de passer à une nouvelle articulation « recherche », « traitement informatique (dont visualisation) et analyse », puis « publication ». La réunion de l'analyse et du traitement informatique est déjà une réalité dans les projets de recherche actuels et ne cesse de se renforcer, en témoigne la place croissante prise par les ingénieurs d'étude et de recherche à côté des chercheurs dans le monde de la recherche.

Annexe A

Analyse du combat du paon et du serpent dans le Beatus de Saint-Sever



93 Combat du paon et du serpent, f. 13r (détail)

Parchemin

1. préparation calcique + orpiment

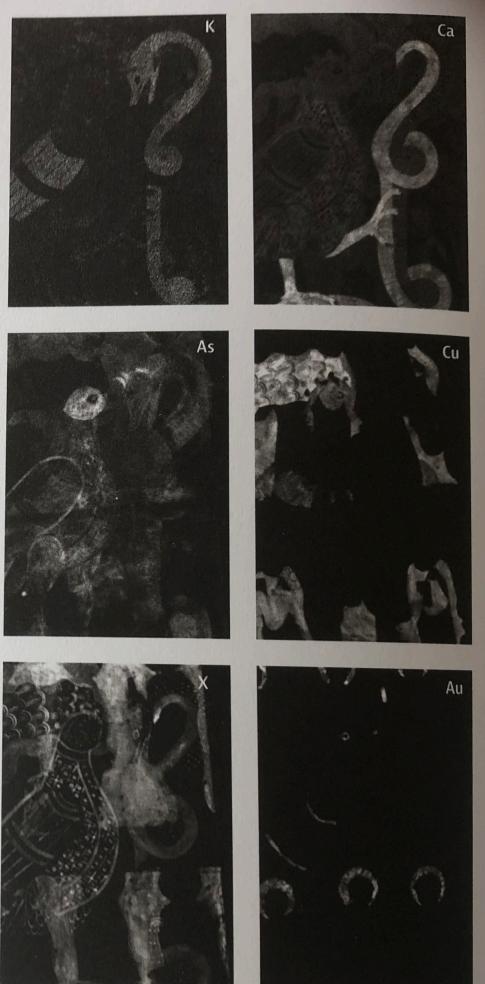
Paon

- 2. bleu (aile) : bleu outremer + grains épars de bleu égyptien
- 3. doré (œil et aile) : feuille d'or
- 4. rouge (aile et corps) : vermillon
- 5. jaune (aile, plumes et tête) : orpiment
- 6. rose (bec, col, aile et pattes) : colorant pourpre organique
- 7. bleu (plumes) : azurite + grains de bleu égyptien + trait vermillon
- 8. vert (corps) : vergaut (mélange bleu indigo+orpiment)
- 9. rouge (queue) : colorant pourpre organique

Serpent

- 10. bleu (corps) : bleu outremer + points vermillon
- 11. rouge (corps) : vermillon
- 12. jaune (tête) : orpiment

13. noir des contours, des décors et des griffes : noir de carbone



94 Imagerie élémentaire par fluorescence de rayons X, f. 13r (détail)
K : potassium (outremer); Ca : calcium ; As : arsenic (orpiment);
Cu : cuivre (azurite); X : mercure (vermillon); Au : or (feuille d'or)
C2RMF

Annexe B

Exemple d'analyse de feuillet en JSON

```
1  {
2      "geometry": {
3          "type": "Point",
4          "coordinates": [
5              6.95,
6              50.93333
7          ]
8      },
9      "properties": {
10         "Titre": "Sacramentaire de Saint-Géréon de Cologne,
11             Monogramme VD avec la Croix, Paris, BNF, ms. Latin
12             817, f. 14v",
13         "Parent": "02389266-ee27-48ea-be1e-0b40e5e0144e",
14         "Type": "manuscrit - sacramentaire - enluminure",
15         "Attribution": "?",
16         "Lieu_de_creation": "Cologne - Allemagne",
17         "Realisation": "1051 - 1000",
18         "Identifiant": "02389266-ee27-48ea-be1e-0b40e5e0144e_001
19             ",
20         "Image": "https://agorha.inha.fr/ark:/54721/ce34ce88
21             -1066-452d-8732-1640d085e348/doc/802fb768-aec9-4d2f-
22             a45a-c409bb9dab1b/original.PNG",
23         "Type_description": "Couche colorée",
24         "Caracteristique": "Couche picturale",
25         "Motif": "décor",
```

```
21     "Technique": "enluminé - gouache",
22     "Couleur": "rouge - rouge foncé",
23     "Materiaux": "orseille",
24     "Certitude": "supposé",
25     "Technologie": "FORS - XRF",
26     "Mesure": "?",
27     "Date": "2023",
28     "Rapport": "?",
29     "Source": "?",
30     "Localisation": "BnF - département des Manuscrits",
31     "Cote": "Latin 817",
32     "Date_filtre": "1051",
33     "Projet": "La fabrique de l'art. Couleurs et matériaux de
34     l'enluminure"
35   },
36   "type": "Feature"
}
```

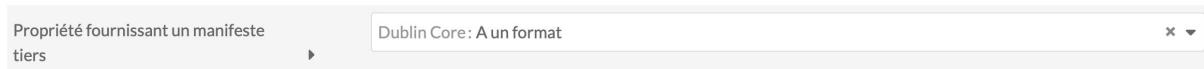
Annexe C

Mirador dans Omeka S

Une visionneuse, comme Mirador, s'implémente par un module propre à Omeka. Plus qu'un module, il convient plus précisément d'en télécharger trois pour que la visionneuse soit en capacité d'afficher convenablement un manifeste IIIF :

- **le module Common¹** : qui permet de gérer les fonctionnalités internes utilisées dans divers modules (fonctions en bloc, éléments de formulaire, assistants de vue, tâches uniques pour l'installation et les paramètres, etc.), de sorte qu'il évite au développeur de copier-coller du code commun entre les modules.
- **le module IIIF Server²** : qui intègre les spécifications IIIF pour permettre de traiter et de partager instantanément des images de toute taille et de tout support (pdf, audio, vidéo, 3D...) dans les formats souhaités.
- **le module Mirador³** : une visionneuse en ligne avancée, qui affiche des images, des livres, des cartes, etc. via la norme IIIF. La visionneuse est multi-fenêtres, avec la possibilité de zoomer, d'afficher, de comparer et d'annoter des images.

Les deux derniers modules doivent être paramétrés pour qu'ils puissent lire un manifeste tiers. Il faut, pour ce faire, définir pour propriété du « Dublin Core : A un format ».



Pour ajouter un manifeste à une collection sur Omeka, il faut se rendre dans l'onglet « ressources », « contenus », puis « ajouter un contenu ».

Ensuite, tout se fait dans l'onglet « valeurs ». Après avoir attribué un titre à l'objet, il convient de rechercher la propriété « a un format » de Dublin Core pour ajouter le manifeste.

1. <https://gitlab.com/Daniel-KM/Omeka-S-module-Common>
2. <https://gitlab.com/Daniel-KM/Omeka-S-module-IIIFServer>
3. <https://gitlab.com/Daniel-KM/Omeka-S-module-Mirador>

The screenshot shows the Omeka S 'Nouveau contenu' (New Content) interface. On the left, there are tabs for 'Valeurs', 'Médias', 'Collections', 'Sites', and 'Avancé'. The 'Avancé' tab is selected, showing fields for 'Modèle de ressource' (Choisir un modèle), 'Classe' (Choisir une classe), 'Titre' (Persan 174), 'Description' (Valeur à supprimer), and 'A un format' (https://theoburnel.github.io/Persan174/manifests/manifest.json). On the right, a sidebar titled 'Cliquez sur une propriété pour l'ajouter au panneau de modification.' lists 'DUBLIN CORE (3)' properties: 'Format', 'A un format', and 'Est un format de'. At the bottom, there are links to 'Omeka S version 4.1.0', 'Informations sur le système', 'Manuel d'utilisation', and 'Forums de soutien'.

Pour faire apparaître le manifeste IIIF sur la page internet, il suffit d'ajouter un bloc « Mirador Viewer » et de mettre comme pièce jointe le manifeste souhaité. Après l'enregistrement, le manifeste et les annotations apparaissent.

The screenshot shows the Omeka S 'Bienvenue - Modifier' (Welcome - Edit) page. The left panel contains fields for 'Titre' (Bienvenue), 'Identifiant pour l'URL' (welcome), and 'Disposition' (Flux normal). The right panel shows a 'Pièces jointes' (Attachments) section with a button 'Ajouter une pièce jointe'. A modal window titled 'Choisir un contenu' lists items such as 'Persan 174', 'Copte 13', 'Arabe 5847', etc. At the bottom, there are links to 'welcome# sur cette page, dans un nouvel onglet.', 'Omeka S version 4.1.0', 'Informations sur le système', 'Manuel d'utilisation', and 'Forums de soutien'.

Ci-dessous, un exemple d'une visualisation d'une image IIIF sur Omeka S avec la visionneuse Mirador.

BnF. Département des Manuscrits. Persan 174

The image shows a page from a Persian manuscript (Persan 174, folio 15r). The central illustration depicts a chimera standing between two trees, with a horse standing below it. The chimera has a lion's head, a goat's body, and a serpent's tail. It is wearing a green tunic with a yellow collar and a blue sash. The horse is brown and is grazing. There is Arabic calligraphy at the bottom of the page.

Annotations

Showing 6 annotations

ITEM: [15R]

Motif : arbre, feuillage
Couleur : vert
Matériaux : terre verte

Motif : animal, chimère, pelage du cou
Couleur : gris
Matériaux : indigo, cochenille

Motif : animal, chimère, plumage du cou
Couleur : rose
Matériaux : indéterminé

Motif : vêtement, tunique, galon du col
Couleur : jaune

37 of 309 • 15r

Annexe D

Retour critique et problèmes techniques

Leaflet

La carte interactive proposée avec la bibliothèque JavaScript Leaflet fonctionne sans aucun défaut. Toutes ses fonctionnalités sont opérationnelles, qu'il s'agisse de l'affichage de la carte, des différents marqueurs et de leurs informations dans le carrousel et la barre latérale, ou de la combinaison des différents filtres. Cependant, je crois qu'une option supplémentaire aurait pu améliorer encore l'expérience de navigation et d'exploration des résultats. Actuellement, la carte interactive dévoile les résultats par feuillet, alors que les filtres portent sur les analyses. Ainsi, si le résultat d'un filtre ne correspond qu'à la troisième analyse d'un feuillet, il est nécessaire de faire défiler le carrousel jusqu'à l'apparition de la bonne analyse. Cet inconvénient ralentit quelque peu l'exploration de la carte, bien que les informations affichées demeurent correctes.

J'aurais souhaité, si la durée du stage me l'avait permis, de réaliser une seconde bascule dans les données, semblable à celle proposée pour le projet. Cette fois-ci, l'utilisateur aurait eu le choix de consulter les résultats par feuillet ou par manuscrit. Le carrousel aurait alors défilé les différentes pages analysées du manuscrit correspondant. Cette option aurait permis à l'utilisateur d'avoir une idée plus précise du nombre de manuscrits répondant à ses critères de sélection, ce qui est peut-être plus pertinent que d'obtenir le résultat par feuillet analysé.

VIKUS Viewer

La frise chronologique développée avec VIKUS Viewer présente un défaut majeur lors de son utilisation. Bien que l'articulation entre les différents paramètres fonctionne correctement et que les images apparaissent comme prévu lors de l'application des filtres, la cohérence de la barre latérale avec le feuillet affiché peut poser problème après un certain temps d'utilisation. Après une période indéterminée de navigation, les informations contenues dans la barre latérale se figent et ne correspondent plus à celles du feuillet

sélectionné. Il est alors nécessaire de rafraîchir la visualisation. Ce dysfonctionnement est problématique pour les utilisateurs : non seulement ils doivent à nouveau renseigner leurs paramètres de recherche, mais ils doivent également faire preuve d'une grande vigilance pour éviter de recopier des informations erronées. J'ai inclus un avertissement dans le guide d'utilisation de la visionneuse à ce sujet. Cependant, ce défaut persiste et pose un réel problème. Il me semble qu'il pourrait être dû à la manière dont VIKUS déclare les variables et les met à jour. Malheureusement, dans le cadre de mon stage de trois mois, je n'ai pas eu l'occasion de pousser plus loin l'investigation ni de proposer une correction aux développeurs.

Annotations IIIF

Les annotations présentes dans les manifestes IIIF ne correspondent pas à celles initialement prévues. Le serveur d'annotations SAS permet de sélectionner un type de marqueur qui me semble bien plus précis que les carrés retenus et qui seront déployés sur le site Omeka S. Après une phase de tests pour quelques annotations, j'ai constaté que les données SVG (Scalable Vector Graphics) qui définissent ce marqueur étaient trop complexes pour la version de Mirador prévue pour Omeka S. J'ai donc dû simplifier ces données et opter pour la représentation des analyses par petits carrés. De même, la version 3 de Mirador affiche le contenu des annotations dans une barre latérale, contrairement à la version précédente où elles apparaissaient dans une info-bulle. Cette modification rend l'information moins lisible.

De plus, l'implémentation de la visionneuse dans Omeka S pose un problème quant à la gestion de certaines fonctionnalités de IIIF. Comme déjà souligné au cours de ce mémoire, il m'était impossible de proposer un affichage par calques, tel que je l'aurais souhaité initialement. Tous les calques se superposaient, et la seule solution a été d'abandonner cette possibilité. En revanche, la migration des données a été satisfaisante et toutes les informations qui devaient apparaître y figurent bien.

Annexe E

Liste des manifestes IIIF

Tous les manifestes IIIF et leurs annotations sont déposés sur le compte Github <https://workbench.gdmrdigital.com/login.xhtml>. Ils sont reliés à mon compte personnel, mais ils sont tous attribués à la Bibliothèque nationale de France. Ils peuvent être appelés par tout le monde à ces adresses :

Grec 139 :

https://theoburnel.github.io/Grec139-v3/manifests/manifest_3b.json

NAL 1203 :

<https://theoburnel.github.io/NAL1203-v2/manifests/manifest2.json>

Latin 8878 :

<https://theoburnel.github.io/Latin8878-v2/manifests/manifest-3.json>

Latin 9389 :

<https://theoburnel.github.io/Latin9389-v4/manifests/manifest2.json>

Syriaque 341 :

<https://theoburnel.github.io/Syriaque341-v2/manifests/manifest2.json>

Copte 13 :

<https://theoburnel.github.io/Copte13-v2/manifests/manifest2.json>

Arabe 5847 :

<https://theoburnel.github.io/Arabe5847-v2/manifests/manifest2.json>

Mexicain 1-10 :

<https://theoburnel.github.io/Mexicain1-10/manifests/manifest.json>

Latin 10514 :

<https://theoburnel.github.io/Latin10514/manifests/manifest.json>

Supplément Grec 1296 :

<https://theoburnel.github.io/SupplementGrec-1286/manifests/manifest.json>

Latin 10525 :

<https://theoburnel.github.io/Latin10525/manifests/manifest.json>

Persan 174 :

<https://theoburnel.github.io/Persan174/manifests/manifest.json>

Annexe F

Liens vers la carte interactive et la frise chronologique

- Lien vers la carte Leaflet :
<https://github.com/TheoBurnel/Stage.git>
- Lien vers la frise VIKUS :
https://github.com/TheoBurnel/vikus_bnf.git

Index

- Beatus de Saint-Sever, vii, 46
Bibliothèque nationale de France, viii, 1, 2, 35, 38, 53
bleu égyptien, vii–ix, 46

carte interactive, x, xi, 8, 14, 19, 25, 26, 32, 35, 36, 41, 63
chronologie, viii, 35, 37, 39, 41, 43, 45–48

frise chronologique, x, xi, 8, 41, 46, 63

géographie, viii, 37, 46

IIIF, xi, 9, 23, 55–57, 61

Leaflet, 14, 15, 26, 28, 38

thésaurus, ix, 2–4, 10–12, 17, 18, 63

VIKUS, 37–41, 43, 46–48
visualisation, ix–xi, 1–3, 8–12, 14–17, 19, 21, 25–27, 32, 35–41, 43–48, 52, 63, 64

Table des figures

1.1	Le média d'une notice Aghora	5
1.2	Les champs structurés sur Agorha	5
2.1	Filtrer par projet sur la carte	18
2.2	Filtrer par année sur la carte	18
2.3	Filtrer par matériau sur la carte	19
2.4	Filtrer par couleur sur la carte	19
2.5	Marqueurs et clusters	20
2.6	Initialisation de la carte	21
2.7	Exemple d'éditorialisation	22
2.8	Une info-bulle avec les premières informations	24
2.9	Une image du carrousel	24
2.10	Un descriptif du carrousel	24
2.11	Une barre latérale avec la totalité des informations	24
3.1	Déplacement d'icônes dans Vikus	40
3.2	Les séquences chronologiques dans la frise	42
3.3	Co-présence de matériaux	42
3.4	Présentation des données sous Vikus	43
3.5	Distribution temporelle du bleu égyptien	46
3.6	Chronologie de la production abbatiale	47
4.1	Structure d'un manifeste IIIF	51
4.2	API Image	52
4.3	API Présentation	52
4.4	Personnalisation des métadonnées descriptives	54
4.5	Lettre d'Alexander Hamilton	56
4.6	Annotations dans Mirador	59
4.7	Addition d'un calque et d'une annotation	60

Table des matières

Résumé	i
Remerciements	iii
Introduction	vii
I. L'exemple du bleu égyptien, ou ce que l'investigation physico-chimique apporte à l'histoire de l'art	vii
II. Une recherche scientifique et culturelle	viii
III. La visualisation : la représentation au service de l'interprétation	x
1 Présentation du projet de recherche	1
I. Les bases de donnée	2
II. La plateforme Agorha	4
III. Deux sites Omeka S	8
IV. L'enjeu du thésaurus	10
2 Ordonner les données dans l'espace, la carte interactive	13
I. Leaflet : une bibliothèque JavaScript open-source et légère	14
II. Éditorialiser une base de données avec une cartographie interactive	17
III. Préparer les données	26
IV. Coder la bibliothèque JavaScript Leaflet	28
V. La cartographie interactive et ses biais	35
3 Ordonner les données dans le temps, la chronologie interactive	37
I. VIKUS Viewer : un système de visualisation web avancé	38
II. Éditorialiser avec Vikus	40
III. Traitement des données	44
IV. Chronologie et interprétation	46
4 Annotations sémantiques d'images	49
I. Un fac-similé numérique	50
II. Enrichir une image numérique	53

III. La sélection du département des manuscrits	57
Conclusion	63
A Analyse du Beatus	65
B Analyse en JSON	67
C Mirador dans Omeka S	69
D Retour critique	73
E Liste des manifestes IIIF	75
F Liens des projets	77