

# Cahier des charges pour le développement du site ArTerm

Version du 19 décembre 2025

## Table des matières

<u>Table des matières</u> .....	1
<u>I. Introduction</u> .....	2
A. Présentation du projet ArTerm.....	2
B. L'objet du cahier des charges.....	2
C. Contexte et objectifs du projet.....	2
D. Planning prévisionnel / jalons.....	3
E. Équipe du projet et contacts.....	3
<u>II. Les données du projet ArTerm</u> .....	3
A. Les données textuelles.....	3
B. Les images.....	7
<u>III. Architecture et apparence du site</u> .....	8
A. Architecture globale du site Web.....	8
B. Mise en page commune.....	9
C. Charte graphique.....	9
D. UX / UI.....	9
<u>IV. Fonctionnalités du site web</u> .....	10
A. Fonctionnalités requises (MVP - minimal viable product).....	10
B. Fonctionnalités supplémentaires.....	15
<u>V. Contraintes techniques</u> .....	17
A. Technologies envisagées.....	17
B. Hébergement et recours aux API.....	18
<u>VI. Annexes</u> .....	18
Annexe 1 : Exemple de TEI header.....	19
Annexe 2 : Exemple du contenu XML encodé en TEI.....	21
Annexe 3 : L'index XML <persName>.....	23
Annexe 4 : L'index XML <placeName>.....	24
Annexe 5 : architecture de la notice JSON œuvre.....	26
Annexe 6 : architecture de la notice JSON monument.....	27

# I. Introduction

## A. Présentation du projet ArTerm

Le projet de recherche ArTerm est financé par l'Agence nationale de la recherche. Il s'inscrit dans les disciplines des humanités numériques, de la linguistique et de l'histoire de l'art et porte sur le lexique de la peinture en Italie et en France à la période moderne. Son objectif est de favoriser l'étude comparatiste des lexiques italien et français de la peinture à l'époque moderne.

## B. L'objet du cahier des charges

Ce cahier des charges est consacré au développement d'un site web pour exposer les transcriptions numériques des imprimés anciens (XV<sup>e</sup> - début XIX<sup>e</sup> siècle) du corpus du projet et permettre la recherche et l'étude du vocabulaire artistique franco-italien de cette période.

## C. Contexte et objectifs du projet

Le site web constitue l'aboutissement du projet ArTerm : une plateforme en ligne dédiée à la diffusion et à l'étude des éditions numériques des ouvrages du corpus. Ce projet s'adresse principalement à un public de chercheurs et de chercheuses en histoire, histoire de l'art et linguistique. La plateforme offrira deux grands axes d'exploitation :

- La consultation des sources, enrichies par la mise en valeur des entités nommées permettant d'identifier et de repérer les personnes, lieux et œuvres mentionnés dans les textes, grâce à leur encodage en XML-TEI.
- L'analyse lexicale et linguistique, permettant d'étudier le vocabulaire français et italien employé dans la littérature artistique de l'époque moderne.

## D. Planning prévisionnel / jalons

- Janvier 2027 : prototype du site terminé (MVP)
- Courant de l'année 2027 : correction et améliorations possibles du site
- Janvier 2028 : Livraison du site définitif

## E. Équipe du projet et contacts

L'équipe du projet est présentée sur le carnet Hypothèse :

<https://arterm.hypotheses.org/equipe-editoriale>. Elle est composée d'**Anna Sconza** maîtresse de conférences à l'Université Sorbonne Nouvelle et porteuse du projet, **Julia Castiglione**, docteure en études italiennes, **Emma Bondoerffer**, ingénierie de recherche en Humanités numériques et **Pierre Husson**, ingénieur d'étude. Pour toutes questions concernant ce cahier des charges, contacter ce dernier à l'adresse suivante : [pierre.husson@efrome.it](mailto:pierre.husson@efrome.it) ou à l'adresse générique du projet : [arterm.paris@gmail.com](mailto:arterm.paris@gmail.com)

## II. Les données du projet ArTerm

### A. Les données textuelles

#### Édition numérique en XML-TEI des textes

Les éditions numériques enrichies en XML-TEI forment le cœur du projet. Elles ont été réalisées à partir de transcriptions d'ouvrages numérisés.

Les éditions numériques du corpus sont organisées selon trois grands thèmes artistiques : **peinture, architecture et perspective**.

Elles sont encodées en XML-TEI par les ingénieur.es d'étude du projet, conformément aux recommandations de la TEI (en ligne : <https://tei-c.org/guidelines/>).

Dans le cadre du site web, en fonction des textes, le balisage sera effectué selon **deux niveaux de finition** différentes :

- éditions complètes : certains textes présentent l'ensemble du balisage TEI, intégrant les informations structurelles ainsi que l'identification des entités

nommées (les personnes, les lieux et les œuvres). Ces textes seront entièrement consultables dans l'interface.

- **éditions synthétiques** : d'autres textes non consultables en intégralité suivront une structuration minimale et normalisée selon les règles du projet, afin de faciliter la comparaison des textes pour l'analyse intertextuelle et les cas de traductions. Ils pourront être requêtés à l'aide du moteur de recherche.

Le corpus complet se compose de **77 textes**, répartis ainsi :

- Peinture : 24 éditions complètes, 25 éditions synthétiques
- Architecture : 10 éditions complètes, 7 éditions synthétiques
- Perspective : 8 éditions complètes, 3 éditions synthétiques

La méthode de mise en relation des éditions complètes avec les éditions synthétiques est encore en cours d'étude par l'équipe du projet Arterm. Elle pourrait s'appuyer soit sur l'utilisation des balises `<ptr>` en XML-TEI, soit grâce à un système de liaison des paragraphes à l'aide d'identifiants.

Voir [Annexe 1](#) pour l'organisation du TEI Header.

Voir [Annexe 2](#) pour un exemple de transcription en XML TEI.

## Les index

Les données issues des différents index constituent le fondement d'une base de données relationnelle, destinée à structurer et relier les entités nommées extraites du corpus. Elles concernent les personnes, les lieux, les œuvres d'arts et les citations bibliographiques présentes dans le corpus.

Ces données sont enrichies par l'équipe du projet grâce aux recherches menées dans le cadre du projet ainsi qu'aux ressources externes disponibles sur Wikidata.

Les index Personnes et Lieux sont stockés dans des fichiers XML distincts et sont liés aux éditions numériques en XML-TEI grâce à `<xi:include>`.

### Personne `<persName>`

L'index des personnes est généré à partir des balises `<persName>` repérées dans le corpus. Il inclut également certaines figures non citées directement dans les textes, mais impliquées dans la création ou la mention d'œuvres présentes dans le corpus.

Les données correspondantes sont saisies dans un index XML comprenant les champs suivants :

- Identifiant (issu de l'identifiant TEI)

- Nom
- Prénom
- Date de naissance
- Date de mort
- Lieu de naissance
- Lieu de mort
- Profession / rôle (plusieurs entrées possibles)
- Courte biographie (optionnelle)

Voir [Annexe 3.](#)

### Lieux <placeName>

L'index des lieux est constitué à partir des balises <**placeName**> identifiées dans le corpus. Les informations sont consignées dans un document XML comportant les champs suivants :

- Identifiant (issu de l'identifiant TEI)
- Nom du lieu
- Type de lieu (ville, région, pays, bâtiment, etc.)
- Coordonnées géographiques (latitude, longitude) (optionnel, pour la carte)
- Pays
- Commentaires ou précisions (optionnel)
- Identifiant Wikidata

Voir [Annexe 4.](#)

### Œuvres <objectName>

L'index des œuvres est élaboré à partir des balises <**objectName**> employées dans l'édition du corpus. Il rassemble les œuvres artistiques mentionnées ou décrites dans les textes. Chaque balise <**objectName**> possède un attribut @ref renvoyant à un identifiant xml:id dans la base de données. Les œuvres mentionnées qui ne possèdent pas cet attribut n'ont pas été identifiées ou ont été perdues.

Les œuvres peuvent être de deux types : les **œuvres** (peintures, sculptures, arts décoratifs, dessins, estampes) et les **bâtiments**.

Pour le moment, l'index est construit à partir de notices JSON pour faciliter leur création et leur modification. Les recommandations de la TEI ne permettent pas de décrire de manière convenable des œuvres artistiques. Le format en JSON pourra être changé en XML ou en base de données relationnelles SQL pour le développement du site.

Pour les œuvres (peintures...), les informations suivantes sont présentes dans la base de données :

- Identifiant [@xml:id](#)
- Identifiant Wikidata (si existant)
- Créateurs
  - [xml:id](#) lié à l'index des personnes <persName>
  - rôle
- Dates de création (minimale, maximale et le texte à afficher)
- Matériaux et techniques
- Lieu de conservation et numéro d'inventaire
- Liens entre les œuvres
- Références bibliographiques
- Lien vers des images (IIIF si disponible)
- Commentaires ou précisions (optionnel)
- Historique de l'édition de la notice d'œuvre

Voir [Annexe 5.](#)

Pour les bâtiments, les informations suivantes sont présentes dans la base de données :

- Identifiant [@xml:id](#)
- Identifiant Wikidata (si existant)
- Créateurs
  - [xml:id](#) lié à l'index des personnes <persName>
  - rôle
- typologie (cathédrale, église, monastère, édifice public, palais, demeure privée)
- Dates de création (minimale, maximale et le texte à afficher)
- Localisation (ville, pays et coordonnées géographiques)
- Liens entre les œuvres
- Références bibliographiques
- Lien vers des images (IIIF si disponible)
- Commentaires ou précisions (optionnel)
- Historique de l'édition de la notice d'œuvre

Voir [Annexe 6.](#)

## Bibliographie

L'ensemble des références bibliographiques mentionnées dans les textes du corpus, ainsi que celles permettant de documenter les personnes, les lieux et les œuvres associées au corpus, est stocké dans une base de données BibTeX, constituée à l'aide du logiciel Zotero. Dans le XML, elles sont indiquées par la balise `<bibl>` avec l'attribut `@key` qui renvoie à la clé Zotero (créeée par le plugin Better BibTeX for Zotero) et l'attribut `@ref` indique la localisation exacte dans l'ouvrage mentionné (page, volume,...) sous la forme d'une chaîne de caractères.

Exemple : `<bibl @key="albertiTTrattatoNobiltaPittura1585" @ref="p. 5">`à la page 5 du Traité de la peinture d'Alberti`</bibl>`.

## B. Les images

Les images sont deux types : les numérisations des sources textuelles et la galerie d'œuvres mentionnées dans les textes. La majorité des images est récupérée via des appels d'API. Leur mention et leurs métadonnées sont homogénéisés afin d'assurer une cohérence.

### Numérisations des sources textuelles

Les éditions TEI sont associées aux numérisations des documents originaux, qui devront également être consultables depuis le site. Le rapprochement entre les pages éditées et les fac-similés s'effectue grâce à l'attribut `@facs` intégré dans les balises `<pb>` des éditions XML-TEI.

L'ensemble des images et numérisations des textes est stocké sur des entrepôts distants, auxquels le site du projet accédera via des API, grâce au protocole IIIF. Les numérisations de textes s'appuient principalement sur les API suivantes :

- Gallica ([documentation de l'API](#))
- [Archive.org \(guide API\)](#)

Lorsque les numérisations ne sont pas accessibles via des appels d'API, l'autorisation de reproduction des textes est sollicitée par l'équipe ArTerm auprès des institutions concernées. Dans ce cas de figure, les images sont hébergées sur l'entrepôt pérenne Nakala (Huma-num) et accessibles grâce au protocole IIIF.

### Les images des œuvres

Pour les œuvres, les images proviennent majoritairement des collections numérisées suivantes, compatibles avec le protocole IIIF, permettant leur affichage dans des visionneuses intégrées aux sites suivants :

- Musée du Louvre (IIIF)
- The MET Museum (IIIF)
- Europeana (IIIF)
- Wikimedia Commons lorsqu'elles sont libres de droit

Comme pour les pages des textes, si certaines œuvres ne sont pas disponibles en accès libre dans ces collections, les droits de diffusion sont systématiquement demandés aux institutions concernées par l'équipe ArTerm. Les images correspondantes seront également hébergées sur Nakala.

## III. Architecture et apparence du site

### A. Architecture globale du site Web

Le site se composera de plusieurs sections principales destinées à présenter le projet et à exposer ses données.

Structure du site :

- [Page d'accueil du site](#) : Présentation synthétique du projet, accès aux principales sections, moteur de recherche
- [Section « Présentation du projet ArTerm »](#) : Présentation du projet, des équipes et institutions impliquées.
- [Section « Accès au corpus »](#) : Accès aux éditions numériques des textes du corpus
- [Section « Index \(personnes, lieux, œuvres\) »](#) : Accès aux différents index (personnes, lieux, œuvres)
- Pages institutionnelles : [Mentions légales et crédits](#), [Contact](#)
- Des sections supplémentaires sont en réflexion en fonction de l'avancement des recherches menées par les équipes et des possibilités que permettent le corpus :
  - [Section « Glossaire des termes artistiques français et italiens »](#).
  - Expositions virtuelles avec les œuvres

## B. Mise en page commune

L'ensemble des onglets du site doit partager une mise en page cohérente pour assurer une navigation fluide et une identité visuelle homogène.

- En tête :
  - Logo du projet (cliquable, renvoyant vers la [page d'accueil](#))
  - Menu principal contenant les liens vers les sections :
    - [Le projet](#)
    - [Corpus](#)
    - [Index \(personnes, lieux, œuvres\)](#)
    - [Glossaire](#)
- Pied de page :
  - Liens vers les pages :
    - [Mentions légales](#)
    - [Contact](#)
    - [Crédits](#)

## C. Charte graphique

Aucune charte graphique n'a été définie. Arterm possède un logo.

## D. UX / UI

Concernant l'UI et l'UX, les équipes du projet Arterm ne disposent pas d'expertise spécifique dans ces domaines. Afin de faciliter la compréhension de l'organisation du site, ce cahier des charges emploie les termes **sections**, **pages** et **onglets** pour hiérarchiser les contenus : les sections regroupent plusieurs pages, et les pages peuvent elles-mêmes être subdivisées en onglets, affichés au sein d'une même page. Toute proposition visant à améliorer l'organisation et la mise en page des contenus est donc bienvenue.

## IV. Fonctionnalités du site web

Les fonctionnalités souhaitées du site web sont réparties en 3 catégories :

- **Fonctionnalités requises** : ensemble des fonctionnalités indispensables à la présentation et à la diffusion de la production scientifique du programme de recherche.
- **Fonctionnalités supplémentaires** : fonctionnalités souhaitées qui dépendent de leur faisabilité, spécifiques au projet et dont le développement nécessitera une collaboration étroite entre les développeurs et les équipes du projet.

### A. Fonctionnalités requises (MVP - minimal viable product)

Le site (MVP) doit permettre une navigation complète à travers les pages principales du projet, la consultation des données et l'accès aux outils de recherche et d'indexation.

#### Bandeau

- Barre de recherche simple, permettant d'effectuer une recherche sur l'ensemble du site.
- Logo
- Menu de navigation dans les différentes sections

#### Page « Accueil du site »

- Présentation succincte du projet ArTerm, avec un lien vers la page « À propos » pour une description détaillée du projet.

#### Section « Présentation du projet ArTerm »

#### Page « À propos »

- Présentation du projet (objectifs et méthodologie)
- Introduction à la structure du site et aux fonctionnalités

#### Page « Équipe »

- Liste des membres du projet

- rôle et biographie
- format similaire au carnet hypothèses du projet :  
<https://arterm.hypotheses.org/equipe-editoriale>

#### Page « Liste des partenaires »

- Institutions impliquées avec logos, brèves descriptions et liens externes
- Liste des collaborateurs, soutiens et contributeurs au projet
- Remerciements

#### Page « Documentation des recherches »

- Protocoles et recherches développées dans la thèse d'Emma Bondorffer, utilisant le corpus ArTerm pour arriver à l'identification de l'intertextualité au sein du corpus.

#### Page Contact

- Adresse e-mail de contact générique du projet
- Formulaire de contact simple (nom, e-mail, message)

### Section « Accès au corpus »

Cette section donne accès à l'ensemble des éditions numériques constituant le corpus du projet.

#### Page « Accès au corpus »

La page « Accès au corpus » présente la liste des [éditions complètes](#) disponibles. Chaque ouvrage de cette liste comprend un lien vers la page « [Présentation de l'édition](#) » dédiée.

#### Page « Présentation de l'édition »

La page d'accueil de chaque [édition complète](#) présente les informations principales relatives à l'ouvrage, indiquées dans le [header](#) de l'édition en XML-TEI :

- Visuel de la page de titre (ou du frontispice)
- Auteur
- Titre complet
- Date de publication
- Lieu d'édition
- Commentaire introductif

- Ouvrages en rapport (traduction,...)
- Bibliographie sélective
- Informations sur la numérisation (bibliothèque, cote de l'ouvrage)
- Lien vers la numérisation
- Auteur.ices de l'encodage en XML-TEI
- Listes des références bibliographiques citées dans l'édition en XML-TEI (signalées par la balise **<bibl>**)
- Une table des matières, permettant d'accéder directement à la visualisation et à l'édition enrichie du document.

## Page « Visualisation de l'édition enrichie »

La page de visualisation d'une édition numérique doit être adaptée à chaque page du texte en utilisant les balises **<pb>** et permettre une consultation interactive et enrichie. Les différents éléments constituant le contenu de la page peuvent être affichés ou masqués en cliquant dessus.

### **Contenu de la page**

#### 1. Menu d'informations de l'édition

- Menu rétractable
- Rappel des informations mentionnées dans la [page « Présentation de l'édition »](#) :
  - titre
  - auteur
  - date de publication et lieu d'édition

#### 2. Numérisation du document

- Affichage de l'image de la page du texte, requête via API IIIF et visualisé à l'aide d'un framework de visualisation comme OpenSeadragon ou Mirador (les liens vers les numérisations sont indiquées par l'attribut @facs, [voir ch. II, B.](#))

#### 3. Transcription enrichie

- Le texte transcrit avec un marquage spécifique pour les entités nommés, chacune soulignée et accompagnée d'un marqueur visuel (couleur/icône) :
  - Personnes
  - Lieux
  - Œuvres
- Fenêtres pop-up d'information contextuelle

- Cliquer sur une entité nommée ouvre une fenêtre pop-up présentant les informations principales et un lien vers la notice complète dans [l'index](#) correspondant.
- Contenu des pop-ups :
  - Personnes : nom, date de naissance, date de mort
  - Lieux : nom du lieu, pays
  - Œuvres : prévisualisation, nom de l'artiste, dates de création (texte), technique, lieu de conservation
  - Bibliographie : mentionnée par la balise [`<bibl>`](#), citée de manière normalisée
- **Texte en surbrillance pour mettre en valeur les emprunts**
  - Certains paragraphes apparaîtront en surbrillance pour indiquer un emprunt à un autre texte
  - en cliquant sur ce texte, la fenêtre « Visualisation des emprunts ou de l'intertextualité » sera affichée pour montrer le paragraphe emprunté (issu d'un autre texte) (voir ci-dessous, n°6)

#### 4. Données structurées du document en XML-TEI

- L'encodage en XML-TEI du texte pourra être téléchargé

#### 5. Galerie d'images

- galerie des images des œuvres ([`<objectName>`](#)) mentionnées dans la page visualisée.
- sous la forme d'un bandeau rétractable ou d'une fenêtre sur le côté

#### 6. Visualisation des emprunts ou de l'intertextualité

- Certains paragraphes de l'édition ont des emprunts identifiés et encodés dans le XML de l'[édition complète](#), qui doivent apparaître **en surbrillance**. Les passages empruntés proviennent des [éditions synthétiques](#) ou d'autres éditions complètes. En cliquant sur ces paragraphes, ces passages empruntés seront affichés.

#### Pages « Critères d'édition »

Cette page documente les choix effectués pour sélectionner les textes à éditer et la manière dont a été réalisée l'édition numérique en XML-TEI.

## Section « Index (personnes, lieux, œuvres) »

### Listes d'index

La section index est organisée autour d'un moteur de recherche qui permet de filtrer facilement parmi les intitulés d'œuvres, de personnes ou de lieux et les métadonnées qui y sont associées.

Chaque entrée d'index doit renvoyer vers le ou les textes du corpus où l'entité nommée apparaît.

### Page d'entrée d'index

Des pages dédiées aux entrées d'index permettent ensuite de consulter les informations relatives aux entités nommées :

- Pour les personnes [`<persName>`](#) :
  - Informations principales relatives à la personne (voir la partie [Personnes](#) [`<persName>`](#) dans le chapitre II. sur les Données du Projet)
  - Liste des œuvres dont la personne est l'autrice
- Pour les lieux [`<placeName>`](#) :
  - Informations principales relatives aux lieux (voir la partie [Lieux](#) [`<placeName>`](#) dans le chapitre II. sur les Données du Projet)
  - Une carte avec le lieu
  - Liste des œuvres présentes dans ce lieu
- Pour les œuvres [`<objectName>`](#) :
  - Informations principales relatives aux œuvres (voir la partie [œuvres](#) [`<objectName>`](#) dans le chapitre II. sur les Données du Projet)
  - Une carte pour les bâtiments
  - Les liens possibles entre les œuvres (par ex: copie,...)

## Pages institutionnelles et légales

### Page « Mentions légales » et « Crédits »

- Contenu obligatoire pour la conformité légale du site à préparer avec le service juridique de l'université.
- Copyright synthétique accessible depuis tous les onglets en pied de page
- Informations à inclure dans les “Mentions légales” :
  - Conditions de réutilisation
  - Protection des données

- Informations à inclure dans les “Crédits” :
  - Mentions des contributeur.ices, des partenaires du projet, du concepteur du site et des sources mobilisées

## Page « Contact »

- Permet aux utilisateur.ices de prendre contact avec l’équipe du projet.
- Contenu :
  - Formulaire de contact avec champs :
    - nom
    - adresse e-mail
    - objet du message
    - message
    - bouton d’envoi
  - Coordonnées : adresse email

## Accessibilité multi-plateforme

Le site Web doit être accessible sur tous types d’appareils, y compris smartphones et tablettes, grâce à une version mobile *responsive*.

Pour les éditions numériques, un menu de navigation dédié permettra de basculer entre les différentes vues du document :

- Numérisation (image du texte)
- Édition enrichie (texte avec annotations et pop-ups) : par défaut
- Encodage XML-TEI (données structurées)

Seule la version sélectionnée sera affichée à l’écran. Par défaut, c’est l’édition enrichie qui sera présentée.

## Administration du site

Le site doit permettre la mise à jour aisée des éditions numériques, soit par l’injection directe des documents XML (textes et index), soit via une interface d’édition simple.

## B. Fonctionnalités supplémentaires

Ces fonctionnalités souhaitées nécessitent des développements supplémentaires par rapport au MVP.

## Moteur de recherche général

Un moteur de recherche global devra être intégré au site afin d'interroger l'ensemble des données. Un développement approfondi est souhaité, incluant les fonctionnalités suivantes :

- présentation hiérarchisée et organisée des résultats par grandes sections :
  1. recherche dans les titres et les auteurs des textes ;
  2. recherche à l'intérieur des textes, avec accès aux résultats dans leur contexte pour les **éditions complètes**
    - prise en charge de requêtes comprenant plusieurs mots-clés (séparés par des virgules ou par des booléens), notamment au niveau des paragraphes/pages/éditions (en fonction de ce qui est possible)
  3. recherche au sein des différents index.

La possibilité d'appliquer des filtres par type de contenu est également souhaitée.

À titre d'exemple, pour une recherche sur « Léonard de Vinci », le moteur devra :

1. afficher en priorité les textes dont le titre contient Léonard de Vinci ou rédigés par lui présents dans les éditions du corpus
2. proposer ensuite les occurrences des termes « Léonard de Vinci » dans les textes des éditions complètes, puis dans les éditions synthétiques (seules les éditions complètes étant intégralement consultables)
3. présenter enfin les résultats issus des trois index : Lieux, Œuvres et Personnes.

## Suggérer des corrections

Un bouton interactif permettant d'afficher en pop-up la page du formulaire de contact, dont les informations de l'édition déjà remplies et la possibilité d'ajouter un message pour préciser la correction.

## Fonctionnalités du Backend

Voici les fonctionnalités envisagées pour la gestion des données, des contenus et des utilisateurs :

- Gestion et modifications légères des éditions xml:tei
- Gestion des index et du glossaire
- Gestion des utilisateurs
- Suggestions des modifications, validation

- Gestion des contenus multimédias : images, API externes (IIIF, Gallica)
- Maintenance

## Section « Glossaire des termes artistiques français et italiens »

Une section supplémentaire est en réflexion : elle comprendrait un glossaire bilingue des termes artistiques français et italiens. Les entrées y seraient classées par ordre alphabétique et présentées en deux colonnes (français / italien). Lorsque cela sera possible, chaque terme serait accompagné de sa traduction dans l'autre langue et de renvois vers ses occurrences dans les textes. Les termes choisis pour le glossaire seront annotés probablement à l'aide de la balise `<term>` dans les fichiers XML.

## Datavisualisations

Les équipes du projet travaillent à des datavisualisations qui pourront être intégrés sur le site, en particulier reposant sur la balise `<bibl>`, permettant de visualiser le réseau de citations bibliographique interne aux œuvres, sous la forme d'un diagramme de réseau dynamique.

# V. Contraintes techniques

## A. Technologies envisagées

### Format des données

Les textes sont encodés en XML-TEI et doivent être téléchargeables dans ce format depuis l'interface du site. Les index des personnes et des lieux reposent sur le même format, mais pourront également être modifiés ou exploités dans d'autres formats, tels que des bases de données relationnelles (SQL) ou non relationnelles (NoSQL, par exemple en JSON). De même, l'index des œuvres est rédigé sous la forme de documents JSON, puisque l'XML-TEI était inadapté à la description d'œuvres.

### Framework et montage du site

Différentes technologies sont envisagées pour construire le site :

- XQuery / XSLT avec base de données XML pour le requêtage et la transformation des données issues des textes et de la base de données.
- Application python Flask et SQL (ou mySQL) pour un site dynamique, bien que cette solution puisse poser des limitations pour la gestion directe des textes TEI.
- WordPress / Omeka s'ils permettent d'intégrer les données en XML-TEI
- L'outil TEI Publisher peut également être envisagé pour la visualisation

Le protocole de partage de texte *Distributed Text Services* (DTS) pourrait être envisagé, peut-être avec l'implémentation DoTS BaseX DTS Tools (en XQuery), si l'outil permet de gérer facilement l'organisation par page <pb> (page beginning).

## B. Hébergement et recours aux API

L'hébergement se fait en majorité sur les infrastructures mises à disposition par Huma-Num :

- Hébergement principal du site permis par Huma-Num
  - base de données
  - Textes TEI conservés dans la base de données hébergée sur Huma-Num.
- Images stockées sur Nakala.

L'appel des images via API sur certains sites, notamment Gallica, est soumis à des limitations afin d'éviter la surcharge des serveurs. Plusieurs solutions peuvent donc être envisagées en concertation avec les équipes du projet Arterm, notamment :

- le recours à des vignettes, avec affichage de l'image en pleine résolution uniquement après un clic
- un défilement volontairement lent des images afin de permettre leur chargement progressif

## VI. Annexes

Toutes les annexes sont disponibles sur un dépôt github à l'adresse suivante :  
[https://github.com/ANR-ArTerm/cahier\\_des\\_charges](https://github.com/ANR-ArTerm/cahier_des_charges)

# Annexe 1 : Exemple de TEI header

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="tei_arterm.rng" type="application/xml"
schematypens="http://relaxng.org/ns/structure/1.0"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xmlns:xi="http://www.w3.org/2001/XInclude">
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title xml:lang="fra"><!--Titre de l'oeuvre--></title>
      <respStmt><!--Ce bloc peut être répété -->
        <resp><!--Indiquer la responsabilité : encodage, transcription, etc.--></resp>
        <name><!--Nom de la personne : Prénom puis Nom--></name>
      </respStmt>
    </titleStmt>
    <editionStmt> <!--On indique ici des informations sur l'édition et le cadre dans lequel elle se place (ex première édition numérique, dans le cadre d'un projet, etc.)-->
      <edition><!--Description du projet et contexte de l'édition--></edition>
    </editionStmt>
    <publicationStmt><!--Ajouter ici les infos relatives à l'édition électronique : date de publication, éditeur, licence-->
      <publisher><!--Nom de l'organisme responsable de la publication ou de la distribution--></publisher>
      <pubPlace>Paris</pubPlace>
      <date when="2028">2028 ?</date> <!--Date de publication du corpus et du site à la fin de l'ANR ?-->
      <availability>
        <licence target="a completer"></licence> <!--Ajouter ici la licence-->
      </availability>
    </publicationStmt>
    <sourceDesc> <!--Informations bibliographiques sur la source à partir de laquelle l'édition numérique a été produite-->
      <bibl><!--Cette section correspond à la référence bibliographique du document-->
        <author ref="a completer"><!--Nom de l'auteur--></author>
        <title xml:lang="fra"><!--Titre de l'oeuvre--></title>
        <publisher> <persName ref="#PrenomNomEditeur"><!--Nom de l'éditeur--></persName><!--Autres informations sur l'éditeur, le cas échéant--></publisher>
        <pubPlace><!--Lieu de publication--></pubPlace>
        <date when="date"><!--Date de publication--></date>
      </bibl>
      <msDesc> <!--Cette section correspond à la description de l'imprimé original-->
        <msIdentifier> <!-- Informations requises pour identifier le document-->
          <country><!--Nom du pays dans lequel le document est conservé--></country>
          <settlement><!--Nom de la ville dans laquelle le document est conservé--></settlement>
          <institution><!--Nom de l'institution dans laquelle le document est
```

```

conservé--></institution>
    <repository><!--Nom du dépôt ou département dans lequel le document est
conservé--></repository>
        <idno><!--Identifiant standardisé, par exemple l'identifiant ark pour les documents
conservés à la BnF--></idno>
            </msIdentifier>
            <msContents><!--Cette section est réservée à la description du contenu intellectuel du
document-->
                <summary><!--Informations diverses sur le contenu, l'histoire, la mise en page, etc. du
document--></summary>
                </msContents>
                </msDesc>
                </sourceDesc>
                </fileDesc>
            <profileDesc><!--On inclut ici les index des personnes et des lieux du projet-->
                <xi:include href="IndexPersonnes.xml" xpointer="element(1/1)"/>
                <xi:include href="IndexLieux.xml" xpointer="element(1/1)"/>
            </profileDesc>
        </teiHeader>
        <text>
            <front>
                <titlePage type="titre">
                    <docTitle><!--Titre du document tel qu'il apparaît sur la page de titre-->
                        <titlePart type="main"><!--Titre principal de l'ouvrage--></titlePart>
                        <titlePart type="sub"><!--Sous titre s'il existe--></titlePart>
                    </docTitle>
                    <byline><!--Informations sur l'auteur--><docAuthor ref="#PrenomNomAuteur"><!--Nom de
l'auteur tel qu'il apparaît sur la page de titre, ex : P. MONIER--></docAuthor><!--Autres
informations sur l'auteur, ex : Peintre du Roi et Professeur en l'Académie Roiale de Peinture et
Sculture.--></byline>
                    <figure type="ornement" source="" /><!--Si nécessaire-->
                    <docImprint><!--Informations relatives à l'édition et à l'éditeur-->
                        A <pubPlace><!--Lieu d'édition--></pubPlace>
                        <publisher><!--Informations sur l'éditeur--></publisher>
                        <pubPlace><!--Si nécessaire, précisions sur le lieu d'édition--></pubPlace>
                    </docImprint>
                    <docDate when="date"><!-- Date d'édition --></docDate>
                    <imprimatur><!--Mention autorisant la publication de l'ouvrage type 'Avec Privilege de Sa
Majesté'--></imprimatur>
                </titlePage>
            <!-- Exemple
            <front>
                <figure type="gravure"
source="https://prod-arkindex-iiif.europe.iiif.teklia.com/iiif/2/ee88c0f5-211c-4508-8c13-242ee227
2eed/1141,136,2003,3268/max/0/default.jpg"/>
                <titlePage type="titre">

```

```

<docTitle>
    <titlePart type="main"><hi rend="capitals">HISTOIRE
<lb/> DES ARTS QUI ONT RAPORT AU DESSEIN, <lb/> DIVISEE EN TROIS
LIVRES</hi></titlePart>
    <titlePart><hi rend="capitals">OÙ IL EST TRAITÉ DE
SON ORIGINE</hi>, de son Progrés, de sa Chute, et de son Rétablissement. Ouvrage utile au
public pour savoir ce qui s'est fait de plus considérable en tous les ages, dans la Peinture, la
Sculture, l'Architecture et la Gravure; et pour distinguer les bonnes manieres des
mauvaises.</titlePart>
</docTitle>
<byline>Par<docAuthor ref="#PierreMonier"> P.
MONIER</docAuthor> Peintre du Roi et Professeur en l'Académie Roiale de Peinture et
Sculture.</byline>
<figure type="ornement"
source="https://prod-arkindex-iiif.europe.iiif.teklia.com/iiif/2/7b568a57-4004-4aee-aef7-b1689a9d
3adf/1566,2050,1097,577/max/0/default.jpg"/>
<docImprint>
    A <pubPlace>PARIS</pubPlace>, <lb/><publisher>Chez
PIERRE GIFFART, Libraire et Graveur du Roi</publisher>, <pubPlace>rue S. Jacques à l'Image
sainte Therese, vis-à-vis S. Yves.</pubPlace>
</docImprint>
<docDate when="1698">M DC. LXXXVIII.</docDate>
<imprimatur>Avec Privilege de Sa Majesté.</imprimatur>
</titlePage> -->
</front>
<body>
    <p>Some text here.</p>
</body>
</text>
</TEI>

```

## Annexe 2 : Exemple du contenu XML encodé en TEI

```

<body>
    <div type="lettre">
        <pb n="3" facs="https://gallica.bnf.fr/ark:/12148/btv1b8454358x/f5.item"/>
        <figure type="ornement"/>
        <opener>
            <salute>
                <hi rend="capitals">A TUTTI COLORO.</hi>
                <lb/>Che della professione ingegnosissima del Disegno si dilettano.</salute>

```

<lb/>  
<signed>  
  <hi rend="capitals">  
    <persName ref="#GiovanniAntonioMassani">GIOVANNI ATANASIO  
MOSINI</persName>  
  </hi>  
  </signed>  
  <lb/>Salute.</opener>  
<p>  
  <figure type="letrine"/>  
  <persName ref="#AnnibaleCarracci">  
    <hi rend="capitals">ANNIBALE</hi> Carracci</persName> Pittore de' nostri tempi  
dell'eccellenza, che a voi (amatori di così bello artificio) può esser manifesta, fu riputato da coloro,  
che in vita lo conobbero, esser dotato di una felicità d'ingegno maravigliosa; con la quale,  
accompagnando egli con suo gran gusto lo studio, e la fatica, arrivò ad havere così pronta, e  
ubbidente la mano ad esprimer col disegno gli oggetti, che vedeva, e s'imaginava; che non si  
pose mai a far cosa, che felicemente non gli riuscisse, e che, per picciola che ella si fosse, non  
venisse da gl'intendenti oltremodo stimata. E quanto all'**<reg>**isquisitezza**</reg>** dell'imitatione,  
fine principale del Pittore, la quale a lui ottimamente sempre riusciva, alcuni lo paragonavano a  
que' tali, che, per grande habilità naturale, e per istudio particolare, essendo attissimi ad imitare i  
linguaggi, la voce, li gesti, e altre singolari, e proprie qualità di alcuno, fanno credere, che  
ascoltando, o veggendo essi, si ascoltino, o veggano que' medesimi, che da loro vengono imitati.  
E perché al tempo di <persName ref="#AnnibaleCarracci">Annibale</persName> viveva il  
valentissimo <persName ref="#GiovanniGabrielli">Sivello</persName>, il quale superò forse gli  
antichi Histrioni celebrati per bravissimi, e maravigliosi dagli Scrittori, ed era amico de' medesimi  
<persName ref="#Carrache">Carracci</persName>; quindi è, che, in riguardo dell'ottima  
imitatione, specialmente a lui veniva <persName ref="#AnnibaleCarracci">Annibale</persName>  
paragonato. Percioché tra l'altre cose molte, che 'l <persName  
ref="#GiovanniGabrielli">Sivello</persName> mirabilmente imitava egli da per sé solo  
rappresentava un congresso di sei persone differenti di linguaggio, di voce, di età, e di conditione.  
Egli si poneva in luogo, dove da niuno potesse esser veduto, ma ascoltato benissimo da molti: e  
fingendo, che le sei persone cominciassero a ragionare fra loro di cose piacevoli, e curiose; dipoi,  
proponendo di giuocare a carte, risolvevano tra di loro di fare una primera, e gli altri di starsene a  
vedere. Principiato il giuoco, e per alquanto continuato pacificamente, nasceva <pb n="4"  
facs="<https://gallica.bnf.fr/ark:/12148/btv1b8454358x/f6.item>">poi fra essi contesa, come suole  
accadere tra i giuocatori; e dicendo il lor parere li tre altri, che a vedere se ne stavano,  
s'interponevano ancora per acchetarli, mentre dalle contese, alle parole mordaci, e indi alle mani  
anco finalmente venivano. Rappresentava il <persName  
ref="#GiovanniGabrielli">Sivello</persName> tutto ciò da per sé solo, imitando isquisitamente la  
diversità de' linguaggi, e della voce, appropriandovi le parole, i motti, facendo sentire alcuno  
strepito, e altre circostanze proprissime a quel congresso, e all'attione, che imitar voleva; di modo,  
che da persone di grandissimo sapere, e giuditio, che l'hanno udito, io ho sentito affermare, che in  
quel genere d'imitatione era quell'huomo arrivato al colmo, e certamente non si potea far più. E  
bene spesso avvenne, che molti tale attione per finta non credettero, sinché ancora col vedere non  
se ne resero sicuri. <persName ref="#AnnibaleCarracci">Annibale</persName> similmente in  
vece di fingere con le parole, o con la voce, o co' gesti, imaginandosi i lineamenti, e tutto quello,  
che appare alla veduta in qual si voglia corpo, o oggetto visibile, e imitabile, l'ha rappresentato in

modo col disegnare, e col dipingere; che, applicando mirabilmente a ciò che faceva, le particolari, e proprissime qualità di contorni, di lumi, e ombre, e colori, secondo che a ciascuna cosa più conviene, ha in tutte le sue opere chiaramente mostrato, quanto possa la mutola favella della mano esprimere all'humano intelletto, col mezzo della facoltà visiva. E perciò assai acconciamente si adattava la comparatione di <persName ref="#AnnibaleCarracci">Annibale</persName> al <persName ref="#GiovanniGabrielli">Sivello</persName>; perché sì come questi ingannava l'uditore, che non vedeva, così quegli ingannò più volte il riguardante, che non toccava.</p>

## Annexe 3 : L'index XML <persName>

```
<?xml version='1.0' encoding='utf-8'?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <particDesc>
    <listPerson>
      <person xml:id="HansVonAachen" role="peintre allemand (1552-1615)"
source="http://www.wikidata.org/entity/Q152835">
        <persName
ref="https://www.getty.edu/vow/ULANFullDisplay?find=Joseph+Hains&role=&nation=&mp;page=1&subjectid=500127905">
          <forename>Hans von</forename>
          <surname>Aachen</surname>
          <addName>Jean Dac</addName>
          <addName>JEAN DAC</addName>
        </persName>
        <birth when="1552-01-01">
          <placeName>Köln</placeName>
        </birth>
        <death when="1615-03-04">
          <placeName>Praha</placeName>
        </death>
        <trait/>
        <nationality>Saint-Empire romain germanique</nationality>
        <idno type="ulan">500127905</idno>
        <idno type="bnf">12564530t</idno>
        <idno type="viaf">41957298</idno>
        <idno type="isni">0000000066298373</idno>
        <idno type="vat">495/376415</idno>
        <idno type="wikidata">Q152835</idno>
        <floruit>
          <objectName>Le Jugement de Pâris</objectName>
        </floruit>
      </person>
      <person xml:id="Aaron" role="personnage de la Bible, dans les livres de l'Exode, du Lévitique
```

et des Nombres"

```

source="https://www.wikidata.org/wiki/Q51676">
<persName>
  <forename>Aaron</forename>
  <surname>Aaron</surname>
  <addName>Aaron</addName>
  <addName>Aronne</addName>
</persName>
<birth when="1396">
  <placeName>Égypte antique</placeName>
</birth>
<death when="1274">
  <placeName>djebel Haroun</placeName>
</death>
<trait>
  <p>cohen, prophète ou prophétesse</p>
</trait>
<idno type="viaf">86951058</idno>
<idno type="vat">495/294384</idno>
<idno type="wikidata">Q51676</idno>
<floruit>
  <objectName>veau d'or</objectName>
</floruit>
</person>
```

## Annexe 4 : L'index XML <placeName>

```

<?xml version='1.0' encoding='utf-8'?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
<settingDesc>
<listPlace>
  <place xml:id="Augsbourg" source="https://www.wikidata.org/wiki/Q2749">
    <placeName xml:lang="fra">Augsbourg</placeName>
    <placeName xml:lang="ita">Augusta</placeName>
    <location>
      <country>Allemagne</country>
      <geo>48.3688888888889 10.897777777777778</geo>
    </location>
    <idno type="wikidata">Q2749</idno>
  </place>
  <place xml:id="Cologne" source="https://www.wikidata.org/wiki/Q365">
    <placeName xml:lang="fra">Cologne</placeName>
    <placeName xml:lang="ita">Colonia</placeName>
    <location>
```

```

<country>Allemagne</country>
<geo>50.942222222222 6.9577777777778</geo>
</location>
<idno type="wikidata">Q365</idno>
</place>
<place xml:id="Heidelberg" source="https://www.wikidata.org/wiki/Q2966">
<placeName xml:lang="fra">Heidelberg</placeName>
<placeName xml:lang="ita">Heidelberg</placeName>
<location>
<country>Allemagne</country>
<geo>49.412222222222 8.71</geo>
</location>
<idno type="wikidata">Q2966</idno>
</place>
<place xml:id="Ingolstadt" source="https://www.wikidata.org/wiki/Q3004">
<placeName xml:lang="fra">Ingolstadt</placeName>
<placeName xml:lang="ita">Ingolstadt</placeName>
<location>
<country>Allemagne</country>
<geo>48.76305555555555 11.425</geo>
</location>
<idno type="wikidata">Q3004</idno>
</place>
<place xml:id="Kalkar" source="https://www.wikidata.org/wiki/Q243554">
<placeName xml:lang="fra">Kalkar</placeName>
<placeName xml:lang="ita">Kalkar</placeName>
<location>
<country>Allemagne</country>
<geo>51.738888888889 6.2925</geo>
</location>
<idno type="wikidata">Q243554</idno>
</place>
<place xml:id="Mainz" source="https://www.wikidata.org/wiki/Q1720">
<placeName xml:lang="fra">Mayence</placeName>
<placeName xml:lang="ita">Magonza</placeName>
<location>
<country>Allemagne</country>
<geo>49.99944444444444 8.2736111111111111</geo>
</location>
<idno type="wikidata">Q1720</idno>
</place>
<place xml:id="Nuremberg" source="https://www.wikidata.org/wiki/Q2090">
<placeName xml:lang="fra">Nuremberg</placeName>
<placeName xml:lang="ita">Norimberga</placeName>
<location>
<country>Allemagne</country>
<geo>49.4538888888889 11.0775</geo>

```

```
</location>
<idno type="wikidata">Q2090</idno>
</place>
```

## Annexe 5 : architecture de la notice JSON œuvre

```
{
  "id": "",
  "QID_wikidata": "",
  "title": "",
  "creator": [
    {
      "xml_id": "",
      "role": ""
    }
  ],
  "entry_type": "peinture",
  "materialsAndTechniques": "huile sur toile",
  "dateCreated": {
    "startYear": "",
    "endYear": "",
    "text": ""
  },
  "holding_institution": {
    "place": "",
    "name": "",
    "inventory_number": "",
    "URL": ""
  },
  "related_works": [
    {
      "link_type": "",
      "xml_id_work": ""
    }
  ],
  "bibliography": [
    {
      "zotero_key": "",
      "location": ""
    }
  ],
}
```

```

"illustrations": [
  {
    "id":0,
    "url":"",
    "copyright":"",
    "caption":"",
    "storage":""
  }
],
"commentary":"",
"history": [
  {
    "date": "",
    "type": "created",
    "author": ""
  },
  {
    "date": "",
    "type": "modified",
    "author": ""
  }
]
}

```

## Annexe 6 : architecture de la notice JSON monument

```

{
  "id": "",
  "QID_wikidata": "",
  "title": "",
  "creator": [
    {
      "xml_id": "",
      "role": ""
    }
  ],
  "entry_type": "architecture",
  "typology": "",
  "dateCreated": [
    {
      "startYear": ""
    }
  ]
}

```

```
        "endYear": "",  
        "text": ""  
    },  
    "location": {  
        "city": "",  
        "country": "",  
        "coordinates": {  
            "latitude": "",  
            "longitude": ""  
        }  
    },  
    "related_works": [  
        {  
            "link_type": "",  
            "xml_id_work": ""  
        }  
    ],  
    "bibliography": [  
        {  
            "zotero_key": "",  
            "location": ""  
        }  
    ],  
    "illustrations": [  
        {  
            "id": 0,  
            "url": "",  
            "copyright": "",  
            "caption": "",  
            "storage": ""  
        }  
    ],  
    "commentary": "",  
    "history": [  
        {  
            "date": "",  
            "type": "created",  
            "author": ""  
        },  
        {  
            "date": "",  
            "type": "modified",  
            "author": ""  
        }  
    ]  
}
```