



HAL
open science

Plan de gestion des données du RIIG (Recueil informatisé des inscriptions gauloises) ANR RIIG 19-CE27-0003

Coline Ruiz Darasse, Nathalie Prévot

► **To cite this version:**

Coline Ruiz Darasse, Nathalie Prévot. Plan de gestion des données du RIIG (Recueil informatisé des inscriptions gauloises) ANR RIIG 19-CE27-0003. UMR 5607 Ausonius. 2022. halshs-03894186

HAL Id: halshs-03894186

<https://shs.hal.science/halshs-03894186>

Submitted on 12 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Plan de gestion des données à 24 mois
Version 2 • janvier 2022

RIIG
ANR19-CE27-0003
Recueil informatisé des inscriptions gauloises
Édition, contexte archéologique, analyse linguistique, étude socio-linguistique.

coordonné par Coline Ruiz Darasse
UMR 5607 Université Bordeaux-Montaigne

Le projet ANR JCJC RIIG vise à une *editio maior* et pérennisée des inscriptions en langue gauloise du territoire français connues à ce jour, à la mise à jour et à la modernisation des éditions précédentes, à la publication renouvelée de chaque inscription avec une contextualisation précise, à la préparation d'une analyse sociolinguistique, et à la mise à disposition d'une bibliographie archéologique et linguistique actualisées. Commencé en janvier 2020, il est actuellement à mi-parcours.

I. DESCRIPTION DES DONNÉES ET COLLECTE OU RÉUTILISATION DE DONNÉES EXISTANTES

1a. Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?

La méthodologie de collecte des données est double :

- d'une part un travail de collecte et de veille bibliographique, permettant la documentation et la contextualisation la plus précise possible des inscriptions. Elle se fait en bibliothèque, en ligne et aussi dans les institutions que nous visitons au cours du projet (musées, centre d'études, services municipaux d'archéologie etc.) ;
- d'autre part un travail d'autopsie des inscriptions qui permet de constituer une archive photographique du corpus.

Un contrat post-doctoral a débuté le 1^{er} janvier 2021 et se termine fin avril 2022.

Le projet vise également à préparer le corpus déjà existant dans le format TEI-XML de manière à pouvoir l'informatiser et le consulter directement en ligne.

Nouvelles données produites :

- indexation et préparation en format Epidoc des volumes des RIG (*Recueil des inscriptions gauloises*), publiés entre 1985 et 2002. Préparation d'une base de données et d'un site internet pour pouvoir réaliser des requêtes complexes, également disponibles dans le cadre du programme PETRAE et préparée en *Linked Open Data* (LOD). Cette étape a été réalisée grâce au contrat de Nolwenn Chevalier (AI UBM, janvier-novembre 2020) et à Nathalie Prévôt (IE CNRS, Ausohnum).
- commentaires et remarques d'ordre archéologiques, linguistiques, historiques et sociolinguistiques produits par l'équipe qui complètent et enrichissent le corpus initial. Ce travail est toujours en cours. Les différents membres du projet fournissent des documents word

qui sont intégrés par le PI et le post-doc dans le schéma EpiDoc. L'ensemble est suivi par Nathalie Prévôt et documenté dans un GitLab (<https://gitlab.huma-num.fr/ausohnum/riig> ;
- mise à disposition d'une bibliographie concernant l'épigraphie en langue gauloise complète et à jour par le biais d'un groupe Zotero (<https://www.zotero.org/groups/1822453/petrae/collections/NAM9XSNC>), avec les liens vers les documents disponibles en ligne. Au 9 décembre, cette bibliographie comptait 1265 documents ;
- images et documentation photographique (incluant photogrammétrie et 3D quand c'est possible) produite au cours des missions du PI et du post-doc afin de disposer de clichés de travail pour l'inventaire en ligne. Elles sont stockées sur le serveur IIF (International Image Interoperability Framework, <https://iiif.io>) adossé à l'entrepôt Nakala de la TGIR Huma-Num (<https://www.nakala.fr>)

Pour les deux premiers points, le travail est réalisé à partir du logiciel Oxygen (licence accordée à l'UMR 5607 Ausonius). Les données sont stockées sur le serveur de la TGIR Huma-Num et également disponibles dans PETRAE (<http://petrae.huma-num.fr>) élaboré à partir du gestionnaire de base de données XML open-source eXist-db. (<http://exist-db.org/>).

La bibliographie est préparée avec le logiciel Zotero, stockée sur le serveur Zotero et synchronisée sur tous les ordinateurs des membres du projet (<https://www.zotero.org/groups/1822453/petrae/collections/NAM9XSNC>). Toutes les nuits, la bibliographie Zotero est exportée au format TEI-XML et intégrée sur le site du RIIG pour consultation et en interopérabilité avec les inscriptions.

Dans le dernier cas, les images sont produites avec le matériel disponible dans le service Ausohnum et le service de matériel archéologique de l'Institut Ausonius (scanner Artec Space Spider, studio photo mobile, lampes à variations de lumière etc.) et l'acquisition des clichés se fait avec l'appareil photo propre au projet (Nikon D-5600 acheté en mai 2021). Le traitement des images se fait sous Photoshop et Illustrator (licences budgétées dans le cadre du projet initial). Le travail de modélisation numérique par traitement photogrammétrique a pu se faire au cours de l'année 2020 grâce à Florent Comte dans le cadre de son poste d'IE (Ausohnum LaScArBx). Son poste n'ayant pu être pérennisé en 2021, Florent Comte s'est impliqué de manière bénévole pour soutenir la suite de l'acquisition et des traitements afin de ne pas créer un déséquilibre dans les rendus du projet.

Réutilisation de données existantes :

Les données ont publiées initialement dans le supplément 45 à la revue *Gallia* aux éditions du CNRS, actuellement épuisé.

La direction de la revue est informée de notre démarche et nous avons convenu d'une présentation du projet au sein de leur équipe au début de l'année 2022. Contact a été pris avec les archives de la revue, avec les archives nationales et les archives disponibles au service des archives de la MSH Mondes notamment pour retrouver et rassembler les illustrations déjà présentes dans les publications des *RIG*. Il est également probable que plusieurs documents soient retournés aux mains des éditeurs principaux et/ou de leurs ayant-droits. Le projet a la chance de compter avec l'un des éditeurs initiaux et nous avons également pris contact avec la famille de Michel Lejeune qui est informée de la nature et des progrès de nos travaux et qui y est particulièrement favorable.

1b. Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

Les données produites sont principalement numériques, que ce soit la base de données en EpiDoc qui sera consultable sur le site du projet (<https://riig.huma-num.fr/>) à partir de l'été 2022,

comme à partir de PETRAE (<http://petrae.huma-num.fr/corpus/RIIG>) ou les images et les modèles 3D produits dans le cadre des missions. Des dessins et des relevés d'inscriptions seront également numérisés (en .tiff). Le format Epidoc (TEI-XML) est pensé pour l'informatisation des données épigraphiques, donc spécifiquement pour le matériau documentaire de notre projet.

Les minutes des réunions et des workshops de l'équipe seront disponibles en .pdf dans les archives du projet.

L'ouvrage final du projet, synthèse sur l'écriture dans le monde gaulois, sera disponible en format papier et en open access ([OpenBooks éditions](#)).

Les données produites sont disponibles en .xml et exportables à partir de la plateforme de consultation.

L'ensemble de la base de données représente actuellement 8,3 Mo de xml sur les serveurs Huma-Num. La documentation (markdown, etc.) occupe 1,52 Go et le code informatique de l'interface de visualisation, 100 Mo (xquery, html, css, javascript). Pour la journée de lancement du projet, nous avons choisi une inscription test pour évaluer les points à traiter et toutes les données à renseigner. Il s'agissait de l'inscription RIG I, G-224 (RIIG HER-03-01). Elle a nécessité 15 Mo pour le modèle 3D ce qui constitue la moyenne des fiches des inscriptions sur pierre dans le projet.

Toutes les images sont stockées sur Nakala et occupent actuellement 1,1 Go (ce qui correspond à 1026 images en décembre 2021). Elle est composée de différents fichiers que l'on peut distinguer de la manière suivante :

- Paléographie : 33,9 Mo ;
- images initiales du RIG I : 16,5 Mo ;
- Compléments : 53,3 Mo ;
- éléments produits depuis le début du projet : 1 Go
 - Bourgogne-France-Comté : 26 Mo
 - Nouvelle Aquitaine : 7,1 Mo
 - Bretagne : 1,7 Mo
 - Occitanie : 229,6 Mo
 - PACA : 739,6 Mo

La 3D représente environ 500 Mo.

La documentation qui reste à collecter concerne les inscriptions du Bassin parisien et celles conservées au Musée d'Archéologie Nationale. On peut envisager qu'à la fin du projet, l'ensemble représentera un volume de stockage de 5 Go (la 3D n'est pas systématique et ne concerne que les inscriptions sur pierre). Tout dépendra des documents mis en ligne pour la partie de contextualisation car ils sont soumis à des autorisations complexes.

2. DOCUMENTATION ET QUALITÉ DES DONNÉES

2a. *Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?*

Pour ce qui est de la base de données du RIIG, les métadonnées sont celles de la TEI. Plusieurs *Read me* préparés en Markdown ont été mis en place dès le début de l'informatisation du corpus initial pour documenter les choix qui ont été faits au cours du traitement des données et les choix d'encodage (<https://gitlab.huma-num.fr/ausohnum/riig/-/blob/master/ReadMe.md>).

Toute la documentation est disponible sur une forge GitLab préparée pour le projet qui fait état de chaque *issue* en cours ou clôturée (<https://gitlab.huma-num.fr/ausohnum/riig>). Les minutes des réunions relatives à la partie Humanités numériques du projet y sont également stockées.

Le reste de la documentation générée par le projet est partagé sur un espace ShareDocs sécurisé et administré par le service Ausohnum.

Afin de pouvoir collecter les données de recontextualisation ainsi que les données linguistiques et sociolinguistique de manière uniformisée, une fiche à l'usage des chercheurs a été préparée. Les chercheurs participant préparent les données qui sont ensuite insérées dans le XML par Coline Ruiz Darasse et Hugo Blanchet.

Les référentiels utilisés pour décrire les données seront ceux utilisés habituellement pour le LOD :

- Personnes : [ISNI](#) et [ORCID](#)
- Lieux : [Geonames](http://www.geonames.org/) (<http://www.geonames.org/>)
- Textes : Trismegistos (<https://www.trismegistos.org/index.php>)
- Sujets du thesaurus : PACTOLS (<https://pactols.frantiq.fr/opentheso/>) ; EAGLE (<https://www.eagle-network.eu/resources/vocabularies/>)

Chronologie

Protohistoire (<https://ark.frantiq.fr/ark:/26678/pcrtHlenwSnkDM> ;
<https://hdl.handle.net/20.500.11859/26678.pcrtHlenwSnkDM>)

Période gallo-romaine (<https://ark.frantiq.fr/ark:/26678/pcrtGrQ5rGj2I3> ;
<https://hdl.handle.net/20.500.11859/26678.pcrtGrQ5rGj2I3>)

Lieux

Gaule (<https://ark.frantiq.fr/ark:/26678/pcrtxb8Z7Qr4aa> ;
<https://hdl.handle.net/20.500.11859/26678.pcrtxb8Z7Qr4aa>)

Gallia Narbonensis (<https://www.geonames.org/8299477>)

Gallia Aquitania (<https://www.geonames.org/8378499>)

Gallia Lugdunensis (<https://www.geonames.org/8378501>)

Sujet ou thème

épigraphie (<https://ark.frantiq.fr/ark:/26678/pcrtJrInD7k8G1> ;
<https://hdl.handle.net/20.500.11859/26678.pcrtJrInD7k8G1>)

langue gauloise (<https://ark.frantiq.fr/ark:/26678/pcrtSZbNprpAai> ;
<https://hdl.handle.net/20.500.11859/26678.pcrtSZbNprpAai>)

inscriptions gauloises (<https://ark.frantiq.fr/ark:/26678/pcrt2hRPKwPUOA> ;
<https://hdl.handle.net/20.500.11859/26678.pcrt2hRPKwPUOA>)

inscriptions lapidaires (<https://ark.frantiq.fr/ark:/26678/pcrtK8IkAtd45Y> ;
<https://hdl.handle.net/20.500.11859/26678.pcrtK8IkAtd45Y>)

inscriptions sur céramiques (<https://ark.frantiq.fr/ark:/26678/pcrti43gSp7SKT> ;
<https://hdl.handle.net/20.500.11859/26678.pcrti43gSp7SKT>)

inscriptions sur plomb (<https://ark.frantiq.fr/ark:/26678/pcrtQIoupOcrjR> ;
<https://hdl.handle.net/20.500.11859/26678.pcrtQIoupOcrjR>)

gallo-grec (<https://ark.frantiq.fr/ark:/26678/pcrtbfhE9c1363> ;
<https://hdl.handle.net/20.500.11859/26678.pcrtbfhE9c1363>)

gallo-latin

Structure des dossiers

Pour l'heure, le classement des fichiers disponibles sur l'espace ShareDocs est le suivant :

Le dossier **Administratif** rassemble toutes les données relatives à l'**ANR**, au **budget**, aux **missions**, aux recrutements relevant des **Ressources humaines** et enfin au **site internet**.

Le dossier **Bibliographie** permet de conserver les PDF des documents préparés sous Zotero et qui doivent être partagés au sein de l'équipe de travail. S'y ajoutent les documents scannés au cours des différentes missions de collectes de données.

Le dossier **Évènements** rassemble les documents produits lors des réunions (minutes et handouts) ou les textes et présentations PPT des communications, conférences et autres manifestations auxquelles participent les membres du projet.

Le dossier **Iconographie** ne conserve actuellement que les documents permettant d'identifier la provenance des documents, ainsi que les mentions à porter pour chaque clichés pris.

Le dossier le plus important reste celui de l'**Inventaire**, qui est classé par **grandes régions** françaises puis par **départements** et enfin par **sites** ayant livré des inscriptions gauloises. Dans chacun de ces dossiers on pourra trouver les fiches musées qui permettent de collecter les inscriptions et de disposer des coordonnées des gens à contacter pour organiser les missions documentaires.

Enfin le dossier **Numérique** regroupe tous les documents nécessaires au suivi du traitement EpiDoc du projet. Pour ce point, toutefois, il existe une forge GitLab qui rassemble toutes les données (*issues*) relatives à l'encodage lui-même.

Toute la documentation et les scripts du projet sont centralisés sur <https://gitlab.humanum.fr/ausohnum/riig>

2b. Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?

Les données initiales sont celles du *Recueil des inscriptions gauloises* (RIG) et des publications qui lui ont succédé. Toutes sont traitées, préservées en insistant sur leur citation exacte et référencement le plus précis et le plus exact possible. Leur mise en ligne garantit également leur préservation.

La conformité avec le schéma EpiDoc est vérifiée à tout moment par Nathalie Prévôt, puis contrôlés par la coordinatrice du projet. Ponctuellement, sont consultés d'autres spécialistes familiers de ce type de bases de données épigraphiques qui font partie du projet (notamment Noemí Moncunill et Maria José Estarán qui ont travaillé sur Hesperia (<http://hesperia.ucm.es/index.php>) et AELAW (<http://aelaw.unizar.es/>), ou David Stifter qui a dirigé l'entreprise du LexLep (https://www.univie.ac.at/lexlep/wiki/Main_Page).

La qualité des données produites est de la responsabilité de chaque membre du projet, chacun étant expert dans le domaine des inscriptions gauloises d'une manière ou d'une autre. Dans l'équipe, des référents pour les types de données et les différents domaines de compétences ont été désignés (archéologie, linguistique, humanités numériques, sociolinguistique, etc.). Le format EpiDoc permet de renseigner toutes les personnes responsables d'une lecture ou d'un commentaire tout au long du projet.

L'ensemble du projet compte actuellement 428 fiches qui sont relues à 70%. Les relectures se font par vagues (et par ensemble d'inscriptions par département). Des révisions sont planifiées selon les spécialistes et des réunions d'étapes ont eu lieu au cours de l'année 2021 : d'abord les archéologues, puis les linguistes, puis les historiens et numismates (le cas échéant) puis les

sociolinguistes. Chaque inscription fera l'objet d'une double validation (deux passages par chaque membre de l'équipe) de manière à ce que toutes les données soient vérifiées, validées et complétées par chacun. Toutes les remarques seront dûment consignées.

Toute la documentation et les scripts du projet sont centralisés sur <https://gitlab.humanum.fr/ausohnum/riig>. Toutes les versions de toutes les données depuis le début du projet y sont disponibles.

3. STOCKAGE ET SAUVEGARDE PENDANT LE PROCESSUS DE RECHERCHE

3a. Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?

Les données du projet seront préparées sur l'ordinateur des membres du projet puis stockées sur la plateforme Huma-Num que ce soit la base de données elle-même ou les documents qui s'y rapportent (stockés sur ShareDocs) : archives bibliographiques, plans et images des sites archéologiques concernés ; .

La sauvegarde est garantie par les serveurs de la TGIR.

Par ailleurs, la forge GitLab du projet conserve et versionne le code source des interfaces de visualisation web et tous les xml de données.

Les images produites dans le cadre du projet sont stockées sur le serveur IIIF (International Image Interoperability Framework, <https://iiif.io>) adossé à l'entrepôt Nakala de la TGIR Huma-Num (<https://www.nakala.fr>).

Le serveur Huma-Num effectue des sauvegardes incrémentables qui permettraient de récupérer les données en cas d'incident.

3b. Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?

Les ordinateurs utilisés pour la saisie des données (ordinateur portable du projet, ordinateur de la coordinatrice du projet et celui de la responsable des Humanités numériques à Ausonius) sont tous chiffrés.

Au cours du projet, les données seront accessibles par connexion sécurisée avec mot de passe sur le site du RIIG tant qu'elles ne sont pas mises à la disposition du public. Les données partagées sur ShareDocs sont soumises également à une inscription via un HumanID qui permet un accès protégé aux données. L'administration de chacune de ces plateformes est gérée par Nathalie Prévôt du service Ausohnum.

4. EXIGENCES LÉGALES ET ÉTHIQUES, CODES DE CONDUITE

4a. Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?

Il n'y a pas de données à caractère personnel dans le projet RIIG.

Les commentaires et remarques des membres du projet seront dûment renseignés pour garder une trace des hypothèses proposées à chaque étape de la réflexion.

4b. Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?

Le propriétaire des données sera l'équipe du projet dans son ensemble. C'est le chef du projet qui en contrôle l'accès avec l'accord de l'ensemble de l'équipe et l'aide de Nathalie Prévôt qui est la responsable des Humanités numériques à Ausonius et la directrice technique de PETRAE.

Pour chaque visite de musée ou de centre de conservation, un accord préalable avec le responsable de collections est établi. Pour chaque objet documenté, il est précisé quels sont les crédits à mentionner pour l'utilisation des images. Par exemple dans le cas d'une image prise dans le musée lapidaire d'Avignon, pour laquelle nous avons obtenu l'autorisation de la part de l'administratrice et de la part de la conservatrice du musée, le cliché appartiendra au projet RIIG mais il sera dûment fait mention que l'objet est « propriété de la Fondation Calvet d'Avignon ».

Les clichés pris dans le cadre du RIIG suivront la licence : Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International (CC BY-NC-ND 4.0 <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr>).

Les données encore inédites seront conservées au sein des archives du RIIG, afin de ne pas les divulguer avant que leur étude ne soit menée à bien par l'un (ou plusieurs) des membres de l'équipe.

4c. Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?

Il n'y a pas de collecte de données personnelles dans le cadre du projet.

Il n'y a pas de questions d'éthique actuellement soulevée.

Des autorisations seront demandées aux ayants droit pour l'utilisation de documents ou la reproduction de parties de documents dans les livrables (illustration ou usage pour l'analyse par exemple).

Les clichés pris dans les musées et le cas échéant dans des collections particulières font l'objet d'un accord préalable. Si certains objets ne doivent pas être diffusés, les images seront protégées par un mot de passe et réservées à l'usage interne du projet.

5. PARTAGE DES DONNÉES ET CONSERVATION À LONG TERME

5a. Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?

Les données collectées dans le cadre du RIIG sont déposées sur les serveurs de la TGIR Humanum. Elles seront également disponibles dans PETRAE dans un corpus à part entière. De ce point de vue, leur consultation se fera de manière pérenne et évolutive (et au-delà des 48 mois du projet lui-même).

À l'issue du projet, on envisage d'archiver les données au CINES (<https://facile.cines.fr/>).

Les données de synthèses du projet seront proposées dans le cadre de l'ouvrage final, qui proposera une réflexion autour de la place de l'écriture dans les sociétés gauloises puis gallo-romaines. L'ouvrage sera diffusé en format papier et en Open Edition selon les critères demandés par l'ANR.

Les images seront disponibles au fur et à mesure du projet et proposées sous licence CC BY-NC-ND 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) sur le serveur IIF Nakala.

Les dossiers étudiés feront l'objet de présentation dans des colloques ou des congrès internationaux par les membres de l'équipe. Par exemple :

- [XVI^e Congrès international d'épigraphie grecque et latine](#) qui aura lieu à Bordeaux du 29 août au 2 septembre 2022 (session sur les langues d'attestation fragmentaire)
- colloque des études celtiques à Utrecht (proposition d'une session sur le celtique continental) au cours de l'été 2022.

Enfin, des mises en avant des objets dans les musées avec lesquels le projet aura été en contact donnera lieu à des conférences ou à des articles dans les médias de médiation et de valorisation des collections. Une conférence à la BNF est prévue le [9 février 2022](#) et un dossier de 4 pages doit paraître dans le magazine Sciences et Avenir de décembre 2021.

5b. Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?

Le partage des données et des outils est rendu possible par l'environnement technique du programme PETRAE. La structure même de cet outil garantit la durabilité, la sauvegarde et l'interopérabilité des données et son évolution régulière est liée à l'utilisation du TGIR HumNum.

Les métadonnées seront exposées dans un entrepôt OAI-PMH de manière à ce qu'elles soient accessibles par [Isidore](#).

5c. Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?

Les données seront accessibles directement depuis un simple navigateur web.

Toute la documentation et les scripts du projet sont centralisés sur <https://gitlab.humnum.fr/ausohnum/riig>

5d. Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?

Chaque inscription dispose d'un DOI, d'un identifiant pérenne (URI). La fiche disponible et exportable en PDF constituera un livrable et une édition renouvelée du document épigraphique.

6. RESPONSABILITÉS ET RESSOURCES EN MATIÈRE DE GESTION DES DONNÉES

6a. Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?

Dans un premier temps, la saisie des données dans le cadre de l'informatisation du RIG a été faite par l'assistant-ingénieur du projet (Nolwenn Chevalier), supervisé par Nathalie Prévôt et Coline Ruiz Darasse. Le stockage et l'archivage sont assurés par Nathalie Prévôt.

Par la suite, la saisie des données ont été à la charge de chaque membre de l'équipe mais l'intégration dans la base de données en TEI-XML est faite par Coline Ruiz Darasse et Hugo Blanchet secondés, le cas échéant, par Nathalie Prévôt et certains membres du projet formés à la TEI.

La responsable de la mise en œuvre du PGD, sa révision et sa mise à jour est la coordinatrice du projet, Coline Ruiz Darasse.

Le PGD sera mis à jour à la fin du projet ainsi qu'il est requis par l'ANR.

6b. Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?

Le projet vise à préparer le corpus déjà existant en format papier dans le format EpiDoc (TEI-XML) de manière à pouvoir l'informatiser et le consulter directement en ligne. Un contrat d'assistant-ingénieur de 10 mois (janvier-octobre 2020) s'est chargé de ce volet.

Par la suite, l'intégration des données est faite par la coordinatrice du projet en collaboration avec le chercheur post-doctoral à l'issue de ses missions et lorsque chaque expert fournit ses éléments d'enrichissement des données. On peut évaluer à **quatre jours par mois pendant les 2 années suivantes le travail à fournir.**

Pour éviter un coût concernant le stockage et éviter des problèmes liés à la pérennité de ce stockage sur le serveur Zotero, les PDF liés à la bibliographie sont rassemblés dans un dossier dédié sur la plateforme Sharedocs (<https://sharedocs.huma-num.fr/>)

Les données suivront les principes FAIR :

facilement trouvables : car disponibles via une plateforme internet et gratuitement ; les métadonnées seront préparées en suivant les principes du LOD, alignées sur des ontologies bien référencées ;

accessibles : car elles seront disponibles en Open Access et à l'aide d'un simple navigateur.

interopérables avec plusieurs autres bases de données existantes (Petrae, Trismegistos, Eagle, etc.) ;

réutilisables directement via la possibilité de télécharger les fiches préparées en xml ou en PDF, ce qui constituera un renouvellement de l'édition des inscriptions gauloises en temps réel.