

## **Projet ArchivesBaelo**

#### Rapport scientifique final

Rapport scientifique rédigé par Gwladys Bernard, Maëla Le Péron, Bruno Morandière, avec la collaboration d'Elisabeth Collantes.



Septembre 2020-Novembre 2022

## Sommaire

| I. L'équipe du projet Baelo4   |
|--|
| 1. Des partenaires franco-espagnols complémentaires et historiques 4   |
| 2. Pilotage et répartition des tâches 5  |
| II. Récolement et numérisation des fonds archéologiques relatifs à Baelo Claudia7  |
| 1. Méthodologie d'inventaire et de numérisation 7  |
| 2. Mouvements et transports internationaux des archives 8  |
| 3. Adaptations de la première étape d'inventaire et numérisation 9   |
| III. Guide des sources sous AtoM, consolidation, stockage et mise en ligne des métadonnées9  |
| 1. Partenaires impliqués et organisation générale 9  |
| 2. Choix d'un guide des sources et de l'application AtoM 10  |
| 3. Plan de classement 11   |
| 4. Licence et préparation de la mise en ligne12  |
| 5. Le versement des métadonnées dans Huma-Num 12   |
| IV. Les résultats au regard des objectifs15  |
| 1. Un récolement et une numérisation terminés, un guide des sources vivant et inscrit pleinement dans une perspective de Science ouverte |
| 2. Des perspectives toujours ouvertes 16   |
| 3. Valorisation du projet et exploitation scientifique   |

#### <u>Introduction : les objectifs du projet CollEx ArchivesBaelo</u>

Bénéficiant d'une approche interdisciplinaire et d'une dimension internationale, le projet CollEx-Persée ArchivesBaelo avait pour objet la conservation, l'archivage, la numérisation et la valorisation de la documentation archéologique française produite depuis 1917 par les chercheurs liés à la Casa de Velázquez de Madrid sur le site antique de Baelo (Tarifa, Cadix). Visitée par 150 000 visiteurs par an, la cité antique de Belo¹ a été d'abord fouillée entre 1917 et 1921 par Pierre Paris, archéologue français fondateur et premier directeur de la Casa de Velázquez en 1928. C'est toujours sous l'égide de cette institution que des fouilles scientifiques ont repris en 1966, sous la conduite de Claude Domergue et de Manuel Pellicer. Les fouilles de la Casa de Velázquez y ont continué jusqu'en 1991, avant de passer le relai à la fois à la *Junta de Andalucía* pour la conservation du site ainsi que, pour les opérations archéologiques, à l'Institut de recherche en architecture antique de Pau (IRAA) et aux universités de Séville et Cadix. En 2022, plus de seize équipes espagnoles, françaises et internationales interviennent à des degrés divers sur le site, dont la conservation et l'exploitation touristique sont assurées par le *Conjunto arqueológico de Baelo Claudia*.

L'objectif central du projet était de regrouper les documents des fouilles françaises aujourd'hui dispersés dans différents fonds sur une même base numérique, afin d'être consultables et de constituer un réservoir de données scientifiques. Les ensembles d'archives conservés à la Casa de Velázquez et à l'Institut de recherche sur l'architecture antique de Pau (IRAA de Pau) sont constitués d'une grande variété d'éléments papier qui représentent au total huit mètres linéaires. Les tâches étaient de :

- poursuivre et finaliser la numérisation des fonds détenus par la Casa de Velázquez et l'IRAA de Pau
- -intégrer les nouveaux versements reçus à l'occasion du Centenaire des fouilles de Baelo en 2017
- récupérer et mettre en forme les métadonnées des institutions espagnoles
- déposer les fichiers numérisés dans Nakala
- mettre à disposition du public et des chercheurs un instrument de recherche sous AtoM
- parallèlement développer une plateforme numérique bilingue permettant d'éditorialiser l'ensemble de la documentation archéologique du site de Baelo
- intégrer le corpus dans l'espace « chantiers de fouilles » de la future Perséide des Écoles françaises à l'étranger

Les objectifs du projet étaient donc ambitieux, comme l'avait souligné à juste titre le jury du projet Collex Archives Baelo.

3

<sup>&</sup>lt;sup>1</sup> L'identification des ruines présentes sous le village de Bolonia avec la cité de *B(a)elo Claudia* citée par l'*Itinéraire Antonin* a été définitivement permise par la découverte en 1973 du piédestal de Quintus Pupius Vrbicus, duumvir du municipe de Baelo Claudia.

#### I. L'équipe du projet Baelo

# 1. Des partenaires franco-espagnols complémentaires et historiques

L'équipe initiale, composée de deux services de la Casa de Velázquez, c'est-à-dire l'École des Hautes Etudes Hispaniques et Ibériques dirigée pour sa section ancienne et médiévale par Laurent Callegarin jusqu'en septembre 2020, puis par Gwladys Bernard à partir du démarrage du projet, et le service Bibliothèque et archives, dirigé par Elisabeth Collantes, ainsi que de l'Institut de recherches sur l'Architecture Antique de Pau et du Centro documental de Arqueología y Patrimonio (CeDAP) de la Universidad Autónoma de Madrid, a commencé dès septembre 2020 les premières réunions de travail, en visioconférence essentiellement étant donné le contexte pandémique.

L'Institut de recherches sur l'Architecture antique de Pau a travaillé sur le site de Baelo Claudia pendant 35 ans, en prenant progressivement la relève de la Casa de Velázquez à partir des années 1980 sur la fouille du site et pour la conservation des archives. L'IRAA a notamment fouillé le temple d'Isis et le théâtre de Baelo Claudia. La documentation archéologique conservée à l'IRAA de Pau comprend 3 mètres linéaires de documents de nature diverse, inventoriés dans un gestionnaire de données Filemaker mais non numérisés. L'IRAA de Pau, grâce au travail de Myriam Fincker, a constitué un travail de chronotopie numérique, qui permet de suivre le dégagement sur un siècle des principaux édifices de Baelo Claudia et qui est une des premières exploitations scientifiques terminées du projet dédié aux archives de Baelo.

Le Centro documental de Arqueología y Patrimonio ou CeDAP, dirigé par Juan Blánquez, est un satellite de l'Université Autonome de Madrid créé en 2014, héritier d'un premier service créé en 1988, la Aula de Legados documentales. Fort d'une expérience d'inventaire et de numérisation de plus de 50 000 documents archéologiques et 45 000 photographies, le CeDAP, grâce à son expertise en termes de conservation archivistique de documents archéologiques et de legs documentaires, a constitué un partenaire important pour la première partie du travail. En effet, lors du dépôt de l'AAP CollEx, ni la Casa de Velázquez, ni l'IRAA de Pau ne disposaient de poste d'archiviste; Maëla Le Péron, archiviste de la Casa de Velázquez, dont l'implication et le travail ont été essentiels pour le projet, n'était pas encore recrutée lors du premier dépôt de l'AAP et n'avait donc pas été comptée parmi les membres permanents de l'équipe, alors qu'elle a de fait constitué un de ses membres clés.

La Casa de Velázquez, l'IRAA et le CeDAP avaient déjà un premier historique de travail sur les archives de Baelo. Un premier projet mené entre 2016 et 2018 en collaboration entre la Casa de Velázquez, l'Université Autonome de Madrid et l'IRAA de Pau, avait déjà abouti à l'inventaire de 5000 photographies, grâce à la construction d'un outil d'inventaire sous Excel. Ce premier projet, lauréat d'un AAP annuel issu d'une convention spécifique entre la UAM et la Casa de Velázquez, avait également produit une exposition itinérante, entre le site de Baelo et la Casa de Velázquez, ainsi qu'une première publication parue à la Casa de Velázquez à l'occasion du Centenaire des fouilles de Baelo. Ce catalogue de cent images commentées est intitulé Baelo. 100 años de arqueología, 100 imágenes para la memoria, et a été codirigé par Juan Blánquez Pérez, Laurent Callegarin, Lourdes Roldán Gómez, Ángel Muñoz Vicente et Gabriela Polak. Il a constitué à la fois un point de départ pour le travail collectif et les premiers outils d'inventaire comme une première exploitation scientifique des archives de Baelo Claudia.



Fig. 1 Couverture de Baelo. 100 años de arqueología, 100 imágenes para la memoria, 2017, Casa de Velázquez

Cette première équipe avait laissé des jalons que l'équipe du CollEx Persée ArchivesBaelo a d'abord suivis pour faire démarrer le projet. Ensuite, des adaptations ont été nécessaires dans un deuxième temps pour aboutir à un archivage numérique conforme aux attendus du projet et consultable en ligne conformément aux objectifs de Science ouverte.

### 2. Pilotage et répartition des tâches

La réponse à l'Appel à projet Collex avait été élaborée par Laurent Callegarin et Elisabeth Collantes, les deux premiers codirecteurs du projet. Laurent Callegarin, après sept ans passés à la direction des études anciennes et médiévales de l'EHEHI, se devait de revenir à son poste de Maître de conférences à l'université de Pau. Il est donc resté impliqué fortement dans le projet, à un titre différent. Gwladys Bernard, nommée directrice des études anciennes et médiévales à partir du 1<sup>er</sup> septembre 2020, l'a donc remplacé à Madrid.

Les deux services impliqués de la Casa de Velázquez ont assuré pendant deux ans un pilotage bilingue d'une équipe internationale, pour coordonner d'abord les étapes de conservation, transports, inventaire et numérisation, avant l'étape de création et de constitution de l'outil d'archivage en ligne. Le suivi du budget, le recrutement et le suivi des prestataires externes a été également du ressort de la Casa de Velázquez, avec l'appui de l'IRAA de Pau.

Le projet CollEx-Persée ArchivesBaelo se devait de faire appel à de nombreux partenaires pour réaliser à bien les ambitions importantes du projet. La cartographie de l'équipe a évolué au cours des deux ans, en fonction des évolutions de carrière et des restructurations d'activités des équipes.

L'expertise archivistique, avec le suivi du récolement, de la mise en place de l'outil AtoM et de la formation à la première utilisation d'AtoM a été assurée comme prévu par la responsable du service des archives de la MSH Mondes, Elisabeth Bellon (USR 3225).

L'expertise technique et l'ingénierie du projet, initialement réparties entre l'UMS 3602 Persée et l'UMR 5607 Ausonius, ont finalement été intégralement assurées par Bruno Morandière, ingénieur du Réseau des Écoles françaises à l'Étranger.

| Institutions   | Personnels permanents initialement impliqués             | Personnels permanents du projet 2020-2022  |
|--|--|--|
| Casa de Velázquez-pilotage   | 1. Elisabeth Collantes (bibliothèque)                    | 1. Elisabeth Collantes   |
|  | <ol><li>Laurent Callegarin<br/>(EHEHI)- départ</li></ol> | 2b. Gwladys Bernard (EHEHI)  |
| Nouvelles tâches Casa de<br>Velázquez- suivi inventaire et<br>numérisation, mise en place<br>plateforme numérique AtoM |  | 3. Maëla Le Péron (archives)   |
| Nouvelles tâches Casa de<br>Velázquez : suivi juridique et<br>administratif  |  | 4. Renforts Casa de Velázquez : assistance scientifique et administrative, service juridique : Soledad Duran, Clémence Genet, Charlotte Krzyzanski |
| UAM (CeDAP) : suivi inventaire et numérisation   | 5. Juan Blánquez   |  |
| UAM (CeDAP) : suivi inventaire et numérisation   | 6. Lourdes Roldán  |  |
| IRAA de Pau : suivi inventaire et numérisation   | 7. Myriam Fincker-<br>retraite                           | 7b Laurent Callegarin et Véronique Picard  |
| MSH Mondes-expertise archivistique et formation AtoM   | 8. Elisabeth Bellon                                      | Elisabeth Bellon   |
| (Ausonius) (retrait du projet)<br>mise en place plateforme<br>numérique  | 9. (Nathalie Prévôt)                                     |  |
| ResEFE : ajout au projet   |  | 9b. Bruno Morandière   |
| Persée : expertise technique et  | 10. (Sébastien Mazzarese)                                | 10b et 11b. Maëla Le Péron et  |
| numérique  | <ul> <li>nouvelles fonctions</li> </ul>                  | Bruno Morandière   |
| Persée: appui et expertise   | 11. (Eric Astier) –                                      | 10b et 11b. Maëla Le Péron et  |
| numérisation   | nouvelles fonctions                                      | Bruno Morandière   |

Fig. 2 Les personnels permanents initialement prévus et les personnels impliqués dans le projet CollEx ArchivesBaelo

# II. Récolement et numérisation des fonds archéologiques relatifs à Baelo Claudia

#### 1. Méthodologie d'inventaire et de numérisation

Le fonds Baelo Claudia de la Casa de Velázquez englobe des documents liés à la préparation des opérations de terrain, la correspondance des archéologues (lettres et documents administratifs relatifs à la fouille), la production documentaire sur le terrain (carnets de fouilles, documents graphiques et photographiques, l'inventaire des matériels et les mises en forme des résultats pour leur publication et les articles de journaux. Les archives papier représentent cinq mètres linéaires et sont composées d'environ 650 dossiers avec un nombre divers de pièces. La moitié de cette documentation concerne les campagnes de fouilles, un quart concerne les publications et la part restante est ventilée entre les études topographiques et architecturales pour lesquels on compte plus d'un millier de plans, relevés et résultats de sondages. Avaient été versés en 2017 et 2018 dans le fonds de la Casa de Velázquez trois ensembles documentaires qui avaient été conservés par des chercheurs associés à la Casa de Velázquez (Ariane Bourgeois, François Didierjean, Jean-Claude Richard) ou par leurs ayants droits (Mme Jean-Noël Bonneville).

Ce fonds avait fait l'objet d'un inventaire rapide par dossier, afin d'évaluer le temps nécessaire au traitement et à la numérisation de l'ensemble. Il n'était pour autant pas consultable, ayant été déplacé dès 2018 au CeDAP où les conditions de place et de personnel disponibles ne sont pas réunis pour la consultation directe.

Le conditionnement des archives de la CVZ et de l'IRAA de Pau, l'inventaire et la numérisation des 9 mètres d'archives linéaires papier ont donc été effectués au CeDAP de la UAM pour mutualiser l'espace et le personnel disponible au récolement et à la numérisation, avant un retour dans les institutions propriétaires.

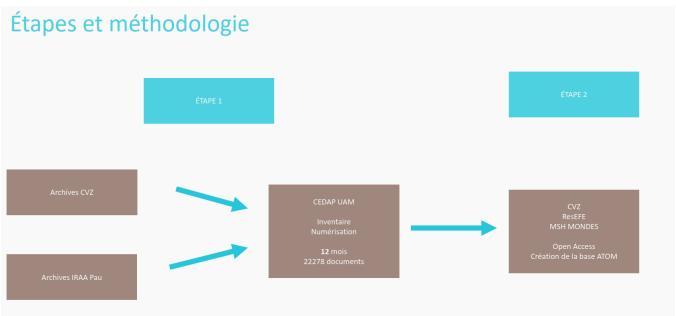


Fig. 3 Schéma des étapes et de la méthodologie

Une chargée d'inventaire a été recrutée dès début novembre 2020 : il s'agit de Gabriela Polak, archéologue spécialisée en histoire de la photographie archéologique, qui avait déjà participé à la publication du Centenaire des Fouilles de Baelo Claudia. Elle a continué l'inventaire pièce à

pièce, qui avait été déjà commencé et avait commandé une première étape de numérisation de 5000 photographies. Ce premier inventaire photographique avait été réalisé à l'aide d'un tableur type Excel bilingue franco-espagnol, constitué de 24 champs, articulé autour de trois rubriques principales :

- l'identification du document, avec 3 numéros d'inventaire (le numéro d'inventaire général au sein du projet, le numéro d'inventaire dans le fonds d'archives initial -CVZ ou IRAA de Pau, la localisation temporaire au CeDAP)
- l'analyse du support photographique : positif, négatif, diapositive, dimensions, couleur etc...
- la description de la photographie : sa date, son auteur, la localisation de son contenu (site, secteur), l'édifice ou le matériel archéologique concernés

Les autres types de documentation archéologique présente dans les fonds de la Casa et de l'IRAA ont été inventoriés avec un outil Excel similaire.

**Bilan de la première phase** (10 mois de travail équivalent temps plein). Entre novembre 2020 et décembre 2021, Gabriela Polak, sous la supervision de la professeur Lourdes Roldán a pu reconditionner et inventorier pièce à pièce :

22278 documents (soit 2000 de plus que le décompte initial) dont :

17523 photographies (11904 CVZ et 5349 IRAA Pau)

5025 documents (3499 documents CVZ et 1527 IRAA Pau)

Quasiment simultanément, la numérisation a repris, toujours dans les locaux madrilènes du CeDAP: le travail a été confié à Carlos Comas-Mata, en prestataire externe via l'entreprise Trebede Ediciones, déjà chargée de la maquette et de l'impression du catalogue *Baelo*, *100 imagenes*. Le travail a été ici réévalué par rapport aux prévisions originelles: il s'est étendu sur 16 mois, a concerné plus de 10 000 numéros d'inventaire et un total de presque 20 000 pages numérisées.

### 2. Mouvements et transports internationaux des archives

- CVZ (Madrid) → CeDAP (Madrid) avant septembre 2020 : Les archives CVZ avaient déjà été déplacées au CeDAP, où elles étaient entreposées dès le début du projet.
- IRAA (Pau) → CVZ (Madrid) Mars 2021: Les archives de l'IRAA de Pau ont été convoyées
  à Madrid par une enseignante-chercheuse de l'université de Pau, Françoise des Boscs,
  rattachée au projet et bénéficiant d'un séjour au Madrid Institute for Advanced Study, à
  la Casa de Velázquez.
- CeDAP (Madrid) → CVZ (Madrid) mai 2021: Au terme de leur inventaire et de leur numérisation, les archives de la CVZ sont revenues à la Casa, où elles ont été entreposées dans la salle de conservation de la bibliothèque. Les archives de l'IRAA de Pau sont allées au CeDAP.
- CVZ (Madrid) → IRAA Pau mars 2022 : Françoise des Boscs a convoyé à Pau la partie inventoriée, mais non numérisée des archives de l'IRAA pour une fin de numérisation au

printemps-été 2022. Restent à ce jour deux mètres linéaires d'archives de l'IRAA de Pau au CeDAP, qui seront renvoyées en janvier 2023 toujours par le biais de la même enseignante-chercheuse.

## 3. Adaptations de la première étape d'inventaire et numérisation

Les critères de numérisation et d'inventaire ont été partiellement adaptés en début et en cours de la première phase :

- La numérisation en très haute définition, utilisée pour les documents et les photographies présentées dans l'exposition de 2017 (600 dpi) a été abandonnée pour une réduction de temps et d'espace.
- A l'occasion du transfert des archives de Pau à Madrid en mars 2021, un premier tri a été effectué: un certain nombre de pièces administratives destinées à terme à l'élimination, ont été conservées, mais signalées et exclues de l'inventaire et de la numérisation du projet.
- La partie du budget consacrée aux réunions a été réinvestie dans la fin de la numérisation, ainsi que dans d'autres tâches initialement non évaluées (cf infra)
- La fin de la numérisation d'une partie des archives de l'IRAA a été assurée à Pau, par un prestataire externe, afin de laisser davantage de temps aux travaux d'import dans l'outil AtoM, temps qui avait été sous-évalué dans le projet initial.

# III. Guide des sources sous AtoM, consolidation, stockage et mise en ligne des métadonnées

#### 1. Partenaires impliqués et organisation générale

Fin avril 2022, le travail de numérisation au CeDAP a dû s'arrêter, pour laisser place au processus de création de répertoire sous AtoM, de versement dans Nakala et d'import dans ces outils en accès ouvert et répondant aux principes FAIR (données trouvables, accessibles, interopérables, réutilisables).

Les autres partenaires sont intervenus au cours de cette deuxième étape, commencée à la fin 2021.

L'investissement d'Elisabeth Bellon, archiviste de la MSH Mondes, a permis de valider un plan de classement solide, détaillé plus bas. Il s'agissait de passer d'un inventaire pièce à pièce, fondé sur l'examen d'un document ou d'une photographie dans un classeur donné, à une collection virtuelle réunissant l'ensemble des archives relatives à Baelo Claudia des deux producteurs. Cette collection virtuelle est présentée sous la forme d'un guide des sources réalisé sous AtoM. Le personnel de la MSH a formé à l'usage d'AtoM Maëla Le Péron, archiviste de la Casa de Velázquez et les imports ont été réalisés avec le concours de Bruno Morandière, ingénieur du Réseau des Écoles françaises à l'étranger.

Parallèlement à l'organisation de cet import des métadonnées et des vues numérisées dans AtoM, deux tâches supplémentaires ont été conduites lors de la deuxième année du projet :

- La normalisation et la correction des inventaires par une chercheuse francophone:
   Marie de Jonghe, docteure en archéologie de l'université Paris Nanterre, a réalisé deux mois de prestation au printemps 2022, pour corriger et normaliser les fichiers Excel avant le passage au format csv.
- La consolidation juridique des droits de publication des documents, tâche signalée par les experts et menée à bien depuis la Casa de Velázquez, avec les personnels permanents.

En effet, les fonds réunis appartiennent à la Casa de Velázquez et à l'IRAA de Pau ; les deux institutions ont financé les travaux et rémunéré les chercheurs qui sont intervenus dans les missions archéologiques françaises à Baelo Claudia. Cependant, les auteurs des documents, chercheurs, enseignants-chercheurs et architectes, conservent leurs droits d'auteurs sur les documents signés et doivent donner l'autorisation à la publication des vues numériques dans la base de données en ligne.

Ont donc été extraits les noms d'auteurs de l'inventaire, pour tous les documents signés et identifiés et chaque auteur ou ayant droit identifié a reçu un contrat de cession de droits, accompagné d'une annexe désignant précisément les documents destinés à la mise en ligne. Ce contrat a été élaboré par le service juridique de la Casa de Velázquez; aucun refus de communication n'a été reçu, les chercheurs ayant bien compris et perçu l'importance de leur accord pour la mise en œuvre du projet.

#### 2. Choix d'un guide des sources et de l'application AtoM

Plusieurs acteurs ont travaillé sur le site archéologique de Baelo Claudia. Chacun a produit des documents conservés dans des administrations différentes au sein des fonds d'archives auxquels ils sont rattachés. Entre la Casa de Velázquez et l'IRAA de Pau, ce sont plus de 22 000 documents qui ont été inventoriés et tous ont pratiquement été numérisés (une minorité était en mauvais état et impropre à la manipulation et à la numérisation).

Il a donc fallu structurer ces données de façon à répondre au mieux aux projets de réutilisation à des fins scientifiques, archéologiques et historiques. Le choix du guide des sources s'est alors imposé. Par définition, un guide des sources est un instrument de recherche décrivant au niveau approprié les fonds ou collections relatifs à un thème donné, indépendamment des lieux de conservation et des systèmes de classement. Il permet donc de répondre de façon structurée à la problématique de l'éparpillement des données chez plusieurs producteurs, via un plan de classement « virtuel » et tout en s'assurant donc du respect des fonds.

AtoM (contraction de l'appellation Access to Memory) est une application en open source pour la description archivistique et l'accès aux instruments de recherche et aux archives mises en ligne. Elle permet de réaliser et de diffuser tout type d'instrument de recherche conforme aux directives du Conseil international des archives. L'environnement multilingue, la possibilité d'associer aux notices des archives en ligne (pièce à pièce ou par lot), les fonctionnalités liées à la gestion des droits et des localisations ont été déterminants quant au choix de cet outil.

#### 3. Plan de classement

Le plan de classement qui structure l'inventaire est organisé par thématiques et secteurs, du plus général (la région) au particulier (les bâtiments). Chaque niveau de titre est rédigé en français et en espagnol.

Une partie est consacrée au site. Elle contient les documents relatifs au site en général, les documents concernant plusieurs bâtiments, les documents n'ayant pu être rattachés à un bâtiment en particulier. De plus, pour faciliter la recherche, nous avons aussi intégré dans cette partie les journaux de fouilles, les inventaires de matériel et les rapports de fouilles.

Chaque partie a ensuite été organisée avec des thématiques intermédiaires telles que «le terrain», «le matériel archéologique» et «la préparation des publications». Le terrain contient tous types de photographies et de documents graphiques organisés en sous-thématiques (bâtiment, procédés de fouilles, conservation, personnes...). Le matériel archéologique comprend également tous types de photographies et des dessins de matériel. Enfin, dans la partie « préparation des publications », on trouve les documents préparés pour la publication, soit dans la revue des *Mélanges de la Casa de Velázquez*, soit dans les 9 ouvrages actuellement publiés dans la collection *Casa de Velázquez* et appartenant à la série « Belo ».

À cet ensemble s'ajoute une partie consacrée à l'organisation des campagnes de fouilles. Cette partie, actuellement vide, sera enrichie plus tard (mais les documents sont inventoriés et numérisés). En effet, il reste à vérifier la communicabilité de ces documents, notamment qu'ils ne contiennent pas de données à caractère personnel (cf supra).

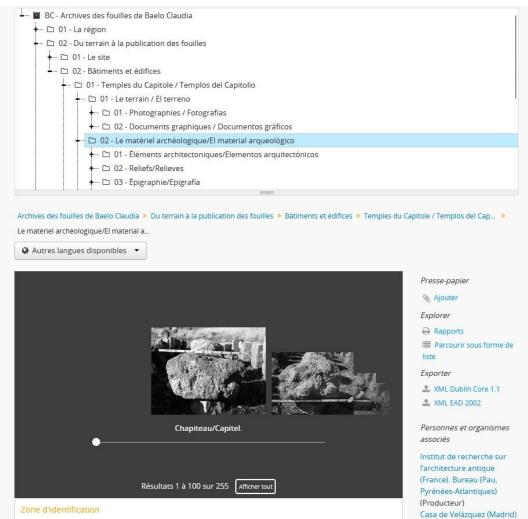


Fig. 4 Capture d'écran du guide des sources AtoM

### 4. Licence et préparation de la mise en ligne

Les documents mis en ligne sont sous licence Creative Commons.

L'inventaire des archives a débuté avant le financement du projet par CollEx-Persée et la décision d'utiliser la solution AtoM. Or les fichiers d'inventaires des documents avaient été commencés sur des fichiers Excel conformes à l'utilisation faite par le CeDAP, mais non normés pour l'import dans AtoM. Aussi, une fois les inventaires achevés, un premier travail de relecture et de normalisation a été effectué sur l'ensemble des fichiers à l'exception de ceux, plus sériels, des photographies. Cette phase terminée, il a ensuite fallu convertir les fichiers Excel en fichiers .csv compatibles pour des imports dans la plateforme et les répartir selon leur destination dans le plan de classement. Cela a été possible grâce à une formation à la manipulation du logiciel et à la préparation des imports. Cette tâche, quoique simple et répétitive, a demandé beaucoup de temps et les données contenues dans les fichiers Excel n'ont pas encore été toutes intégrées dans des fichiers .csv à la fin 2022.

Parallèlement, un travail de contrôle des images a été effectué afin de s'assurer de la possibilité de leur mise en ligne puis la création des liens images a permis de lier les notices descriptives aux documents numérisés.

Une fois les fichiers .csv complets (notices + liens images), il a été possible de les importer dans le plan de classement créé au préalable dans AtoM.

#### 5. Le versement des métadonnées dans Huma-Num

A l'issue du travail initial de récolement des données, de leur numérisation et le renseignement des métadonnées que les institutions ont à disposition, une longue phase de curation et de consolidation des métadonnées (à l'aide de fichiers Excel) a été nécessaire.

Toutes ces données et métadonnées sont versées au sein de l'infrastructure de recherche Huma-Num.

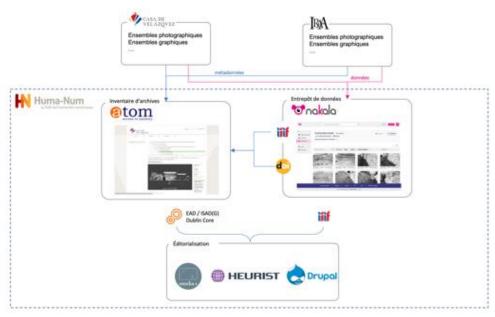


Fig. 5 : schéma du versement des métadonnées

Nous avons pu automatiser (à l'aide de l'outil Nakala Workflow) le versement des données et des métadonnées (issues des fichiers Excel) dans l'entrepôt de données Nakala qui permet d'enregistrer tout type de données, de les décrire en vue de les exposer et les rendre réutilisables.

Pour chaque ressource déposée un DOI a été récupéré.

Ces DOI nous permettent d'accéder à toutes données du fonds Baelo (18 500 ressources pour l'instant soit 80 % de l'ensemble du fonds) via le protocole IIIF. Les URL IIIF ont été associées aux métadonnées de chaque ressource et intégrées dans les fichiers Excel qui restent les fichiers pivots du projet.

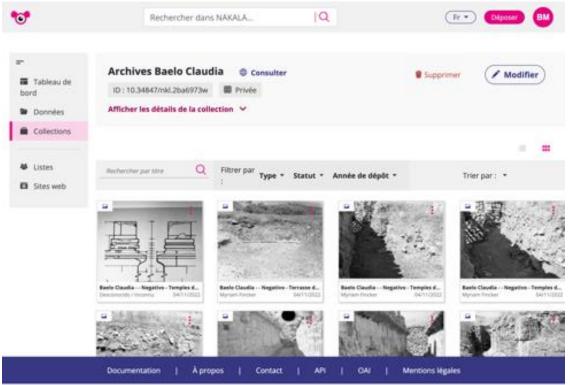


Fig. 6 Capture d'écran du versement dans Nakala

Ces fichiers Excel ont pu ensuite être versés directement (après une conversion en CSV) dans la solution AtoM. Les données dans cet outil respectent ainsi les normes archivistiques en particulier ISAD(G) et ISAAR et reposent sur le standard XML EAD. Deux instances AtoM ont été déployées, la première sur un serveur de la Casa de Velázquez pour la préproduction, la seconde au sein de l'infrastructure Huma-Num pour les données en production.

À partir des métadonnées normalisées en EAD et/ou Dublin Core et des ressources de Nakala accessibles via le protocole IIIF, d'autres interfaces de consultation (avec des plateformes comme Omeka, Heurist ou Drupal toujours hébergées par Huma-Num) pourraient éventuellement répondre aux besoins des chercheurs avec un point de vue moins archivistique et des navigations différentes sur ces données. On pourra envisager cette question selon l'usage des chercheurs et leurs premiers retours sur l'interface AtoM.

#### IV. Les résultats au regard des objectifs

### Un récolement et une numérisation terminés, un guide des sources vivant et inscrit pleinement dans une perspective de Science ouverte

À l'issue de ces deux années, le projet CollEx-Persée a permis d'achever le récolement et la numérisation des archives consacrées aux fouilles de Baelo Claudia de 1917 à la fin des années 1980/début des années 1990, la création d'un instrument de recherche et la mise en ligne d'une partie des documents numérisés. Le travail présenté ici n'est pas encore achevé mais 50 % des documents sont actuellement présents dans la base. Si tous les documents ont bien été inventoriés et numérisés et si la moitié est désormais accessible en ligne, il reste encore à préparer la mise en ligne des documents liés aux autres bâtiments du site et ceux liés à l'organisation des campagnes de fouilles.

Ce projet répond à une demande des chercheurs à l'échelle internationale. Des équipes d'archéologues continuent de fouiller le site et le libre accès aux archives est un enjeu important. À terme, ces chercheurs pourront à leur tour décider d'alimenter l'instrument de recherche afin de l'actualiser avec de nouvelles données. En effet, ce guide des sources n'est pas un instrument de recherche figé, c'est un outil vivant qui peut être enrichi par l'apport de documents/données et qui s'inscrit parfaitement dans la dynamique de la science ouverte.

Au terme de ce projet CollEx, nous avons l'objectif d'intégrer l'ensemble des données et métadonnées au sein d'une Perséide « Écoles françaises à l'étranger ».

Persée décline son savoir-faire en termes de valorisation du patrimoine documentaire scientifiques (en assurant sa diffusion, son enrichissement et sa préservation) pour les corpus via la création de Perséides au sein desquelles il accompagne les communautés scientifiques afin d'assurer la production de corpus, les rendre accessibles et visibles et ainsi permettre leur exploitation scientifique.

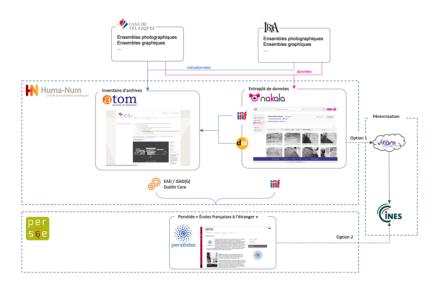


Fig. 7. Schéma d'intégration dans la Perséide

L'idée d'une Perséide « Écoles françaises à l'étranger » a germé dès 2019 avec l'objectif de regrouper dans un même espace les publications des Écoles qui sont sur Persée et travailler à des passerelles entre celles-ci et les données des Écoles (notamment les fonds d'archives). Un des points d'entrée de cette Perséide serait les sites de fouilles, et plusieurs projets (Collex-Persée, FNSO, ...) vont dans ce sens. Le projet Baelo alimentera donc naturellement cet espace. Enfin cette Perséide permet d'envisager la numérisation d'ouvrages relatifs aux opérations de fouilles qui ne sont pas disponibles numériquement. Les ouvrages papier concernant la fouille de Baelo pourraient y trouver un espace de diffusion pertinent.

Pour l'instant la Perséide EFE est toujours à l'état de gestation pour des contraintes techniques mais la volonté qu'elle voit le jour est toujours affichée de la part de Persée comme du côté du Réseau des Écoles.

Concernant la question de la conservation des données, elles sont pour l'instant hébergées par le service Nakala qui, au-delà de l'exposition, leur assure une sauvegarde. La question de leur archivage au CINES (Centre informatique national de l'enseignement supérieur) dépend de notre avancée sur la Perséide. La solution de l'archivage via Persée (qui a développé depuis son début des mécanismes de versement au CINES) aurait notre préférence mais on pourrait également s'appuyer sur la solution en cours de développement par Huma-Num sur l'interface interministérielle Vitam avec le CINES.

ArchivesBaelo est un projet assez simple techniquement. Il peut être répliqué assez facilement. Il s'appuie massivement sur les infrastructures nationales qui garantissent l'exposition et l'accessibilité des données ainsi que leur sauvegarde sur le temps long.

Le projet propose à l'issue de ces 2 ans :

- Des données numérisées dans des formats réputés archivables (Tiff, Pdf-A, ...)
- Des métadonnées qui respectent les standards (.xml, .ead, .dublin core)
- Une licence qui explicite l'utilisation qui peut être fait des données et des métadonnées Le projet respecte donc les principes FAIR : Facile à trouver, Accessible, Interopérable, Réutilisable et s'inscrit ainsi entièrement dans la dynamique de la science ouverte.

#### 2. Des perspectives toujours ouvertes

Au niveau des partenaires, une des équipes initiales, l'UMR Ausonius, avec la collaboration prévue de Nathalie Prévost, n'a pas donné suite : la participation prévue d'Ausonius était la constitution et la maintenance d'une plateforme d'éditorialisation, sous Omeka-S, plus intuitive pour le public et les chercheurs que le guide de sources, mais le temps contraint de Nathalie Prévost ne lui a pas permis de s'investir sur le projet, au-delà du prototype conçu en 2018. La piste de l'outil Omeka n'a pas été continuée pour l'instant, pour se concentrer sur le guide des sources AtoM.

Les prolongements du projet au-delà des fonds papier, avec l'intégration des métadonnées d'autres archives en Espagne sur Baelo, est un chantier qui reste ouvert : les premiers contacts pris montrent qu'en fait, bien souvent, si des documents existent et seraient disponibles dans les fonds d'archives identifiés, ils ne sont ni complètement inventoriés, ni indexés, ni numérisés. Intégrer des métadonnées qui pour l'heure, ne sont pas existantes, est un objectif qui s'est avéré trop ambitieux pour les deux ans du projet. La collection virtuelle des archives Baelo est

constituée et pourra s'élargir bien entendu à d'autres fonds espagnols ou internationaux, mais quand ceux-ci auront entrepris l'effort d'inventaire et de numérisation préalable à un partage de ressources numériques et interopérables.

### 3. Valorisation du projet et exploitation scientifique



Fig. 8 Programme de la journée du 16 novembre de présentation des résultats du projet CollEx-Persée Archives Baelo



Les partenaires se sont réunis le 16 novembre 2022 à la Casa de Velázquez, pour **une journée internationale d'études** en format hybride dédiée au projet à la communauté des chercheurs, avec la présentation des résultats et des partenaires du programme Persée-CollEx, des recherches archéologiques actuelles sur le site et les premiers exemples d'exploitation scientifique.

Le couplage de cette journée avec l'hommage, à l'occasion de son 90<sup>e</sup> anniversaire, à M. Claude Domergue, premier codirecteur français de la renaissance du chantier en 1966, a permis à de

nombreux archéologues espagnols et français de plusieurs générations de se retrouver autour des travaux archéologiques passés, présents et futurs sur le site de Baelo Claudia.

L'exploitation scientifique des archives, en plus de la publication du premier catalogue en 2017, s'oriente vers trois directions majeures :

- L'évolution du chronotype du site et les travaux de restitution 3D, menés de concert par Myriam Fincker, l'IRAA de Pau et l'équipe du Conjunto de Baelo Claudia, à partir des archives de la Casa de Velázquez et de Pau. Une première publication est prévue dans les Mélanges de la Casa de Velázquez.
- L'appui aux nouvelles fouilles présentes sur le site: les archives du secteur intermédiaire entre le marché et le forum, actuellement fouillé par Laetitia Borau (CNRS Ausonius avec l'appui de la CVZ et du MEAE), sont particulièrement intéressantes pour les nouveaux travaux d'étude de l'hydraulique du site, comme les archives concernant les boulangeries de Baelo, dont l'étude est reprise par Oliva Rodríguez, de l'université de Séville, dans le cadre de son projet I+D+I (financement ministère de la Recherche et de l'innovation espagnol) sur le recyclage des matériaux.
- L'appui aux travaux de corpus concernant Baelo et sa région : Helena Gimeno, dans le cadre de la constitution du *Corpus Inscriptionum Latinarum* de la région de Gadès (Cadix) s'intéresse particulièrement aux archives épigraphiques et aux archives de la publication du volume *Belo V* consacré aux inscriptions (Jean-Noël Bonneville, Sylvie Dardaine, Patrick Le Roux, 1988).

La journée d'études a permis aux chercheurs de se rencontrer et d'échanger avec les archivistes et les partenaires du projet CollEx-Persée Baelo ; de nouveaux versements ont ainsi été réalisés et sont venus enrichir encore les archives.

Grâce au soutien du programme CollEx-Persée, le projet a permis aux partenaires de réaliser des étapes décisives en terme non seulement de récolement et de numérisation, mais également en terme de construction d'un outil de science ouverte accessible au public, durable, réplicable et déjà largement exploitable par la communauté scientifique internationale.

| NOM de l'institution partenaire                                 | ROLE AU SEIN DU PARTENARIAT  |   |
|---|--|---|
| Casa de Velázquez   | <ul> <li>✓ Management du projet</li> <li>✓ Pôle de conservation des archives</li> <li>✓ Opérateur de numérisation (externalisée)</li> <li>✓ Exploitation scientif que du corpus</li> <li>✓ Création d'un guide des sources dans AtoM</li> </ul>                      |   |
| UMS 3602 Persée   | ✓ Expertise technique de la numérisation   |   |
| Relai ResEFE (Bruno Morandière)                                 | √ Intégration du corpus dans la Perséide des EFE à l'issue du programme  |   |
| USR 3155 – Institut de Recherche sur<br>l'Architecture Antique  | <ul> <li>✓ Pôle de conservation des archives</li> <li>✓ Création d'un instrument de classement et de recherche archivistique</li> <li>✓ Création des métadonnées (csv-xml EAD)</li> <li>✓ Exploitation scientif que du corpus</li> </ul>                             | Rôle passé à la CVZ<br>IDEM               |
| CNRS - UMR 5607 (Ausonius, Bordeaux) Partenaire sorti du projet | archéologique de Baelo » (sous Omeka)  X Expertise documentaire et technique sur la plateforme Omeka (gestion des métadonnées, ergonomie, choix des plugins)   | OMEKA non réalisé<br>Rôle passé au ResEFE |
| Service des archives de l'USR 3225, MSH<br>Mondes, Nanterre     | <ul> <li>✓ Suivi du récolement et du traitement intellectuel et matériel des archives archéologiques</li> <li>✓ Aide à la mise en place de l'instrument de recherche sous AtoM</li> </ul>  |   |
| Universidad Autónoma de Madrid<br>(CeDAP)                       | <ul> <li>✓ Aide au conditionnement des archives (conservation)</li> <li>✓ Création d'un inventaire (Excel)</li> <li>✓ Création des métadonnées (Excel)</li> <li>✓ Opérateur de numérisation (externalisée)</li> <li>✓ Exploitation scientif que du corpus</li> </ul> | Rôle passé à la CVZ<br>(outil AtoM)       |

#### Légende

- ✓ Réalisé
- √ En cours de réalisation
- ✓ Réalisé par un autre partenaireX Non réalisé