

Consortium Huma-Num « Paris Time Machine » (PTM)

Lettre d'information 2021 N°1

Cette newsletter fait un point synthétique sur les activités du consortium durant cette année 2020 et sur les projets pour l'année à venir.

Il s'agit d'une première newsletter qui reprend les points essentiels du projet de consortium et précise les résultats obtenus à ce jour.

D'autres numéros seront publiés dans le courant de l'année à venir à raison d'une tous les 6 mois environ. Des informations plus détaillées sont disponibles sur le site web du projet à l'adresse <https://paris-timemachine.huma-num.fr/>

Un consortium pour travailler sur les référentiels géo-historiques.

La constitution d'un référentiel géo-historique doit permettre à des chercheurs de différentes disciplines historiques de travailler ensemble sur un espace géographique commun. C'est le postulat de base qui a amené à la labélisation du consortium Paris Time Machine en 2019 au sein de TGIR Huma-num.

La notion de référentiel géohistorique est conçue dans le cadre des travaux de ce consortium, comme un concept heuristique permettant à des programmes de recherche ou des équipes de développer des pratiques innovantes et partageables en matière d'analyse objets géo-historiques. Il s'agit plus spécifiquement de donner les moyens aux spécialistes des sociétés du passé de spatialiser leurs données et d'accéder, pour des périodes où ces référentiels n'existent pas ou sous une forme difficilement exploitable avec les outils actuels, à des données et des outils qui répondent à leurs problématiques et leurs permettent de partager leurs approches dans les principes FAIR (Findable, Accessible, Interoperable, Reusable) et open data.

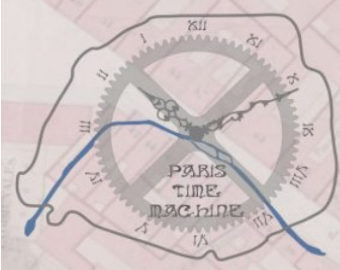
Paris a été pour notre groupe de recherche un espace d'expérimentation, sorte d'archétype, sur lequel nous avons axés nos premiers travaux par la construction d'une série de « chantiers types » permettant d'explorer de manière pragmatique les différentes expressions de la notion de référentiel géo-historique, le but de moyen terme étant d'aboutir à une série de guide méthodologiques sur la constitution et l'utilisation de ces référentiels.

D'un point de vue opérationnel deux ensembles de chantiers avaient été définis dans le projet initial : les chantiers « données » et les chantiers « outils ».



Les principes FAIR

Les acquis au niveau des données :



4 groupes de travail ont produit (et produisent encore) une série de données référentielles traitant différents aspects du territoire parisien, des adresses au parcellaire ancien, de l'altimétrie ancienne au recensement de population.

Le chantier annuaires et adresses a engagé des travaux sur « l'OCRisation » de l'annuaire des Propriétaires, publication annuelle active entre 1894 et 1937 accessibles sur Gallica. La masse de données à récupérer (environ 40 000 propriétaires par volume) exigeait de travailler avec un système d'extraction d'information ayant recours à un modèle d'apprentissage automatique. Conçu à l'origine pour les corpus dictionnaires, il a été tenté d'adapter GROBID-Dictionaries (outil d'encodage automatique en TEI) aux documents de type annuaire avec des premiers résultats qui demandent encore à être développés.

R. 1 RABELAIS (Rue) ^{30,31}

- 1 Eiffel, Paris, r. Rabelais, 1.
- 3 Richelot et Ostheimer, Paris, r. Rabelais, 3.
- 5 Entrée r. Montaigne, 26.
- 2-4 Gérard (B^{no}), Paris, r. Rabelais, 2.
- 6-8 Aligré de Préaulx (M^{re} d'), Paris, fg St-Honoré, 89.

M. Eiffel est le propriétaire du n° 1 de la rue Rabelais, où il réside.

M. 253

MONTAIGNE (Avenue) ²⁹

- 1-3 Hennessy, Paris, r. Bassano, 31.
- 5 Salignac-Fénelon (V^{te} de), av. Rapp, 20.
- 7 Norbert, Paris, r. Lubeck, 42.
- 9 Chateaubriand (C^{esse} de), av. Montaigne, 9.
- 11 Lesseps (C^{esse} de), Paris, av. Montaigne,

M. Hennessy est le propriétaire du n° 1-3 de la rue Montaigne et réside au 31 de la rue Bassano.

Les propriétés sont listées par rues. Les chiffres à gauche indiquent le numéro de l'immeuble, suivis par le nom du propriétaire et son adresse

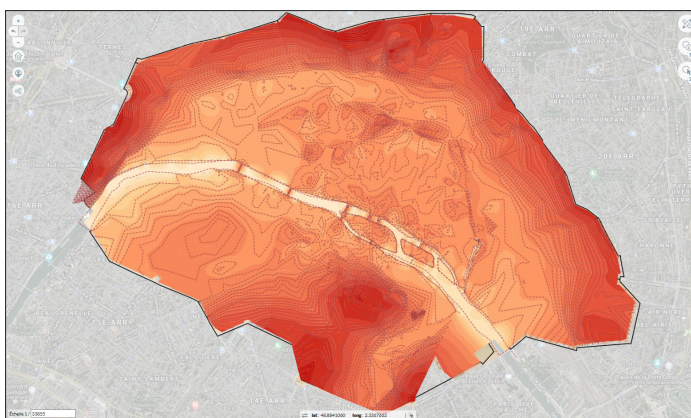
Le chantier parcellaire ancien s'est fixé comme objectif de construire un référentiel parcellaire pour la ville de Paris intermédiaire entre les plans Vasserot (premier « cadastre » levé entre 1810 et 1836) et le plan actuel, en géoréférençant les plans du cadastre municipal établi à la fin du XIX^e siècle. Ces 142 plans, divisés en 732 feuilles régulières pour former un Atlas mis en vente à la feuille à partir de 1903 sont maintenant géoréférencés et ont été mis en ligne sous la forme de flux de données intégrables dans un SIG. Leur vectorisation dans le cadre d'un nouveau groupe de travail a été engagée.



Cadastre municipal : 732 feuilles assemblées (à droite) et zoom sur le premier arrondissement (à gauche)



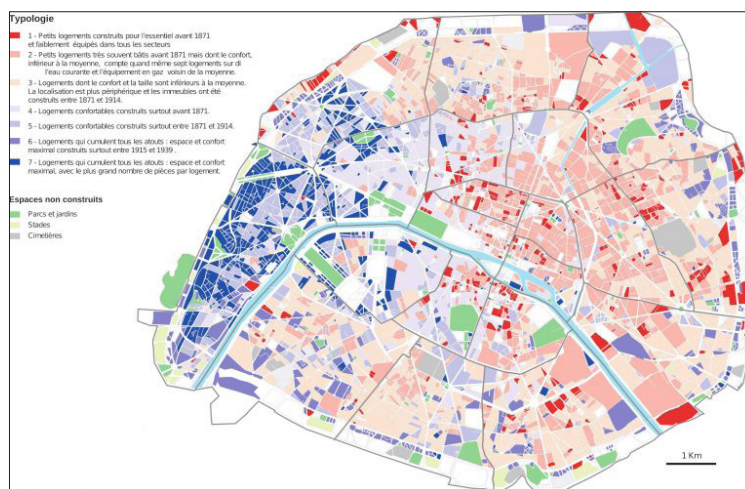
Le chantier modèles numérique de terrain s'est fixé comme objectif de travailler sur l'aspect tridimensionnel des données pour obtenir une série de couches topométriques sur cartes de Paris et le département de la Seine permettant de restituer les différents états du sol avant les grandes phases d'aménagement de la fin du XIXe s. Ces données ont permis d'ajouter au MNT actuel, trois MNT de référence historiques : (1) Le relief au début XIXe siècle sur l'emprise des Fermiers Généraux ;



Le relief au début du XIXe s. sur l'emprise de fermier généraux

;(2) Le relief de la seconde moitié du XIXe siècle sur l'emprise du département de la Seine, issus de l'interpolation des données historiques mesurées par les corps d'ingénieurs à ces époques ; (3) le paléorelief sur l'emprise de Paris, modélisé par l'interpolation l'altimétrie du toit du substrat géologique mesurée par l'IGC au fil de divers sondages et travaux menés dans Paris. Ces données seront prochainement mises à disposition dans le cadre de l'open data et du projet Fabrique numérique du passé.

Le Chantier Seine 1954 concerne la mise en place d'une base de référence sur le recensement de la population française de mars 1954, un travail de collecte des données statistiques sur la population et les logements des communes de la Seine a été conduit conjointement par l'INSEE et l'Institut d'Aménagement et d'Urbanisme de la Région Parisienne. Ce travail a abouti à la publication d'un fascicule pour 72 des 81 communes du département de la Seine. Ce recensement a été dépouillé par îlot. Grâce au soutien de la Bibliothèque de l'Hôtel de Ville et de sa directrice Valérie Alonzo, nous avons pu disposer des quatre volumes regroupant les fascicules concernant la Seine-banlieue. Ceux-ci ont été numérisés et spatialisés pour constituer un nouveau référentiel qui sera prochainement disponible sous la forme de flux de données. Le chantier a été soutenu financièrement par le projet *Archival city*.

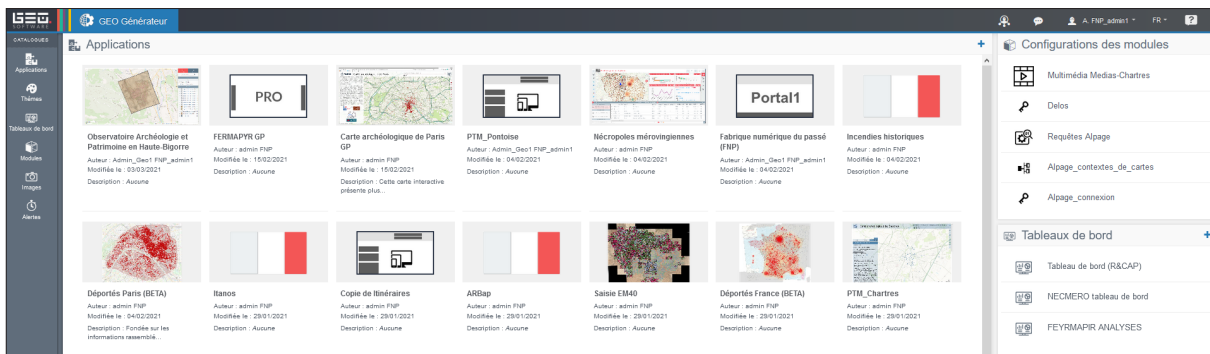


Chantier Seine 1954 : Exemple de traitement, typologie du confort des logements en 1954 à Paris

Les acquis au niveau des des outils :

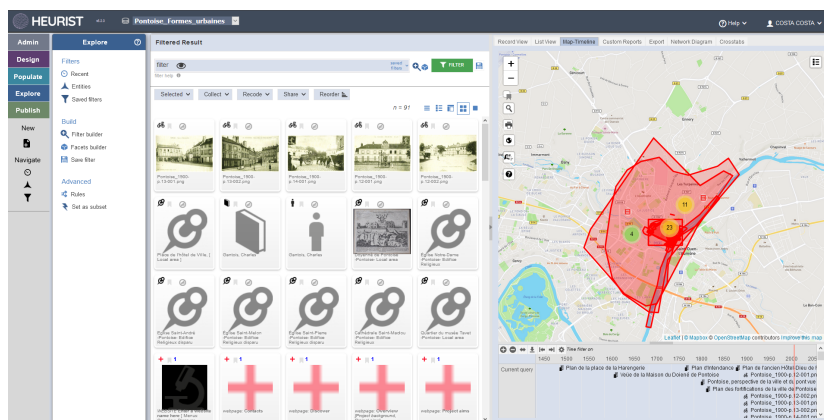
4 chantiers ont été développés durant la première année d'existence du consortium.

Le chantier GEO, du consortium Paris Time Machine, mis en place initialement dans le cadre du programme ANR Alpage puis repris dans le cadre du programme transversal ArcheoFab – Bassin parisien de l'UMR 7041 Archéologies et Sciences de l'Antiquité (ArScAn), constitue un outil clé du consortium PTM. Il s'agit d'une plateforme de webmapping dite transactionnelle dont les différentes applications ont été mises en place sur les serveurs de la TGIR. Elle permet de disposer des fonctionnalités d'un SIG en ligne pour les projets de recherche. Associé à cet outil de manipulation et de partage, une réflexion sur les processus de dissémination et de diffusion des couches de données géo-historiques a été lancée. L'idée étant de constituer des bases de métadonnées accessibles sur les sites institutionnels (géocatalogue, data.gouv.fr...) permettant à un public plus large d'accéder aux sources et données construites dans les programmes de recherche.



Le backoffice de GEO

Le Chantier Heurist s'articule autour du développement d'une base de données open source (MySQL) permettant de répondre à la plupart des besoins des programmes de recherche. Cette application a été élaborée depuis 13 ans dans le cadre des travaux de recherche du professeur Ian Johnson (Univ. Sydney) en collaboration avec des dizaines de projets de recherche en sciences humaines. Un chercheur peut de manière autonome dans un navigateur web et sans programmation construire et même publier sur le web ses ensembles de données avec un jeu de fonctions de traitement, de recherche et de représentation complexes. Un nœud français a été mis en place sur les serveurs de la TGIR Huma-num et des ateliers sont organisés régulièrement.



Exemple d'interface Heurist

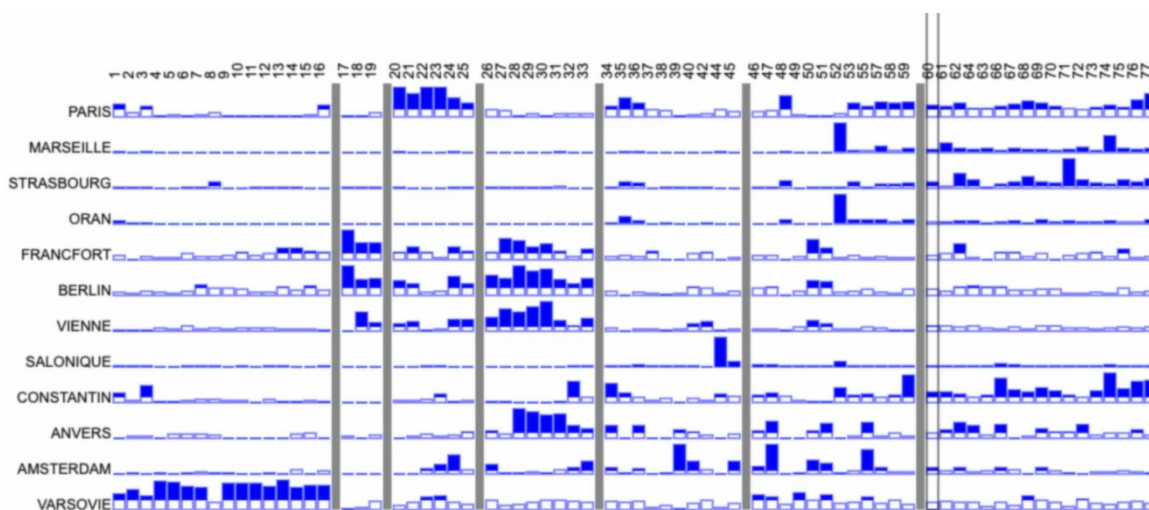


Le chantier Oronce fine est une plateforme de valorisation de collections de documents numériques dont le contenu repose sur des informations spatio-temporelles (cartes et plans anciens, cartes postales, photos de terrains, images satellites, etc.) développée par la plateforme géomatique de l'EHESS avec un financement PSL. Le projet repose sur la plateforme Omeka-S. Cette dernière permet de gérer des contenus à la manière d'un CMS. Contrairement à Omeka, la version S gère nativement le web sémantique, permettant dès lors de joindre des vocabulaires spécifiques sur les champs de métadonnées des contenus mis en ligne.



Exemple d'interface Oronce fine

Le Chantier **Amado Online** engagé en 2019 vise à produire une application permettant de visualiser des tableaux croisés (tableaux de contingence) selon les principes élaborés par Jacques Bertin dans la Sémiologie Graphique. Amado est un puissant outil d'analyse visuelle exploratoire qui permet de mettre en évidence la structure intelligible des phénomènes représentés par des données numériques. Le développement de ce logiciel et une première version de la documentation ont été livrées en ce début d'année. Le produit est maintenant opérationnel et des formations seront organisées prochainement afin de développer la communauté des utilisateurs (<https://paris-timemachine.huma-num.fr/amado-online/>).



Amado online : Exemple d'un traitement sur les convois des juifs déportés de France

Les travaux en cours...

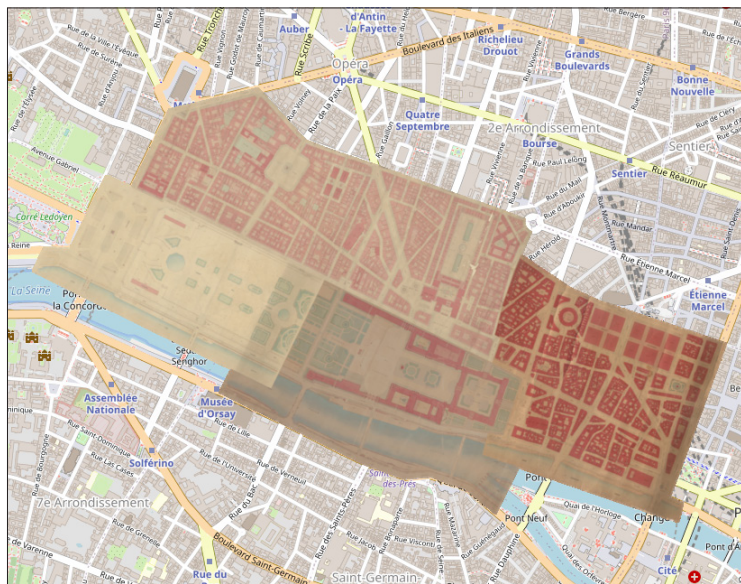
Le développement de nouveaux ensembles de données référentielles.

Le géoréférencement des fonds se poursuit activement avec trois chantiers : Le premier concerne la diffusion du plan numérisé et géoréférencé en 2019 du plan parcellaire municipal (fin XIX^e) autrement nommé « draps de lit ».

Ces 142 documents établis entre 1871 et 1896 dont les originaux sont actuellement aux archives de Paris ont été géoréférencés et mosaïqués à partir de 2019 pour être déployés sous la forme de flux WMS. Ils sont visibles en ligne sur le site de la carte archéologique de Paris mise en ligne en 2020.

Le second concerne le cadastre municipal de la ville de Paris, dit « Atlas 1900 » ou « Plan 1900 ». Ce cadastre, imprimé à l'échelle 1/500^e, dans un format plus maniable est une révision des « draps de lit » (levés entre 1871-1896). L'Atlas 1900 est composé de 732 feuilles gravées en 1901, en format 55x80 cm, en noir et blanc, de bonne qualité de conservation (peu de feuilles altérées) et d'une bonne homogénéité (pas de lacune contrairement aux « draps de lit ») conservé à la BHVP. Son géoréférencement et sa mise en ligne sont prévus pour 2021.

Le troisième concerne le géoréférencement du cadastre napoléonien des communes annexées. Ce cadastre des communes annexées en 1860 réalisé dans les années 1840 couvre la zone des arrondissements périphériques de Paris. Il est divisé par communes en 140 feuilles et permet, pour une chronologie similaire au cadastre Vasserot, d'étendre la zone cartographiée. Les feuilles ont été numérisées et sont accessibles sur le site de la BHVP. Ils sont actuellement en cours de geo-référencement et seront mis à disposition de la communauté courant 2021.



Les draps de lit assemblés : ici le 1er arrondissement



Atlas 1900 : exemple d'une feuille



Les cadastres napoléoniens des communes annexées, vue des communes de l'est Parisien

La poursuite des travaux sur les **gazettiers et annuaires** : l'Annuaire des propriétaires et des propriétés de Paris et du département de la Seine est une publication annuelle active entre 1894 et 1937 recensant les adresses parisiennes et leurs propriétaires. La collection de la Bibliothèque nationale en possède trente-huit unités et il n'existe aucune numérisation disponible. Il s'agit en particulier de rassembler, numériser, transcrire, structurer, publier, spatialiser, et analyser cette source portant une dimension spatio-temporelle importante, pour mieux représenter un moment précis de la géographie de la ville de Paris. Pour ce faire, il est nécessaire de mettre en œuvre une chaîne de traitement permettant d'obtenir, à partir d'un document numérisé, un document en format XML-TEI ainsi que dans un format tabulaire, plus apte au traitement quantitatif. Entre mai 2019 et octobre 2020 dans le contexte du groupe Annuaires et adresses du consortium PTM un contrat de vacation pour une durée d'environ 2 mois, a permis d'initier le projet. Par la suite un contrat postdoctoral financé par le MESRI et géré par l'INRIA, a démarré en avril 2020 et a pu permettre de poursuivre ces travaux. Parallèlement 2020 a vu la création d'un document au format CSV comprenant la numérisation et l'OCRisation de l'**annuaire des rues de Paris** supprimées. Cette base de donnée sera rendue accessible dans le courant de l'année 2021.

Le développement des **annotations des cartulaires** en partenariat avec l'École des Chartes : Depuis des siècles, de nombreuses sources écrites mettent en lumière l'Histoire de la ville de Paris et ses environs. Les Cartulaires d'Île-de-France constituent un exemple remarquable. L'École des chartes poursuit depuis 10 ans l'édition numérique en XML/TEI de ces textes qui répertorient en extraits, des titres relatifs aux biens et aux droits d'une personne physique ou morale du Royaume dans le temps long. Un balisage des noms de lieux et de personnes existe à l'heure actuelle pour l'ensemble des sources, et une étude exploratoire a laissé entre-



voir un nombre considérable d'occurrences exploitables spatialement. Néanmoins, un encodage homogène des noms s'impose afin de pouvoir analyser systématiquement ces textes, ce qui ouvrirait la voie à de nombreuses exploitations historiques, via l'analyse de réseaux ou spatiale. Par exemple, il serait intéressant de tracer dans le temps, les lieux d'origine des personnes répertoriées et les relier aux lieux où se trouvent leurs biens dans le Royaume. Aussi, ces documents abondent de connaissances sur des noms de lieux anciens qui pourraient servir à enrichir un gazetier géo-historique.

Le développement des outils et des communautés :

La nouvelle version de **Heurist** a été installée sur le nœud français (V6). Elle offre une nouvelle interface graphique qui améliore l'expérience de l'utilisateur novice en restant entièrement compatible avec les bases de données existantes. Par ailleurs un travail sur la documentation associée à Heurist est en cours.

L'outil **AMADO** en version outil online est aujourd'hui finalisé. Une première version de la documentation est maintenant disponible et des formations ciblées vers la communauté des historiens et des archéologues sera organisée dans le courant de l'année 2021.

Le développement d'un nouvel outil de traitement, **MorphAL** a été engagé en 2020. MorphAL est à l'origine un outil qui a été développé dans le cadre du projet ANR ALPAGE. L'objectif principal de MorphAL est de permettre de réaliser des analyses morphologiques à partir de données vectorielles, historiques ou non. L'ensemble des développements logiciel propres au consortium ont été placés sous licence libre OdBI et le code a été déposé sur le Github du consortium. Les codes des applications Amado et MorphAL sont d'ores et déjà accessibles sur le lien <https://github.com/paristimemachine/>

Paris Time Machine
http://paris-timemachine.huma-num... @TimeParis ptm-copil@listes.huma-num.fr

Repositories 3 Packages People Projects

Find a repository... Type: All Language: All

ptm4qgis-morphal
MorphAL plugin for QGIS
Python GPL-2.0 Updated on 1 Feb

amado-online
JavaScript MIT Updated on 13 Jan

morphal-openjump
MorphAL extension for OpenJUMP
GPL-2.0 Updated on 27 Nov 2020

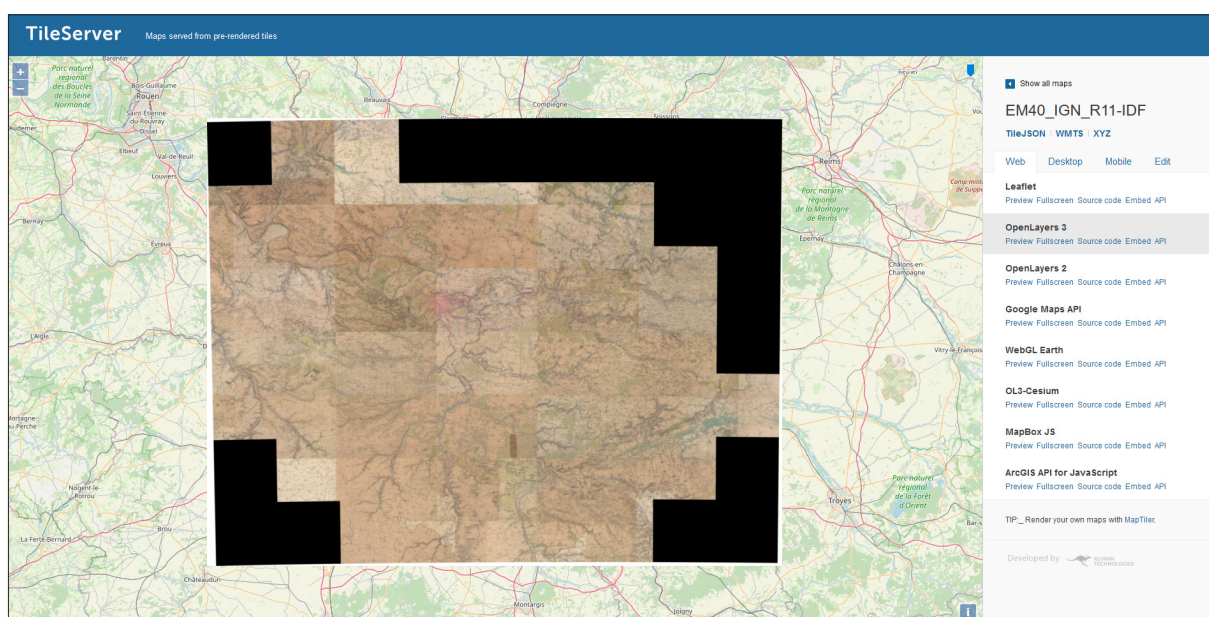
Top languages
JavaScript Python

People
This organization has no public members. You must be a member to see who's a part of this organization.

La plateforme Github du Consortium



Signalons aussi, le développement d'un serveur de tuiles, dispositif de diffusion des données qui permet de déposer les fonds géoréférencés et de les rendre disponibles sous la forme de flux XYZ compatible avec la plupart des SIG. Actuellement 140 fonds différents sont accessibles et déployés pour la plupart dans des interfaces de webmapping. Cet outil qui constitue une brique supplémentaire dans notre gamme d'outil de gestion des données géohistoriques est utilisé activement par 5 projets de recherches différents (Alpage, PTM, PCR ARBap, FEDER FEYRMAPIR). Voir : <https://tile.maps.huma-num.fr/uc2usU/>

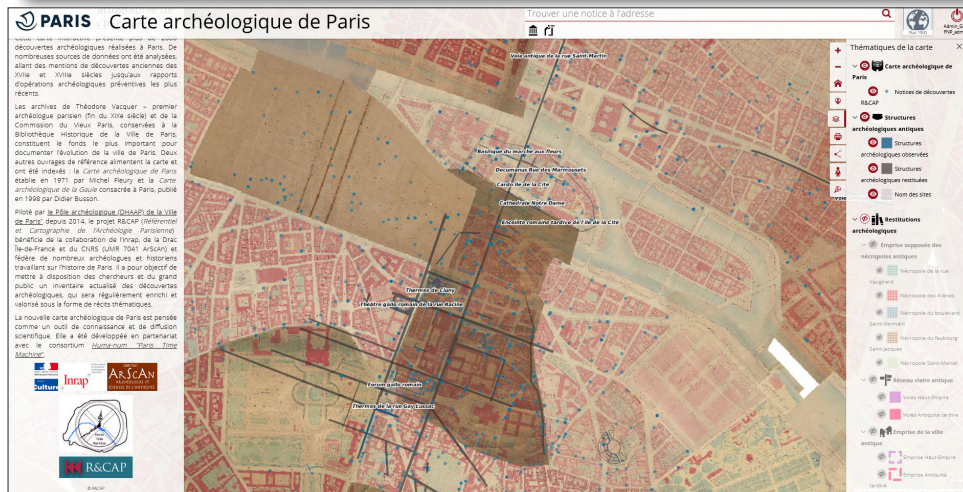
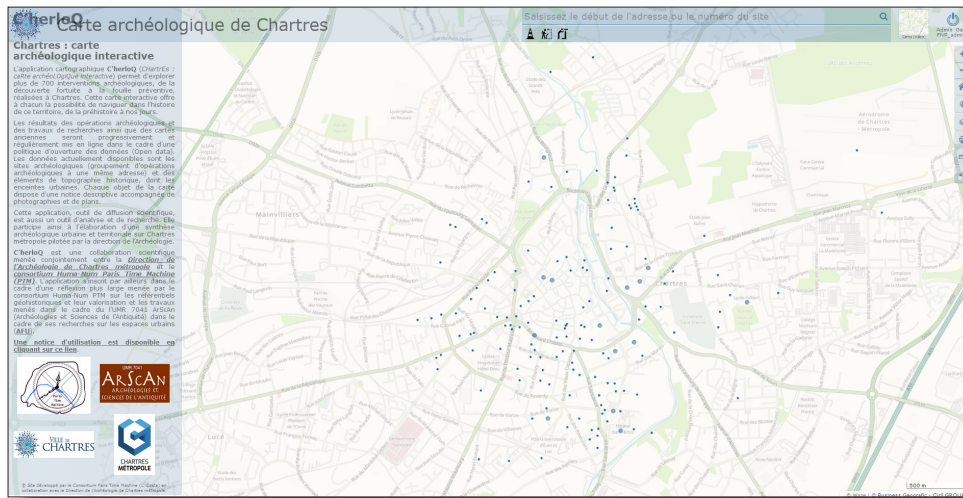


L'interface Tile maps

Le développement de partenariats avec les projets de recherche associés :

Le développement des communautés associées aux outils du consortium rendue complexe par la situation sanitaire [Oronce Fine (35 collections, 14824 documents, 10000 annotations, 19 sites au 01/10/2020), GEO (23 projets, 15 conteneurs, 198 utilisateurs référencés au 01/10/2020) et Heurist (146 bases de données ouvertes sur le nœud français au 01/02/2021) ainsi que Amado et MorphAL] est devenu une priorité du consortium pour 2021 - 2022.

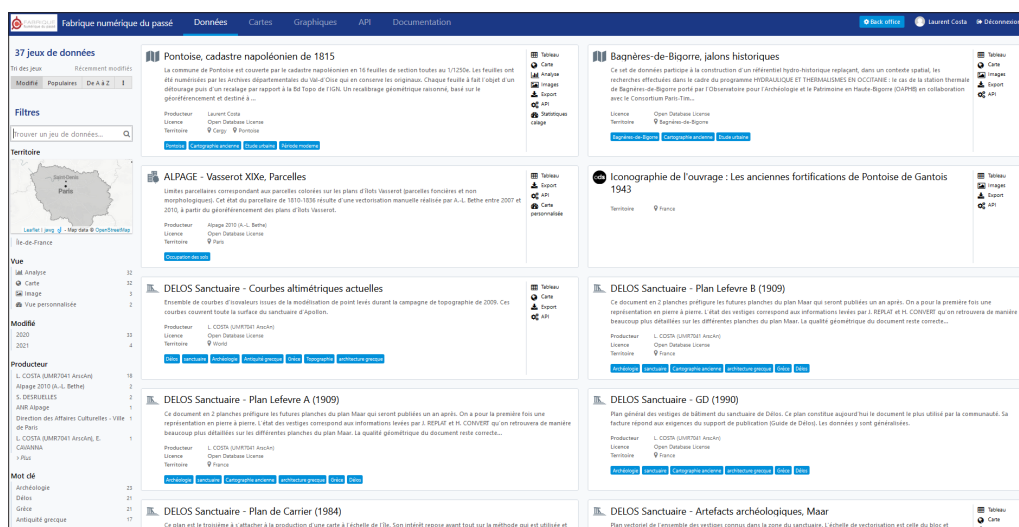
C'est pourquoi le consortium a créé la notion de projet associé. Ces « projets associés » constituent des entités autonomes administrativement, budgétairement et scientifiquement mais qui souhaitent s'associer aux travaux de notre consortium. Ces projets associés peuvent prendre des formes diverses telles que PCR, programme ANR, programme ERC... Le Consortium peut alors jouer un rôle de conseil et d'accompagnateur sur les questions méthodologiques et peut le cas échéant jouer le rôle de coordinateur de moyens. Cette approche permet de travailler dans le cadre de projets possédant des problématiques très différentes sur les be-



Vues des interfaces de deux projets associés : la carte archéologique de Chartres en haut et la carte archéologique de Paris

soins transversaux et permet de fédérer les actions et les ressources humaines tout en valorisant les bonnes pratiques et en mettant en avant les services développés par la TGIR ou ses Consortiums (Sharedocs, Humanum Box, Nakala, GEO, Heurist, etc...) le but étant de développer des outils partagés et des méthodologies communes. Un zoom sur ces différents projets associés au consortium sera proposé dans la prochaine édition de cette newsletter.

Parallèlement l'obtention du projet Ministériel « **Fabrique numérique du passé** » associé aux travaux du consortium permettra de développer une plateforme ouverte et en ligne spécifique aux besoins de la donnée géohistorique permettant le moissonnage de l'existant, le dépôt, la visualisation, la diffusion et l'analyse de ces données pour les rendre compatibles avec les plateformes open data généralistes existantes (data.gouv.fr, <https://data.europa.eu/euodp/fr/home>) et avec les outils d'archivage de la TGIR Huma-num (Nakala) selon les principes FAIR. Ce projet associé directement aux travaux du Consortium et qui profite du soutien du



Vue de l'interface de la plateforme.opendata «Fabrique numérique du passé»

Groupe de Travail Systèmes de Peuplement dans le Temps Long (SPTL), du Labex Dynamite s'articule autour de trois actions : le recrutement d'un post doc d'un an qui a débuté à partir du 1er Janvier 2021, l'organisation d'un Workshop qui se tiendra du **28 juin – 2 juillet 2021**, « **Des données aux plateformes géohistoriques en accès libre** », Abbaye de Royaumont (Asnières-sur-Oise, Val d'Oise) et le développement d'une plateforme Open data pour le dépôt et l'accès aux données géo-historiques qui ouvrira fin Mars. Cette plateforme sera adossée aux services de la TGIR (Nakala) et aux principales plateformes open data (Data.gouv.fr par exemple).

Le rapprochement avec le **Labex DYNAMITE** et le groupe de travail «Système de Peuplement sur le temps long» (A. Bretagnolle, Géographie-cités - UMR 8504 ; S. Robert, Géographie-cités - UMR 8504 ; N. Verdier, Géographie-cités - UMR 8504 et L. Costa, ArScAn UMR 7041. voir <http://labex-dynamite.com/fr/recherches/enjeux-scientifiques-groupes-travail/systemes-de-peuplement-sur-le-temps-long/>) auquel plusieurs membres du Copil du Consortium participent permet d'inscrire les travaux du consortium dans la dynamique des travaux de ce Labex et ainsi d'étendre le public potentiellement touché par les travaux de notre consortium (<http://labex-dynamite.com/fr/recherches/enjeux-scientifiques-groupes-travail/systemes-de-peuplement-sur-le-temps-long/>).

Signalons par ailleurs, que le consortium a poursuivi son investissement dans la Time Machine Organisation (TMO) à laquelle il est affilié comme membre fondateur en participant aux Web séminaires qui ont été mis en place en cette fin d'année. Dans le même esprit, le Consortium s'est impliqué dans le réseau RDA France (Réseau data Alliance) en suivant les différents événements virtuels qui ont pu être organisés (réunion du 16 septembre).



Informations pratiques et agenda

11 mars 2021 (après-midi) - Assemblée générale (en distanciel)

Printemps 2021 (dates à confirmer) : Ateliers « Amado Online »

Printemps 2021 (dates à confirmer) : Ateliers « Heurist »

28 juin – 2 juillet 2021 : Atelier « **Des données aux plateformes géohistoriques en accès libre** », Abbaye de Royaumont (Asnières-sur-Oise, Val d'Oise). Cet atelier est organisé à l'initiative du projet « *Fabrique Numérique du Passé* » (FNP), et profite du soutien du Groupe de Travail Systèmes de Peuplement dans le Temps Long (SPTL), du Labex Dynamite et du consortium de la TGIR Humanum Paris Time Machine.

Contacts :

Email : ptm-copil@listes.huma-num.fr

Adresse Web : <https://paris-timemachine.huma-num.fr/>

