

<p>Cartographier des revues en situation coloniale</p> <p>Séminaire TAIS 2020-2021</p>
--

Au fil du plan nous ferons référence aux dossiers trouvables dans notre repository Github :

<https://github.com/LilyG98/MappingJournals>

Table des matières

1. Introduction	2
2. Analyse des cartes créées.....	5
3. Démarche de constitution des cartes et présentation du repository	6
1. Récupération des entités nommées par article	6
2. Le monde au début du XXe siècle : pays souverains et situations de dépendance	7
Désambiguïsation.....	8
3. Exploration des bases de données et récupération des coordonnées	8
Extraction des coordonnées.....	9
Mise en carte	10

1. Introduction

“No other scientific specialty [than medicine] relied so heavily on geographical location”¹.

« La pathologie exotique offre un vaste champ d'études où il y a encore beaucoup à glaner. Notre domaine colonial s'est tellement étendu dans ces dernières années que presque tout est à faire, au point de vue de la géographie médicale. »²

À l'aube du XXe siècle, alors que l'empire français se stabilise, maîtriser la population devient un enjeu de taille. Se prémunir des maladies tropicales, contrôler les naissances, gérer les relations avec les populations indigènes... les autorités françaises déploient les investissements financiers, humains et institutionnels pour développer une politique de santé publique spécifique aux colonies. La médecine devenant un outil d'empire, les médecins, « sources de savoir et acteurs du pouvoir à partir du XIXe siècle »³, deviennent alors, du Sénégal aux rives du Mékong, des acteurs clés de la colonisation.

Si les disciplines médicales se sont très tôt munies de revues⁴, l'importante distance physique des médecins en situation coloniale, leur confrontation à des milieux pathogènes peu connus et les enjeux politiques dont ils avaient la charge (notamment assurer la santé des européens puis celle des ouvriers et enfin des populations locales) firent des revues médicales un objet clé de la colonisation. Développées dans l'ensemble des pays colonisateurs, les revues médicales dédiées aux situations coloniales ont permis à la fois de mettre en valeur les avancées de chaque pays dans une optique de compétition mais aussi de permettre une coopération transimpériale sur des thématiques communes, difficilement retenu par des frontières politiques.

¹ « Introduction : Tropical Medicine Before Manson.” In *Warm climates and Western Medicine : The Emergence of Tropical Medicine .1500-1900*.

² Alexandre Kermorgant, “Introduction”, *Annales d'Hygiène et de médecine coloniales*, vol.1,n°1, 1898, p.6.

³ Delphine Peiretti-Courtis, « « Corps noirs et médecins blancs. Entre race, sexe et genre : savoirs et représentations du corps des Africain(e)s dans les sciences médicales françaises (1780-1950) », Thèse, sous la direction d'Anne Carol, Aix-Marseille Université, 2014. », *Genre & Histoire* [En ligne], 15 | Automne 2014-Printemps 2015, mis en ligne le 27 septembre 2015, consulté le 17 avril 2019. URL : <http://journals.openedition.org/genrehistoire/2126>.

⁴ Valérie Tesnière et Alina Cantau, « Les revues médicales depuis 1800 », *Revue de Synthèse*, septembre 2014, vol. 135, n° 2-3, p. 203-219.

Dès lors, la circulation des savoirs entre colonies et métropole, mais aussi entre les colonies même ou entre espaces impériaux, était primordiale.

Dans le cadre français, les *Annales d'hygiène et de médecine coloniale* (AHMC) ainsi que les *Archives de médecine navale* (AMN) ont tenu ce rôle d'organe de communication officiel à triple objectif : administratif, corporatiste et scientifique. Ces deux revues étaient chacune dépendante d'un ministère et du corps de santé en dépendant. D'un part le ministère des Colonies et le corps de santé des Troupes coloniales, créés en 1890. D'autre part, celui de la Marine et le corps de santé navale. Pourtant, de 1890 à 1898, les deux corps partageaient une même publication, les *Archives de médecine navale et coloniale*. Un des enjeux de notre mémoire est alors d'en faire la généalogie d'un point de vue thématique, corporatif et géographique.

La géographie a joué un rôle primordial pour la définition des territoires relevant du pouvoir métropolitain tant d'un point de vue ethnologique que pathologique. Via la géographie médicale, le développement de médecines dites « tropicales », « coloniales », « navales », « exotiques », les a fait dialoguer autour d'un même objet d'étude relevant de l'altérité. La géographie médicale alimente alors une « géographie imaginaire »⁵. Pourtant, le développement de nouveaux paradigmes scientifiques, tels que la théorie des germes et le rôle des parasites dans la transmission de certaines maladies, détourne à plusieurs égards l'importance donnée à la topographie et au climat comme unique étiologie.

Les deux revues sont des supports connus des historiens de la médecine ou bien ceux de la colonisation pour étudier un territoire en particulier ou une maladie⁶. Preuve en est leur numérisation en 2009 par la Bibliothèque interuniversitaire de santé (BIUS) pour répondre aux besoins des chercheurs. Pourtant, jusqu'à aujourd'hui peu d'études ont été réalisées pour étudier les conditions de production et de circulation de ces supports.

Par ce rendu, nous avons souhaité considérer la répartition géographique des contributions à nos deux revues afin d'en saisir les territoires d'intérêt. En effet, la publication dans une revue officielle de mémoires originaux et de rapports effectués à l'administration nous renseigne sur ce que les autorités ministérielles estimaient comme pertinent pour la gouvernance des territoires.

⁵ Edward W. Said, *Orientalism*, 1st Vintage Books ed., New York, Vintage Books, 1979, 368 p.

⁶ **DONNER DES EXEMPLES**

Nous avons concentré notre étude sur les titres de contributions originales des années s'écoulant de 1898 à 1908 comme échantillon pour mettre en place une méthodologie de cartographie des contributions faites aux revues. Si cette période débute avec la création de l'une d'entre elles (les AHMC), elle se termine par la création de la Société de pathologie exotique dont nous étudions aussi le Bulletin dans le cadre de notre mémoire. Cette société a eu pour effet de capter les principaux médecins coloniaux et de la marine ainsi que de promouvoir la recherche bactériologique et en parasitologie, introduisant alors un nouveau rapport au terrain.

A plus long terme, l'objectif est en effet de nous munir de bases de données et de fonds de cartes exploitables pour notre rendu final de mémoire en M2. En effet, nous souhaitons pour l'an prochain cartographier l'ensemble des contributions aux trois revues que nous étudions et de détailler plus encore le cadre dans lequel ces territoires sont évoqués : compte rendu d'expédition, géographie médicale, clinique, rapports sur d'autres empires etc.). L'idée étant aussi de coupler ces recherches avec du topic modelling sur l'ensemble de notre période (1898-1940) pour saisir plus précisément le jeu des échelles qui transparaissent dans notre revue, allant du micro au macroscopique, cherchant tantôt dans l'environnement, tantôt dans les cellules, les conditions d'apparition, de développement de certaines pathologies et les moyens pour les endiguer.

Retour critique sur notre démarche.

Il nous importe de même de mettre cette carte en perspective avec les problématiques relatives à l'appropriation coloniale de l'espace au travers de la construction des savoirs géographiques et des pratiques spatiales. En effet, en aillant recours à une représentation à l'échelle mondiale et datant de 1914, non seulement nous nous conformons à une vision téléologique des appropriations coloniales mais participons de même à supposer que les puissances coloniales avaient un pouvoir continu alors même que certaines régions, comme le Sahara, n'étaient que très peu maîtrisées. Cette différenciation, primordiale entre représentation de la domination de l'espace par les savoirs coloniaux et l'espace de la domination en pratique n'a pas pu être développé dans ce premier rendu et nous considérerons pour l'année prochaine une manière d'en rendre mieux compte.

2. Analyse des cartes créées

Cf [./Carto/Cartes créées](#)

Les deux cartes de chaleur créées, l'une avec et l'autre sans le détail des lieux mentionnés par revue, nous permettent de saisir les lieux d'intérêt dans chacune des revues. Ainsi, on observe trois pôles majeurs : la France, l'Indochine et l'Afrique de l'Ouest. La France est essentiellement traitée par les AMN qui rapportent des faits survenus dans les ports français tandis que l'Indochine et l'Afrique de l'Ouest font notamment l'objet de comptes rendus de la part des AHMC.

Enfin, la carte de l'Afrique de l'Ouest a été créée pour être comparée avec une carte historique datant du début de notre période. Ainsi, nous pouvons observer l'important biais cognitif imposé par notre carte. Donnant à voir une entité géopolitique uniforme et donc, un pouvoir continu sur l'ensemble de l'espace teinté de bleu, notre carte diffère grandement de celle produite à l'époque. Cette différence nous mène à considérer, pour notre étude l'an prochain, des échelles plus restreintes afin de mieux saisir les degrés de maîtrise du territoire.

3. Démarche de constitution des cartes et présentation du repository

1. Récupération des entités nommées par article

Cf : le dossier `./data/df_annotes/` et le repository `Rendu-TAL`

Dans le cadre du master Humanités numériques, nous avons souhaité proposer un rendu tirant aussi parti de l'enseignement de Traitement automatique de la langue proposé par M. Thierry Poibeau. Ainsi, pour cartographier nos deux revues de 1898 à 1908, nous avons procédé à une reconnaissance d'entités nommées. Cette démarche nous a entre autres permis d'extraire de 1198 titres d'article deux types d'entités nommées aisément cartographiables : les *Geopolitical Entities* (GPE) et *Location* (LOC). Les GPE correspondent à l'ensemble des noms de lieux dépendant d'une entité apparentée à une administration (pays, région, province, ville, village etc.). Les LOC désignent pour leur part les entités nommées « purement » géographiques, n'étant pas par le résultat d'une intervention humaine (cours d'eau, montagne, mer, etc.).

En raison de problèmes techniques nous ayant fait perdre la quasi-totalité de nos annotations manuelles sur Doccano, nous avons eu recours au premier de notre 5-fold cross validation set pour entraîner un NLP et l'appliquer à nos sources⁷. Au total, 559 entités GPE et LOC ont été reconnues par notre NLP parmi lesquelles nous avons décompté 225 lieux différents. Néanmoins, n'ayant pu entraîner totalement notre NLP, nous avons donc eu quelques faux positifs parmi cette liste :

Si certains font référence à des notions scientifiques (Caldwelle-Luc, Danaïde, Surra, Phlégéton), d'autres sont en fait des épithètes liées à des entités GPE ou LOC. Par exemple : « Annamite » se référant à l'Annam, « Bambaras » à la région située entre l'actuel Mali, le Sénégal et le Burkina Faso. D'autres enfin sont le résultat d'une séparation des tokens d'une même EN comme « Nord » qui se trouve être en fait issu de « Cochinchine du Nord » ce qui n'est donc pas totalement incorrect mais tout de même faux au regard des attentes d'une démarche de reconnaissance d'entités nommées.

L'exploration de nos entités nommées GPE nous ont confirmé la nécessité d'avoir un fond de carte permettant de rendre compte de la situation politique de chacune d'entre elles au cours de la période que nous étudions.

⁷ Nos sources annotées se trouvent dans notre repository au chemin suivant : `./data/df_annotes/`

2. Le monde au début du XXe siècle : pays souverains et situations de dépendance

Cf : `./data/1-QGIS_GPH_HistNatBound.ipynb`

Un enjeu principal de notre démarche fut la constitution d'une carte du monde reflétant les frontières d'époque, notamment entre États souverains et territoires dépendants. Pour cela, nous avons eu recours à deux bases de données dont nous avons fait correspondre les entités pour obtenir le fond de carte de notre projet (cf Annexe 2). Nous avons ainsi eu recours aux données de deux projets accessibles librement en ligne.

Le premier est « Historic National Boundaries » de l'Université de Minnesota, disponible sur ArcGIS que nous avons téléchargé en format GeoJSON grâce au SQL d'ArcGIS online⁸. Cette base nous a permis d'avoir les shapes des entités géopolitiques. Le second est les données historiques fournies par le projet GeoPolHist du Medialab porté par Paul Girard⁹. Très complète, cette base de données permet de multiples exploitations grâce au travail de systématisation des données de 1815 à nos jours. Cette base nous a permis de saisir le statut précis des territoires étudiés à la date de 1914.

Limites : imprécisions et illusions

Si ces deux bases de données sont en anglais, elles diffèrent par leur degré de précision. Au regard de notre volonté de cartographier les données, nous avons priorisé les données de l'université de Minnesota au détriment de la précision des données du Médialab. Ainsi par exemple, il n'existe dans la base de données ArcGIS qu'une seule shape pour l'ensemble de l'Algérie, le Haut-Sénégal, le Soudan français et le Burkina Faso alors même que leurs statuts de dépendance à la France n'étaient pas identique sur toute la période que nous étudions.

Enfin, il nous semble que notre carte est téléologique à plusieurs égards. Premièrement, le fond de carte, trouvé sur ArcGIS, date de 1914 tandis que nos entités débutent en 1898. Deuxièmement, la carte donne l'illusion d'une maîtrise continue du territoire tandis qu'il est nécessaire de considérer les empires comme une "collection éparse de territoires de tailles variables"¹⁰ sur laquelle le pouvoir est exercé de manière discontinue.

⁸ University of Minnesota, Historic National Boundaries, https://services.arcgis.com/8df8p0NILFESh10r/arcgis/rest/services/Historic_National_Boundaries_NEW/FeatureServer. Carte mise jour pour la dernière fois le 13 oct. 2016

⁹ <https://github.com/medialab/GeoPolHist>

¹⁰ *Territoires impériaux: Une histoire spatiale du fait colonial*, s.l., 2020, p. 12.

3. Désambiguisation

Comme mentionné précédemment, le principal enjeu de la constitution de cette carte fut la désambiguisation des noms d'entités géopolitiques et des noms de lieux géographiques.

Pour cela, nous avons eu recours à trois bases de données que nous avons enrichies mutuellement. La raison pour ce recours multiple est que deux de ces bases de données étaient déjà affiliées à des coordonnées géographiques tandis que la troisième nous servait de complément pour les noms de lieux n'apparaissant pas dans les précédentes listes.

Orthographe des noms :

Forme retenue	Autres formes de nom
Guinée Française	Guinée française
Côte d'Ivoire	Côte d'Ivoire ; côte d'Ivoire
Pak-Hoi	Pak-Hoï , Pakhoï
Chengdu	Tchen-Tou ; Tchentou
Yunnan	Yun-Nam ; Yun-Nan ; Yunnam
Hanoï	Hanoi
Laokay	Lao-kay

4. Exploration des bases de données et récupération des coordonnées

Cf : 2-Exploration_GPH-IREL.ipynb et 3-Reconstitution_URL_IREL

Dans cette partie, nous avons eu recours à trois bases de données pour récupérer les coordonnées décimales de nos entités GPE et LOC :

- Les Q_code indiqués dans le tableur GeoPolHist_entities.csv
- La base de données en ligne des Archives nationales d'Outre-mer extraite par Webscrapping.
- Geonames, pour les noms de lieux n'ayant pas été reconnus par les deux précédentes et que nous ne traiterons donc pas en détail dans cette partie.

Nous avons décompté au total 87 lieux étaient présents dans les deux bases de données, parmi lesquelles 23 correspondaient à notre période.

BDD	GPH		IREL	
Nombre total de lieux de la base	1228		14571	
	EN en français	EN en anglais	IREL via notre reconstitution d'URL	IREL avec ajout manuel
Nombre d'EN de nos revues trouvables par la base de données	23	41	103	57

Afin de généraliser notre démarche à toute possibilités futures, nous avons récupéré l'ensemble des coordonnées géographiques décimales des bases de données IREL et GPH¹¹.

Toutefois, il sera nécessaire avant cela de revoir la reconstitution des liens URL car 5805 liens URL des 14571 ne sont pas fonctionnels (soit plus d'1/3). De nos observations empiriques, cette grande proportion de liens non exploitables est notamment dû à des problèmes de ponctuation, d'espace et d'apostrophe. Ces liens défectueux sont notamment le résultat du Webscrapping et de l'encodage url avec notre fonction `url_encode`.

5. Extraction des coordonnées

Cf : `3-Préparation-Cordonnees-AHMCAMN-QGIS.ipynb`

Pour cette partie, nous avons notamment output

En plus des 5805 liens URL non fonctionnels de l'IREL, on décompte pour la base de données GeoPolHist du Médialab 5059 lieux n'ayant pas de coordonnées indiquées sur WikiData alors que dans le tableur GeoPolHist_entities.csv, il est indiqué que seuls 72 lieux n'ont pas de géolocalisation.

¹¹ Ces dernières se trouvent dans `./data/output_finaux/`

6. Mise en carte ¹²

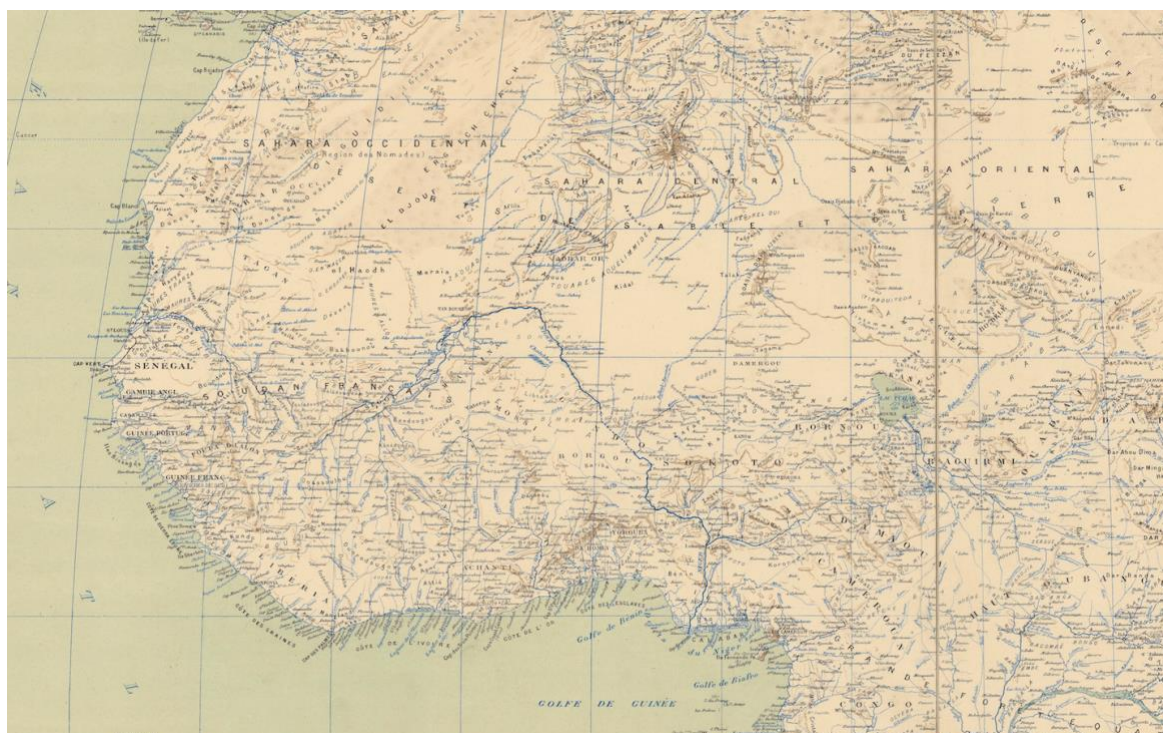
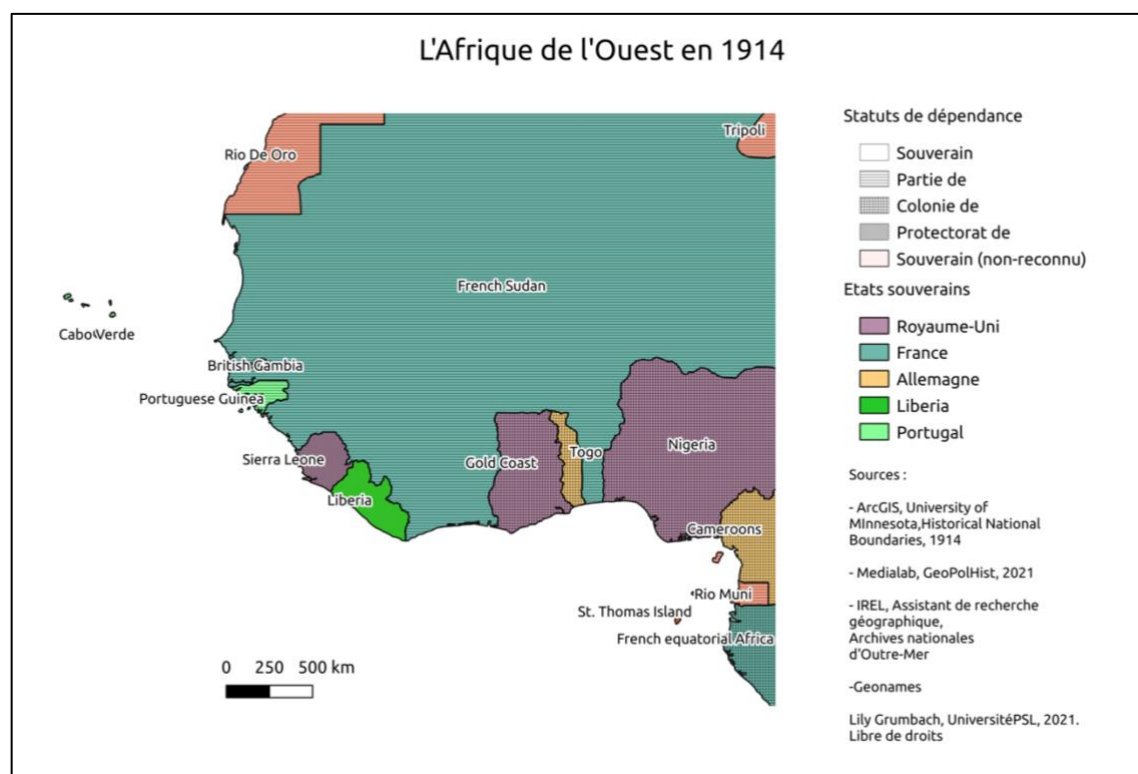
Nous avons récupéré l'ensemble de nos données dans le csv `ENSEMBLE_occurences-coord.csv` que nous avons importé dans QGIS. Si nous avons déjà le fond de carte tel que précisé précédemment, l'exploitation des données à notre disposition a été l'occasion de plusieurs essais dont nous avons retenu les principaux dans le présent rendu.

Outre la symbologie et l'analyse par carte de chaleur, nous avons notamment eu recours à une calculatrice de champs pour déterminer les co-occurrences d'un nom de lieu dans les deux revues. Le champ booléen créé (0 si pas de co-occurrences, 1 si co-occurrence) nous a permis d'avoir une meilleure vision des territoires intéressant chacune des revues comme nous l'avons analysé en début de rendu.

Ajout final des cartes créées. Our m'assurer de bien avoir toutes les dernières versions (avec échelles, etc.)

¹² Nous souhaitons ici remercier Emilien Arnaud pour son aide face aux bugs absurdes de QGIS sur notre ordinateur.

ANNEXE 1 – Illusions d’une continuité de pouvoir. La carte de l’Afrique de l’Ouest en 1914



Carte de l’Afrique, Société de Géographie de Paris, 1895.

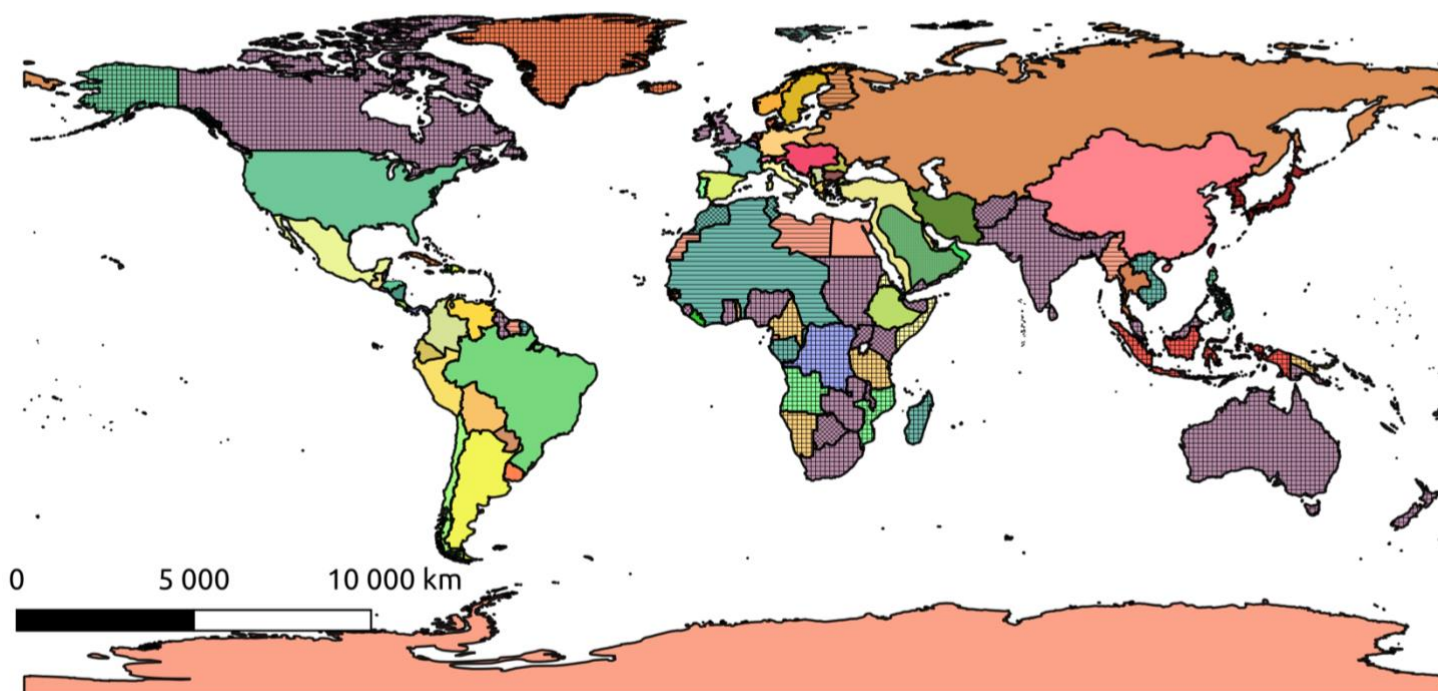
Nom de la BDD	Acteur produisant la BDD	Intérêt de la base de données	Désavantages	Utilisé pour
GeoPolHist URL interface graphique Github Taille :1228	Médialab – Sciences Po	<ul style="list-style-type: none"> Base de donnée historique Comporte les statuts de chaque entité géopolitique de 1816 à nos jours Présence des liens wikidata des entités 	<ul style="list-style-type: none"> En anglais (problème par rapport aux titres d'article en français) Nécessité de se familiariser avec l'API de Wikidata 	Fond de carte Désambiguïsation Géolocalisation (via Wikidata)
IREL URL Taille : 14570	Archives nationales d'Outre-mer	<ul style="list-style-type: none"> Grande quantité de lieux-dit (donc à une échelle plus locale) 	<ul style="list-style-type: none"> Déséquilibré en fonction des territoires (Exemple : on compte plus de 50 lieux pour l'Algérie) 	Désambiguïsation Géolocalisation
Historical National Boundaries URL	Université du Minnesota	<ul style="list-style-type: none"> Comporte des cartes à plusieurs dates clés : 1800,1914,1918,1939,1945,1990,2000 Libre accès. Possibilité de télécharger en plusieurs format 		Fond de carte
Geonames URL	Collaboratif	<ul style="list-style-type: none"> Contient à peu près toutes les géolocalisations nécessaires 	<ul style="list-style-type: none"> L'API n'ayant pas fonctionné, nous avons du rechercher terme à terme dans la version web. La quantité d'informations pour certaines 	Géolocalisation

Annexe 2 : Carte du monde en 1914

Dans le repository : `./carto/Cartes_crées/Fond_de_carte`

Carte du monde en 1914

Etats souverains et territoires dépendants



Statuts de Dépendance

-  Souverain
-  Partie de
-  Possession de
-  Colonie de
-  Protectorat de
-  Dépendant de
-  Revendiqué par
-  Découvert
-  Souverain (non-reconnu)

Principaux Etats ayant des possessions coloniales Outre-mer

-  Etats-Unis
-  Royaume-Uni
-  Pays-Bas
-  Belgique
-  France
-  Espagne
-  Portugal
-  Allemagne
-  Japon

Sources:

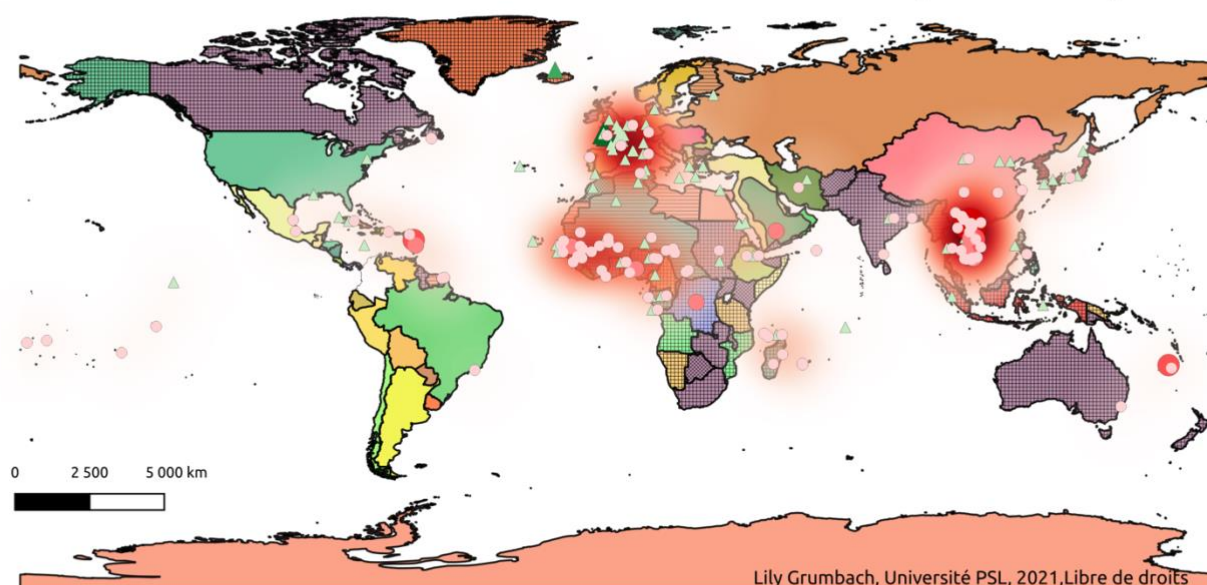
- Medialab, GeoPolHist, GitHub, 2021

- ArcGIS, University of Minnesota, Historical National Boundaries, 2019

Lily Grumbach
Université PSL, 2021
Libre de droits

Annexe 3 :

Carte de chaleur des contributions aux Annales d'Hygiène et de médecine coloniales et aux Archives de médecine navale (1898-1908)



Statuts de dépendance

- Souverain
- Fait partie de
- Possession de
- Colonie de
- Protectorat de
- Dépendant de
- Revendiqué par
- Découvert
- Souverain (non-reconnu)

Etats souverains et leurs dépendance

- Etats-Unis
- Royaume-Uni
- Pays-Bas
- Belgique
- France
- Espagne
- Portugal
- Allemagne
- Japon

AMN par occurrences

- 1 - 5
- 5 - 10
- 10 - 15

AHMC par occurrences

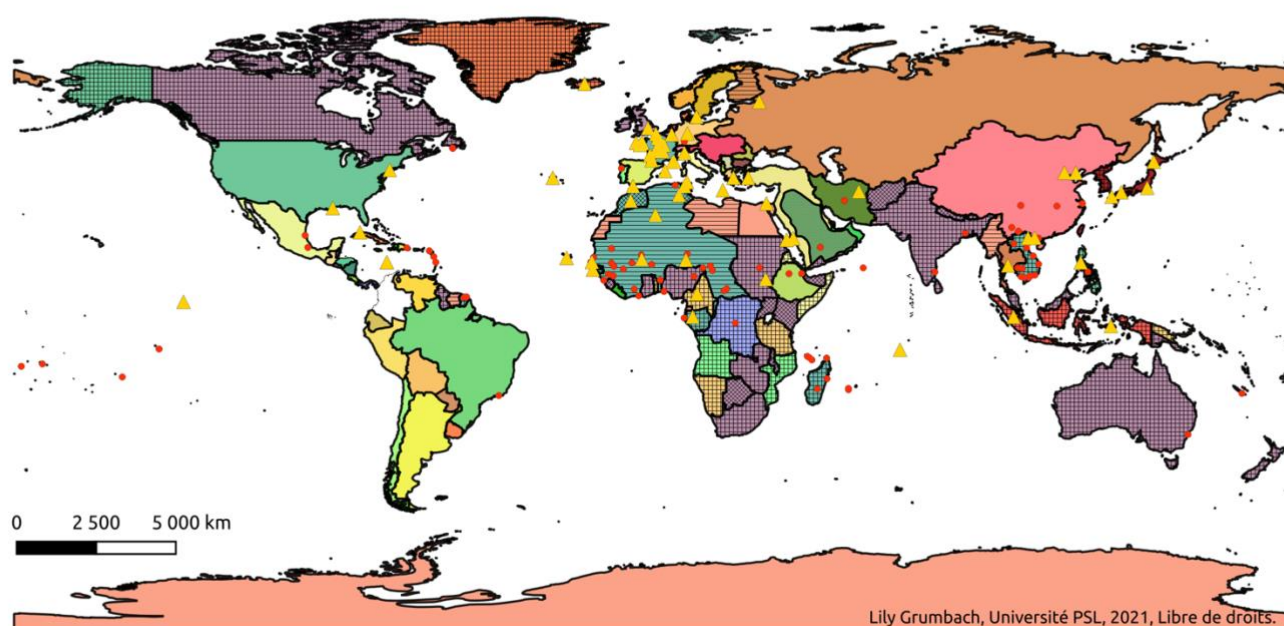
- 1 - 5
- 5 - 10
- 10 - 15
- 15 - 20
- 20 - 23

Sources :

- ArcGIS, University of Minnesota, Historical National Boundaries, 1914
- Medialab, GeoPolHist, 2021
- IREL, Assistant de recherche géographique, Archives nationales d'Outre-Mer
- Geonames

Annexe 4 :

Lieux traités uniquement dans les Archives de médecine navale ou dans les Annales d'hygiène et de médecine coloniale (1898-1908)



Statuts de Dépendance

- Souverain
- Partie de
- Possession de
- Colonie de
- Protectorat de
- Dépendant de
- Revendiqué par
- Découvert
- Souverain (non-reconnu)

Etats souverains et leurs dépendance

- Etats-Unis
- Royaume-Uni
- Pays-Bas
- Belgique
- France
- Espagne
- Portugal
- Allemagne
- Japon

Revue traitant du lieu

- AHMC
- AMN

Sources :

- ArcGIS, University of Minnesota, Historical National Boundaries, 1914
- Medialab, GeoPolHist, 2021
- IREL, Assistant de recherche géographique, Archives nationales d'Outre-Mer