



Corpus Réfugié·es

“Je suis doctorant·e en sociologie, je débute une thèse sur les réfugié·es à travers le monde au 21e siècle. Je souhaite construire une bibliographie sur le sujet, distinguer la part de documents traitant des réfugié·es climatiques et des réfugié·es politiques et enfin identifier leur origine géographique.”

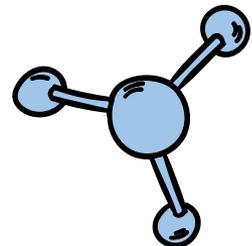
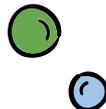
Description générale

Pour répondre à cette problématique, il faut constituer un corpus de documents traitant des réfugié·es grâce au réservoir Istex. Ce corpus va ensuite être exploré et enrichi dans Lodex.

Objectif : repérer les études sur les réfugié·es et débiter leur analyse.

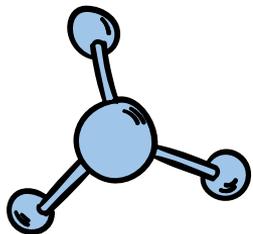
Outil de TDM utilisé : Lodex et les web services associés.

Contraintes imposées par l'outil : Istex propose un format de sortie adaptée à Lodex. Pour utiliser les web services qui m'intéressent, il faut vérifier que les documents sont en anglais et contiennent des résumés.



Exercice 1

Construire une requête
Istex



Étape 1

Se rendre sur [Istex Search](#).

Étape 1

Se rendre sur [Istex Search](#).

Étape 2

Rechercher les formes anglaises et françaises *refugee*, *réfugié*, *asylum seeker* et *demandeur d'asile*.



◦◦. Combien de ressources sont retournées par Istex ?

ISTEX Search

Créez et téléchargez votre corpus scientifique
Première étape de votre projet de fouille de textes

Résultats de votre requête

78 248 documents trouvés

refugee réfugié "asylum seeker" "demandeur d'asile"

RECHERCHER

Requête: refugee réfugié "asylum seeker" "demandeur d'asile"

BOUQUET (48)

Rechercher

- wiley 18 698
- oup 12 631
- elsevier 11 098
- cambridge 8 334
- sage 8 028

APPLIQUER

LANGUE (13)

Rechercher

Indicateurs sur votre corpus (78 248 documents)

Résumé



33679 doc.

PDF texte



75637 doc.

Texte nettoyé



36671 doc.

Langue de publication



anglais : 75148 doc. (96 %)
français : 2742 doc. (3,5 %)
allemand : 281 doc. (0,4 %)
autres : 74 doc. (0,1 %)

Compatibilité avec les passerelles (78 248 documents)

afficher : 10 trier par : pertinence & qualité

Meeting the health and social needs
of pregnant asylum seekers

Cultural competence among nurse
practitioners working with asylum

Étape 3 : Limiter le bruit et le silence

Étape 3.1 : supprimer le silence

- Rechercher les variantes (singulier / pluriel et féminin / masculin) de chacun des termes de la requête.



- ○ ○ *Cette requête a permis de rajouter des documents. Combien ?
Que risque-t'on si l'on constitue un corpus sans cette étape ?*

ISTEX Search

Créez et téléchargez votre corpus scientifique
Première étape de votre projet de fouille de textes

Résultats de votre requête

145 862 documents trouvés



refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile"
"demandeuse d'asile" "demandeuses d'asile"

RECHERCHER

Requête: refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "d..."

BOUQUET (49)

Rechercher

- wiley 32 071
- oup 20 188
- elsevier 20 136
- cambridge 19 133
- sage 15 059

APPLIQUER

LANGUE (15)

Rechercher

Indicateurs sur votre corpus (145 862 documents)

Résumé



63865 doc.

PDF texte



141062 doc.

Texte nettoyé



67263 doc.

Langue de publication



anglais : 136759 doc. (93,8 %)
français : 8136 doc. (5,6 %)
allemand : 783 doc. (0,5 %)
autres : 180 doc. (0,1 %)

Compatibilité avec les passerelles (145 862 documents)

afficher: 10 trier par: pertinence & qualité

Une clinique de l'accueil et de l'accompagnement. La prise en charge d'une famille de demandeurs

La mobilisation du clinicien dans les procédures d'asile

ISTEX Search

Créez et téléchargez votre corpus scientifique
 Première étape de votre projet de fouille de textes

Résultats de votre requête

146 137 documents trouvés



refugee* réfugié* "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile"

RECHERCHER

Requête : refugee* réfugié* "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'a...

BOUQUET (49)

Rechercher

- wiley 32 125
- oup 20 241
- elsevier 20 183
- cambridge 19 157
- sage 15 084

APPLIQUER

LANGUE (15)

Rechercher

Indicateurs sur votre corpus (146 137 documents)

Résumé



63962 doc.

PDF texte



14328 doc.

Texte nettoyé



67350 doc.

Langue de publication



anglais : 137018 doc. (93,8 %)
 français : 8169 doc. (5,6 %)
 allemand : 786 doc. (0,5 %)
 autres : 360 doc. (0,1 %)

Compatibilité avec les passerelles (146 137 documents)

afficher : 10 trier par : pertinence & qualité

La rédaction d'un certificat médical pour un demandeur d'asile : enjeux thérapeutique et social

Meeting the health and social needs of pregnant asylum seekers, midwifery students' perspectives

Étape 3.2 : Limiter le bruit

- Pour supprimer le bruit, on peut préciser les champs les plus à même de renvoyer des résultats pertinents (soit le *titre*, le *résumé* et les *mots-clés d'auteur·ices* dont la dénomination est *title*, *abstract* et *subject.value*). La requête prend alors la forme *champ:()*. La liste des champs est accessible dans la recherche assistée.
- Pour limiter le bruit, on doit s'assurer de ne sélectionner que des articles de recherche (filtre : *Type de contenu* ; nom technique : *genre*).



○ ○ ○ *Après suppression du bruit, combien de ressources écrites en français sont retournées ?*

ISTEX Search

Créez et téléchargez votre corpus scientifique
Première étape de votre projet de fouille de textes

Résultats de votre requête

3 829 documents trouvés



(title:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile") abstract:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile") subject.value:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile")) AND genre.raw:"research-article"

RECHERCHER

Requête: (title:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'a...

BOUQUET (16)

Rechercher

- oup 1 035
- elsevier 812
- sage 597
- cambridge 445
- springer-journals 374

APPLIQUER

LANGUE (4)

Indicateurs sur votre corpus (3 829 documents)

Résumé



3347 doc.

PDF texte



3770 doc.

Texte nettoyé



3468 doc.

Langue de publication



anglais : 3780 doc. (98,7 %)
français : 40 doc. (1 %)
allemand : 7 doc. (0,2 %)
autres : 1 doc. (0 %)

Compatibilité avec les passerelles (3 829 documents)

afficher : 10 trier par : pertinence & qualité

TÉLÉCHARGER LE CORPUS (3 829)

Démocratisation syndromes des nat

Étape 4

Répondre aux contraintes **scientifiques** et **techniques**



En vous aidant du filtre *date de publication* et du sujet de thèse du cas d'usage, donnez une **contrainte scientifique** non respectée.

Rappel de la problématique :

“Je suis doctorante en sociologie, je débute une thèse sur les réfugié-es à travers le monde au 21e siècle. Je souhaite construire une bibliographie sur le sujet, distinguer la part de documents traitant des réfugié-es climatiques et des réfugié-es politiques et enfin identifier leur origine géographique.”

ISTEX Search

Créez et téléchargez votre corpus scientifique
Première étape de votre projet de fouille de textes

Résultats de votre requête

2 435 documents trouvés



```
(title:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile")) abstract:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile") subject.value:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile")) AND genre.raw:"research-article" AND publicationDate:[2000 TO *]
```

RECHERCHER

Requête: (title:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'a...)

BOUQUET (15)

Rechercher

- oup 590
- elsevier 570
- sage 363
- springer-journals 332
- oa-plos 250

APPLIQUER

LANGUE (4)

Rechercher

Indicateurs sur votre corpus (2 435 documents)



Compatibilité avec les passerelles (2 435 documents)

afficher : 10 trier par pertinence & qualité

TÉLÉCHARGER LE CORPUS (2 435)

Demoralisation syndrome does not

Étape 4

Répondre aux contraintes **scientifiques** et **techniques**

Pour répondre aux **contraintes imposées par l'outil**, il faut s'assurer de la présence de résumés dans les documents et s'assurer qu'ils sont en anglais.



... Enfin, quelle est la taille du corpus sélectionné ?

ISTEX Search

Créez et téléchargez votre corpus scientifique
 Première étape de votre projet de fouille de textes

Résultats de votre requête

2194 documents trouvés



(title:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile") abstract:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile") subject.value:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'asile" "demandeuse d'asile" "demandeuses d'asile")) AND genre.raw:"research-article" AND publicationDate:[2000 TO *] AND abstract:* AND language:eng

RECHERCHER

Requête: (title:(refugee refugees réfugié réfugiée réfugiés réfugiées "asylum seeker" "asylum seekers" "demandeur d'asile" "demandeurs d'a...

BOUQUET (13)

Rechercher

- elsevier 542
- oup 493
- sage 349
- springer-journals 326
- oa-plos 249

APPLIQUER

LANGUE (1)

Rechercher

Indicateurs sur votre corpus (2 194 documents)



Compatibilité avec les passerelles (2 194 documents)

afficher : 10 trier par : pertinence & qualité

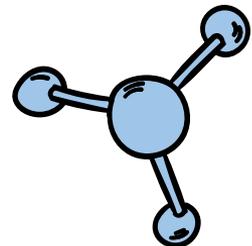
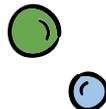
Comparison of two strategies to

TÉLÉCHARGER LE CORPUS (2 194)

International Journal of Infectious Diseases

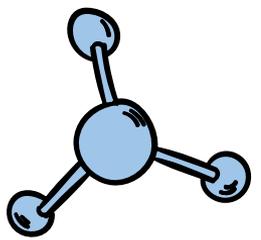
Demoralisation syndrome does not explain the psychological profile of community-based asylum-seekers

Comprehensive Psychiatry



Exercice 2

Premiers pas vers le TDM



Étape 1

Télécharger le corpus : extraire le corpus Réfugié-es en utilisant l'équation corrigée et en choisissant le format adapté pour un import dans Lodex.

Configurez votre téléchargement



USAGE PERSONNALISÉ

LODEX

CORTEXT

GARGANTEXT

NOOJ

- | | | |
|---|---|--|
| <input type="checkbox"/> Texte intégral | <input checked="" type="checkbox"/> Métadonnées | <input type="checkbox"/> Enrichissements |
| <input type="checkbox"/> PDF | <input checked="" type="checkbox"/> JSON | <input type="checkbox"/> multicat |
| <input type="checkbox"/> TEI | <input type="checkbox"/> XML | <input type="checkbox"/> nb |
| <input type="checkbox"/> TXT | <input type="checkbox"/> MODS | <input type="checkbox"/> grobidFulltext |
| <input type="checkbox"/> CLEANED | | <input type="checkbox"/> refBibs |
| <input type="checkbox"/> ZIP | <input type="checkbox"/> Annexes | <input type="checkbox"/> teeft |
| <input type="checkbox"/> TIFF | <input type="checkbox"/> Couvertures | <input type="checkbox"/> unitex |

Trier par : **pertinence & qualité** ▾

Télécharger

2 194

/ 2 194

TOUT

Format de l'archive : **ZIP** ▾

Compression : **moyenne** ▾

TÉLÉCHARGER

Lodex

Application web open-source dédiée aux données structurées qui permet de visualiser et d'enrichir ses données puis de les transformer en site web.

[EN SAVOIR PLUS](#)

Requête

```
(title:(refugee refugees réfugié réfugiée
réfugiés réfugiées "asylum seeker" "asylum
seekers" "demandeur d'asile" "demandeurs
d'asile" "demandeuse d'asile" "demandeuses
d'asile") abstract:(refugee refugees réfugié
réfugiée réfugiés réfugiées "asylum seeker"...
```

Requête brute complète

[https://api.istex.fr/document?q=\(title:\(refugee refugees réfugié réfugiée réfugiés réfugiées \"asylum seeker\" \"asylum seekers\" \"demandeur d'asile\" \"demandeurs d'asile\"...](https://api.istex.fr/document?q=(title:(refugee refugees réfugié réfugiée réfugiés réfugiées \)

Étape 2

Importer le corpus dans Lodex.

Cette étape est la plus sensible aux erreurs de manipulations des données. Nous allons donc la faire tous ensemble, nous récapitulons la procédure dans la suite.

Étape 2

Importer le corpus dans Lodex.

- Se rendre sur votre instance Lodex : se connecter avec votre nom d'utilisateur et votre mot de passe.
- Aller dans l'interface administrateur en cliquant sur *Voir plus > Admin*.
- Importer le corpus en glissant le fichier .zip téléchargé **sans décompression préalable**.
- Choisir le loader *ZIP résultat de dl.istex.fr*.
- Cliquer sur Importer les données.
- Importer le modèle fourni, cliquer sur le menu en haut à droite *Modèle > Importer un modèle*.
- Publier votre site en cliquant sur *Publier* en haut à droite.
- Cliquer sur l'icône en forme d'œil pour voir le résultat.
- Explorer les différents graphiques à partir de l'onglet *Graphiques* en bas à gauche. Consulter quelques ressources grâce à l'onglet *Recherche*.



Étape 2

Importer le corpus dans Lodex.



◦ ◦ *Combien y-a-t'il de documents publiés en 2009 dans la revue Journal of Refugee Studies ?*

Vous pouvez saisir votre recherche ici

Ressources visitées
 Toutes les ressources

Annotations
 Toutes les ressources

Date de publication 2009 Titre de revue Journal of Refugee Studies Tout supprimer

Auteur-ice(s) v

Titre de revue (618) ^

exclure

Titre de revue filtre

VALEUR	COMPTER ↓
<input type="checkbox"/> PLOS ONE	188
<input checked="" type="checkbox"/> Journal of Refugee Studies	180
<input type="checkbox"/> International Journal of Refugee Law	134
<input type="checkbox"/> Refugee Survey Quarterly	84
<input type="checkbox"/> Journal of Immigrant and Minority Health	69

18 ressources trouvées sur un total de 2194

TRIER | ASCENDANT EXPORTER

Representing Sahrawi Refugees' 'Educational Displacement' to Cuba: Self-sufficient Agents or Manipulated Victims in Conflict?

Mainstream accounts of refugee women and children have habitually portrayed their objects of study as 'generic' passive victims of war and/or famine. In stark contrast, however, since the early 1980s Sahrawi refugee women and children have been invoked as active agents constructing and maintaining their camps. In the first part of this article I explore the nature of a selection of mutually-reinforcing images produced by the Polisario Front (the Sahrawi refugees' 'representatives' and camp managers) for a European audience. Their distinctive content could appear to be diametrically opposed, and perhaps designed to offer a...

Journal of Refugee Studies

2009

Contemporary Compulsory Dispersal and the Absence of Space for the Restoration of Trust

This paper investigates the issue of trust, or mistrust, specifically in relation to single adult asylum seekers and asylum seeker families compulsorily dispersed across England. It draws upon doctoral research on the social exclusion of asylum seekers as a result of dispersal and their separation from mainstream welfare provision due to the creation of the National Asylum Support Service (NASS) following the Immigration and Asylum Act 1999. Trust is an ambiguous term and four forms of trust are delineated to assist conceptualizing the experience of forced migration: social, political, institutional and restorative trust. This paper provide...

Journal of Refugee Studies

2009

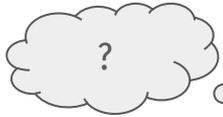
The Faint Footprint of Man: Representing Race, Place and Conservation on the Mozambique-South Africa Borderland

The integration of South Africa's existing Kruger National Park with Mozambique's newly-established Limpopo National Park signified an important milestone in the development of the Great Limpopo Transfrontier Park. Marketed as an example of the African Renaissance-in-action, this move towards a transnational 'mega-park' has promised an enlightened approach to conservation and regional economic development. For...

Étape 3.1

Extraction de mots-clés des résumés via le web service [Teeft](#).

- Aller dans *Données > Enrichissements* et cliquer sur *+ Ajouter*.
- Donner le nom *Mots-clés (WS)*, aller chercher l'url du web service **Teeft** dans le catalogue en cliquant sur le bouton vert à droite du champ *URL du web service*.



Sur quels types de champ peut-on lancer **Teeft** ? A quoi faut-il faire attention ?

Étape 3.1

Extraction de mots-clés des résumés via le web service [Teeft](#).

- Aller dans *Données > Enrichissements* et cliquer sur *+ Ajouter*.
- Donner le nom *Mots-clés (WS)*, aller chercher l'url du web service **Teeft** dans le catalogue en cliquant sur le bouton vert à droite du champ *URL du web service*.



Sur quels types de champ peut-on lancer **Teeft** ? A quoi faut-il faire attention ?

- Choisir *Résumé* dans la *Colonne de la source*, cliquer sur *Sauvegarder*. Cliquer enfin sur *Lancer*.

Étape 3.2

Création du graphique associé à Teeft.

- **Dans Lodex, avant de faire un graphique, il est nécessaire de déclarer la colonne comme une ressource.** Aller dans *Affichage > Ressource principale*, cliquer sur *+ Nouveau champ*. Pour paramétrer ce nouveau champ : dans *Étiquette* nommer le champ *Mots-clés*, sélectionner *Colonne(s) existante(s)* et aller chercher la colonne nommée *Mots-clés (WS)*. Ajouter une opération pour appliquer l'opération de transformation GET, renseigner *term* dans le *path*. Cliquer sur *Sauvegarder*.



◦◦ A quoi sert l'opération de transformation ?

Étape 3.2

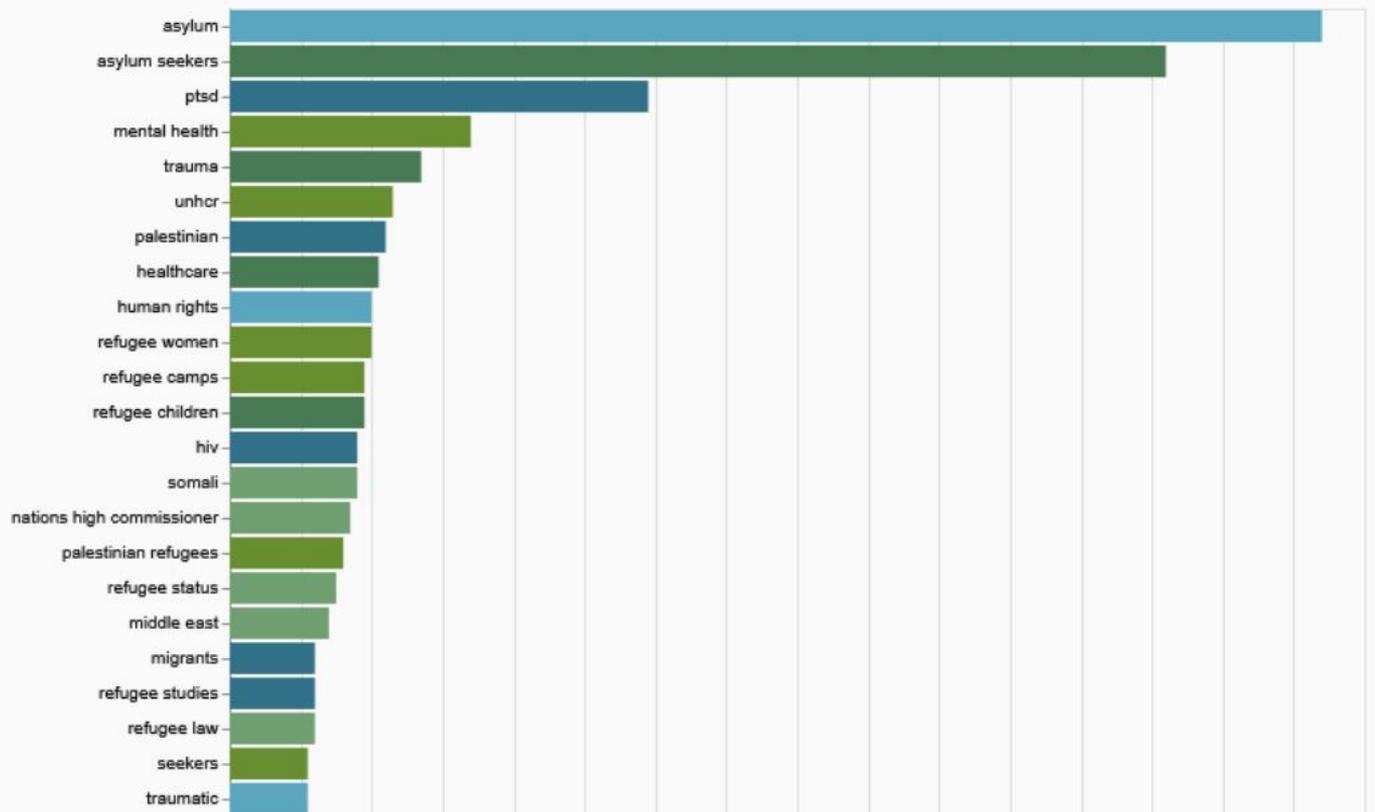
Création du graphique associé à Teeft.

- Pour créer le graphique : aller dans *Affichage > graphiques* cliquer sur *+ Nouveau champ*, nommer le graphique en renseignant *Mots-clés les plus représentés dans les résumés* dans le champ *Étiquette*. Choisir la routine *distinct-by* puis choisir le champ sur lequel la routine va s'appliquer grâce au menu déroulant (soit le champ *Mots-clés*). Enfin, dans *Affichage*, choisir le format *Diagramme en barres* en filtrant les résultats : dans *Paramètres des Données* mettre valeur minimum à afficher à 5 et choisir le tri *Descendant*. Cliquer sur *Sauvegarder*.



Quels sont les mots-clés les plus représentés dans le corpus ?

Mots-clés les plus représentés dans les résumés ⚙️



Étape 4 : Création d'une carte en fonction des pays mentionnés dans les articles

Étape 4.1 : Extraction des entités nommées Unitex

- Certains enrichissements sont déjà effectués sur l'entièreté de la base Istex. C'est le cas des noms de lieux renseignés dans la colonne *Entités nommées (Unitex)* avec les autres entités nommées.

Entités nommées (...)	Types d'enrichisse...	Score qualité
{ "date": ["year 1", "1951"] }	["grobidFulltext", "multi"]	7.389
undefined	["grobidFulltext", "multi"]	6.819
undefined	["nb", "teft"]	9.827
{ "date": ["In the year 20"] }	["grobidFulltext", "multi"]	8.007
undefined	["nb", "teft"]	10
undefined	["nb", "teft"]	10
{ "date": ["2010-06-01"] }	["grobidFulltext", "multi"]	8.667
{ "date": ["1980s"] }	["grobidFulltext", "multi"]	8
undefined	["nb", "teft"]	10
{ "date": ["the end of ye"] }	["grobidFulltext", "multi"]	7.599

```

4 : string "Mandy Savage"
]
  "placeName" : [ 7 items
    0 : string "Utrecht"
    1 : string "Europe"
    2 : string "America"
    3 : string "Amsterdam"
    4 : string "Arnhem"
    5 : string "England"
    6 : string "Netherlands"
  ]
  "ref_url" : [ 2 items
    0 : string "http://www.quantitativesk
    1 : string "http://www.r-project.org"
  ]
  "geogName" : [ 2 items
    0 : string "Netherlands Region"
    1 : string "Eastern Region"
  ]
}

```


Étape 4.2 : Création de la carte

- Pour construire la carte, nous avons besoin du code pays à 3 lettres, retourné par le web service *Associer un terme au vocabulaire Pays et Subdivisions*. Aller dans *Données > Enrichissements > + Ajouter*, puis donner le nom *Pays et subdivisions (WS)*, en utilisant le web service *Associer un terme au vocabulaire Pays et Subdivisions* sur la colonne nouvellement créée *Entités nommées de géographie*. Cliquer sur *Sauvegarder*. Cliquer enfin sur *Lancer*.

Nom

Pays et subdivisions (WS) ▶ LANCER

0 erreurs RÉESSAYER

Statut : En cours [VOIR LES LOGS](#)

Mode avancé

URL du web service

https://loterre-resolvers.services.istex.fr/v1/9SD/identify 📄

Colonne de la source

Entités nommées de géographie ▼ Sous-chemin ▼

SUPPRIMER ANNULER SAUVEGARDER

Étape 4.2 : Création de la carte

Accueil > Web-services > Associer un terme au vocabulaire Pays et Subdivisions

Associer un terme au vocabulaire Pays et Subdivisions

Description

Utilisation

Niveau d'utilisation : Débutant

Niveau de validation : Expérimental

Objectif

Le web service repère des termes dans le vocabulaire [Pays et Subdivisions de Loterre](#) pour récupérer le concept associé. Par exemple, il récupère la forme normalisée du nom d'un pays à partir de ses différentes variations orthographiques.

Ressource : La plateforme Loterre propose des vocabulaires ou terminologies pour différents domaines scientifiques. Autour d'un concept, Loterre propose plusieurs verbalisations et différents liens vers des bases de références. Le vocabulaire propose pour chaque pays et région française des concepts regroupant informations géographiques, variantes syntaxiques, acronymes, et formes normalisées.

Étape 4.2 : Création de la carte

- Dans *Affichage > Ressource principale*, créer une ressource *Code pays* à partir de l'enrichissement *Pays et subdivisions (WS)*. Après avoir sélectionné la colonne, cliquer sur *Ajouter une opération* pour appliquer une transformation GET, renseigner *cartographyCode* dans le *path*. Cliquer sur *Sauvegarder*.

The screenshot displays a configuration interface for a resource named "Code Pays". A modal dialog titled "Ajouter une opération" is centered on the screen. Inside the dialog, a dropdown menu is set to "GET", and the "path" field contains the text "cartographyCode". Below the path field are two buttons: "ANNULER" and "CONFIRMER".

In the background, the main interface shows a table with the following columns:

- Etiquette: Code Pays
- Colonne(s) existante(s): Pays et subdivisions (WS)
- Source de la valeur: Valeur Arbitraire
- Opérations de transformation: TOUT SUPPRIMER

On the right side, a preview of the resource is shown under the heading "Code Pays". It lists several JSON snippets, including:

```
"n/a"
```

```
{"cartographyCode": "LUX", "about": "http://data.lotterre.fr/ark:/67375/9SD-..."}
```

```
{"about": "http://data.lotterre.fr/ark:/67375/9SD-..."}
```

```
{"cartographyCode": "NZL", "about": "http://data.lotterre.fr/ark:/67375/9SD-..."}
```

```
"n/a"
```

A note at the bottom right reads: "* aperçu de la valeur avec les opérations de transformation appliquées".

Cartographie des pays cités ⚙️



Étape 5

Extraction des thématiques du corpus.

- Dans *Données* > *Précalculs* cliquer sur *+ Ajouter*, lancer le précalcul *lda-segment* sur la colonne *Résumé*.



Quelle est la différence entre un web service et un précalcul ?

WS = prend en entrée “une seule ligne”

précalcul = prend en entrée tout un corpus

Étape 5

Extraction des thématiques du corpus.

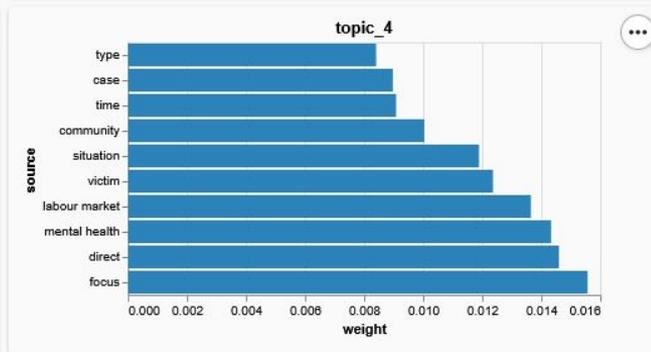
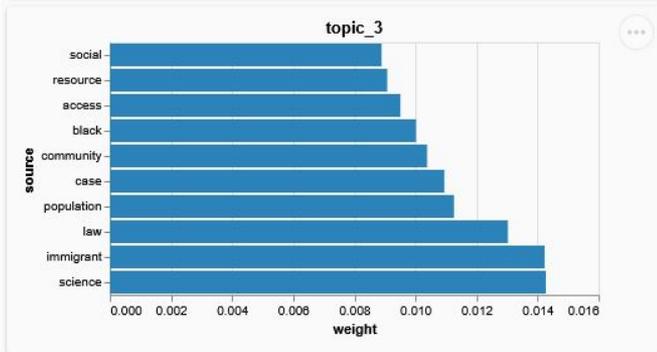
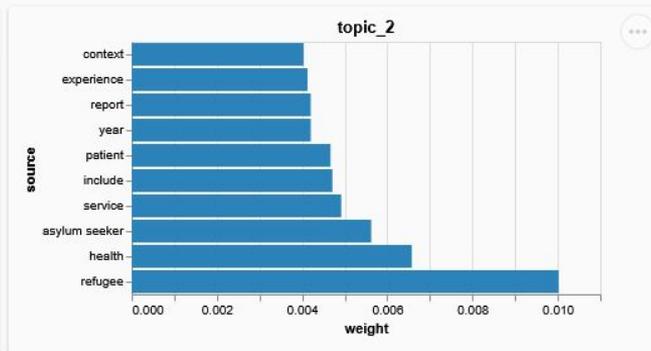
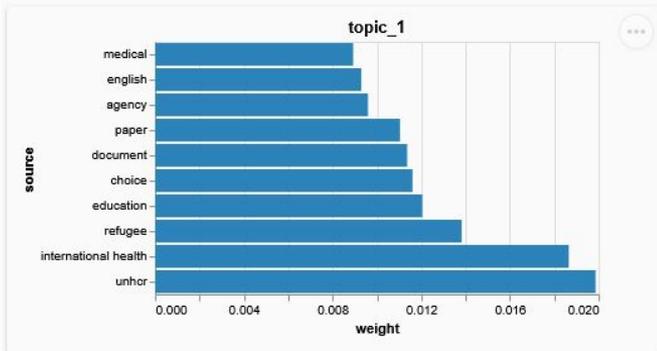
- **Les précalculs peuvent directement être utilisés pour faire des graphiques.** Aller dans *Affichage > graphiques* cliquer sur *+ Nouveau champ*, puis créer le graphique *Thématiques extraites du corpus* à partir du précalcul *lda-segment*. Choisir la routine *segments-precomputed-nofilter*. Enfin, dans *Affichage*, choisir le format *diagramme à barres groupées*.



◦◦. Combien de thématiques sont retournées ?

Étape 5

LDA ⚙



Étape 6 : Pour aller plus loin...

L'objectif ici est d'augmenter l'exhaustivité de la carte. Nous pouvons utiliser un web service de détection d'entités nommées sur le corpus pour extraire des localisations.



En vous aidant d'[Istex TDM](#), quel web service vous semble le plus adapté ?

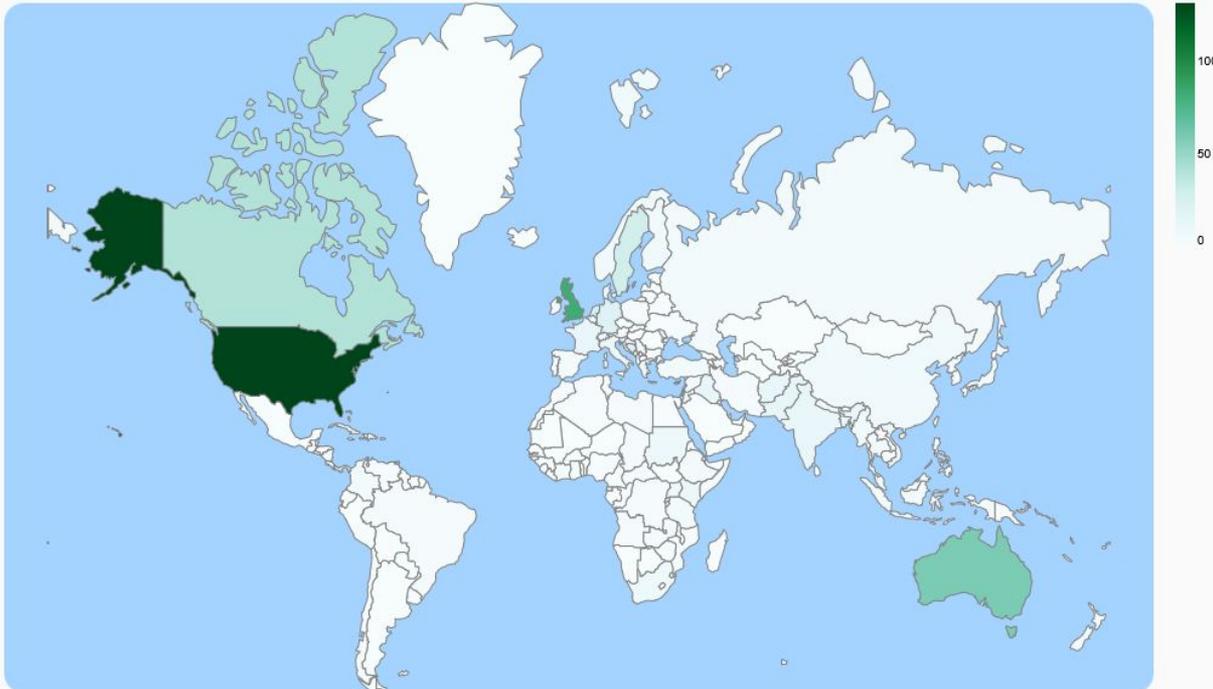
- Utiliser ce web service.
- Extraire les noms de lieux grâce à un nouvel enrichissement avancé en s'assurant de remplacer correctement "XXXXX" :

```
[assign]
path = value
value = get("value.XXXXX.LOC").slice(0,1)
```

- Utiliser de nouveau le web service *Associer un terme au vocabulaire Pays et Subdivisions* sur les noms de lieux extraits à l'étape précédente (LOC). Créer une carte.

Étape 6

Seconde cartographie ⚙



Correction

Le TP corrigé est disponible en ligne :

<https://strasbourg-universite-1591.formation.lodex.fr/instance/correction-tp-refugies/login>

Nom d'utilisateur : *correction*

Mot de passe : *TP2025*