

Approches Sociales et Historiques des Littératures des Suds

Séance du 05 février 2021

Faire émerger le réseau littéraire francophone haïtien à l'ère du numérique

ENS
ENS DE LYON

— université
— Lumière
— LYON 2

Anaïs CHAMBAT - ENS de Lyon
Célian RINGWALD - Université Lumière Lyon 2

Approches Sociales et Historiques des Littératures des Suds

Faire émerger le réseau
littéraire francophone
haïtien à l'ère du
numérique

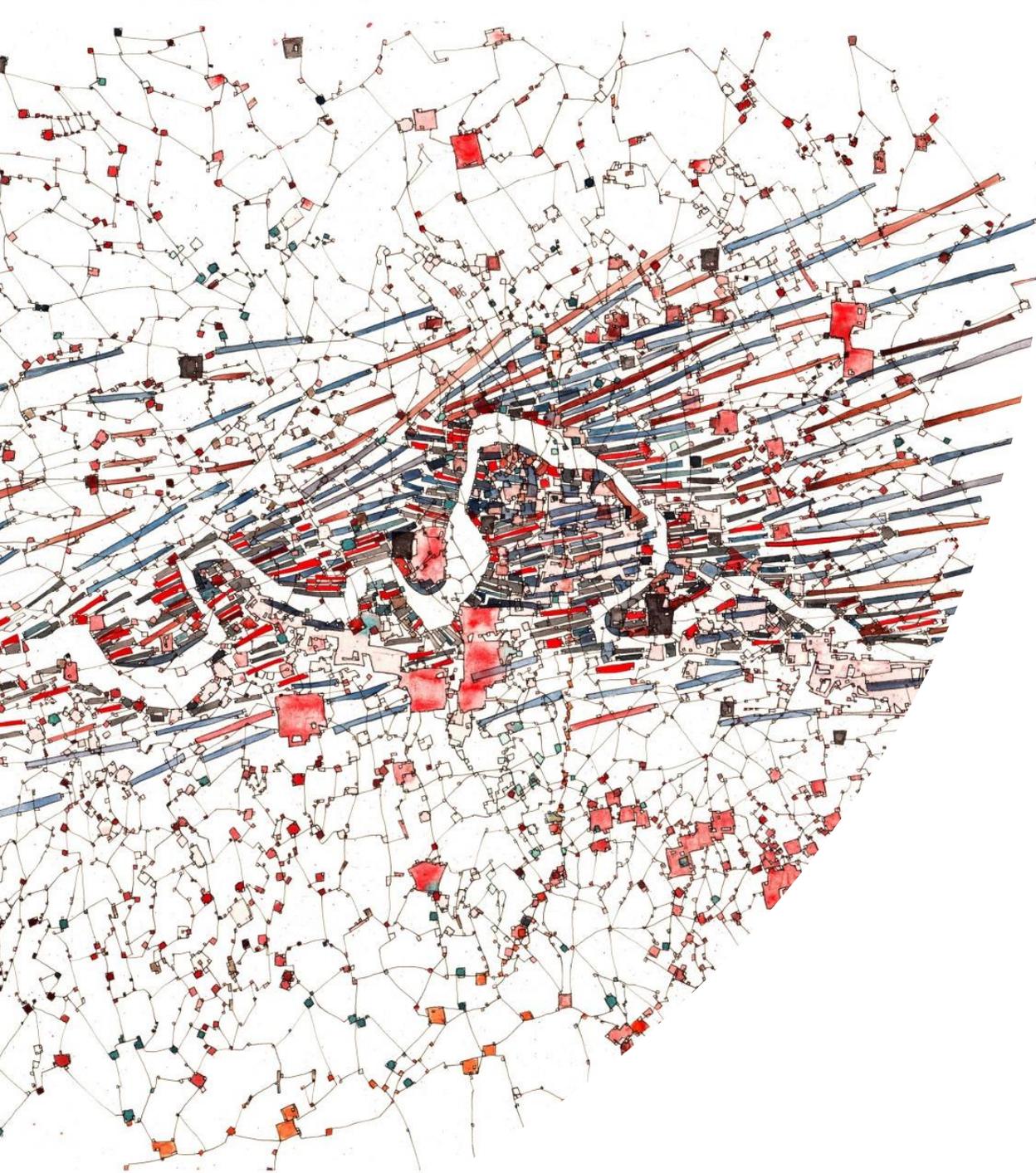
I. Introduction

II. Croiser Web sémantique et contenus participatifs

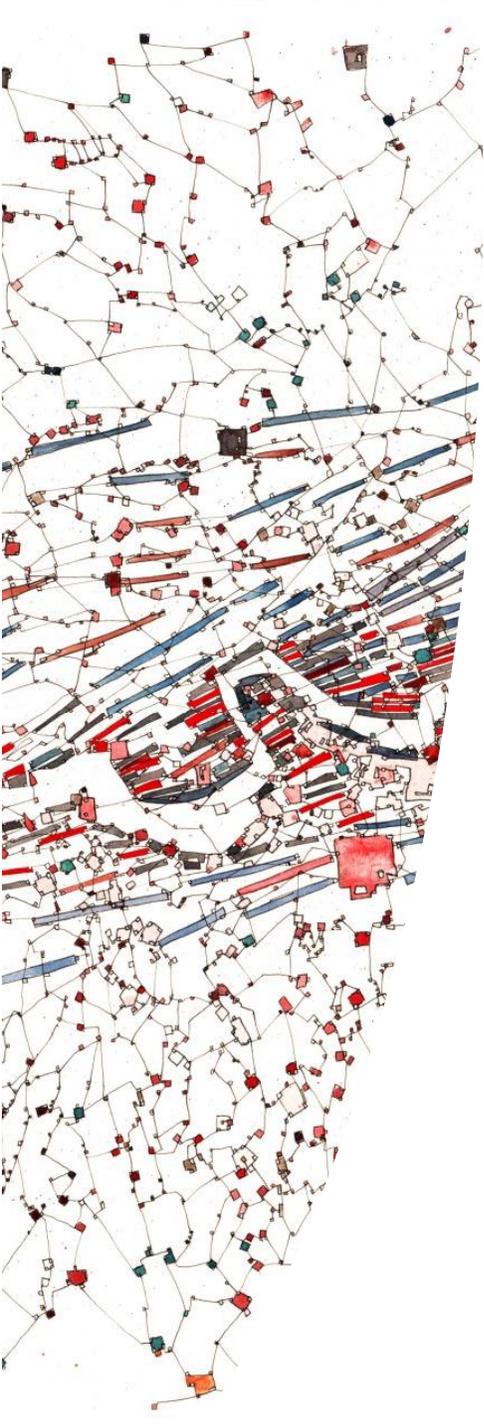
III. Normalisation du corpus et structuration en XML-TEI

IV. Exploitation des données

V. Conclusions



I. Introduction



Des profils complémentaires fédérés autour d'un double cursus en Humanités Numériques

- **Inès Burri** : Infographiste 3D de formation, elle est titulaire d'une licence de lettres et d'un master en histoire ancienne effectués au sein de l'Université Jean Moulin Lyon 3. Elle travaille actuellement au sein du GIRPAM (Groupement International de Recherche sur la Poésie Antique et Médiévale).
- **Anais Chambat** : Diplômée d'un master en histoire des mondes germaniques, elle a étudié à l'ENS de Lyon les applications du numérique et des méthodes quantitatives à l'histoire. Elle est actuellement gestionnaire de données de recherche au sein du service d'histoire de la médecine de la Bibliothèque InterUniversitaire de Santé.
- **Célian Ringwald** : *Datascientist* et spécialiste en traitement automatique du langage. Il est aujourd'hui éditeur bénévole pour la revue *Programming Historian* portant sur les humanités computationnelles et la méthodologie de l'histoire numérique.



Qu'est-ce que le Web littéraire ?

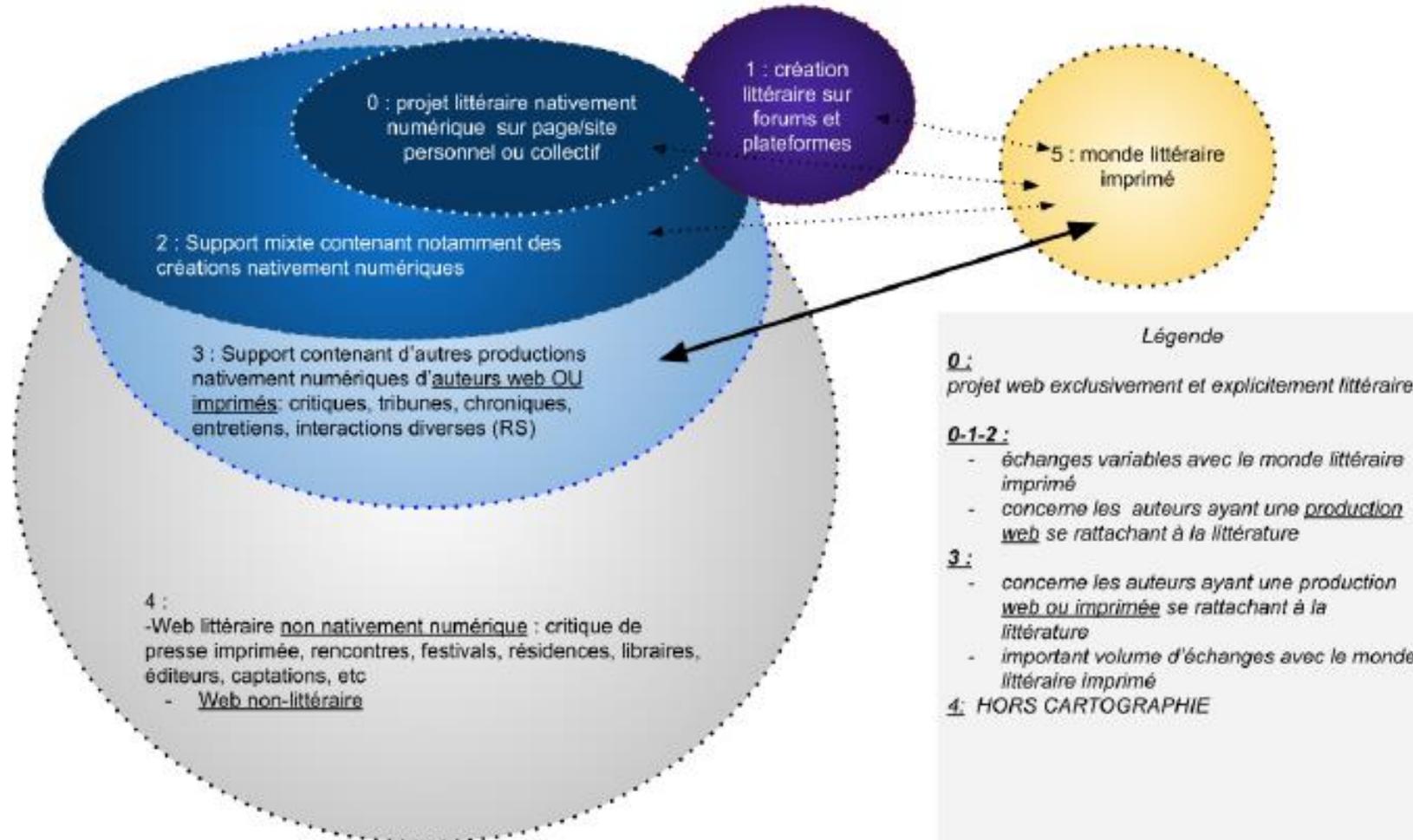
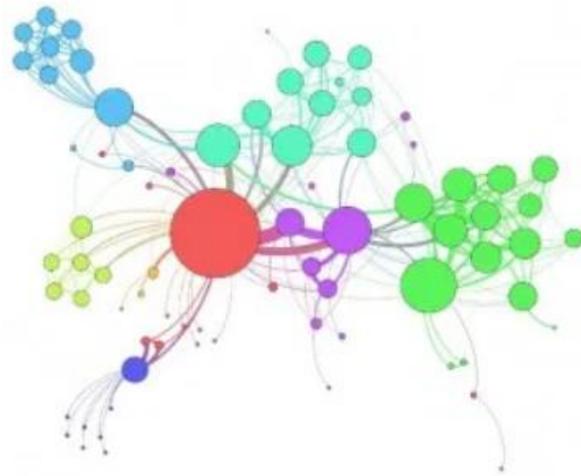


Schéma élaboré par l'équipe MARGE (Université Jean Moulin, Lyon 3) dans le cadre de l'IDEX CartoWeb.

Cartographeur le Web et le visualiser...

Réseaux d'acteurs ?

Contenus, producteurs et autres en milieu Internet



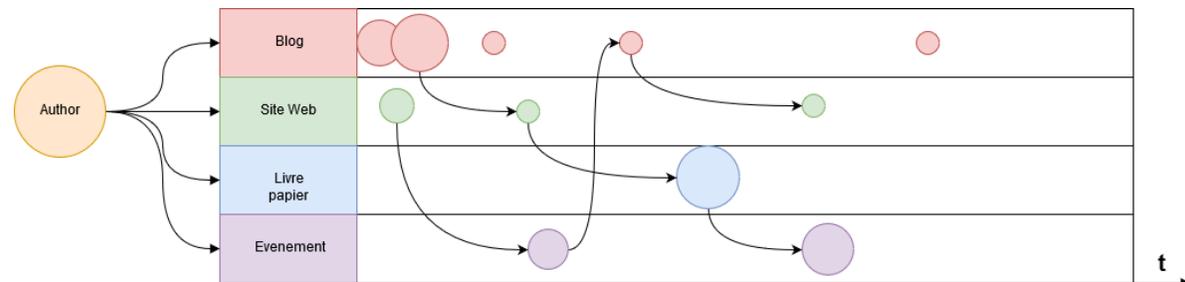
Lecture distante de contenus ?

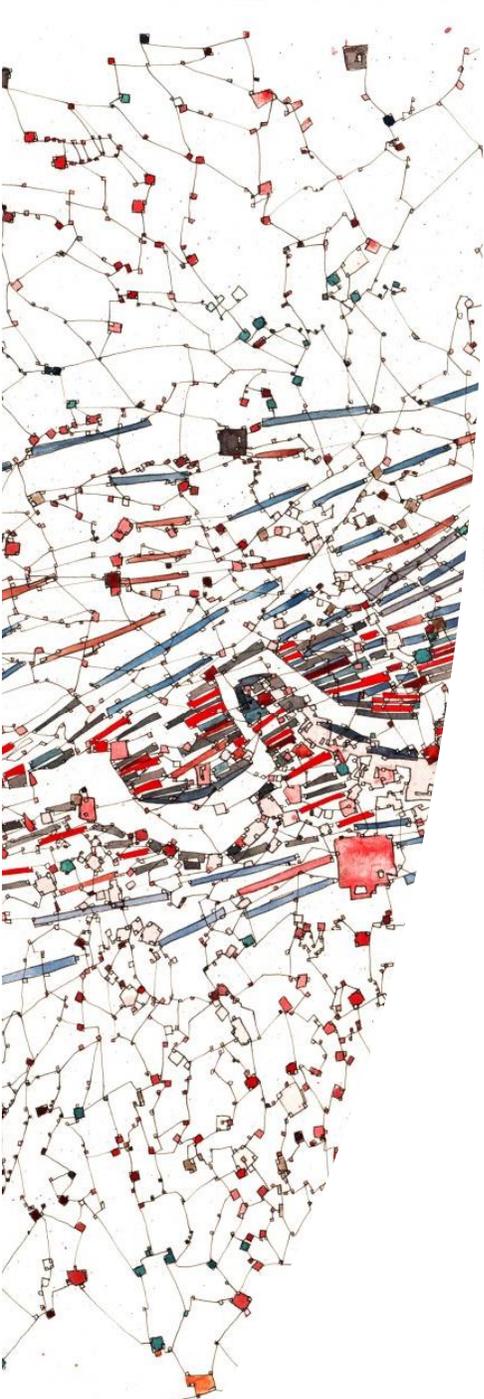


Des cartes ...



Parcours de construction littéraire ?





Communication du 22 janvier 2020

« Ébauche de visualisation des productions littéraires francophones d'Haïti et retour d'expérience »

Colloque international organisé par l'UR MARGE (UR 3712)



**CARTOGRAPHIE
DU WEB
LITTÉRAIRE
FRANCOPHONE**

UNIVERSITÉ LYON III
JEAN MOULIN

marge

IDEX
LYON
UNIVERSITÉ DE LYON

UNIVERSITÉ
JEAN MONNET
SAINT-ÉTIENNE

eric

INSTITUT
INTERNATIONAL
POUR
FRANCOPHONIE

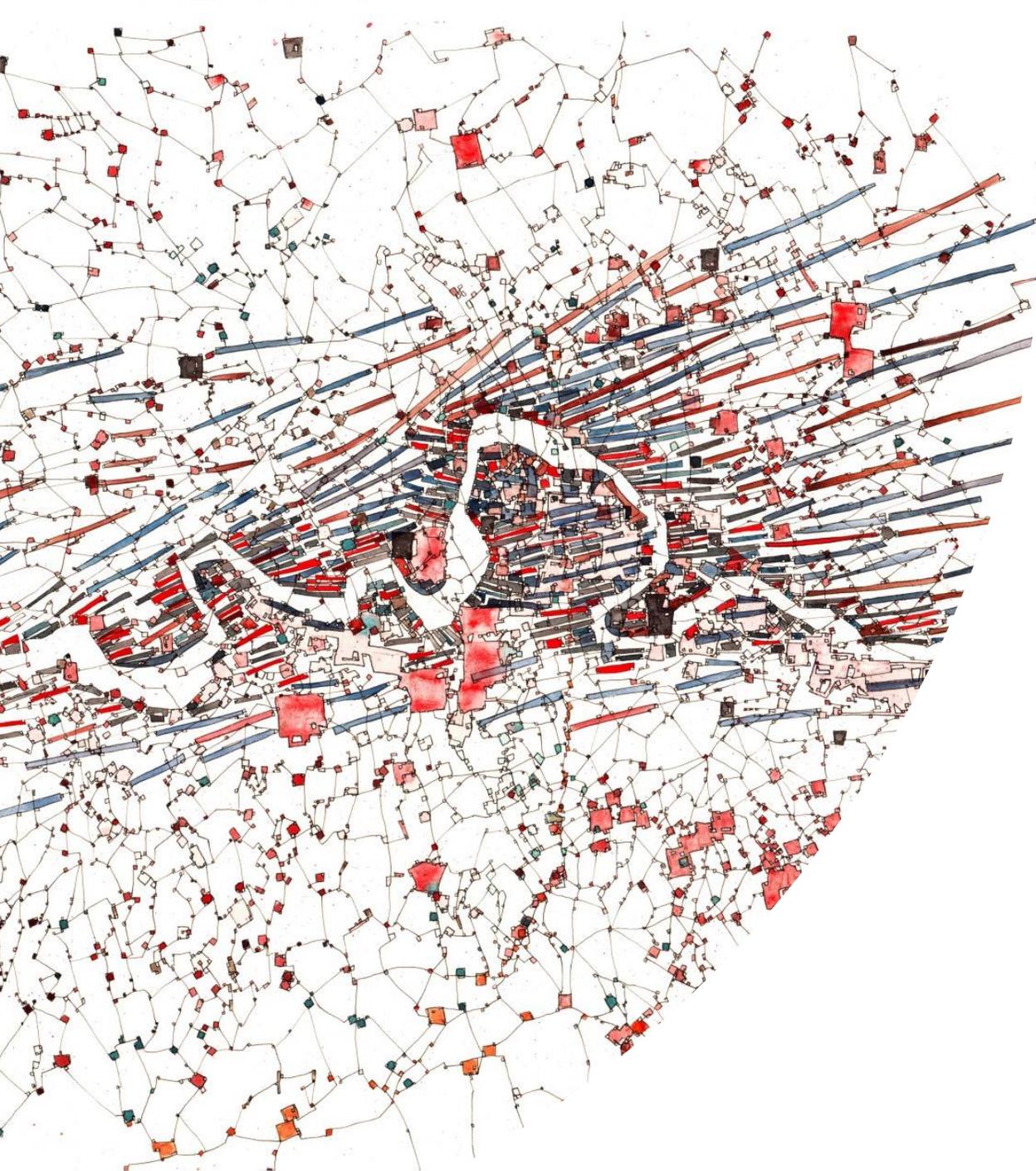
LIFRANUM

msh
Lyon St-Étienne

ANR
PROJET FINANCÉ PAR LE CNRS
PROJET FINANCÉ PAR LE CNRS
PROJET FINANCÉ PAR LE CNRS



<https://webtv.univ-lyon3.fr/videos/?video=MEDIA200205101828908>

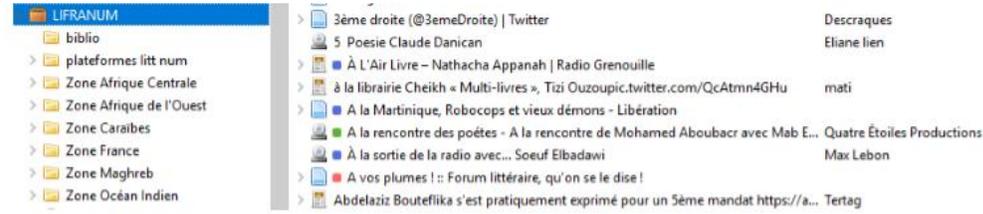


II. Croiser Web sémantique et contenus participatifs

Sur la base d'une bibliothèque de liens Zotéro



Bibliothèque Zotéro
Organisée géographiquement



Extrait RDF des ressources :
Titre site, type, date, auteur,
droits, id, publications,
contributeur, couverture



```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  <rdf:Description rdf:about="https://masiwa.komori.com/fr-km/
  culture/2019-01-17/la-redemption-des-iguanes-est-une-veritable-foire-un-souk-ou-chacun-peut-trouver">
    <dc:title>La Rédemption des Iguanes est une véritable foire. Un souk où chacun peut trouver sa sauce | Masiwa komori</dc:title>
    <dc:type>webpage</dc:type>
  </rdf:Description>
  <rdf:Description rdf:about="https://www.facebook.com/
  DrMohamedAHMEDBACAR?fref=search&__tn__=%2Cd%2CP-R&eid=ARbx1RoMYMy23WabGrOHPtL-3SgWm9U_n2nYTRILBwBF7C6BMwv5FRRA_v6NZ4SwbjatT-w94mMT">
    <dc:title>Mohamed Ahmed Bacar</dc:title><dc:type>webpage</dc:type>
  </rdf:Description>
  <rdf:Description rdf:about="https://www.youtube.com/channel/UC603XzinhMR9bukJdhQzABQ">
    <dc:title>Chaîne Youtube de Mohamed Ahmed Bacar Rezida</dc:title>
    <dc:type>webpage</dc:type>
    <dc:creator>Mohamed Ahmed Bacar Rezida</dc:creator>
  </rdf:Description>
  <rdf:Description rdf:about="http://patpantin.over-blog.com/">
    <dc:title>Le blog de Monique MERABET</dc:title>
    <dc:type>webpage</dc:type>
    <dc:creator>MERABET, Monique</dc:creator>
  </rdf:Description>
```

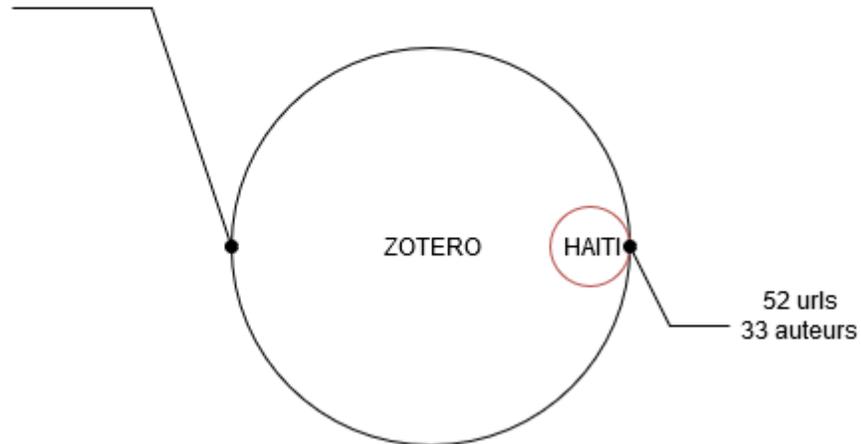
Corpus limité à Haïti



ID	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Accilien, Selmy	Accilien, Selmy	https://ww.facebook.com/selmy_accilien/	français, créole										
2	Anonyme	Des fleurs	https://deliblog.wordpress.com/	français, créole										
3	Anonyme	Poète, slar	https://ww.facebook.com/poete_slar/											
4	Apollon, Pascal	Association	https://www.facebook.com/association/											
5	Bidler, Nel	blog avec	https://ww.facebook.com/blog-avec/											
6	Bidler, Nel	blog d'écri	https://me.blog-supprime.com/											
7	Bonel, Auguste	Poèmes	https://arr.blog/											
8	Bonhomme, Adlyne	Adlyne Bonhomme	https://shv.platforme-editoriale.com/											
9	Bonhomme, Adlyne	Poèmes d'	https://pei.blog/											
10	Cénas, Georges	blog de po	http://peti.blog/											
11	Cocks, Georges	cocks geor	http://coc.blog/											
12	Collectif -	Association	https://www.facebook.com/association/											
13	Collectif -	facebook	https://ww.blog-communautaire-facebook.com/											
14	Collectif -	Société du	https://ww.blog-communautaire-facebook.com/											
15	D'Amérique	blog de po	https://blc.blog-mediapart.com/											
16	Delisca, Al	Changeme	https://shv.platforme-editoriale.com/											
17	Guyve, Fran	guyveco	https://gu.blog/											
18	Jeudi, Jean	Poésie	https://ww.facebook.com/poesie-jeudi/											
19	Jeudi, Jean	Blog	http://gon.blog/											
20	Jeudi, Jean	poésie fac	https://ww.facebook.com/poesie-fac/											
21	Jeudi, Jean													

Premières observations

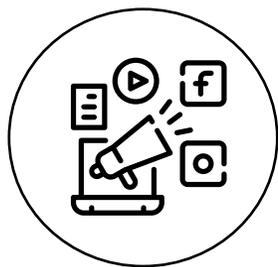
1658 urls
592 ayant un titre
586 typés
240 auteurs identifiés
240 datés



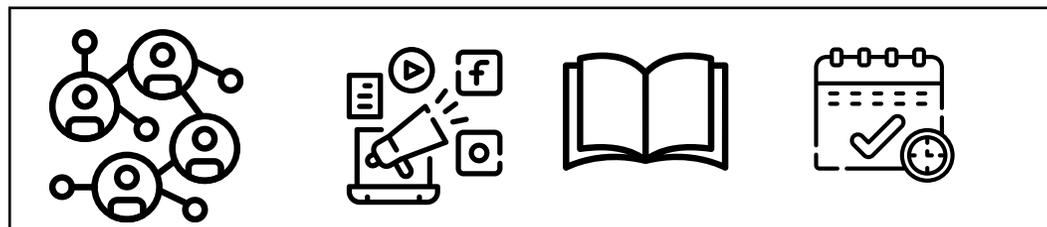
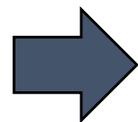
- Visualiser des contenus ?
 - Sources hétérogènes : poèmes, articles de blog ou à caractère journalistique, etc.
- Risque d'un corpus « brouhaha »



Changement de paradigme



Contenu Web

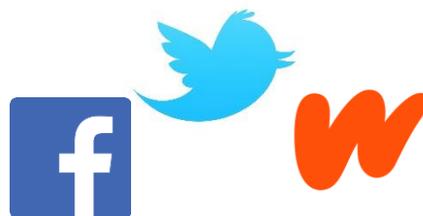


Réseaux d'auteurs → production Web + imprimés + événements

Scrapping de sites de « référence »:
Ile-en-île.org, Spla.pro

Données du Web social

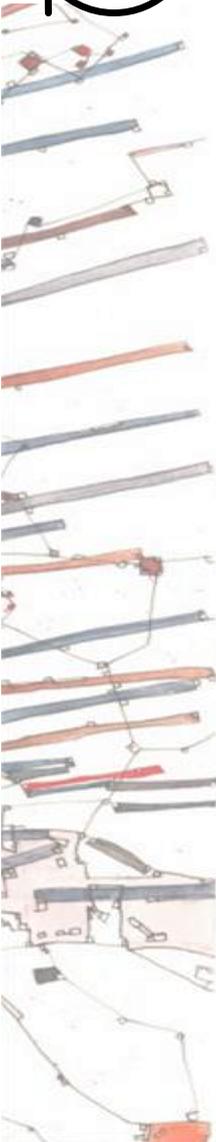
Données du Web sémantique



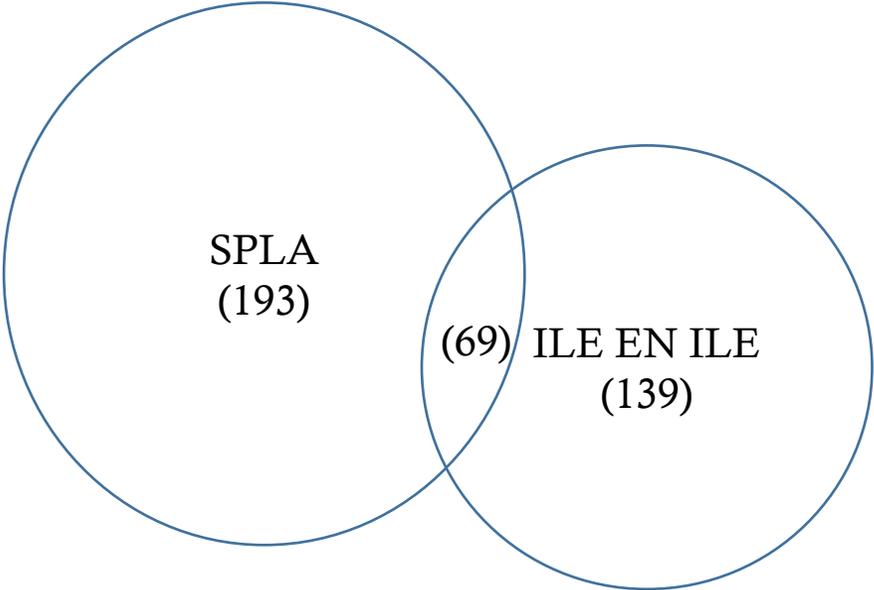
Les sources potentielles et leurs caractéristiques



	<i>Scrapping de sites de « référence »</i>	<i>Données du Web social</i>	<i>Données du Web sémantique</i>
Qualité de formalisation	<ul style="list-style-type: none">- Très peu formatées- Souvent textuelles	<ul style="list-style-type: none">- Formatées	<ul style="list-style-type: none">- Formatées + vérifiées
Informations disponibles	<ul style="list-style-type: none">- Profils auteurs- Bibliographies- Références Web	<ul style="list-style-type: none">- Profils auteurs- Productions- Évènements + liens	<ul style="list-style-type: none">- Profils auteurs- Bibliographies- Liens sémantiques
Accès à l'information	<ul style="list-style-type: none">- Moteur de recherche	<ul style="list-style-type: none">- API + M.R.	<ul style="list-style-type: none">- API + M.R.
Garant éditorial	<ul style="list-style-type: none">- Garant « amateur »	<ul style="list-style-type: none">- L'auteur et son réseau	<ul style="list-style-type: none">- Garant « expert »



Informations obtenues via les sites de « référence »



Meta-auteurs Excel : 31 *auteurs*

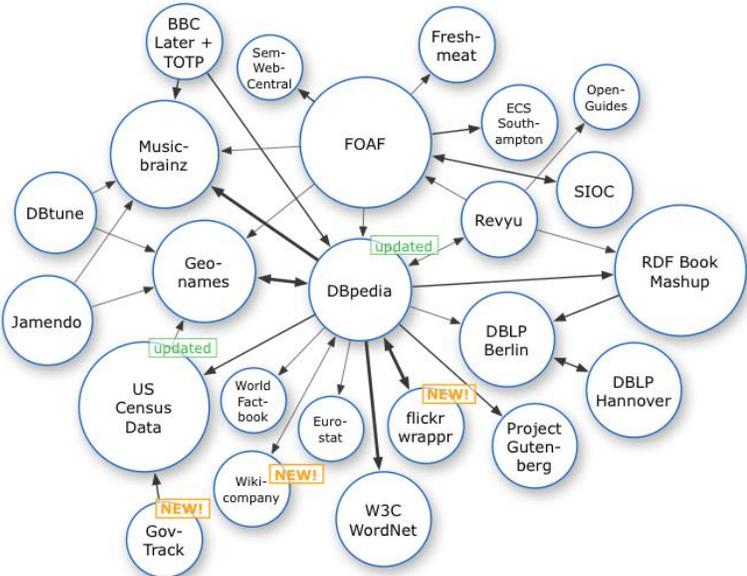
- 17 VIAF
- 14 BNF

Meta-final : 274 *auteurs*

- 74 VIAF
- 86 BNF

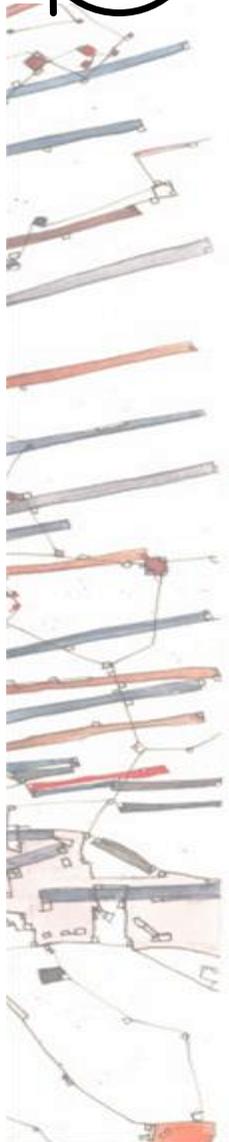
MAIS...

- Présence de 44 doublons
 - Et parfois peu d'informations
- nécessité d'avoir une description minimale et normalisée, indépendante de la langue



Notre proposition : interroger le Wikidata via « OpenRefine » et d'une requête SPARQL

[Introduction aux données liées avec OpenRefine](#)



Explicitation de la requête Wikidata - SPARQL



Wikidata Query Service

Exemples Aide Davantage d'outils français

```
1 #Auteurs, écrivains et poètes... né après 1920 à Haiti
2 SELECT distinct ?s
3 WHERE
4 {
5   { ?s wdt:P106 wd:Q36180 . } UNION { ?s wdt:P106 wd:Q482980 . } UNION { ?s wdt:P106 wd:Q49757 . } UNION { ?s wdt:P106 wd:Q49757 . }
6   { ?s wdt:P19 wd:Q790 . } UNION { ?s wdt:P19 wd:Q34261 . } UNION { ?s wdt:P19 wd:Q923362 . } UNION { ?s wdt:P19 wd:Q986301 . } UNION { ?s wdt:P19 wd:Q921827 . }
7   ?s wdt:P570 ?time0 .
8   FILTER(?time0 >= "1920-01-01T00:00:00Z"^^xsd:dateTime)
9
10 }
11
```

Occupation : écrivain, poète, essayiste...

Né à Haiti

Né après 1920

<https://query.wikidata.org/>

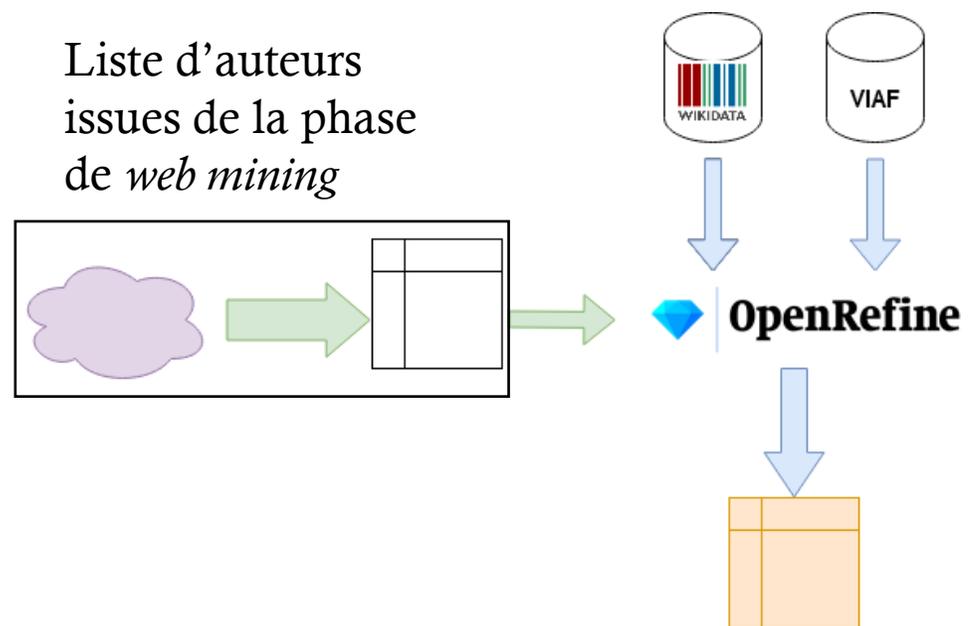
Critères retenus :

- être né après 1920,
- être originaire d'Haïti,
- occuper une fonction littéraire.

→ 146 résultats dont 25 nouveaux auteurs pertinents

Consolidation des données

OpenRefine permet de réaliser des recherches assistées dans des bases de données Ouvertes et Liées mais aussi à partir de répertoires d'autorités



Insertion de la liste d'écrivains haïtiens



 **OpenRefine** *A power tool for working with messy data.*

- Create Project
- Open Project
- Import Project
- Language Settings

Create a project by importing data. What kinds of data files can I import?

TSV, CSV, *SV, Excel (.xls and .xlsx), JSON, XML, RDF as XML, and Google Data documents are all supported. Support for other formats can be added with OpenRefine extensions.

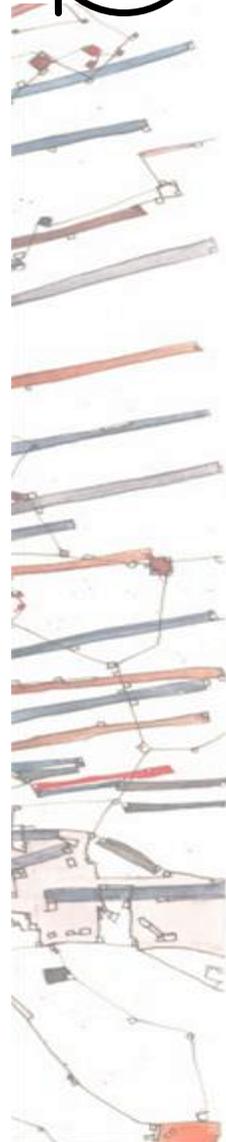
Get data from

- This Computer
- Web Addresses (URLs)
- Clipboard**
- Database
- Google Data

Paste data from clipboard here:

Stanley Péan
Ignace Nau
Jean Métellus
Lenelle Moïse
Edmond Laforest
Raymond Chassagne
Gary Klang
Frederic Marcelin
Antoine Innocent
Yanick Jean
Dantès Bellegarde
Georges Anglade
Jean-Max Calvin
Hervé Fanini-Lemoine
Émile Ollivier
Luckner Desvignes
Frédéric Marcelin
Jean Durosier Desrivières
Yanick Lahens
Roland Morisseau
Gesner Jean Gilles
Wébert Charles
Marc Léo Laroche
Rodolphe Mathurin
Danx Laferrière
Gérard Chenet

Next »



Regroupements d'informations via Wikidata



OpenRefine clipboard [Permalink](#)

Facet / Filter Undo / Redo 0 / 0

243 rows

Show as: rows records

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started?
[Watch these screencasts](#)

Services

Pick a Service or Extension on Left

Wikidata (old) (en) ✕
Getty Vocabularies Reconciliation Service ✕
Wikidata (en) ✕

Recherche d'entités de type Q5

Facet
Text filter
Edit cell
Edit column
Transpose
Sort...
View
Reconcile

Start reconciling...
Reconcile text in this column with items on Fr

Facets
Actions
Copy reconciliation data...
Use values as identifiers

Regroupements d'informations via le VIAF



D'autres services sont disponibles sur <http://refine.codefork.com/> tels que :

- ORCID
- Open Library

Reconcile each cell to an entity of one of these types:

- Person
/people/person
- Work
/book/book
- Expression
/book/book edition

Also use relevant details from other columns:

Column Include? As Property

Pick a Service or Reconciliation Service on Left

Getty Vocabularies Reconciliation Service

Wikidata (en)

Add standard service

Enter the service's URL:

Add Service Cancel

Récupération d'informations clefs



207 matching rows (262 total)

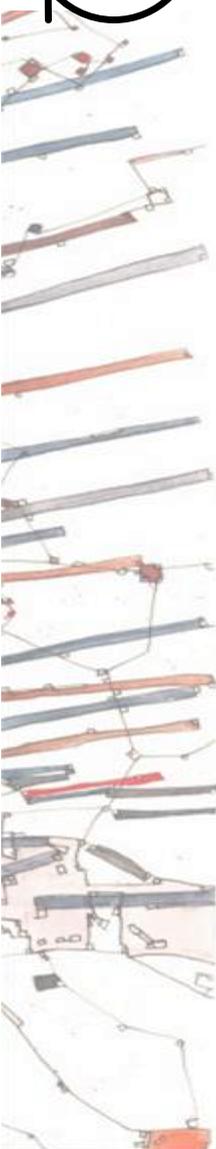
Show as: **rows** records Show: 5 10 25 50 rows Sort ▾

All	Column 1	place of birth	country of citizenship	date of birth	date of death	VIAF
190.	Facet	Mayotte	France	1965-01-06T00:00:00Z		
182.	Text filter		Haiti	1967-01-01T00:00:00Z		
150.	Edit cells			1945-08-30T00:00:00Z		
175.	Edit column					
186.	Transpose			1963-01-01T00:00:00Z		
98.	Sort			1963-01-01T00:00:00Z		
205.	View			1850-10-18T00:00:00Z	1911-09-19T00:00:00Z	
177.	Reconcile			1928-08-25T00:00:00Z		
11.	Antoine Innocent			1874-01-01T00:00:00Z	1960-01-01T00:00:00Z	
151.	Barbara Bastien					
164.	Boisrond-Tonnerre			1776-01-01T00:00:00Z	1806-01-01T00:00:00Z	
107.	Bonel Auguste	Port-au-Prince	Haiti	1973-03-06T00:00:00Z		

Ainsi nous récupérons via Wikidata :

- Date de naissance
- Date de décès
- Lieu de naissance
- Lieu de décès
- Nationalité
- Professions

....



Filtrage du corpus

1) Suppression des doublons

2) Normalisation du corpus

→ 232 individus uniques

3) Application de filtres plus restrictifs:

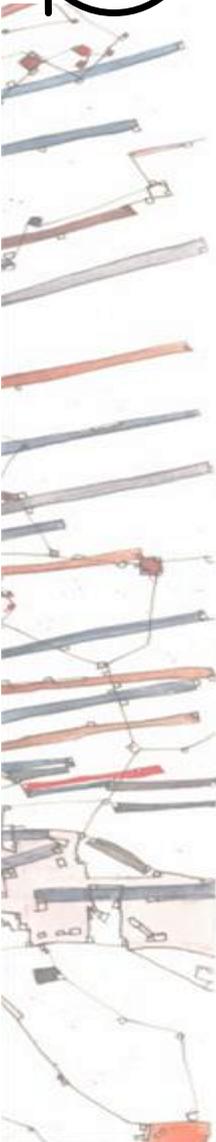
- Débuts d'Internet (à destination du public et des universitaires) : années 1990-2000
- Littéraires ayant pu avoir recours à Internet pour diffuser leurs travaux et faire leur promotion.

Choix corpus : littéraires nés avant 1920, décédés après 2000

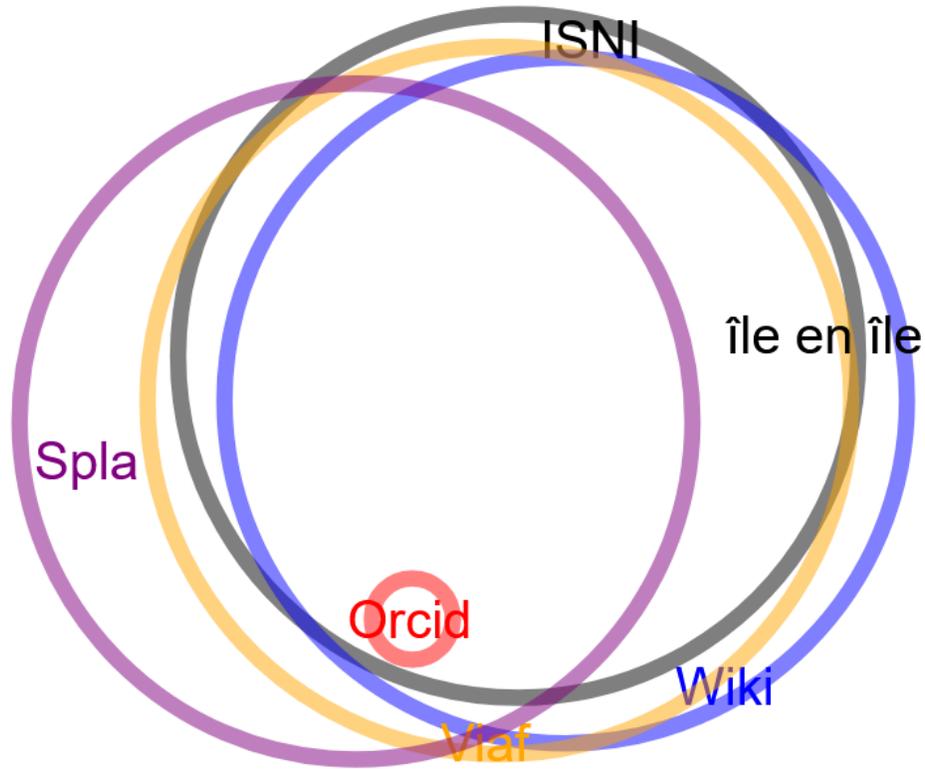
Si la consolidation des données via Wikidata s'est avérée efficace, il demeure que :

- 22 dates de naissances ont pu être récupérées manuellement dans les notices biographiques
- Et 8 autres grâce au récolement du VIAF

→ Il convient donc de mobiliser diverses sources d'enrichissement.



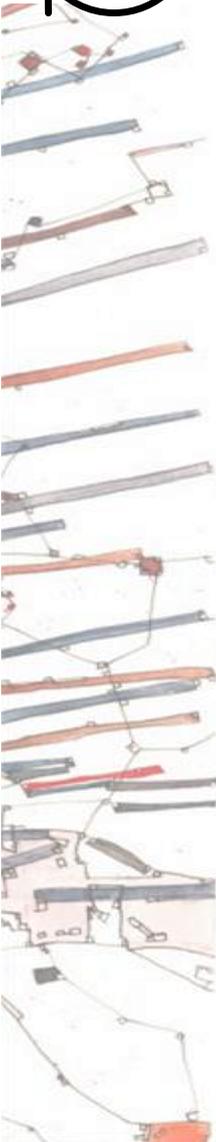
Le corpus final en quelques chiffres clés



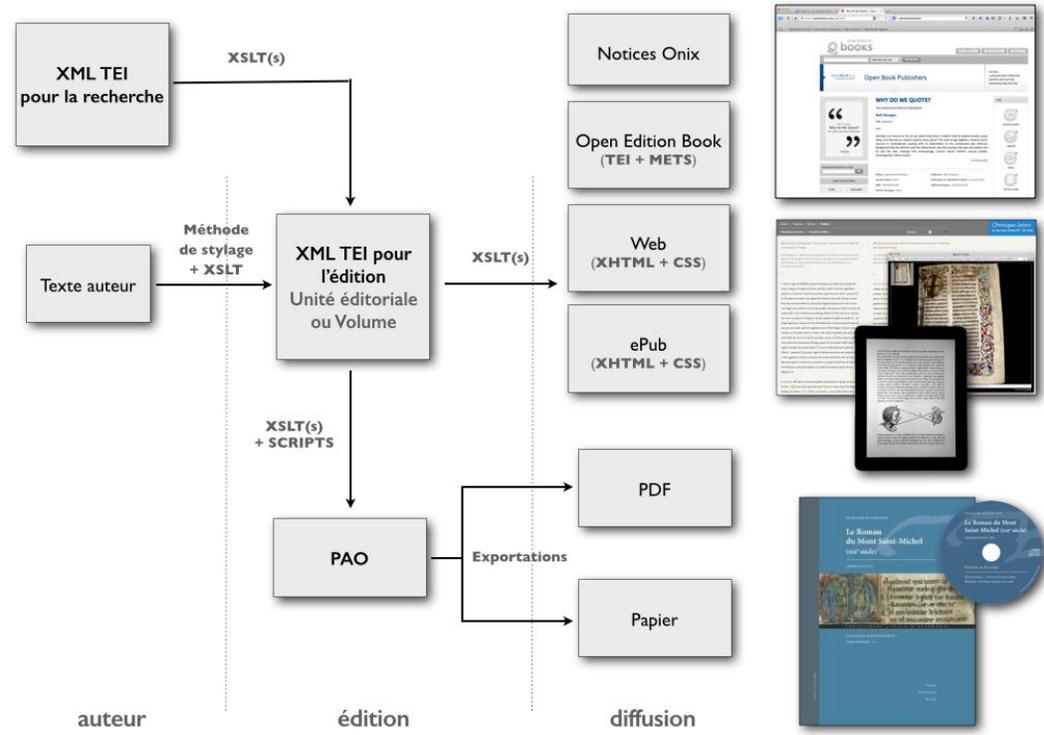
190 individus dont :

- 151 sont renseignés par une notice VIAF
- 142 Wikidata
- 141 ISNI
- 138 Spla
- 108 Ile-en-île
- 2 ORCID

Informations	Nbre de données obtenues
Sexe/genre	190
Nationalité	190
Langue	190
Activités professionnelles	190
Date de naissance	163
Lieu de naissance	129
Date de décès	35
Lieu de décès	26

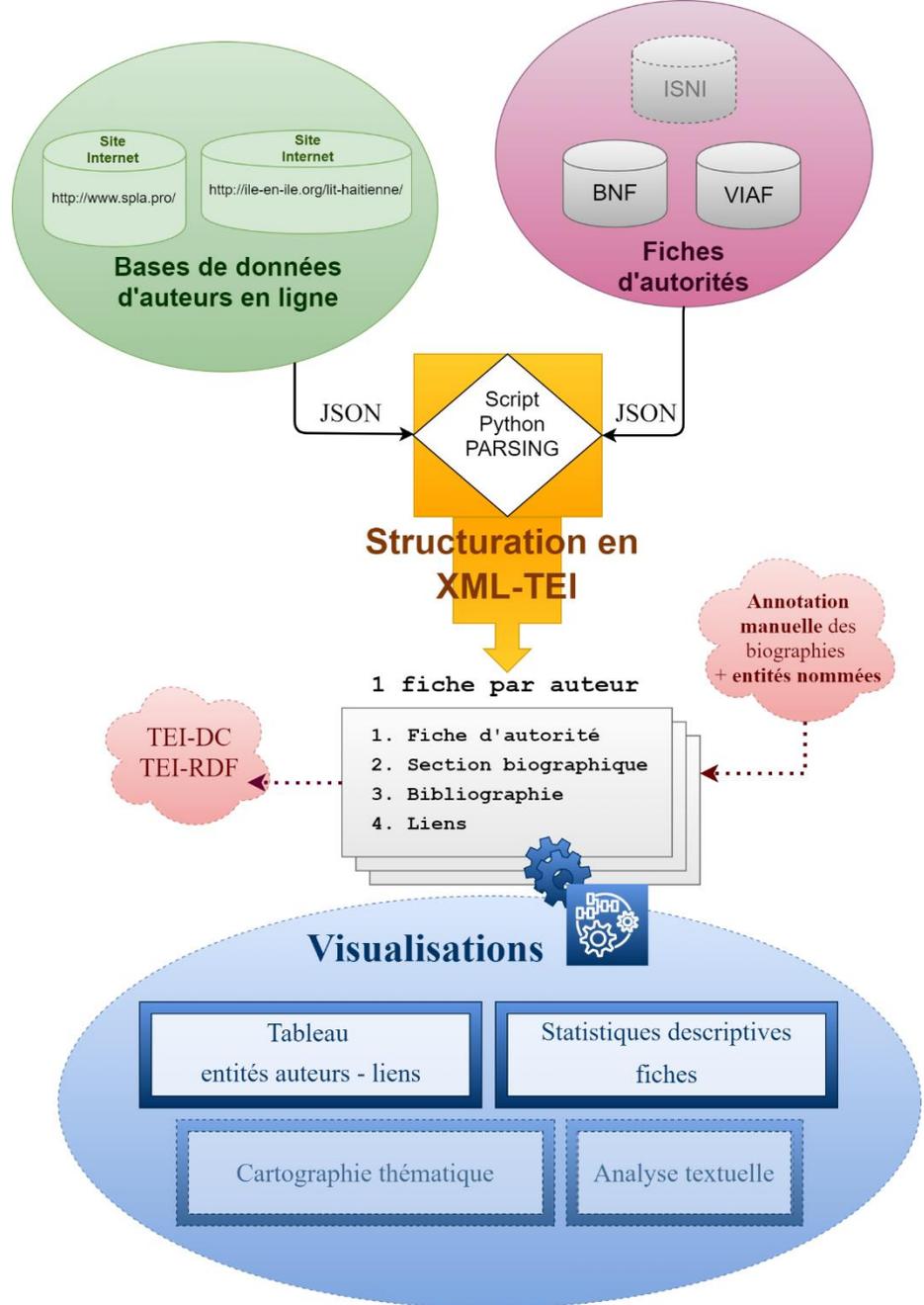


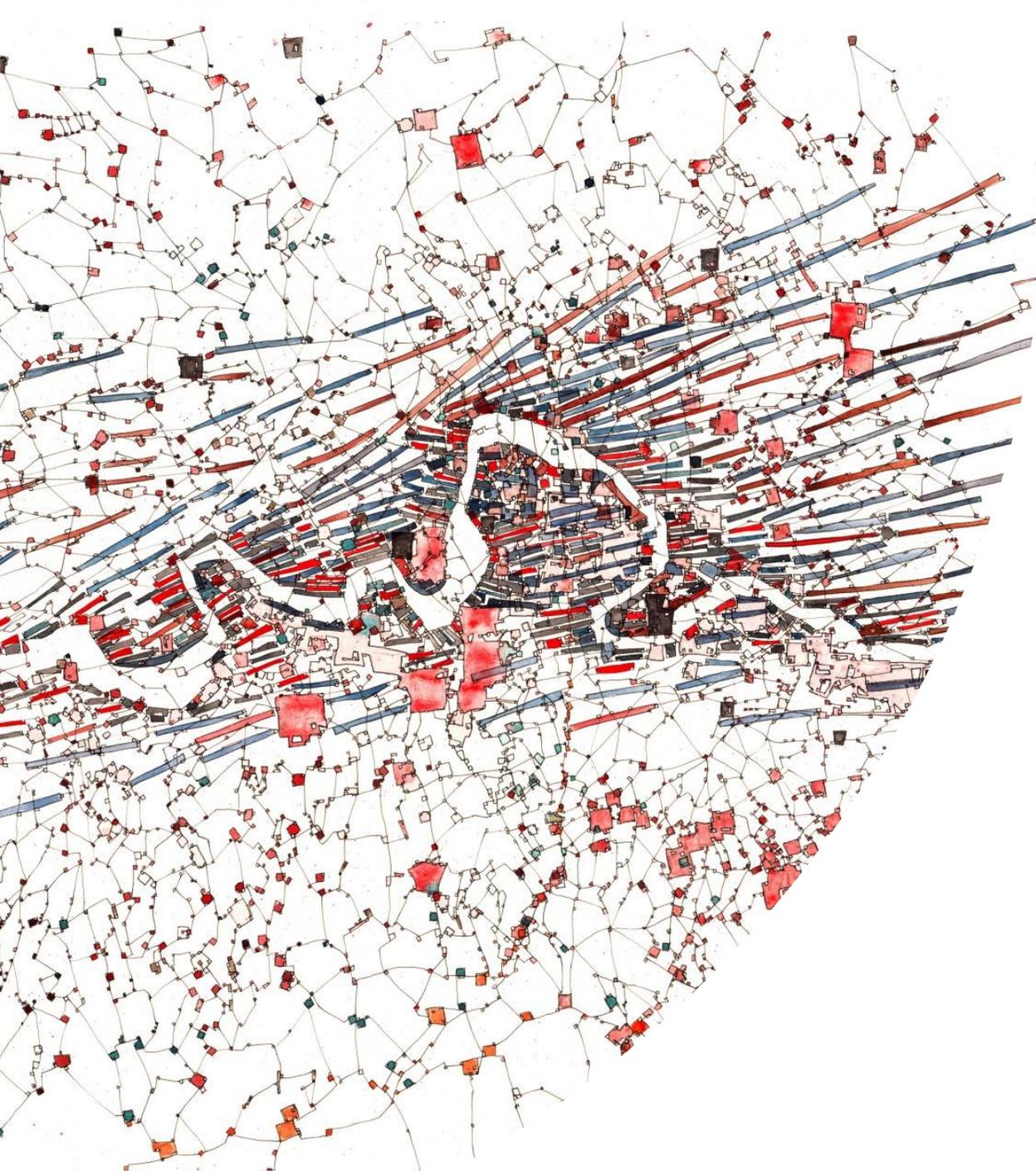
Agréger et normaliser le tout ?



- Il permet l'annotation de document à même le contenu
- Il est lisible par un humain
- Il permet d'accéder à un écosystème orienté documents
- Peut être transformé dans de nombreuses « formes »

Schéma de la chaîne éditoriale (modèle du *Single Source Publishing*)





III. Normalisation du corpus et structuration en XML-TEI

- Avant toute chose ...

Un contenu textuel

```
<persName>Bureau Bruno</persName>
```

Une balise : ouvrante `<balise>`
et fermante `</balise>`

```
<persName>
```

```
  <surname>Bureau</surname>
```

```
  <forename>Bruno</forename>
```

```
</persName>
```

Des balises dans des balises

```
<persName ana="bureau_bruno">Il Maestro</persName>
```

Un attribut typé avec un texte

Un texte qui contient des annotations numériques en vue de l'**enrichir** et sans changer son contenu.



Structure du fichier XML-TEI

TEI

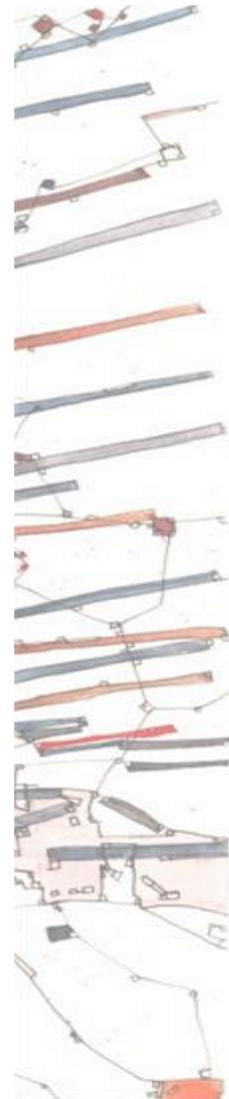
```

1 <?xml version="1.0" encoding="UTF-8"?> <!-- déclaration du fichier TEI et de son encodage -->
2 <TEI xmlns:xi="http://www.w3.org/2001/XInclude" xml:lang="fr" xml:id="Marie-Celie_Agnant">
3
4 <!-- 1 fichier TEI par entité auteur -->
5 <!-- xml:id= correspondant à celui du fichier master.-->
6 <!-- Normalisation de l'id: prenom_nom de entité auteur -->
7
8 <teiHeader> [21 lines]
30
31 <div xml:id="Marie-Celie_Agnant_fiche" n="1"> [65 lines]
97
98 <div xml:id="Marie-Celie_Agnant_bio" n="2"> [73 lines]
172
173 <div xml:id="Marie-Celie_Agnant_biblio" n="3"> [81 lines]
255
256 <div xml:id="Marie-Celie_Agnant_web" n="4"> [11 lines]
268 </TEI>
269

```

↳ Un identifiant *xml:id* unique et normalisé

1. Un `teiHeader`
2. Le corps



Le <teiHeader>

```

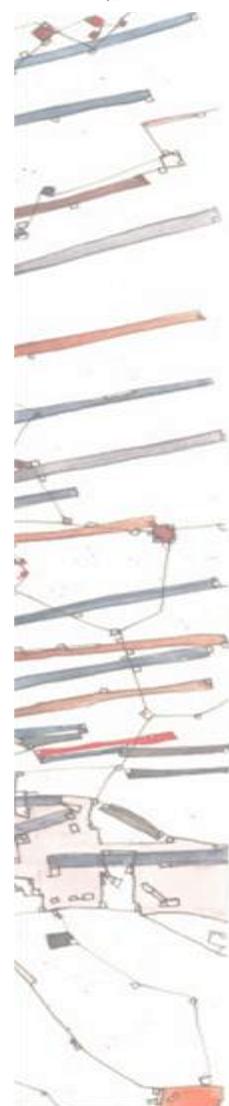
<teiHeader>
  <!-- déclaration des métadonnées -->
  <fileDesc>
    <titleStmt>
      <title>Notice de Marie-Célie Agnant - corpus Haïti</title>
      <!-- prénom, nom de l'auteur -->
      <author>Inès Burri</author>
      <author>Anaïs Chambat</author>
      <author>Célian Ringwald</author>
    </titleStmt>
    <publicationStmt>
      <orgName>Université Jean Moulin Lyon 3</orgName>
      <orgName>Université Lumière Lyon 2</orgName>
      <orgName>ENS de Lyon</orgName>
      <pubPlace>Lyon</pubPlace>
      <date>17.01.2020</date>
      <!-- date de l'extraction -->
    </publicationStmt>
    <sourceDesc>Notice et contenus associés à l'entité auteur dont les données sont issues
      de sites de références et de répertoires d'autorités</sourceDesc>
  </fileDesc>
</teiHeader>

```

Description du fichier, ...

... des données éditoriales

... et des sources



Des <div>

```
TEI
1 <?xml version="1.0" encoding="UTF-8"?> <!-- déclaration du fichier TEI et de son encodage -->
2 <TEI xmlns:xi="http://www.w3.org/2001/XInclude" xml:lang="fr" xml:id="Marie-Celie_Agnant">
3
4 <!-- 1 fichier TEI par entité auteur -->
5 <!-- xml:id= correspondant à celui du fichier master -->
6 <!-- Normalisation de l'id: prenom_nom de entité auteur -->
7
8 <teiHeader> [21 lines]
30
31 <div xml:id="Marie-Celie_Agnant_fiche" n="1"> [65 lines]
97
98 <div xml:id="Marie-Celie_Agnant_bio" n="2"> [73 lines]
172
173 <div xml:id="Marie-Celie_Agnant_biblio" n="3"> [81 lines]
255
256 <div xml:id="Marie-Celie_Agnant_web" n="4"> [11 lines]
268 </TEI>
269
```

Un numéro

Un identifiant unique



<div> fiche d'autorité (1)

```
<div xml:id="Marie-Celie_Agnant_fiche" n="1">
```

```
  <!-- xml:id= prenom_nom de entité auteur_type de fichier (ici fiche)-->
```

```
  <!-- n=1 pour fiche -->
```

```
  <person ana="Marie-Celie_Agnant" sex="f">
```

```
    <!-- ana= prenom_nom de l'auteur -->
```

```
    <!-- sex= f pour féminin; m pour masculin -->
```

```
    <persName>Agnant Marie-Célie</persName>
```

Nom et prénom

```
  <!-- Identifiants liés aux répertoires d'autorités disponibles -->
```

```
  <!-- Il est possible d'indiquer pour chaque balise, la provenance de l'élément grâce à
```

```
  <idno type="BNF"><g ref="https://catalogue.bnf.fr/ark:/12148/cb145884081"
```

```
    />FRBNF14588408</idno>
```

```
  <idno type="VIAF"><g ref="http://viaf.org/viaf/17470529"/>17470529</idno>
```

```
  <idno type="DBPEDIA"><g ref="https://fr.wikipedia.org/wiki/Marie-C%C3%A9lie_Agnant"
```

```
    />Marie-Célie_Agnant</idno>
```

```
  <!-- DB de Wikipédia -->
```

Identifiants des
répertoires autorité

```
  <birth source="BNF">
```

```
    <date>1953</date>
```

```
    <placeName>Haïti</placeName>
```

```
  </birth>
```

```
  <death>
```

```
    <date/>
```

```
    <placeName/>
```

```
  </death>
```

Naissance et décès

```
  <langKnowledge source="BNF">
```

```
    <!-- pour le code de la langue utilisé en XML-TEI voir : http://www.iana.org/assign
```

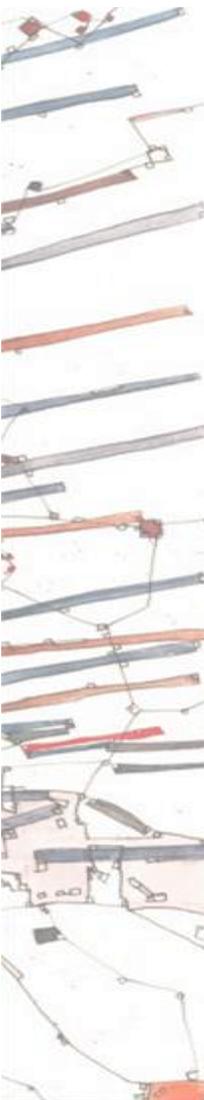
```
    <langKnown tag="en">anglais</langKnown>
```

```
    <langKnown tag="fr">français</langKnown>
```

```
    <langKnown tag="ht">haitien-creole</langKnown>
```

```
  </langKnowledge>
```

Langues parlées



<div> fiche d'autorité (2)

```
<residence>
  <placeName>
    <settlement/>
    <!-- balise qui admet un type="city"(ville) et "town"(village) -->
    <region source="BNF">Québec</region>
    <!-- région -->
    <country source="BNF">Haïti</country>
    <!-- pays -->
    <country source="BNF">Canada</country>
  </placeName>
</residence>
```

Lieux

```
<occupation source="île-en-île">romancière</occupation>
<occupation source="île-en-île">poète</occupation>
<occupation source="île-en-île">nouvelliste</occupation>
<occupation source="île-en-île">conteuse</occupation>
<occupation source="île-en-île">traductrice</occupation>
<occupation source="île-en-île">interprète</occupation>
<occupation source="île-en-île">assistante de recherche</occupation>
<!-- Le contenu textuel des balises "residence" et "occupation" ne pourra être rempli q
```

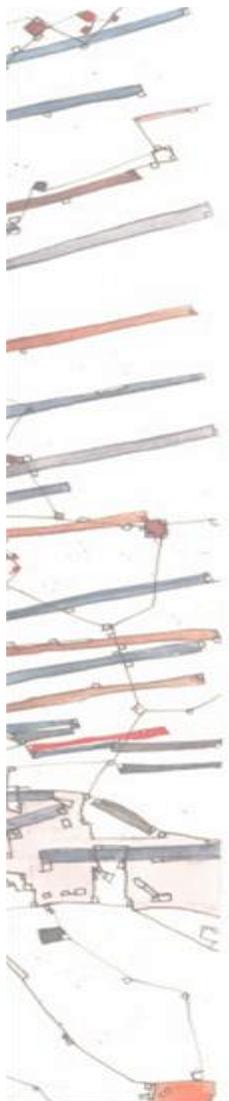
Professions

```
<listRelation>
  <!-- Liste des noms associés à l'auteur -->
  <relation ana="" source="BNF">François Gardant</relation>
</listRelation>
<listEvent>
  <!-- Liste des événements associés à l'auteur -->
  <event type="prix_litteraire" when="01/01/2020"/>
</listEvent>
```

Relations

Événements

```
</person>
<!-- Outre la BNF et le VIAF, l'ISNI peut s'avérer complémentaire en cas de déficit d'infor
</div>
```



Partie 1

```
<div xml:id="Marie-Celie_Agnant_fiche" n="1">
  <!-- xml:id= prenom_nom de entité auteur_type de fichier (ici fiche)-->
  <!-- n=1 pour fiche -->

  <person ana="Marie-Celie_Agnant" sex="f">
    <!-- ana= prenom_nom de l'auteur -->
    <!-- sex= f pour féminin; m pour masculin -->
    <persName>Agnant Marie-Célie</persName>

    <!-- Identifiants liés aux répertoires d'autorités disponibles -->
    <!-- Il est possible d'indiquer pour chaque balise, la provenance de
    <idno type="BNF"><g ref="https://catalogue.bnf.fr/ark:/12148/cb14588408"
    />FRBNF14588408</idno>
    <idno type="VIAF"><g ref="http://viaf.org/viaf/17470529"/>17470529</idno>
    <idno type="DBPEDIA"><g ref="https://fr.wikipedia.org/wiki/Marie-C%27"
    />Marie-Célie_Agnant</idno>
    <!-- DB de Wikipédia -->

    <birth source="BNF">
      <date>1953</date>
      <placeName>Haïti</placeName>
    </birth>
    <death>
      <date/>
      <placeName/>
    </death>

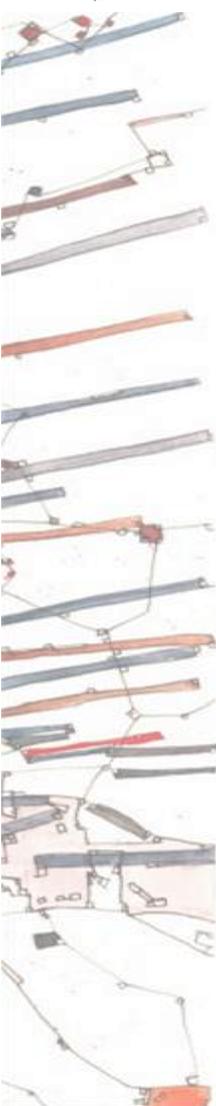
    <langKnowledge source="BNF">
      <!-- pour le code de la langue utilisé en XML-TEI voir : http://
      <langKnown tag="en">anglais</langKnown>
      <langKnown tag="fr">français</langKnown>
      <langKnown tag="ht">haitien-creole</langKnown>
    </langKnowledge>
```

Partie 2

```
<residence>
  <placeName>
    <settlement/>
    <!-- balise qui admet un type="city"(ville) et "town"(village)
    <region source="BNF">Québec</region>
    <!-- région -->
    <country source="BNF">Haïti</country>
    <!-- pays -->
    <country source="BNF">Canada</country>
  </placeName>
</residence>

<occupation source="île-en-île">romancière</occupation>
<occupation source="île-en-île">poète</occupation>
<occupation source="île-en-île">nouvelliste</occupation>
<occupation source="île-en-île">conteuse</occupation>
<occupation source="île-en-île">traductrice</occupation>
<occupation source="île-en-île">interprète</occupation>
<occupation source="île-en-île">assistante de recherche</occupation>
<!-- Le contenu textuel des balises "residence" et "occupation" ne pou

<listRelation>
  <!-- Liste des noms associés à l'auteur -->
  <relation ana="" source="BNF">François Gardant</relation>
</listRelation>
<listEvent>
  <!-- Liste des événements associés à l'auteur -->
  <event type="prix_litteraire" when="01/01/2020"/>
</listEvent>
</person>
<!-- Outre la BNF et le VIAF, l'ISNI peut s'avérer complémentaire en cas d
</div>
```



<div> bio

```
<div xml:id="Marie-Celie_Agnant_bio" n="2">
```

```
<!-- xml:id= prenom_nom de entité auteur_type de fichier (ici bio)-->
```

```
<!-- n=2 pour bio -->
```

La source

```
<text ana="Marie-Celie_Agnant" source="ile-en-ile" ref="http://ile-en-ile.org/agnant/"  
hand="Lucie Lequin">
```

L'adresse Web

```
<!-- ana= prenom_nom de l'auteur -->
```

```
<!-- ref= adresse de la bio -->
```

```
<!-- hand = dernier paragraphe de la bio -->
```

Le nom de l'auteur de la biographie

```
<p>Poète, nouvelliste, romancière, Marie-Célie Agnant quitte Haïti pour le Canada en  
1970. Traductrice et interprète, elle détient un diplôme en enseignement du français  
et a enseigné plusieurs années tout en travaillant comme assistante de recherche.  
C'est ainsi que son premier roman, La dot de Sara, publié en 1995, sera « construit  
à partir de récits de vie de grands-mères d'origine haïtienne, transmis en créole  
dans le cadre d'une recherche en sociologie ».</p>
```

```
<p>Écrivaine présente et attentive au monde qui l'entoure, son écriture porte à la fois  
le sceau de la poésie et de la violence issue des sociétés postcoloniales qui  
naviguent entre misère criante et opulence indécente. Ses textes, dont certains ont  
été traduits en plusieurs langues, trouvent leur ancrage dans la réalité sociale  
contemporaine ; elle aborde les thèmes tels l'exclusion, la solitude, le racisme,  
l'exil. La condition des femmes, le rapport au passé et à la mémoire font aussi  
partie de son champ d'exploration.</p>
```

Contenu de la
biographie dans
des balises <p>
(paragraphe)

<div> biblio

```

<div xml:id="Marie-Celie_Agnant_biblio" n="3">
  <!-- xml:id= prenom_nom de entité auteur_type de fichier (ici biblio)-->
  <!-- n=3 pour biblio -->
  <div ana="Marie-Celie_Agnant" source="ile-en-ile" ref="http://ile-en-ile.org/agnant/"
    n="3.1">
    <!-- ana= prenom_nom de l'auteur -->
    <!-- source= nom du site à partir duquel les données ont été extraites -->
    <!-- ref= adresse de la biblio -->
    <!-- n= numéro itératif qui permet de hiérarchiser suivant le nombre de sites cons
    <!-- Chaque item = une référence unique (les romans ont été développés en exemple)
    <!-- POUR LES LIST : subtype= titre de la section principale -->
    <!-- POUR LES ITEM : type= titre de la section principale ; subtype= titres des soi
    <!-- Les références changent d'une entité auteur à l'autre: l'information extraite
    <list subtype="romans"> [10 lines]
    <list subtype="nouvelles"> [2 lines]
    <list subtype="poésie"> [2 lines]
    <list subtype="romans pour la jeunesse"> [2 lines]
    <list subtype="contes"> [2 lines]
    <list subtype="textes publiés dans des ouvrages collectifs"> [2 lines]
    <list subtype="distinctions littéraires"> [2 lines]
    <!-- oeuvres traduites -->
    <!-- pour le code de la langue utilisé en XML-TEI voir : http://www.iana.org/assign
    <list subtype="catalan" tag="ca"> [2 lines]
    <list subtype="anglais" tag="en"> [2 lines]
    <list subtype="espagnol" tag="es"> [2 lines]
    <list subtype="italien" tag="it"> [2 lines]
    <list subtype="néerlandais" tag="nl"> [2 lines]
    <list subtype="ouvrage collectif"> [2 lines]
    <list subtype="articles sélectionnés"> [2 lines]
    <list subtype="entretiens"> [2 lines]
  </div>
  <div ana="Marie-Celie_Agnant" source="VIAF" n="3.2"> [4 lines]
</div>

```

→ L'adresse Web

→ La source

Un numéro

Des balises <list> avec un attribut *subtype*
 Dans celles-ci des <item> pour chaque entrée bibliographique



<div> adresses Web

```
<div xml:id="Marie-Celie_Agnant_web" n="4">
  <!-- xml:id= prenom_nom de entité auteur_type de fichier (ici liens)-->
  <!-- n=4 pour liens -->
```

```
<list ana="Marie-Celie_Agnant" source="ile-en-ile" subtype="liens">
  <!-- ref target= mettre le lien entre les guillemets ; entre les bal
  <item><ref target="lien"/>text</item>
</list>
```

```
<list ana="Marie-Celie_Agnant" source="spla" subtype="liens">
  <item><ref target="lien"/>text</item>
</list>
```

```
</div>
```

L'adresse Web

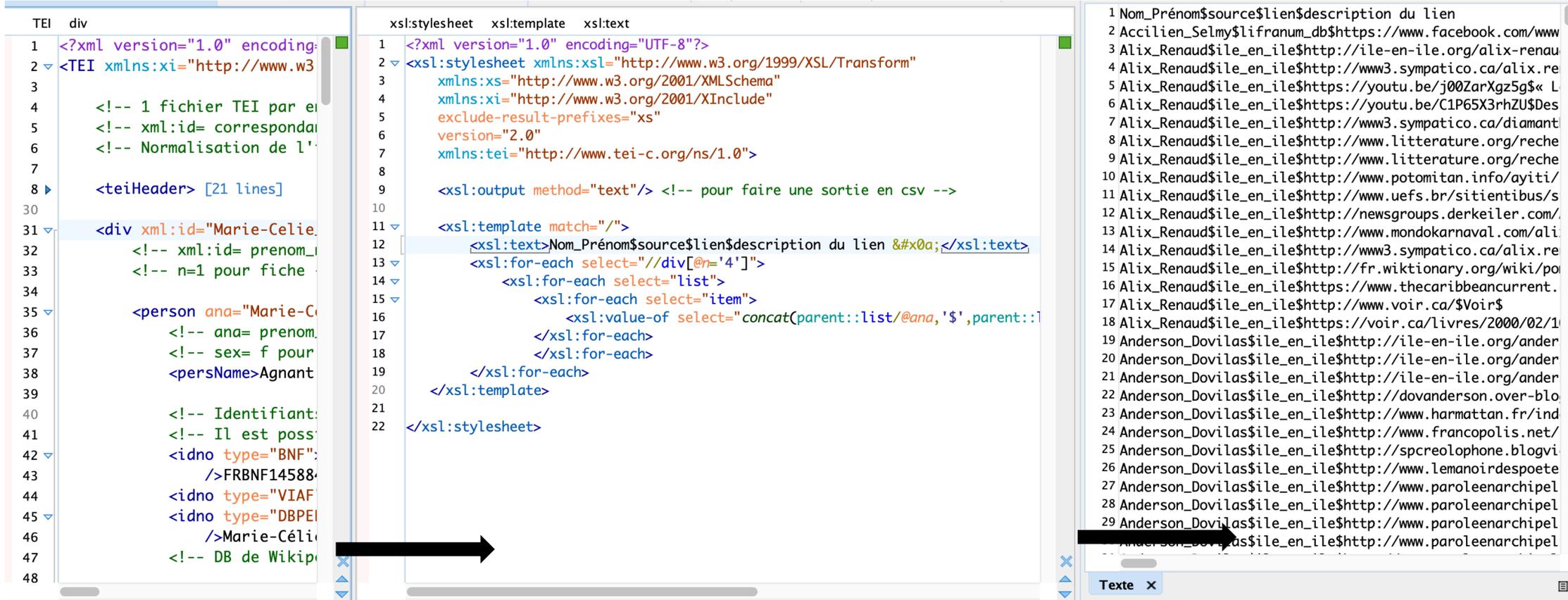
La source

Un sous-type

Description



Les feuilles de transformations XSL



```
TEI div
1 <?xml version="1.0" encoding="UTF-8" ?>
2 <TEI xmlns:xi="http://www.w3.org/2001/XInclude" ?>
3
4 <!-- 1 fichier TEI par ensemble
5 <!-- xml:id= correspondance
6 <!-- Normalisation de l'ensemble
7
8 <teiHeader> [21 lines]
9
10
11 <div xml:id="Marie-Celie" ?>
12 <!-- xml:id= prenom_1
13 <!-- n=1 pour fiche
14
15 <person ana="Marie-Celie" ?>
16 <!-- ana= prenom_1
17 <!-- sex= f pour femme
18 <persName>Agnant
19
20 <!-- Identifiant:
21 <!-- Il est possible de
22 <idno type="BNF" value="FRBNF14588" ?>
23 <idno type="VIAF" value="14588" ?>
24 <idno type="DBPEL" value="Marie-Celie" ?>
25 <!-- DB de Wikipedia
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

xsl:stylesheet xsl:template xsl:text
1 <?xml version="1.0" encoding="UTF-8" ?>
2 <xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" ?>
3   xmlns:xs="http://www.w3.org/2001/XMLSchema"
4   xmlns:xi="http://www.w3.org/2001/XInclude"
5   exclude-result-prefixes="xs"
6   version="2.0"
7   xmlns:tei="http://www.tei-c.org/ns/1.0" ?>
8
9   <xsl:output method="text" /> <!-- pour faire une sortie en csv -->
10
11 <xsl:template match="/" ?>
12   <xsl:text>Nom_Prénom$source$lien$description du lien &#x0a;</xsl:text>
13   <xsl:for-each select="//div[@n='4']" ?>
14     <xsl:for-each select="list" ?>
15       <xsl:for-each select="item" ?>
16         <xsl:value-of select="concat(parent::list/@ana,'$',parent::item)" />
17       </xsl:for-each>
18     </xsl:for-each>
19   </xsl:template>
20 </xsl:stylesheet>

1 Nom_Prénom$source$lien$description du lien
2 Accilien_Selmy$lifranum_db$http://www.facebook.com/www
3 Alix_Renaud$ile_en_ile$http://ile-en-ile.org/alix-renau
4 Alix_Renaud$ile_en_ile$http://www3.sympatico.ca/alix.re
5 Alix_Renaud$ile_en_ile$http://youtu.be/j00ZarXgz5g$« L
6 Alix_Renaud$ile_en_ile$http://youtu.be/C1P65X3rhZU$Des
7 Alix_Renaud$ile_en_ile$http://www3.sympatico.ca/diamant
8 Alix_Renaud$ile_en_ile$http://www.litterature.org/reche
9 Alix_Renaud$ile_en_ile$http://www.litterature.org/reche
10 Alix_Renaud$ile_en_ile$http://www.potomitan.info/ayiti/
11 Alix_Renaud$ile_en_ile$http://www.uefs.br/sitientibus/s
12 Alix_Renaud$ile_en_ile$http://newsgroups.derkeiler.com/
13 Alix_Renaud$ile_en_ile$http://www.mondokarnaval.com/ali
14 Alix_Renaud$ile_en_ile$http://www3.sympatico.ca/alix.re
15 Alix_Renaud$ile_en_ile$http://fr.wiktionary.org/wiki/po
16 Alix_Renaud$ile_en_ile$http://www.thecaribbeancurrent.
17 Alix_Renaud$ile_en_ile$http://www.voir.ca/$Voir$
18 Alix_Renaud$ile_en_ile$http://voir.ca/livres/2000/02/1
19 Anderson_Dovilas$ile_en_ile$http://ile-en-ile.org/ander
20 Anderson_Dovilas$ile_en_ile$http://ile-en-ile.org/ander
21 Anderson_Dovilas$ile_en_ile$http://ile-en-ile.org/ander
22 Anderson_Dovilas$ile_en_ile$http://dovanderson.over-blo
23 Anderson_Dovilas$ile_en_ile$http://www.harmattan.fr/ind
24 Anderson_Dovilas$ile_en_ile$http://www.francopolis.net/
25 Anderson_Dovilas$ile_en_ile$http://spcreolophone.blogvi
26 Anderson_Dovilas$ile_en_ile$http://www.lemanoirdespoete
27 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
28 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
29 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
30 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
31 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
32 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
33 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
34 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
35 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
36 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
37 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
38 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
39 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
40 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
41 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
42 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
43 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
44 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
45 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
46 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
47 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
48 Anderson_Dovilas$ile_en_ile$http://www.paroleenarchipel
```

1. Corpus XML + 2. Le fichier XSL = 3. Un résultat (.xml, .html, .txt, etc)

Les feuilles de transformations XSL

Exemple : création d'un tableau avec *entité auteur/ lien/ source du lien/ description du lien*

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:xi="http://www.w3.org/2001/XInclude"
  exclude-result-prefixes="xs"
  version="2.0"
  xmlns:tei="http://www.tei-c.org/ns/1.0">
```

→ Une sortie .txt

```
<xsl:output method="text"/> <!-- pour faire une sortie en csv -->
```

```
<xsl:template match="/">
  <xsl:text>Nom_Prénom$source$lien$description du lien &#x0a;</xsl:text>
```

→ Création des noms de colonnes

```
  <xsl:for-each select="//div[@n='4']">
    <xsl:for-each select="list">
      <xsl:for-each select="item">
        <xsl:value-of select="concat(parent::list/@ana,'$',parent::list/@source,'$',ref/@target,'$',..,'$', '&#xA;')"/>
      </xsl:for-each>
    </xsl:for-each>
  </xsl:for-each>
```

Remplissage des lignes du tableau

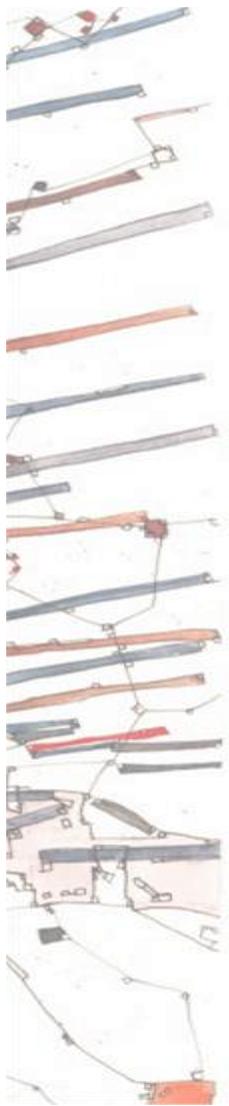
```
</xsl:template>
```

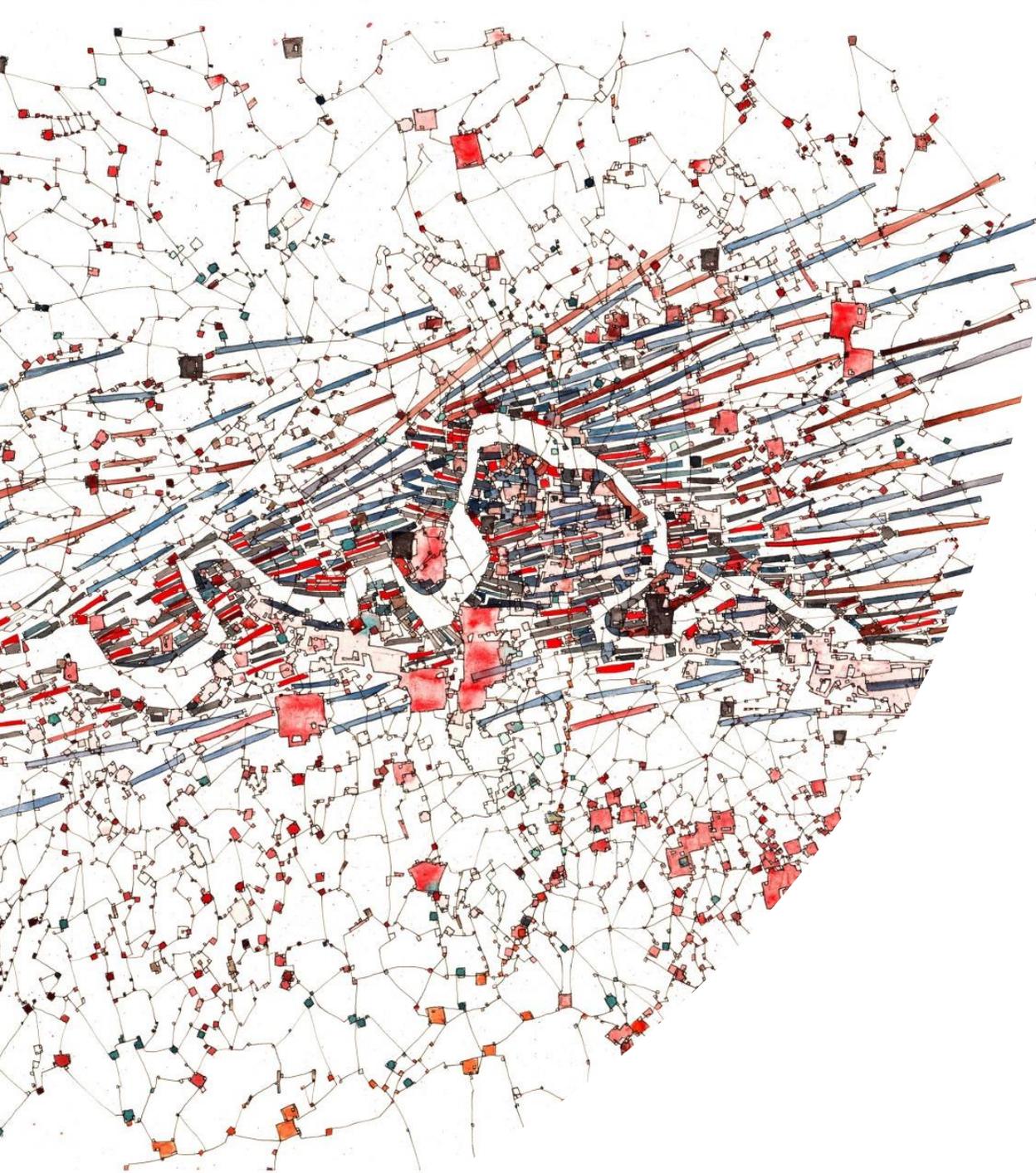
```
</xsl:stylesheet>
```



Pour obtenir :

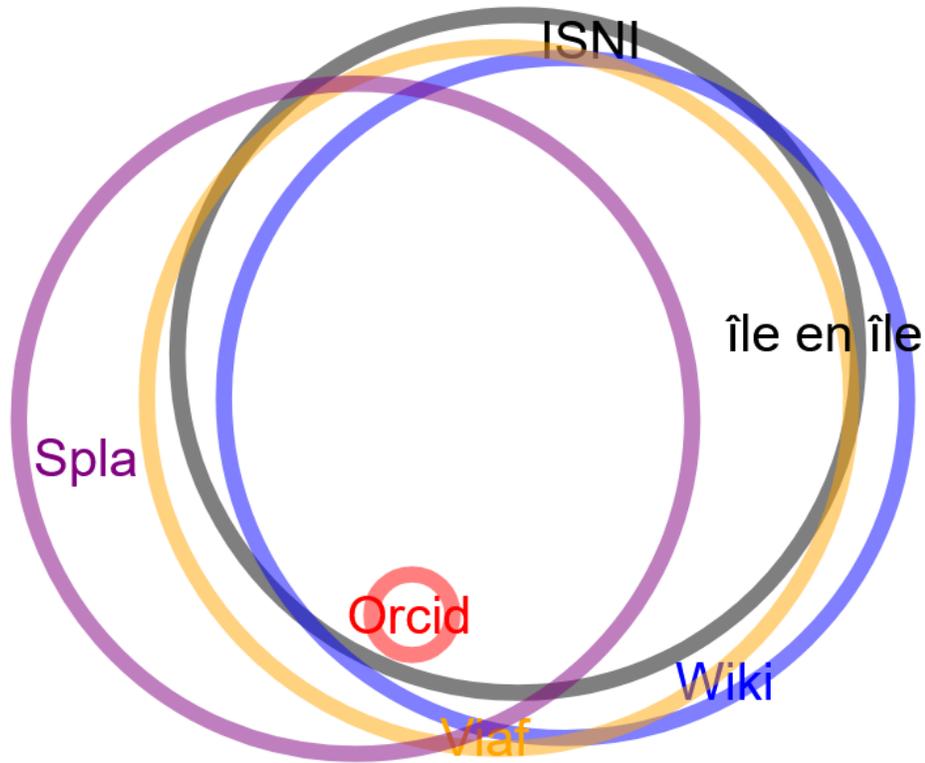
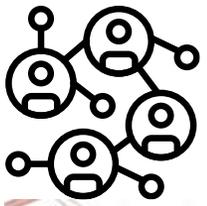
	A	B	C	D
1	Nom_Prénom	source	lien	description du lien
2	Accilien_Selmy	lifranum_db	https://www.facebook.com/www.revedemuse.livre/	Selmy Accilien poète
3	Alix_Renaud	ile_en_ile	http://ile-en-ile.org/alix-renaud-poemes/	Ménage à trois
4	Alix_Renaud	ile_en_ile	http://www3.sympatico.ca/alix.renaud/	Alix Renaud, ou le plaisir du texte
5	Alix_Renaud	ile_en_ile	https://youtu.be/j00ZarXgz5g	« Le Mot et la chose », poème de Lattaignant
6	Alix_Renaud	ile_en_ile	https://youtu.be/C1P65X3rhZU	Des genoux au vertige
7	Alix_Renaud	ile_en_ile	http://www3.sympatico.ca/diamantbleu/accueil.html	À la recherche du Petit Prince
8	Alix_Renaud	ile_en_ile	http://www.litterature.org/recherche/ecrivains/renaud-alix-389/	Présentation sur L'île
9	Alix_Renaud	ile_en_ile	http://www.litterature.org/recherche/ecrivains/renaud-alix-389/	Présentation sur L'île
10	Alix_Renaud	ile_en_ile	http://www.potomitan.info/ayiti/renaud.php	Présentation sur Potomitan
11	Alix_Renaud	ile_en_ile	http://www.uefs.br/sitientibus/sitientibus_9/entrevista.pdf	Entrevista
12	Alix_Renaud	ile_en_ile	http://newsgroups.derkeiler.com/Archive/Soc/soc.culture.haiti/2005-11/msg00159.html	Le don de jouer avec les mots
13	Alix_Renaud	ile_en_ile	http://www.mondokarnaval.com/alix_chronique_2016.html	Les Chroniques d'Alix Renaud
14	Alix_Renaud	ile_en_ile	http://www3.sympatico.ca/alix.renaud/grac_esp.html	Ofrenda de Gracias
15	Alix_Renaud	ile_en_ile	http://fr.wiktionary.org/wiki/pompion	Pompion
16	Alix_Renaud	ile_en_ile	https://www.thecaribbeancurrent.com/quebec-awards-alix-renaud-for-excellence-in-arts-and-culture/	Quebec Awards Alix Renaud for Excellence in Arts and Culture
17	Alix_Renaud	ile_en_ile	http://www.voir.ca/	Voir
18	Alix_Renaud	ile_en_ile	https://voir.ca/livres/2000/02/10/alix-renaud-ovation/	Ovation
19	Anderson_Dovilas	ile_en_ile	http://ile-en-ile.org/anderson-dovilas-pwezi-pefomans/	Anderson Dovilas, Pwezi pèfòmans
20	Anderson_Dovilas	ile_en_ile	http://ile-en-ile.org/anderson-dovilas-la-nuit-nest-pas-le-premier-secret-des-baisers-bleus/	La nuit n'est pas le premier secret des baisers bleus
21	Anderson_Dovilas	ile_en_ile	http://ile-en-ile.org/anderson-dovilas-trois-poemes/	Trois poèmes
22	Anderson_Dovilas	ile_en_ile	http://dovanderson.over-blog.com/	Le blog de Dovilas Anderson
23	Anderson_Dovilas	ile_en_ile	http://www.harmattan.fr/index.asp?navig=catalogue&obj=livre&no=43605	Mémoire d'Outre-Monde
24	Anderson_Dovilas	ile_en_ile	http://www.francopolis.net/revues/Dovilas-mai2012.html	Notes sur « Vingt poèmes pour traverser la nuit »
25	Anderson_Dovilas	ile_en_ile	http://spcreolophone.blogvie.com/	Sosyete Powèt Kreyolofòn
26	Anderson_Dovilas	ile_en_ile	http://www.lemanoirdespoetes.fr/poemes-anderson-dovilas.php	« Je vole », « Je voyage » et « Point »
27	Anderson_Dovilas	ile_en_ile	http://www.paroleenarchipel.org/	Parole en Archipel
28	Anderson_Dovilas	ile_en_ile	http://www.paroleenarchipel.com/article-vingt-poemes-pour-traverser-la-nuit-104458858.html	Vingt poèmes pour traverser la nuit
29	Anderson_Dovilas	ile_en_ile	http://www.paroleenarchipel.com/article-en-hollande-la-france-devient-fran-oise-104871758.htmr	En Hollande, la France devient Française
30	Anderson_Dovilas	ile_en_ile	http://www.paroleenarchipel.com/article-les-iles-en-accent-aigu-105328874.html	Les îles en accent aigu
31	Anderson_Dovilas	ile_en_ile	http://www.paroleenarchipel.com/article-le-sang-de-l-oublie-le-sang-de-l-avenir-105976958.html	Le sang de l'oubli, le sang de l'avenir
32	Anderson_Dovilas	ile_en_ile	http://www.paroleenarchipel.com/article-liminasyon-parcours-d-un-voyageur-106118692.html	Liminasyon: Parcours d'un voyageur
33	Anderson_Dovilas	ile_en_ile	http://issuu.com/pallaikaroly/docs/vents_aliz_s_-_komansman?mode=window&backgroundcolor:	Trois poèmes
34	Anderson_Dovilas	ile_en_ile	http://www.potomitan.info/ayiti/dovilas.php	Pour en finir avec dérogation littéraire en Haïti



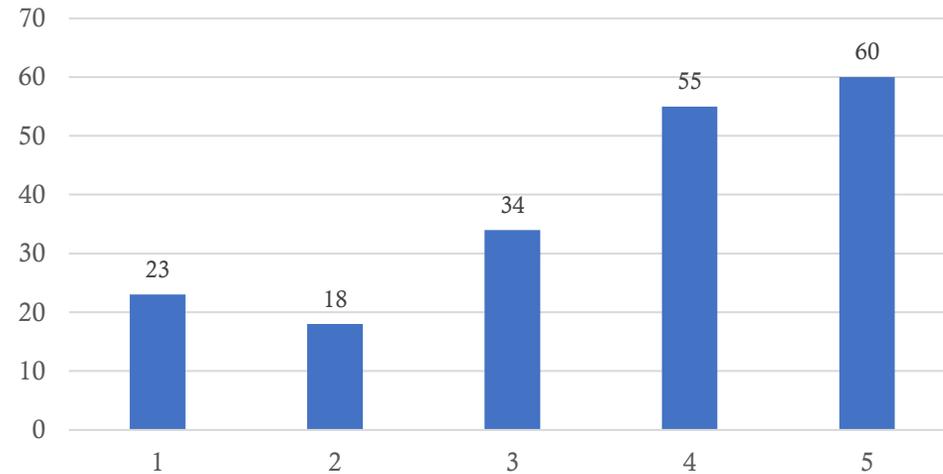


IV. Exploitation des données

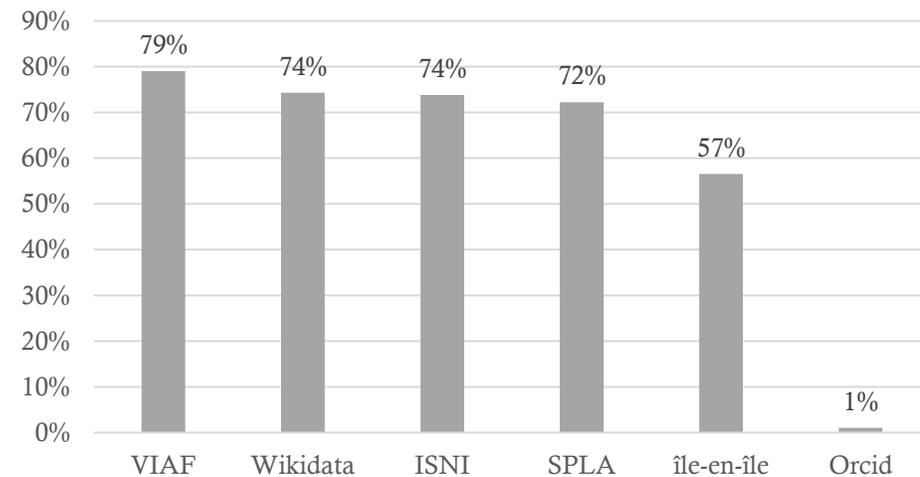
Le spectre couvert par les répertoires d'autorités



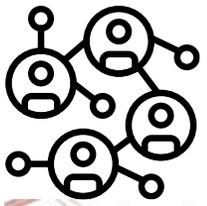
Nombre d'identifiants possédés par entité auteur



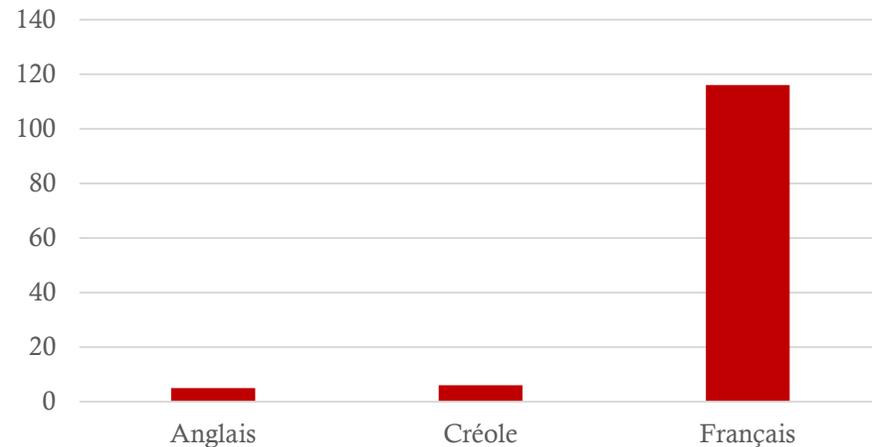
Taux de complétude des références d'autorités



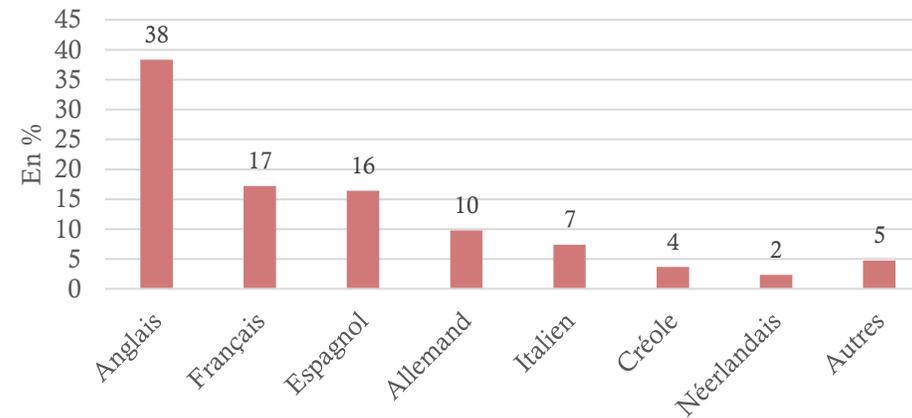
Haïti : une littérature francophone ?



Mentions de langues écrites, parlées ou signées indiquées dans les notices

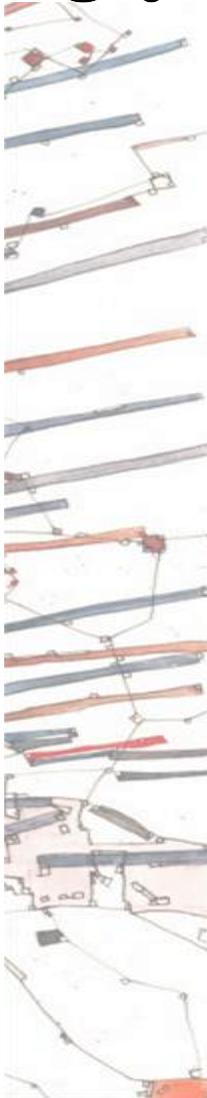


Langue de traduction des ouvrages des entités auteurs présents sur *île-en-île*

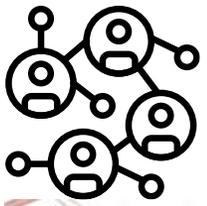


Langues de traduction des ouvrages des entités auteurs

- Biais des répertoires d'autorités et/ou des articles Wikipédia qui ne renseignent que la langue d'écriture de la notice
- Synthèse bibliographique : omission d'ouvrages.

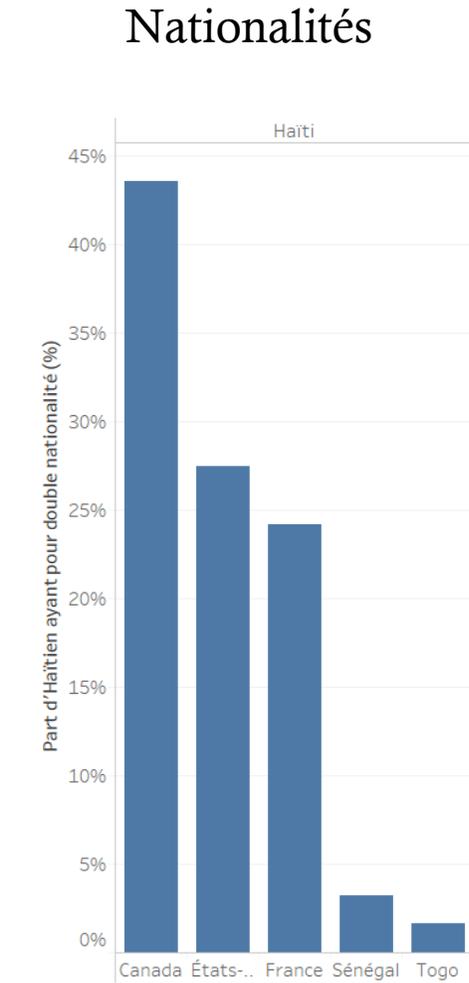


Haiti : un réseau international

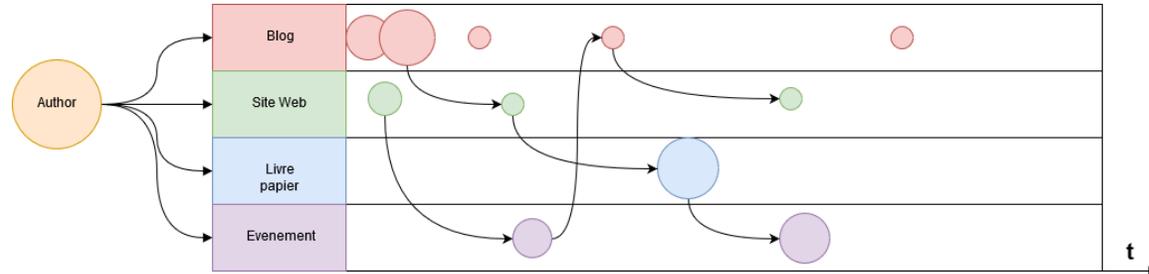
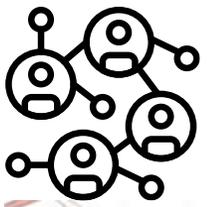


© 2021 Mapbox © OpenStreetMap

Pays Deces	NA	Canada	Comoros	Cuba	Guadelou..	Haïti	Mayotte	Suisse
NA	58	1	2	1	1	98	1	2
Canada						7		
États-Unis						6		
France						4		
Haïti						7		
Sénégal						2		



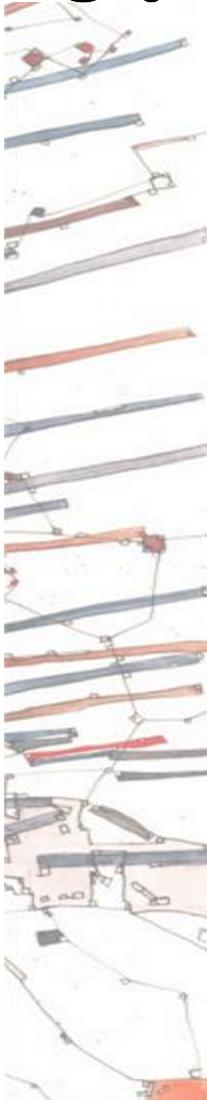
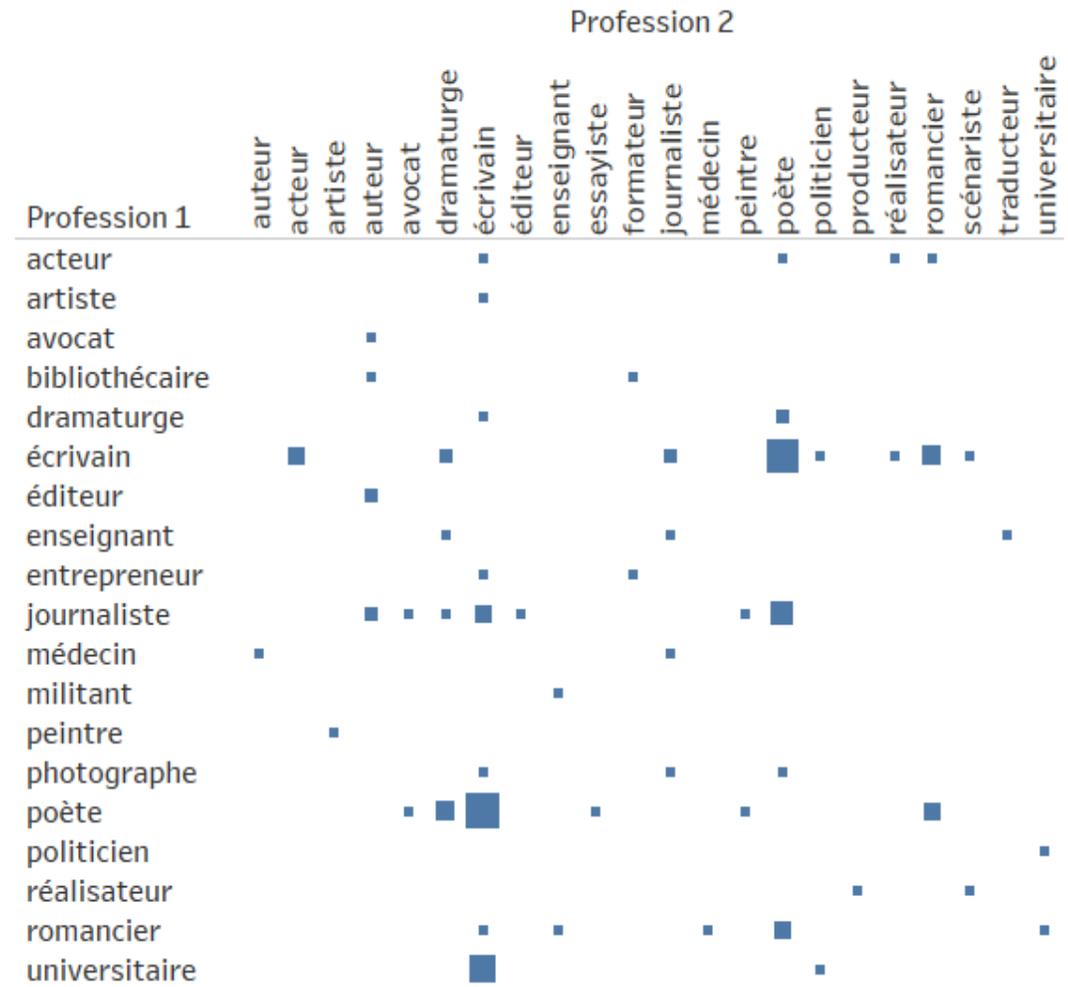
Des professions aux « parcours de vie »



- Un individu peut avoir exercé plusieurs professions au cours de sa vie.
- L'extraction ne nous permet pas d'adopter une approche diachronique des entités auteurs et ne peut qu'en rendre compte quantitativement.

Quelques faits intéressants toutefois :

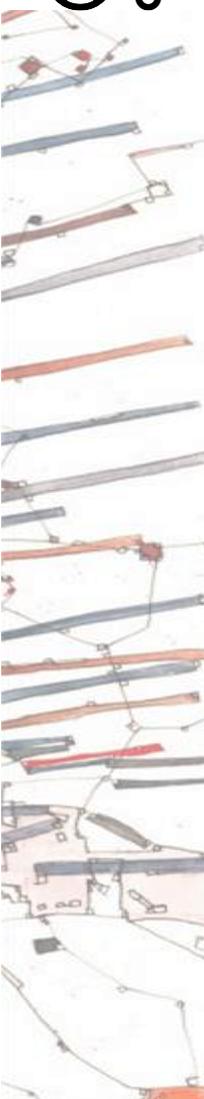
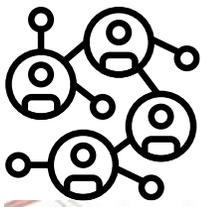
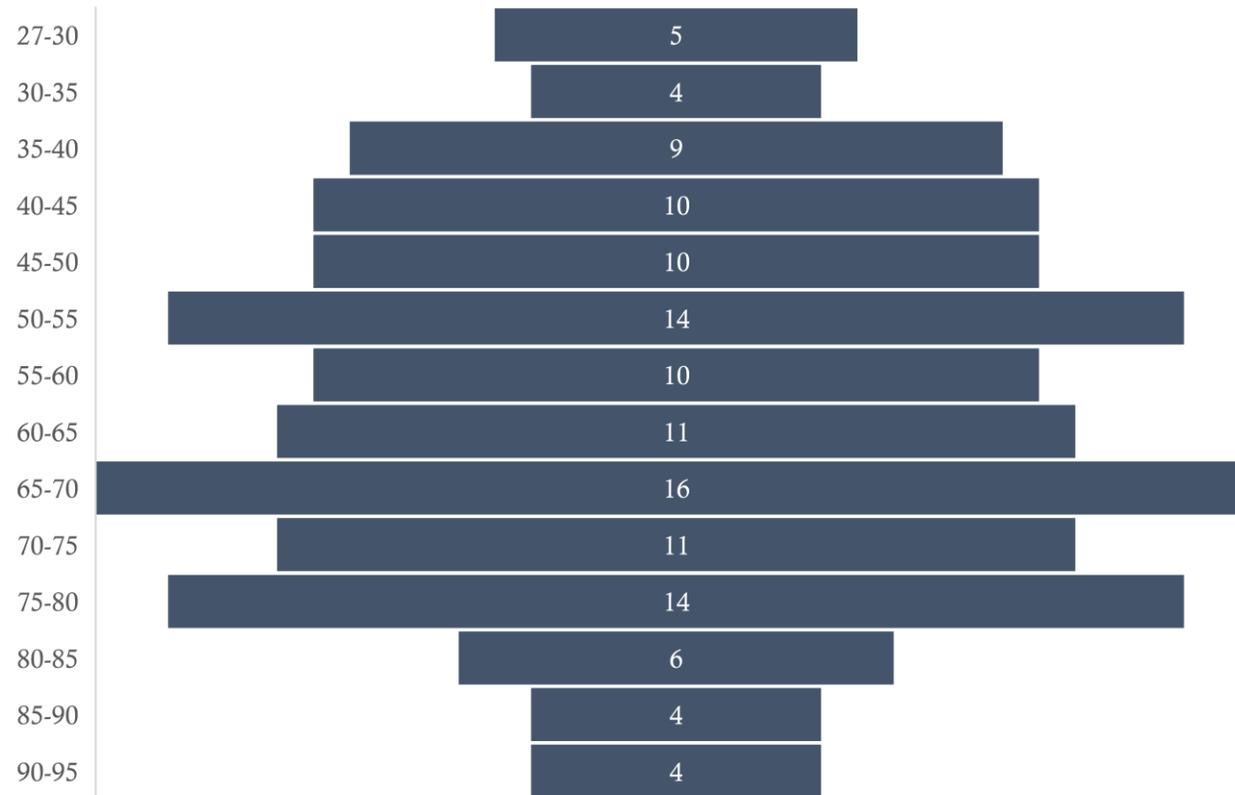
- « auteur » n'apparaît qu'en seconde profession.
- « écrivain » et « poète » sont deux notions très liées.
- Les métiers de la presse, de l'art et de l'enseignement sont des professions pour lesquels l'activité littéraire est secondaire.

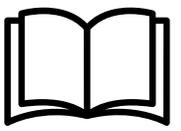


Un corpus d'acteurs « institutionnalisés »

128 dates de naissance
156 auteurs non décédés

Pyramide des âges des acteurs littéraires



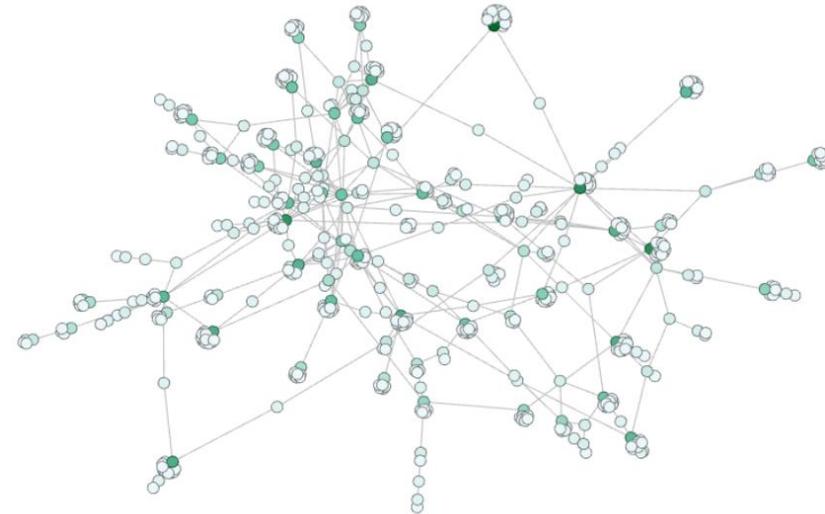


Vers l'identification de nouveaux auteurs : le co-autorat

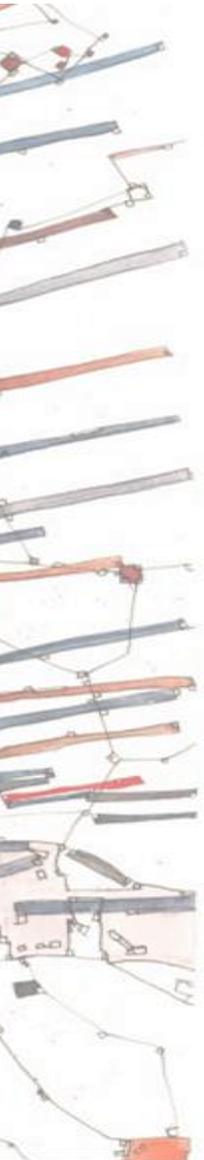
- Analyse textuelle des notices biographiques
- Analyses en réseaux des liens de collaborations identifiés au sein des notices biographiques et des répertoires d'autorités

Réseau des liens de collaborations obtenu sur Gephi

```
<listRelation>
  <relation ana="Tapia_Veronica" source="VIAF">Tapia Veronica</relation>
  <relation ana="Saint_Eloi_Rodney" source="VIAF">Saint-Éloi Rodney</relation>
  <relation ana="Zekina_Daniela" source="VIAF">Zekina Daniela</relation>
  <relation ana="Tiga" source="VIAF">Tiga</relation>
  <relation ana="Pean_Stanley" source="VIAF">Péan Stanley</relation>
  <relation ana="Cousineau_Normand" source="VIAF">Cousineau Normand</relation>
  <relation ana="Mata_Barreiro_Carmen" source="VIAF">Mata Barreiro Carmen</relation>
  <relation ana="Barroux" source="VIAF">Barroux</relation>
  <relation ana="Wiener_Jacqueline" source="VIAF">Wiener Jacqueline</relation>
  <relation ana="Ellis_Zilpha" source="VIAF">Ellis Zilpha</relation>
  <relation ana="Célie_Agnant" source="ile_en_ile_auto">Célie Agnant</relation>
  <relation ana="Alexis_D_Haiti" source="ile_en_ile_auto">Alexis d'Haïti</relation>
</listRelation>
```



- Relations de co-autorat et de liens de collaborations.
 - Près de 326 entités nommées ont été extraites automatiquement depuis île-en-île, 506 du VIAF.
 - Sur ces 832 entités, on note 583 uniques soit 86% du corpus.
 - 65 apparaissent à deux reprises (10%) et 18 à trois reprises (3%).

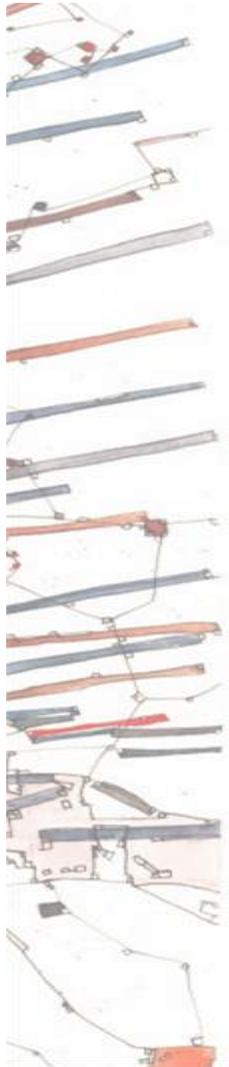


Les littéraires haïtiens et le Web



Nom de domaine	Effectifs
ile-en-ile.org	360
lenouvelliste.com	154
web.archive.org	51
youtube.com	35
potomitan.info	30
tanbou.com	18
csmsmagazine.org	16
lematinhaiti.com	15
hpnhaiti.com	13
alterpresse.org	12

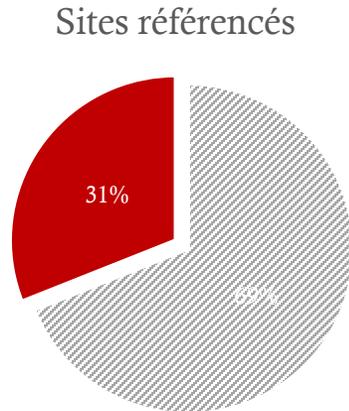
- La majorité des entités auteurs ne renvoient qu'à une à quatre adresses URLs (40 %).
 - 5 liens issus de la base de données RDF d'origine ;
 - 26 liens issus du Spla et 1258 d'*île-en-île*.
 - Les noms de domaines uniques et sites personnels représentent près de 73 % des liens.
- Noms de domaines récurrents
 - Les renvois vers *île-en-île* sont naturellement à écarter (360).
 - *Lenouvelliste* (154) est quant à lui le seul quotidien publié à Port-au-Prince.
 - Site d'archives en ligne (51) et la plateforme Youtube (35).
 - *potomitan* (30) et *tanbou* (18), des sites de promotion des cultures et langues créoles.



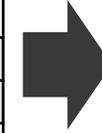
Pages personnelles issues de Spla - Un voyage dans le temps



Sur 9 références, 6 peuvent être observés :

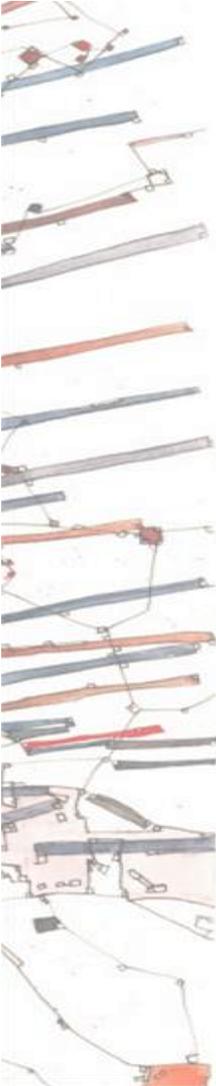


Site Web	Période d'activité
http://josaphat-robert-large	[2004;2014]
http://www.franketienne.com	[2005;2008]
http://www.elifleurant.com	[2013;2017]
http://www.margaretmitchellarmand.com/	[2013;2018]
http://www.elsieaugustave.com/	[2013;2017]
http://jmlominy.com/	[2013;2017]



INTERNET ARCHIVE
WayBackMachine

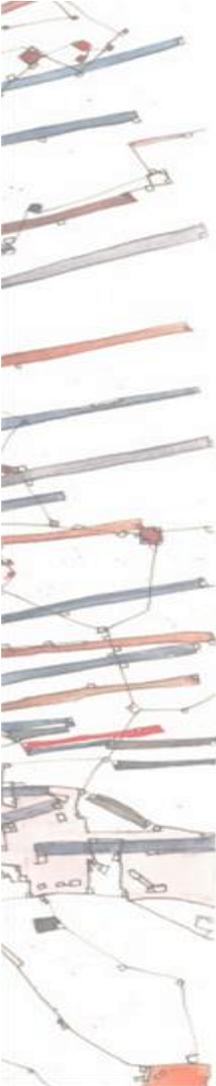
- Clôture des pages liée à l'âge ?
 - décès de certains *auteurs*
 - d'autres actuellement très âgés (plus de 70 ans)
 - mais aussi des jeunes qui ont cessé d'alimenter leur sites afin de se consacrer à d'autres activités.
- Pas seulement : les pages peuvent également disparaître pour des raisons techniques (obsolescences des plugins utilisés non-conformes avec une nouvelle mise à jour).
- Sites très descriptifs.
- « billets d'humeurs », mais ne diffusant pas de contenu littéraire inédit.



18 pages personnelles toujours en activité



- Les *auteurs* ayant des sites renseignés ont entre 35 et 70 ans (soit un âge moyen de 54 ans)
- 3 sites été intégralement rédigés en anglais, 5 en français, 8 sont plurilingues (haïtien + au moins anglais et français)
- 8 ont un format blog, où il est possible « d'échanger » → peu d'échanges à travers ce biais
- Nombre égal de pages promotionnelles et de blogs littéraires
- La moitié sont des sites n'ayant pas été alimentés depuis au moins un an (2013-2017)



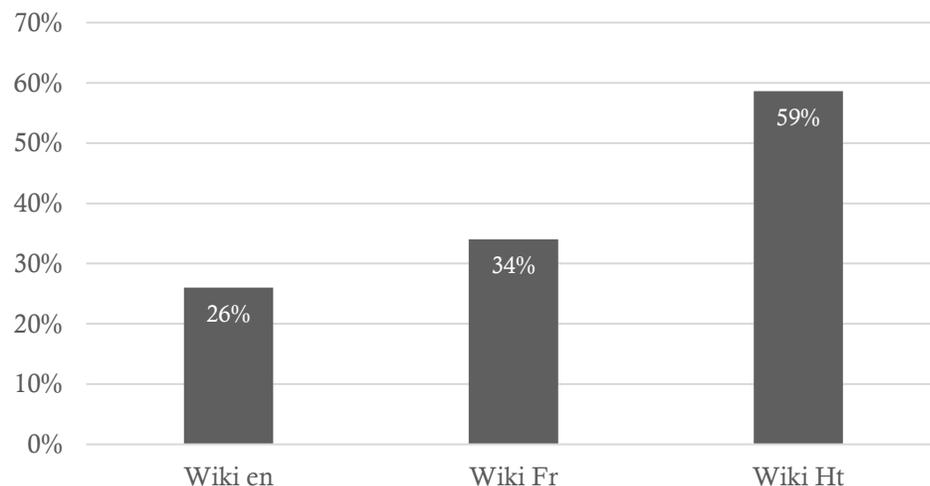
Des entités Wikidata aux pages Wikipédia



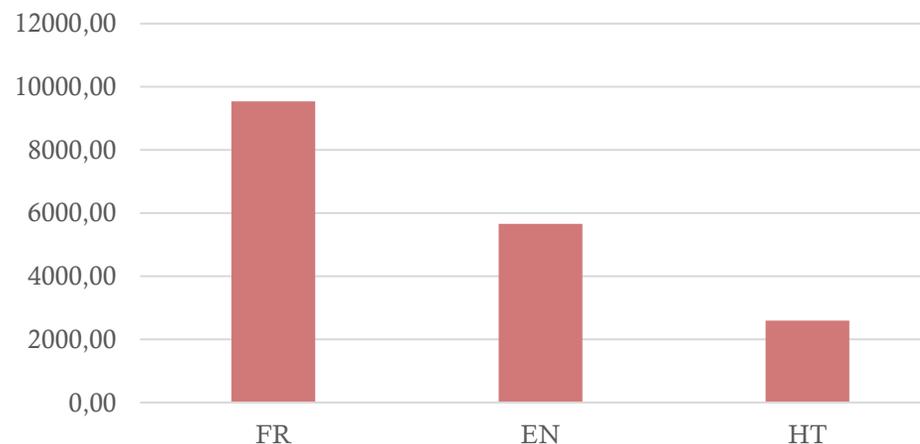
WIKIPEDIA

- 142 entités possèdent un identifiant Wikidata
- 124 au moins une page Wikipédia. Si ce n'est pas le cas, l'entité Wikidata a été identifiée dans une base de données LOD (*Linked Open Data*).

Part d'entités disposant de pages Wikipédia



Longueur moyenne des articles Wikipédia (nombre de caractères)



Wikipédia, un espace à géométrie variable



WIKIPEDIA

Des contributeurs différents par langue :

88% des articles haïtiens ont été créés par le même contributeur

→ l'administrateur de Wikipédia Haïti

4 contributeurs ont écrit 30% des pages en français

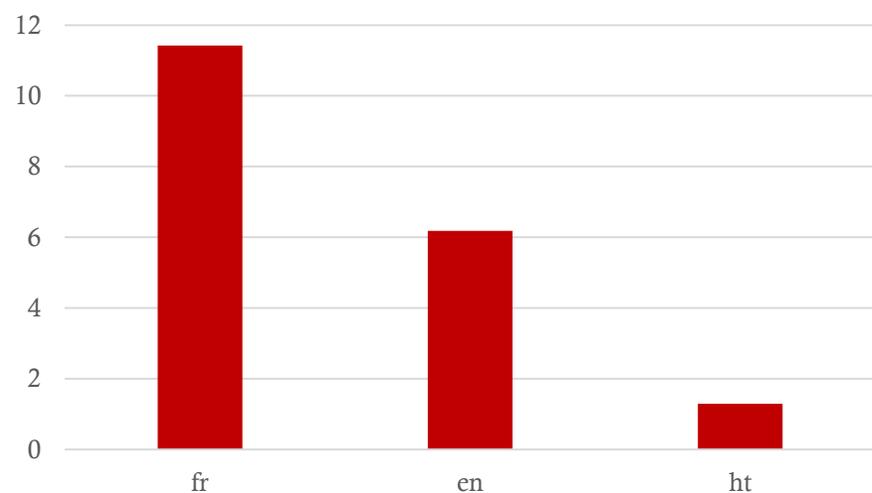
→ visibilité de la francophonie : éditeur belges et québécois

→ souvent des amateurs d'art, de littérature et des universitaires

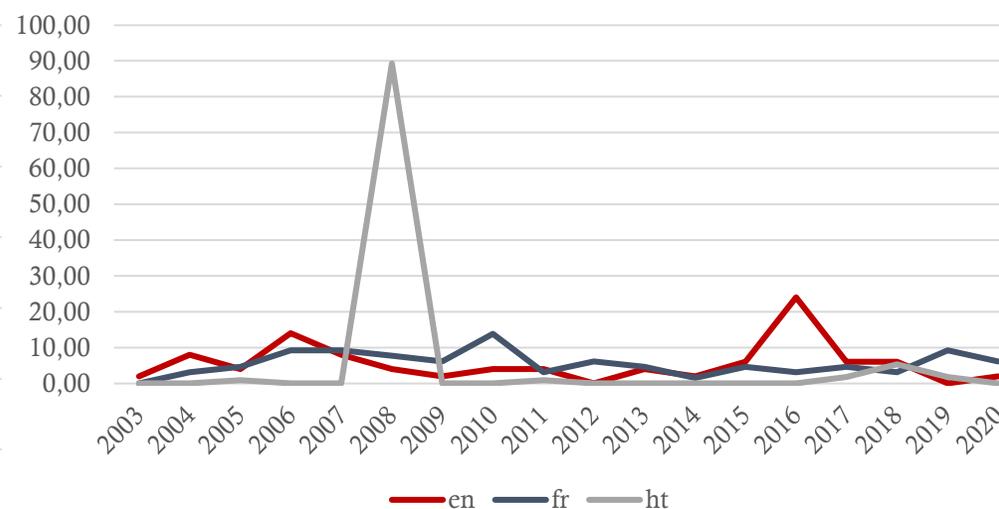
2 contributeurs sont à l'origine de 30% des pages en anglais :

→ contributeurs canadiens et étudiants haïtiens.

Mises à jour annuelles (moy./page)



Répartition de la part d'articles créés par an

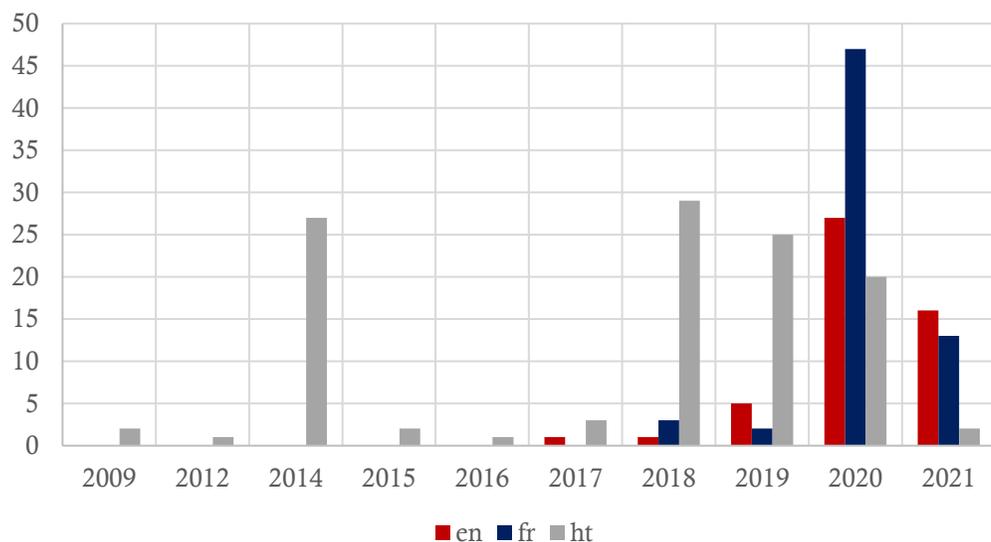




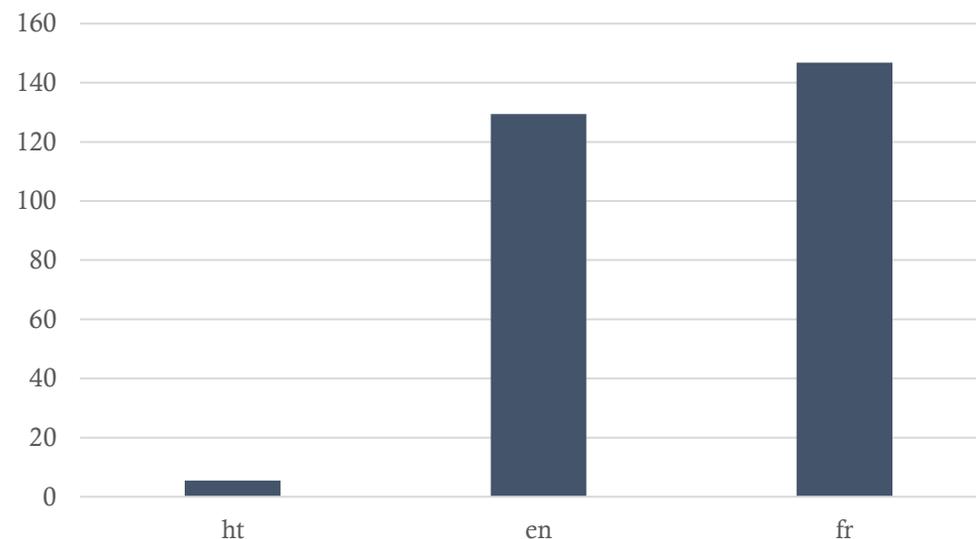
WIKIPEDIA

- Concernant le référencement dans Wikipédia :
 - Aucun article haïtien n'est pointé par d'autres articles
 - Le Wikipédia français réfère plus amplement les articles liés aux entités de notre corpus
- Vis-à-vis du nombre de vues et du nombre de mises à jour, les pages en français sont les plus dynamiques, suivies des pages anglaises, enfin haïtiennes.
- Les articles datent en moyenne d'il y a 9 à 10 ans, seulement 5% n'ont pas été modifiés depuis.

Dernières dates de mise à jours des articles par années et langues



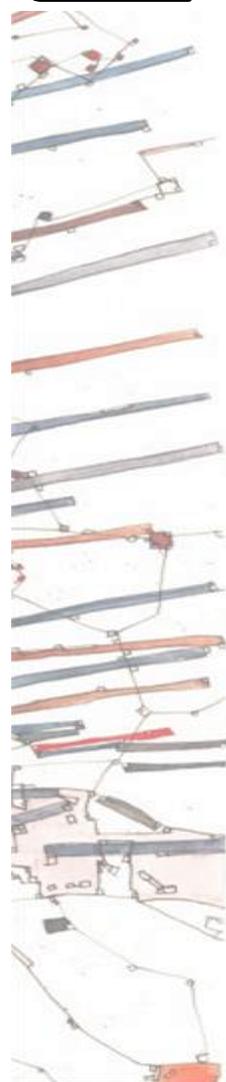
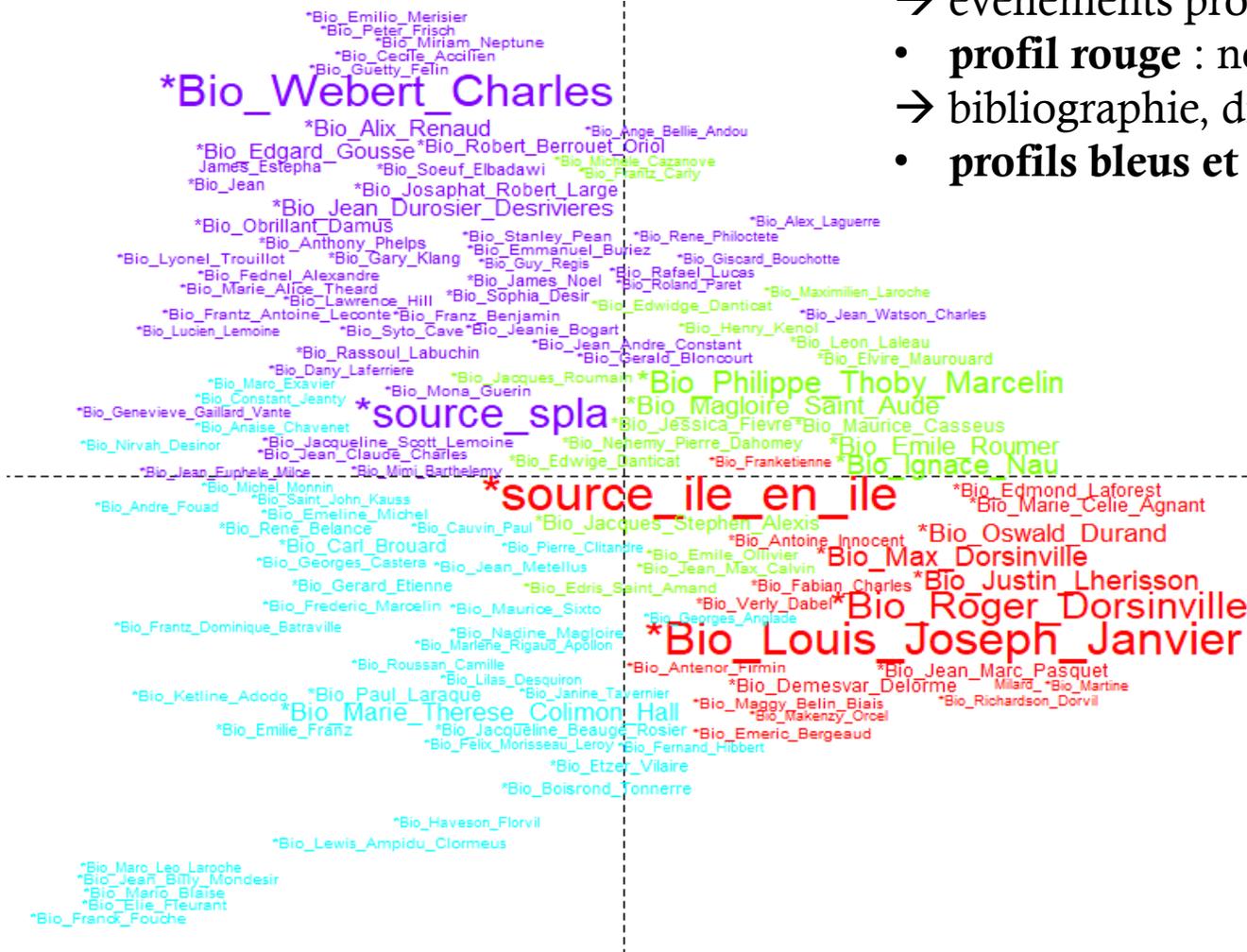
Nombre de vues moyen lors du dernier mois des pages par langues



Des profils variés



- **profil violet** : notices Spla, courtes.
→ événements professionnels mis en exergues.
- **profil rouge** : notices île-en-île, moyennes et longues.
→ bibliographie, distinctions, parcours académique, etc.
- **profils bleus et verts** se mélangent au sein des notices.



Mieux décrire les auteurs émergents

- Une littérature en train de se faire, peu normalisée...
- Un « crawl » est un flot de données qui risque sans indexation de ne pas permettre d'analyses
- Une multitude de canaux, une multitude de discours...
- Une notice biographique traduit un intérêt par les pairs. Elle contient par ailleurs de nombreuses informations textuelles telles que des dates de naissance et de décès.
- Extraire des routines discursives ?

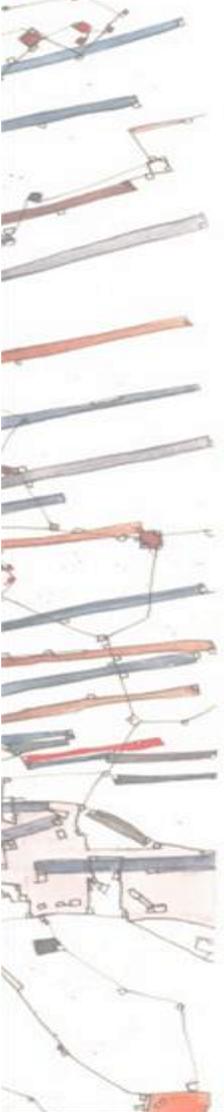
Machiner une variété de textes... pour induire des informations

Quels peuvent être les apports du TAL?

- Repérage d'entités nommées : lieux, noms propres
- Exploration contextuelle : à partir de la recherche des dates d'état civil au sein des notices biographiques



Obtenir des données biographiques...

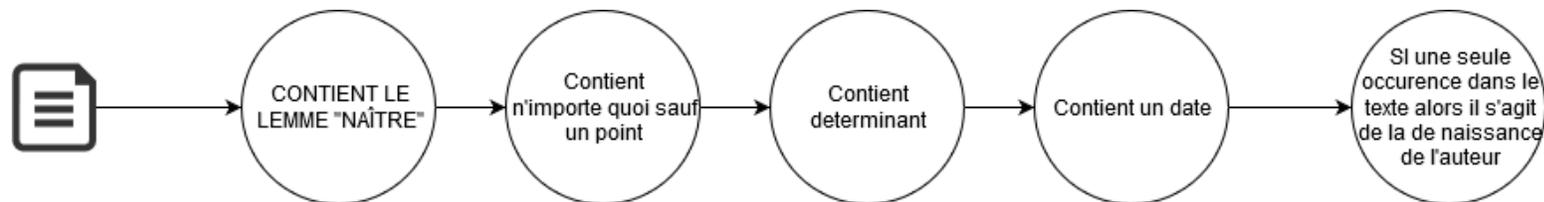


Utiliser les structures récurrentes induite par le Web pour en extraire des informations :

Cléante Desgraves naît le 13 janvier 1891 à Port-au-Prince (Haïti) d'une mère américaine, née Alice Cunningham, et d'un père haïtien, Hector Desgraves, qui était pharmacien et également pianiste.

Bonel Auguste, né à Port-au-Prince, Haïti, le 6 mars 1973, est poète, bibliothécaire et animateur culturel. Il a étudié la linguistique à l'Université d'Etat d'Haïti et a reçu une formation en histoire de l'art à l'Institut français d'Haïti. Fondateur de l'atelier de poésie « Dimanche en poésie » à la bibliothèque Étoile filante, et animateur des « Vendredis littéraires » de l'université Caraïbe, il écrit en français et en créole.

...



Lier les biographies sur des entités plus vastes...



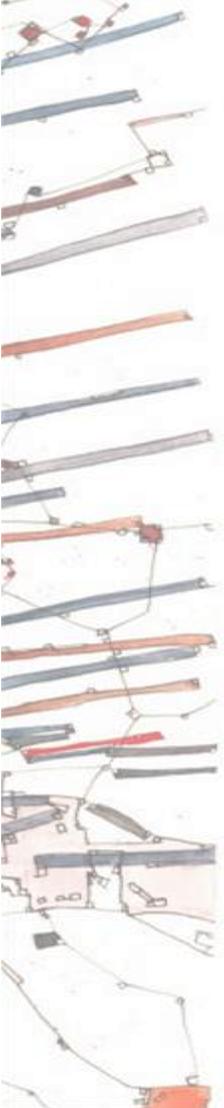
Utiliser des systèmes « entraînés » pour en détecter d'autres:

Bonel Auguste, né à **Port-au-Prince, Haïti**, le 6 mars 1973, est poète, bibliothécaire et animateur culturel. Il a étudié la linguistique à **l'Université d'État d'Haïti** et a reçu une formation en histoire de l'art à l'Institut français d'Haïti. Fondateur de l'atelier de poésie « **Dimanche en poésie** » à la bibliothèque Étoile filante, et animateur des « **Vendredis littéraires** » de **l'université Caraïbe**, il écrit en français et en créole.

Mais....

Né en Haïti en 1994, **Jean D'Amérique** est écrivain, auteur de *Petite fleur du ghetto*, recueil qui lui vaut une mention spéciale du **Prix René Philoctète** 2015 et une sélection au Prix Révélation de Poésie 2016 de la Société des Gens de Lettres.

**Beaucoup d'ambiguïtés,
Demande d'être validé et
mis en contexte....**



Un espace encore à explorer

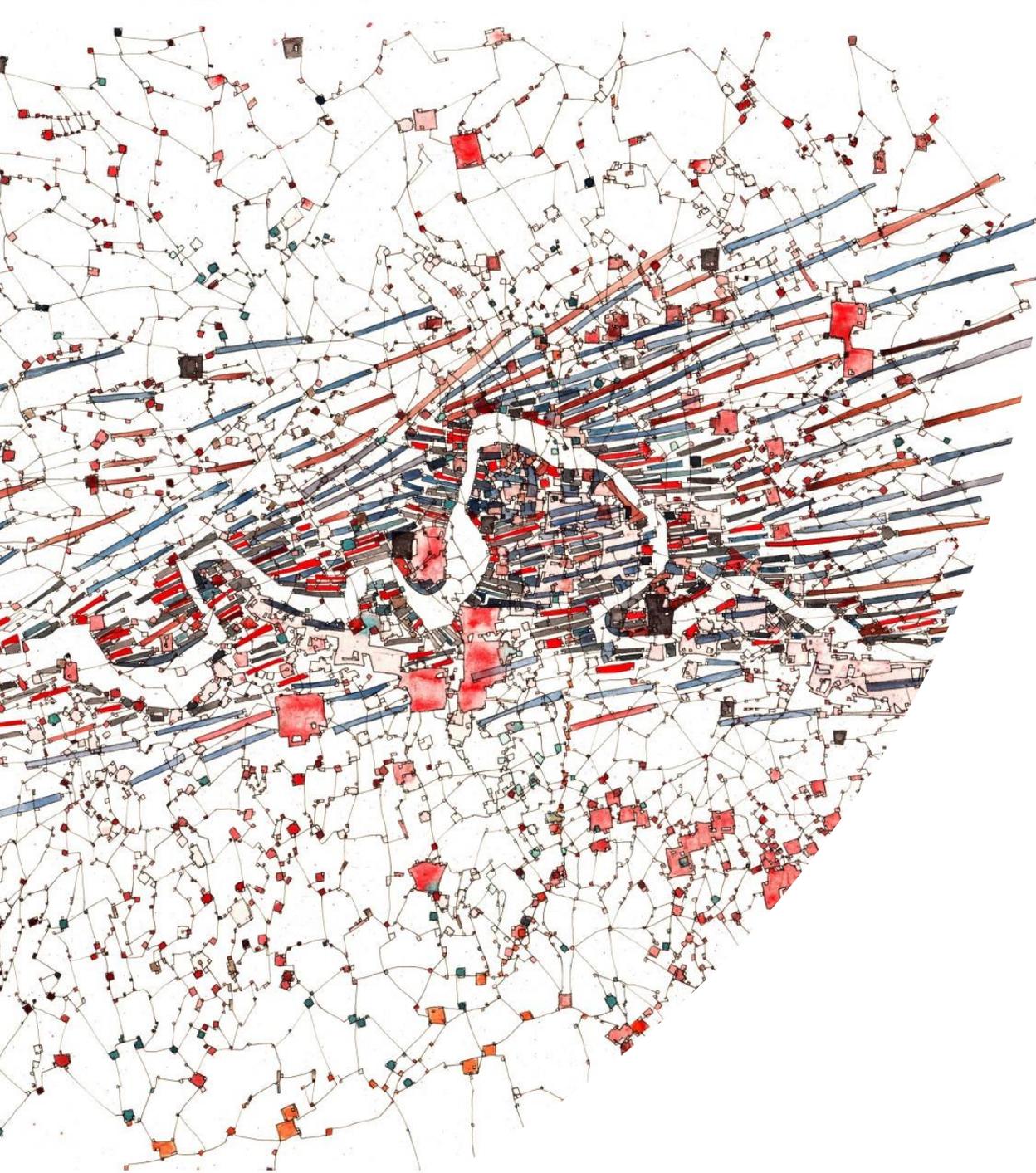
- La porte fermée de Facebook
- Twitter, un espace littéraire à 280 caractères ? [#covidpoème](#)
- Inkitt / Wattpad des plateformes littéraires sociales

C'est ici que notre travail s'arrête et que commence l'ANR LIFRANUM...



<https://projet-lifranum.univ-lyon3.fr/projet>

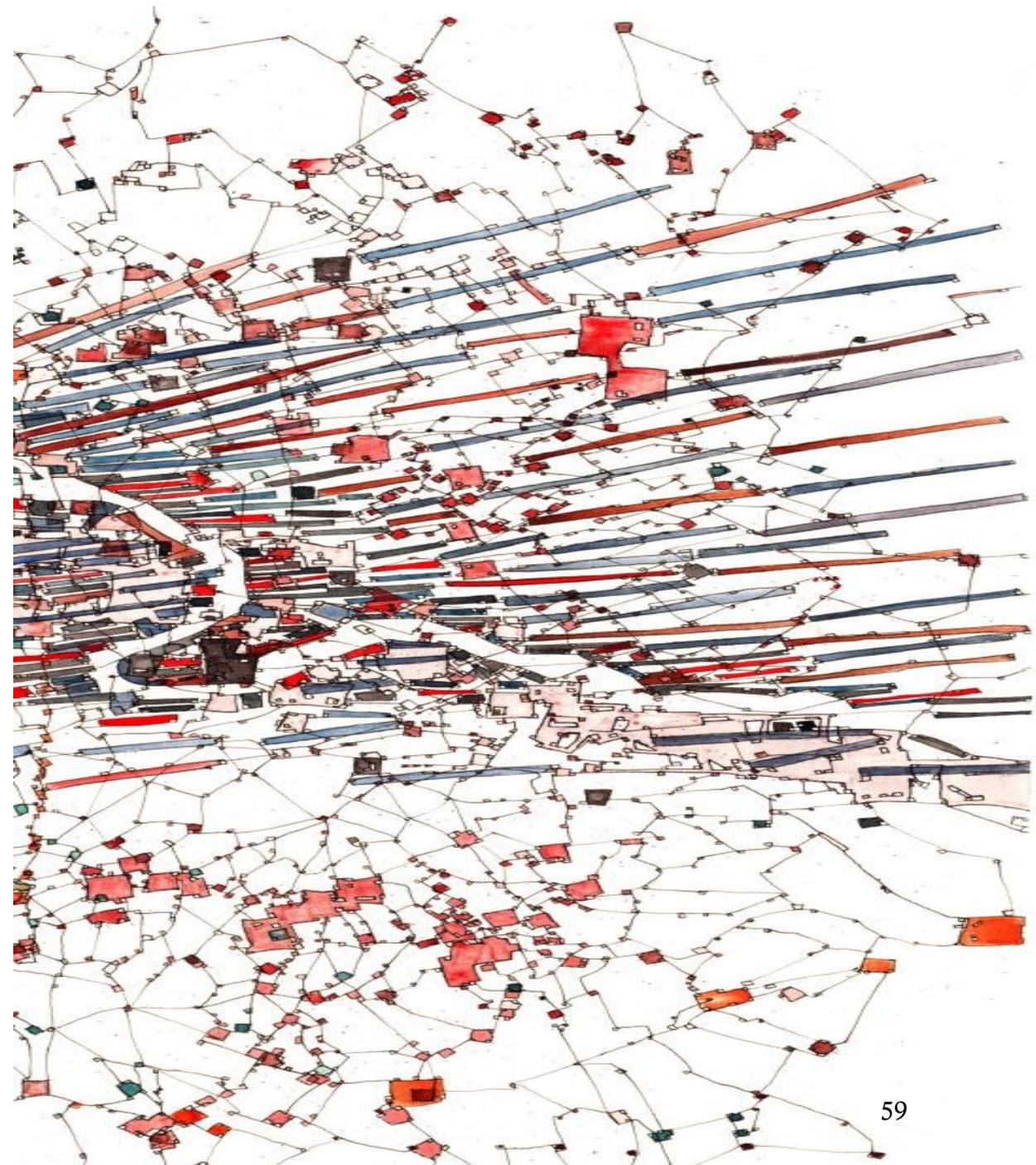


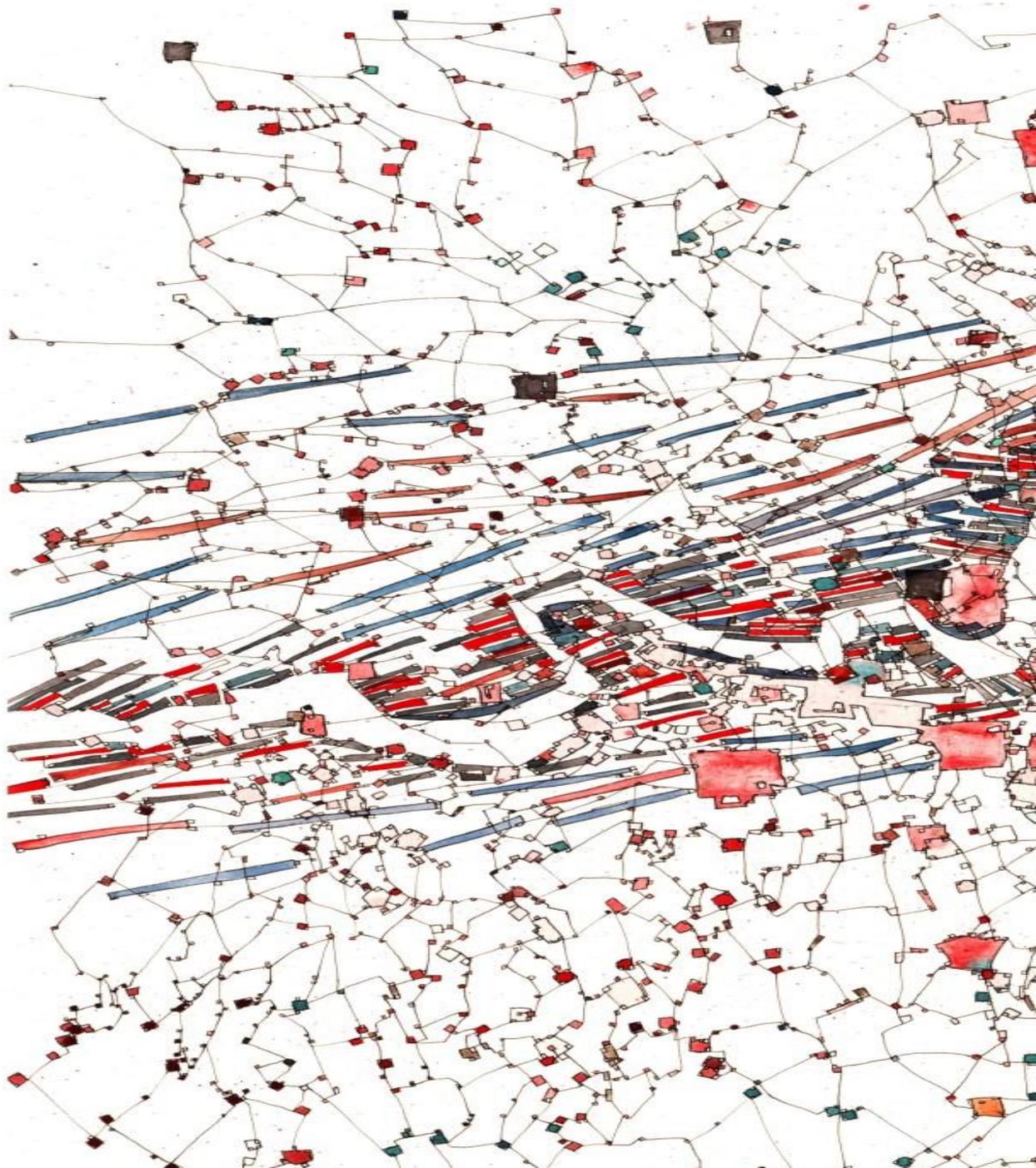


V. Conclusions

Le cas d'Haïti

- Repérage de certaines particularités du réseau : variétés linguistiques, problématiques d'émigration
- Les répertoires d'autorités permettent d'atteindre des individus reconnus par leurs pairs, qui ont souvent déjà publié un premier ouvrage par les canaux officiels.
- Les sites de référence regroupent quant à eux des individus aux profils variés mettant la littérature et les arts au service de la culture haïtienne.
- Les différents degrés de normalisation de ces deux types de sources s'est avérée complémentaire, et préférable, notamment avec un corpus plus restreint.
- Le processus de « Wikipédiatisation » de ces acteurs est à géométrie variable.



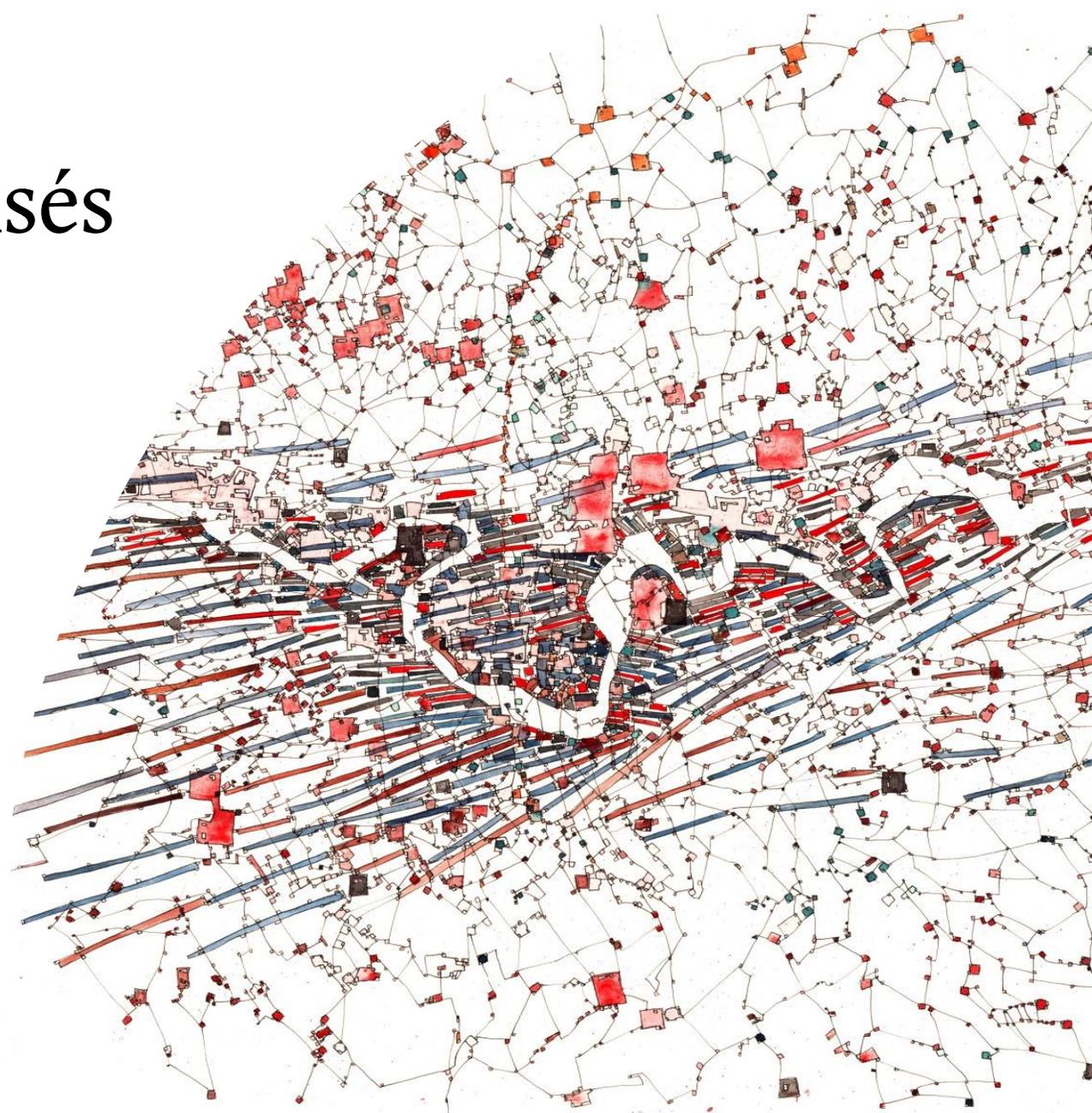


Plus largement

- Une méthodologie numérique adaptable, à une autre échelle, du local au global.
- Modèle flexible.
- Possibilité d'ajouter de nouvelles sources d'informations, notamment issues du Web social.
- Des identifications de nouveaux auteurs grâce au processus d'enrichissement.
- Mais aussi, la nécessité de mobiliser un large spectre de sources hétérogènes afin de capter le plus d'informations possible.
- Et un repérage de certains biais induits par ces sources à destination des personnes intéressées,

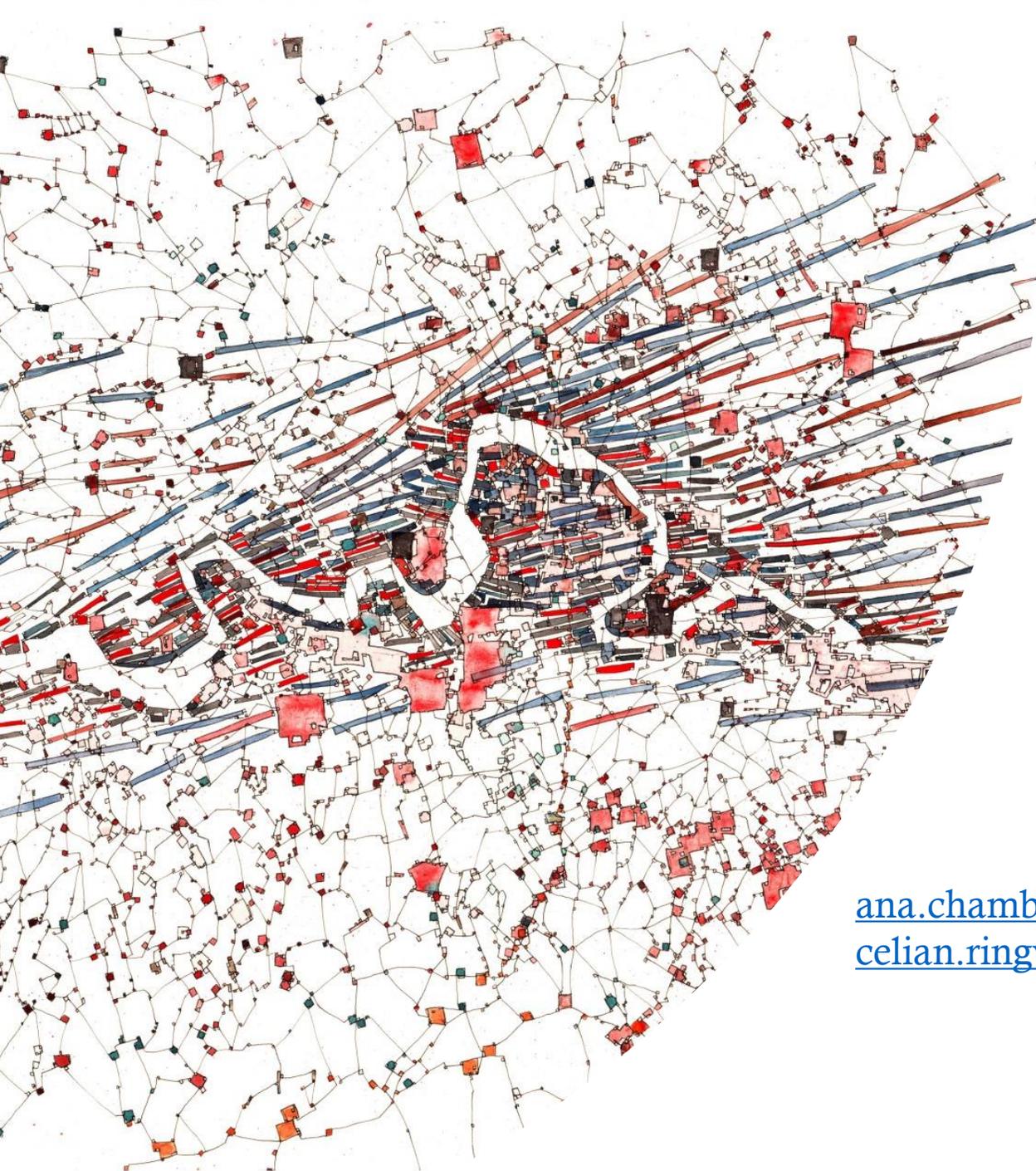
Outils numériques mobilisés

- Gephi : <https://gephi.org/>
- IRaMuTeQ : <http://www.iramuteq.org/>
- OpenRefine : <https://openrefine.org/>
- Oxygen XML Editor :
<https://www.oxygenxml.com>
- Python : <https://www.python.org/>
 - RDFlib,
 - LXML,
 - Beautiful Soup,
 - Scikit-learn,
 - Spacy, etc.
- TXM : <http://textometrie.ens-lyon.fr/>
- Venn JS : benfred.github.io/venn.js



Bibliographie

- BARTHES Roland, « The death and return of the author », Edinburgh : Edinburgh University Press, 1998,
- BÉLISLE Claire, « Du papier à l'écran : lire se transforme », *Lire dans un monde numérique*, Villeurbanne, 2011, pp. 112-162,
- BEN HAZEZ Slim, DECLES Jean Pierre, MINEL Jean-Luc, « Modèle d'exploration contextuelle pour l'analyse sémantique des textes », *HAL*, 2006,
- BONNET Gilles, « L'hypertexte. Poétique de la relecture dans l'œuvre numérique de François Bon », in *Poétique*, n°175, 2014, pp. 21-34.
- -, « L'autoblographie. Écritures numériques de soi », in *Poétique*, n°177, 2015, pp. 131-143.
- BUREAU Bruno, « Méthodologies de construction de corpus », Cours Humanité Numériques, Université Lyon 3, 2019,
- CASATI Roberto, « Contre le colonialisme numérique, manifeste pour continuer à lire », Paris : Albin Michel, 2013,
- COMPAGNON Antoine, « Qu'est-ce qu'un auteur ? », Cours Université de Paris IV-Sorbonne.
- GEFEN Alexandre, « On ne sait jamais rien du sort d'un livre », *Théorie et pratique des best-sellers*, Ghent University & ENS, 2017.
- LEJEUNE Philippe, *Signes de vie. Le pacte autobiographique*, 2, Paris, 2005.
- MERCKLÉ Pierre, *Sociologie des réseaux sociaux*, Paris : La Découverte, 2004.
- MONAT Jean-Baptiste, *A la recherche du Web littéraire*, mémoire de master 2 Humanités Numériques soutenu à l'Université Lumière Lyon 2 sous la direction de Jérôme DARMONT et de Gilles BONNET, 2019, 93 p.
- THOMAS Sue *et al.*, « Transliteracy : Crossing Divides », in *First Monday*, n°12, décembre 2007,



Merci de votre
attention, place
à l'échange !

ana.chambat@gmail.com

celian.ringwald1@univ-lyon2.fr