



***Bruno Gaume, Yannick Chudy,  
Emmanuel Navarro, Pierre Magistry***



***Lundi 3 décembre 2018  
Master mention SDL,  
Linguistique, Informatique, Technologies du Langage***



- **Présentation**
- **Structure des graphes de terrain**
- **Chaîne de traitement de l'information**
- **Démonstration en ligne**



- **Présentation**
- Structure des graphes de terrain
- Chaîne de traitement de l'information
- Démonstration en ligne

# ISTEX

21 millions de documents provenant de 23 corpus de littérature scientifique dans toutes les disciplines, soit plus de 9 207 revues et 344 744 ebooks entre 1473 et 2017 pour l'ESR

Le projet **ISTEX** (Initiative d'excellence de l'Information Scientifique et Technique) s'inscrit dans le programme «**Investissements d'Avenir**», initié par le Ministère de l'Enseignement Supérieur et de la Recherche.

**Le plus vaste réservoir d'archives scientifiques au service de la recherche.**



**Cillex** est un projet Service à valeur ajoutée financé par **ISTEX**. **Cillex** est développé au Laboratoire Toulousain **CLLE** (Cognition Langues Langage Ergonomie) sous la direction de **Bruno Gaume**, avec Yannick Chudy, Emmanuel Navarro et Pierre Magistry.

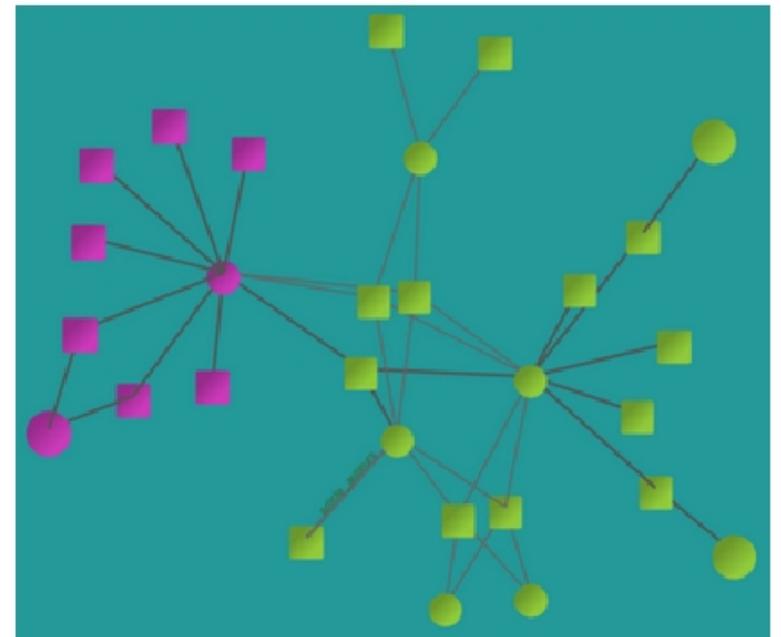


La plus-value de *cilleX* réside dans la **NAVIGATION**, les **CARTES** et les **MESURES** construites sur les archives numériques d'*ISTEX*.

L'utilisateur peut ainsi mieux **cibler** sa recherche ou **explorer** les collections d'*ISTEX* pour les **analyser** ou y **découvrir** de nouvelles connaissances.



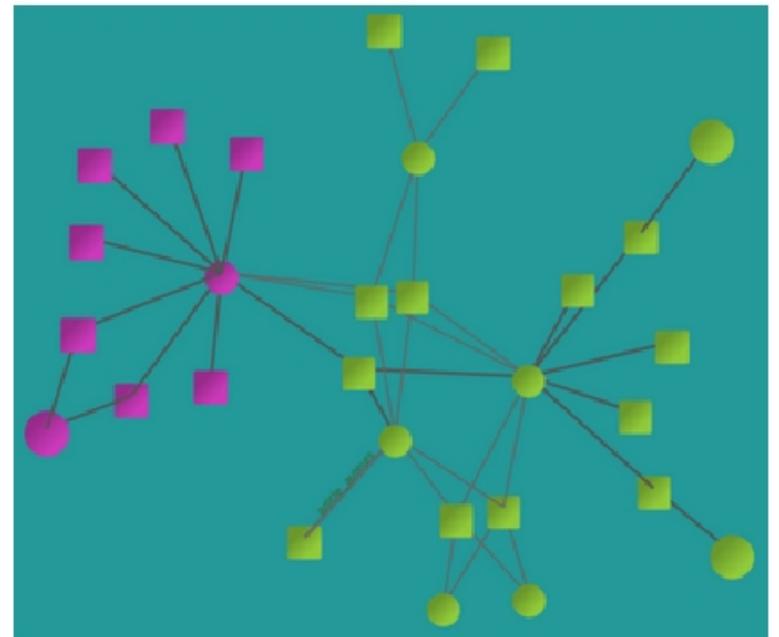
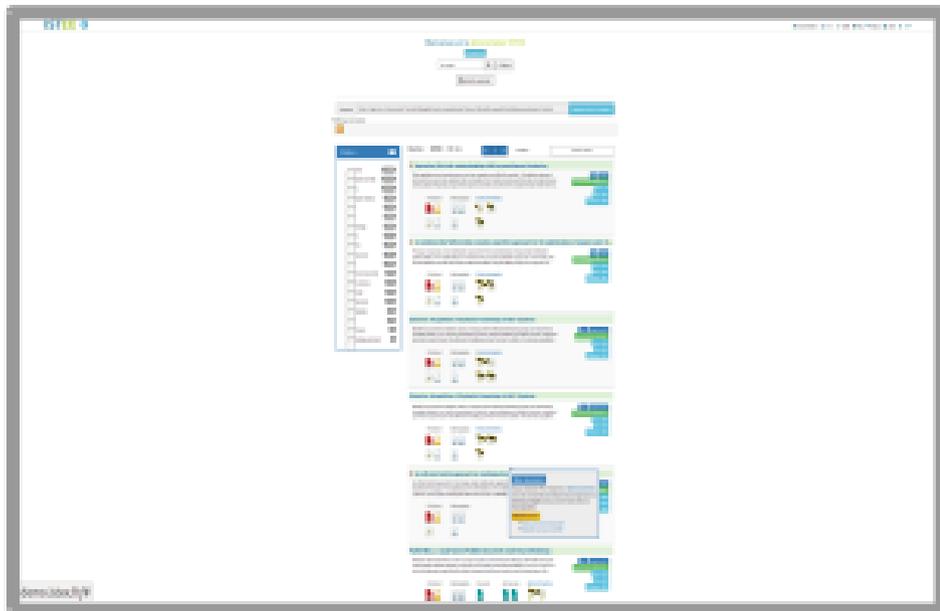
**L'utilisateur peut par exemple comprendre comment neuf auteurs ont collaboré pour rédiger vingt documents.**



# cillex

# CILLEX

**Cartographie** documentaire



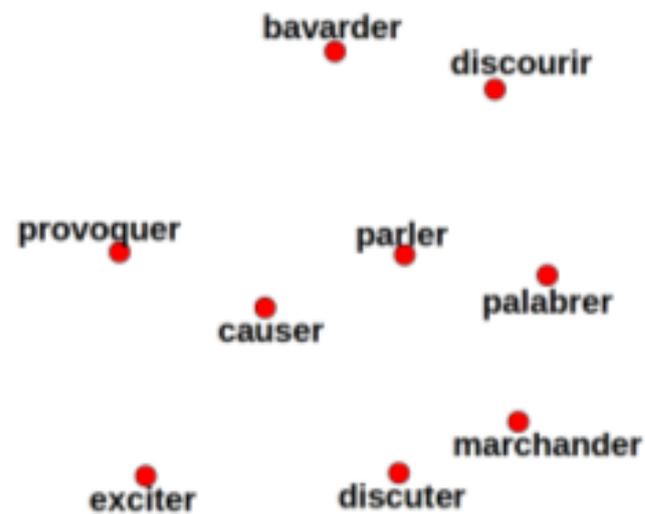




- Présentation
- **Structure des graphes de terrain**
- Chaîne de traitement de l'information
- Démonstration en ligne

$\mathbf{G} = (\mathbf{V}, \mathbf{E})$  où  $E \subseteq P_2^V$  : un graphe de  $n = |V|$  sommets et  $m = |E|$  arêtes

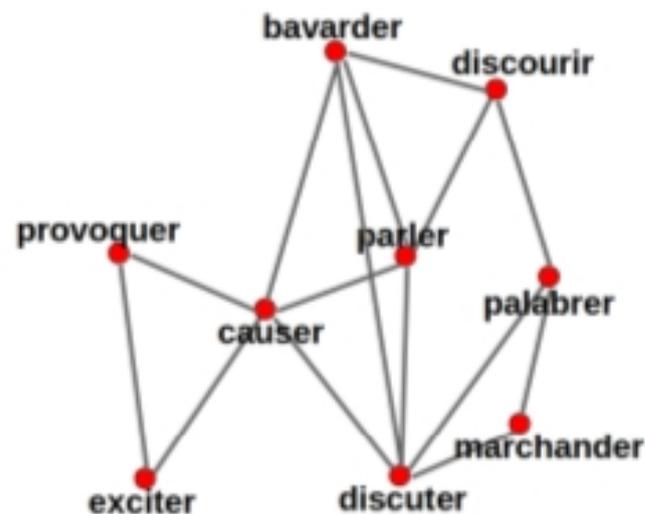
$\mathbf{V} = \{\text{bavarder, causer, discourir, discuter, exciter, palabrer, parler, p rerer, provoquer}\}$



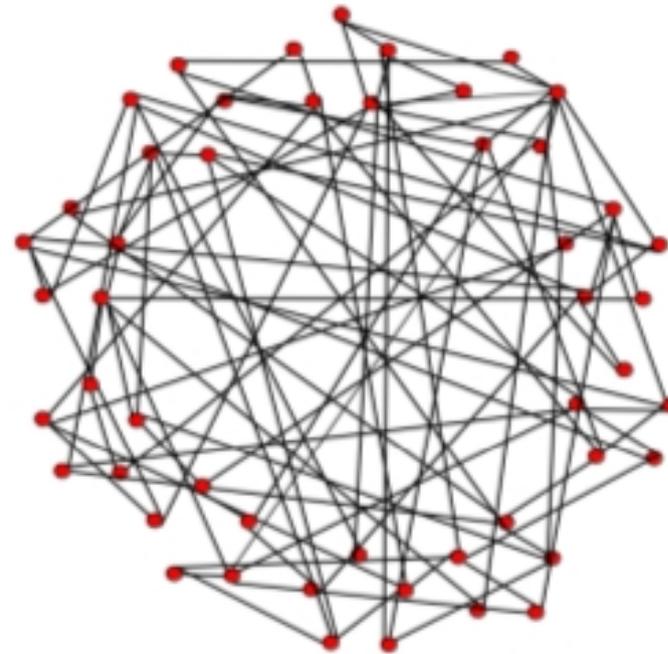
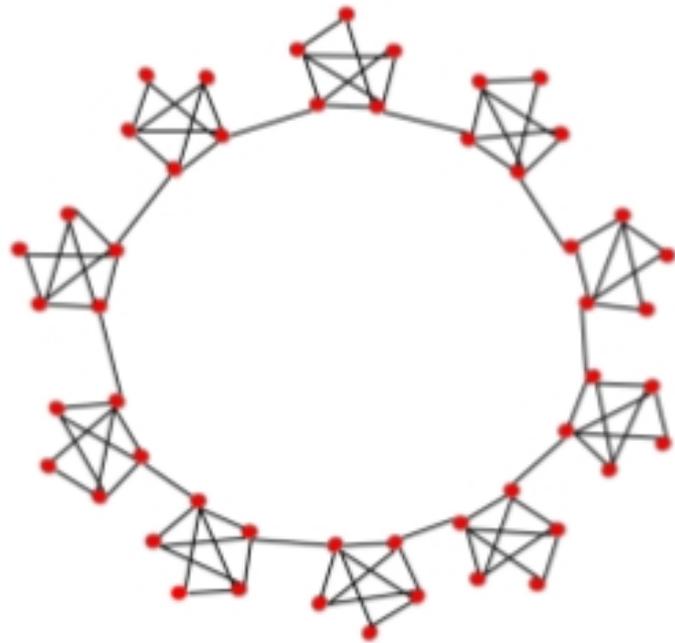
$\mathbf{G} = (\mathbf{V}, \mathbf{E})$  où  $E \subseteq P_2^V$  : un graphe de  $n = |V|$  sommets et  $m = |E|$  arêtes

$\mathbf{V} = \{\text{bavarder, causer, discourir, discuter, exciter, palabrer, parler, pérorer, provoquer}\}$

$\mathbf{E} = \{\{\text{bavarder, parler}\}, \{\text{bavarder, discourir}\}, \{\text{bavarder, causer}\}, \{\text{causer, discuter}\}, \{\text{causer, parler}\}, \{\text{causer, provoquer}\}, \{\text{causer, exciter}\}, \{\text{discourir, parler}\}, \{\text{discourir, palabrer}\}, \{\text{discourir, pérorer}\}, \{\text{discuter, parler}\}, \{\text{discuter, palabrer}\}, \{\text{discuter, pérorer}\}, \{\text{exciter, provoquer}\}, \{\text{palabrer, parler}\}, \{\text{palabrer, pérorer}\}, \{\text{parler, pérorer}\}\}$



# Graphes de laboratoire



## Graphes de terrain

- ▶ Les graphes d'acointance d'un groupe d'humains
- ▶ Le graphe des appels téléphoniques aux USA
- ▶ Le graphe du World Wide Web
- ▶ Le graphe des collaborations cinématographiques
- ▶ Le graphe de *Caenorhabditis elegans*
- ▶ Les graphes lexicaux
- ▶ **Les graphes extraits des bases documentaires**

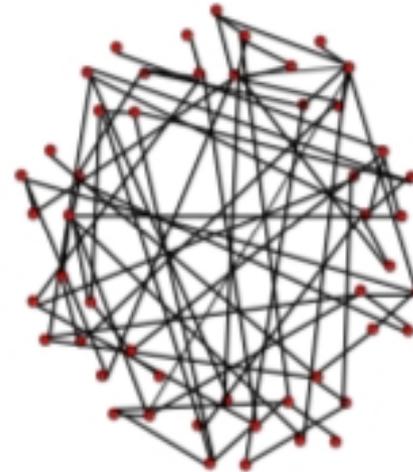
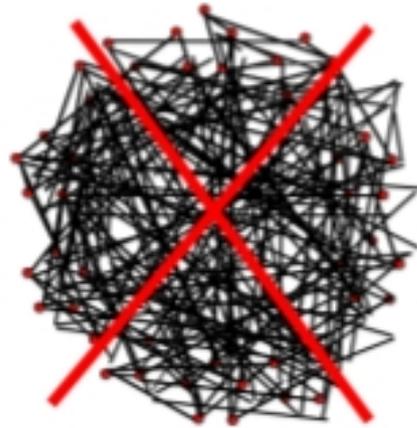
## Graphes de terrain

- ▶ Les graphes d'acointance d'un groupe d'humains
- ▶ Le graphe des appels téléphoniques aux USA
- ▶ Le graphe du World Wide Web
- ▶ Le graphe des collaborations cinématographiques
- ▶ Le graphe de *Caenorhabditis elegans*
- ▶ Les graphes lexicaux
- ▶ **Les graphes extraits des bases documentaires**

## Propriétés topologiques des graphes de terrains

- Faible densité en arêtes
- Chemins courts
- Distribution des degrés à queue lourde (loi de puissance)
- Fort coefficient de clustering

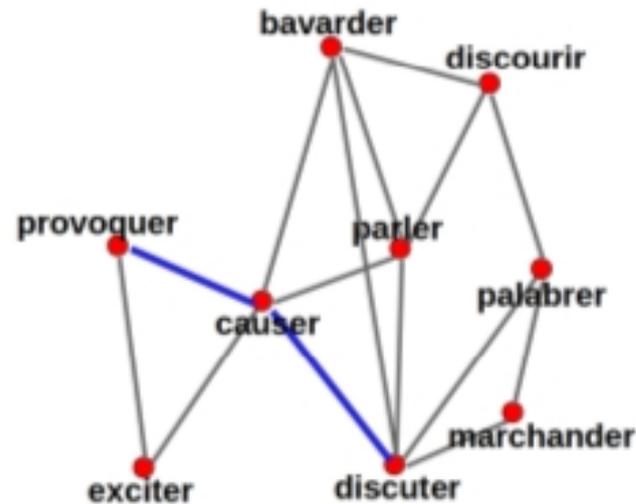
# Graphes de terrain



## Propriétés topologiques des graphes de terrains

- **Faible densité en arêtes**
- Chemins courts
- Distribution des degrés à queue lourde (loi de puissance)
- Fort coefficient de clustering

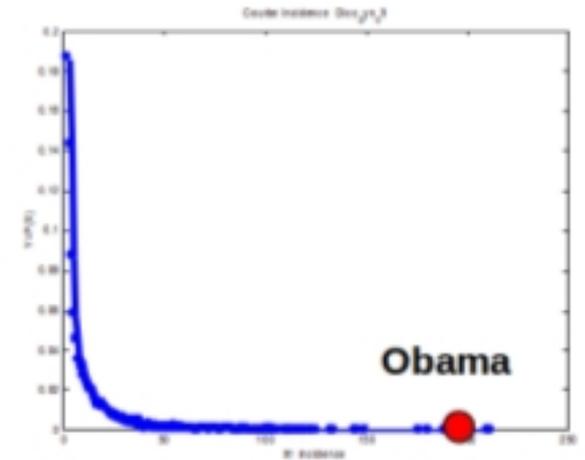
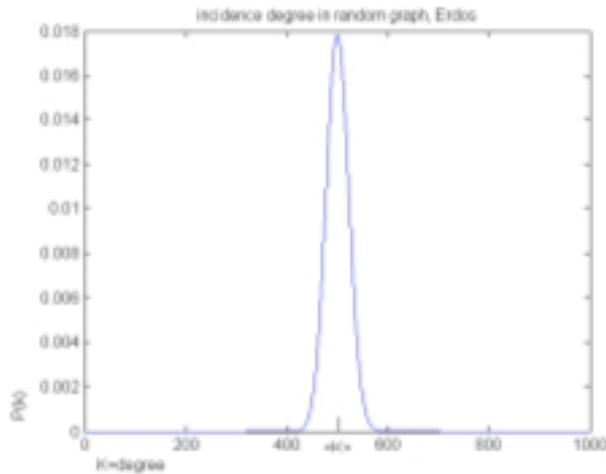
# Graphes de terrain



## Propriétés topologiques des graphes de terrains

- Faible densité en arêtes
- **Chemins courts**
- Distribution des degrés à queue lourde (loi de puissance)
- Fort coefficient de clustering

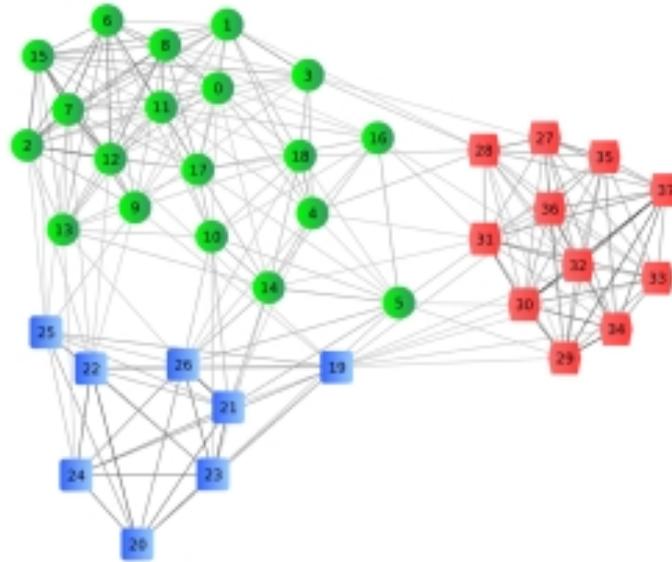
# Graphes de terrain



## Propriétés topologiques des graphes de terrains

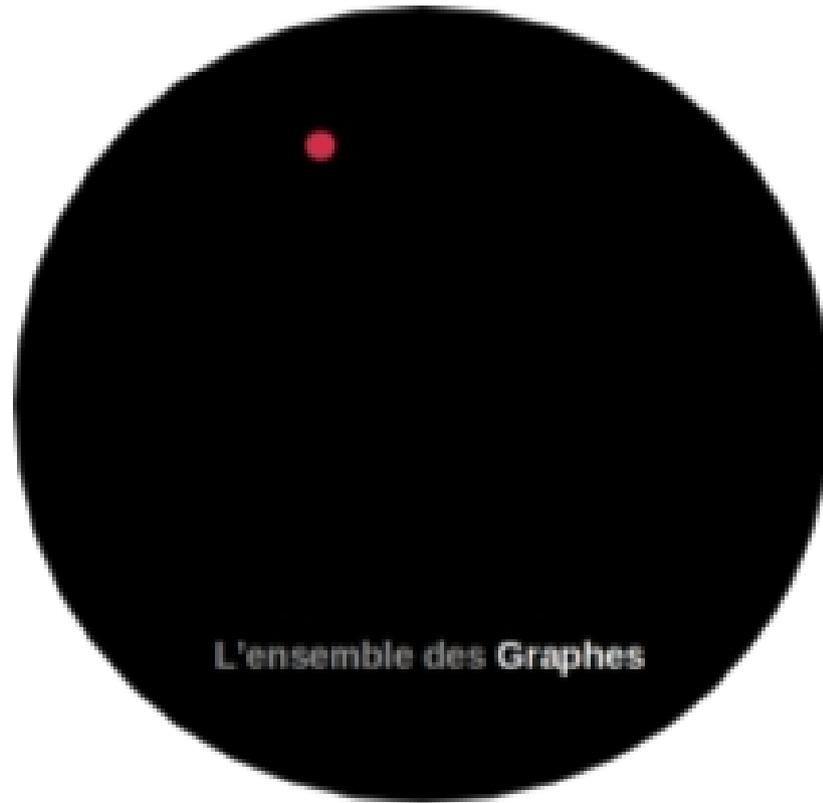
- Faible densité en arêtes
- Chemins courts
- **Distribution des degrés à queue lourde (loi de puissance)**
- Fort coefficient de clustering

# Graphes de terrain



## Propriétés topologiques des graphes de terrains

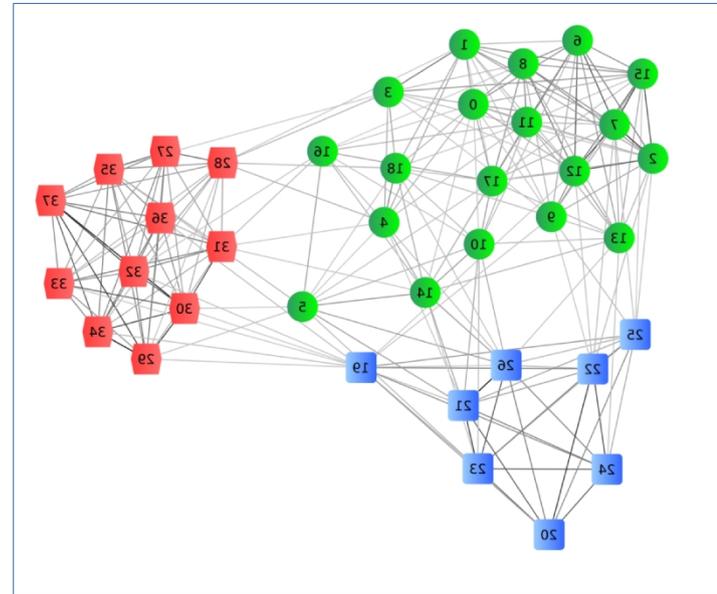
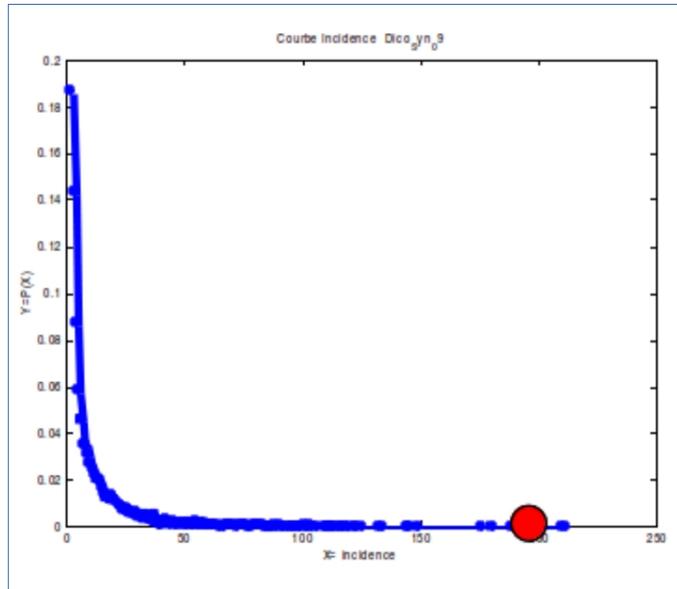
- Faible densité en arêtes
- Chemins courts
- Distribution des degrés à queue lourde (loi de puissance)
- **Fort coefficient de clustering**



## Propriétés topologiques des graphes de terrains

- Faible densité en arêtes
- Chemins courts
- Distribution des degrés à queue lourde (loi de puissance)
- **Fort coefficient de clustering**

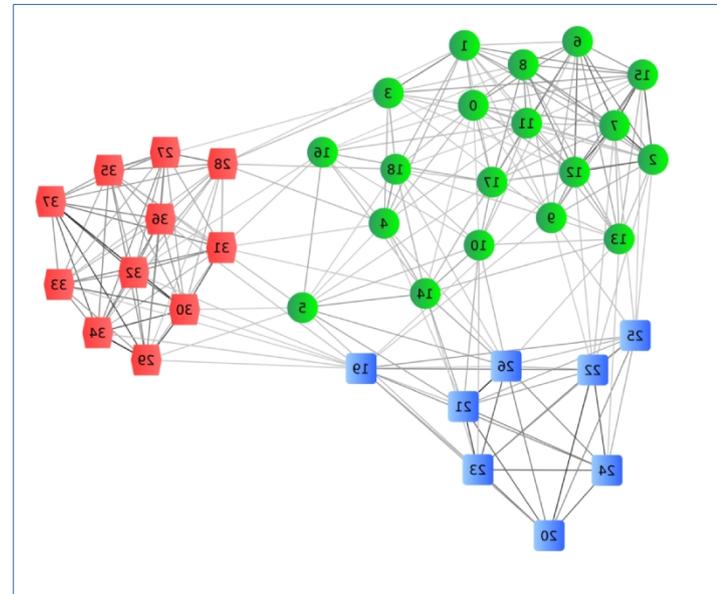
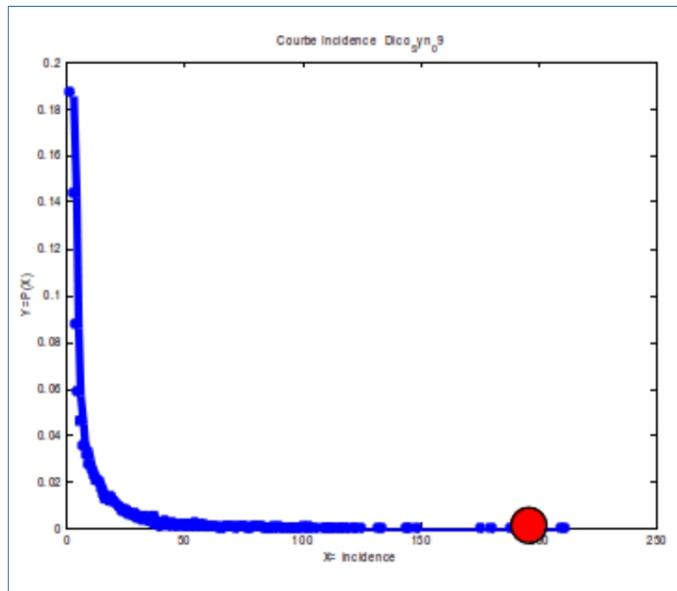
# Graphes de terrain



## Propriétés topologiques des graphes de terrains

- Faible densité en arêtes
- Chemins courts
- **Distribution des degrés à queue lourde (loi de puissance)**
- **Fort coefficient de clustering**

# Graphes de terrain



autourdumot.fr

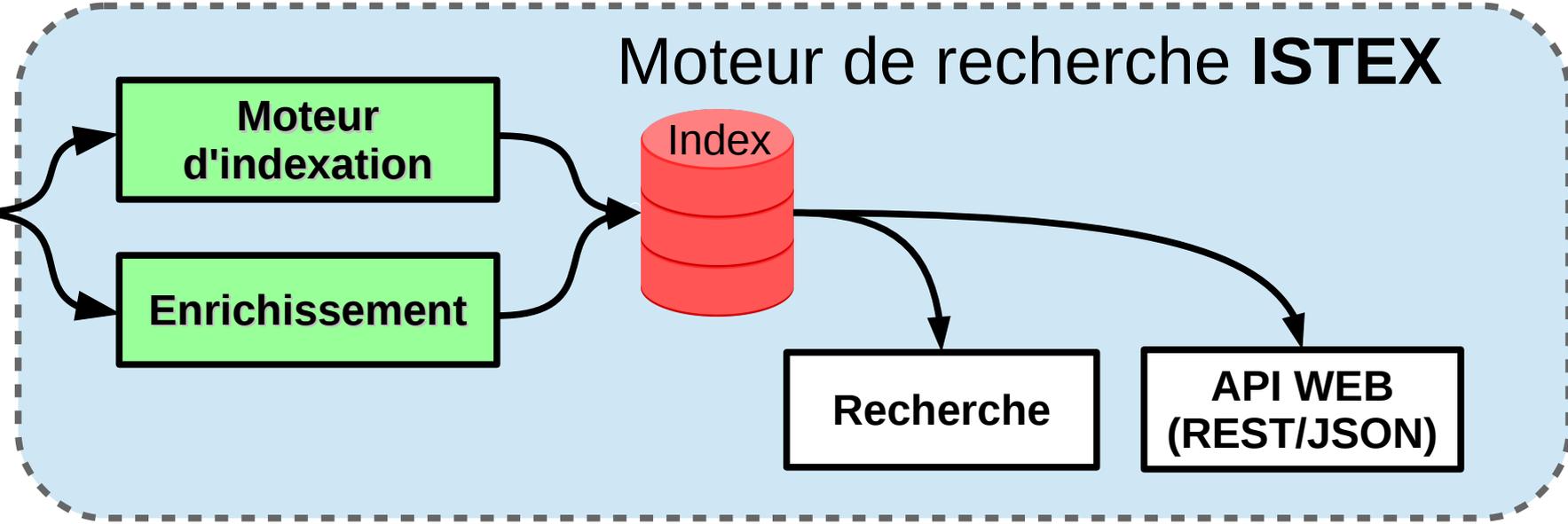
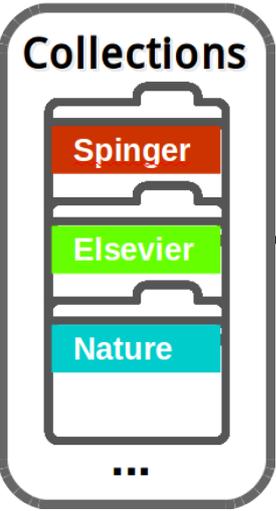
## Propriétés topologiques des graphes de terrains

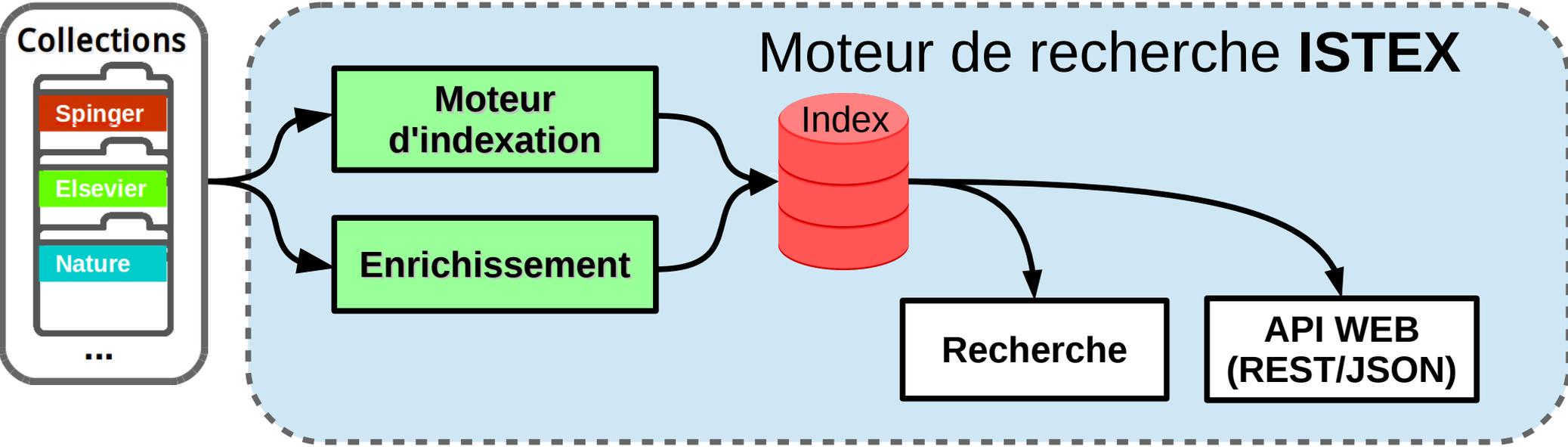
- Faible densité en arêtes
- Chemins courts
- **Distribution des degrés à queue lourde (loi de puissance)**
- **Fort coefficient de clustering**





- Présentation
- Structure des graphes de terrain
- **Chaîne de traitement de l'information**
- Démonstration en ligne



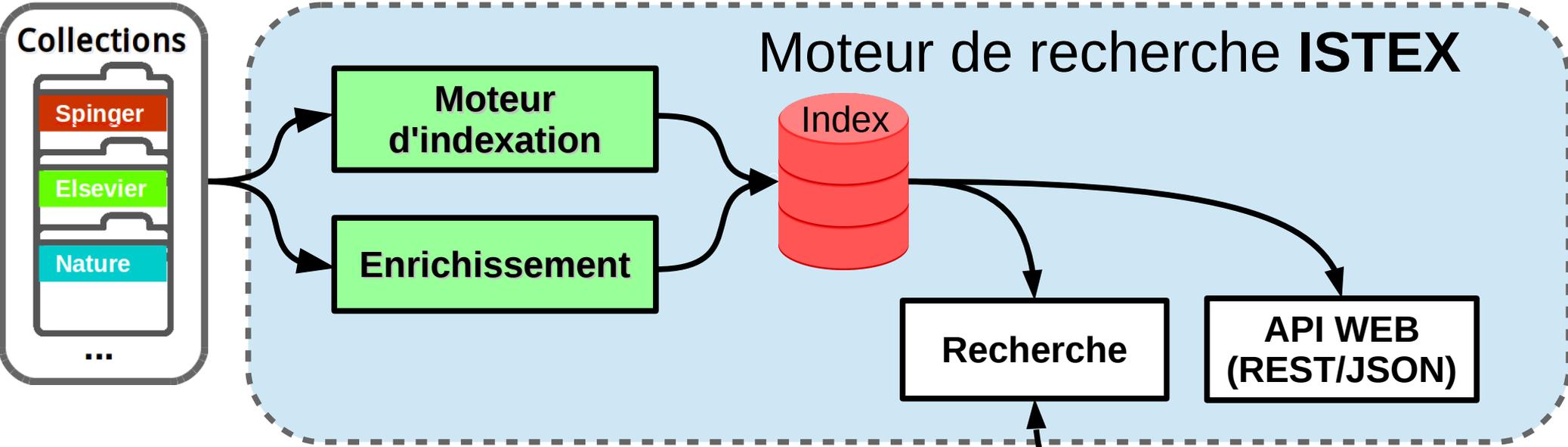


Bienvenue sur le **démonstrateur ISTEM**

[En savoir plus](#)

**Gérard Mourou laser** 🔍 [Options](#)

Recherche avancée

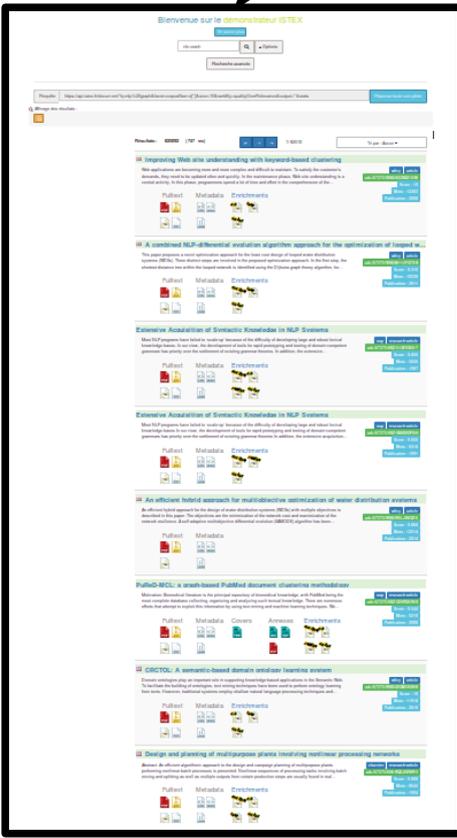
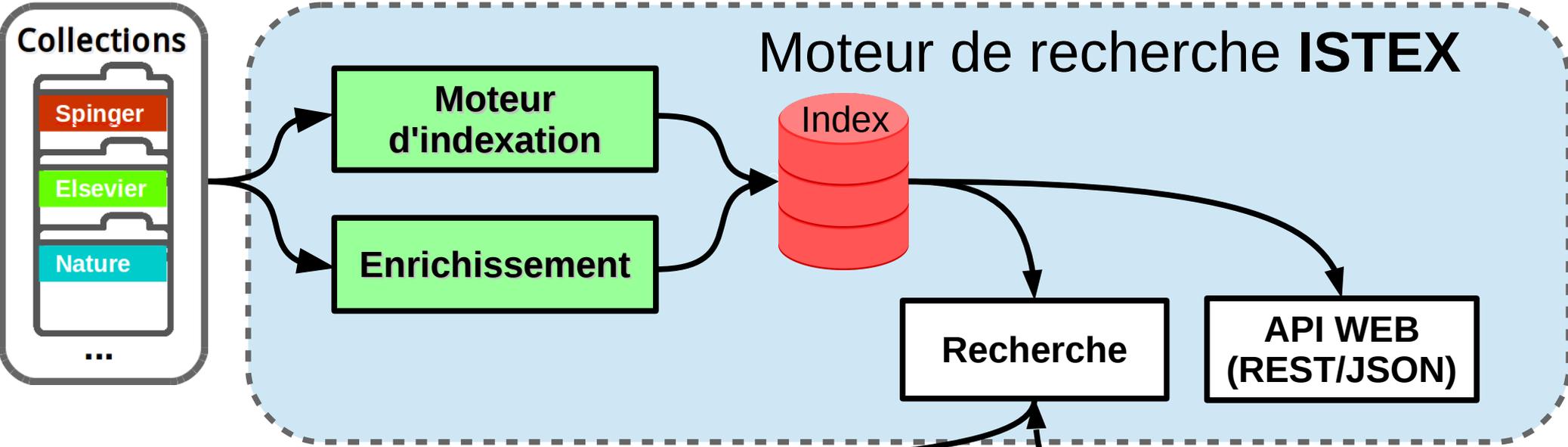


Bienvenue sur le **démonstrateur ISTEX**

[En savoir plus](#)

**Gérard Mourou laser** 🔍 ▼ Options

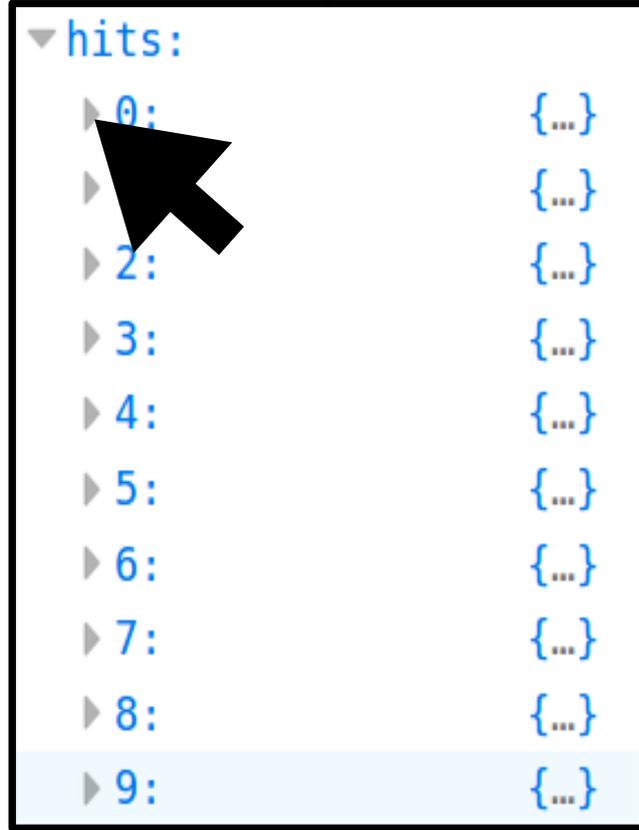
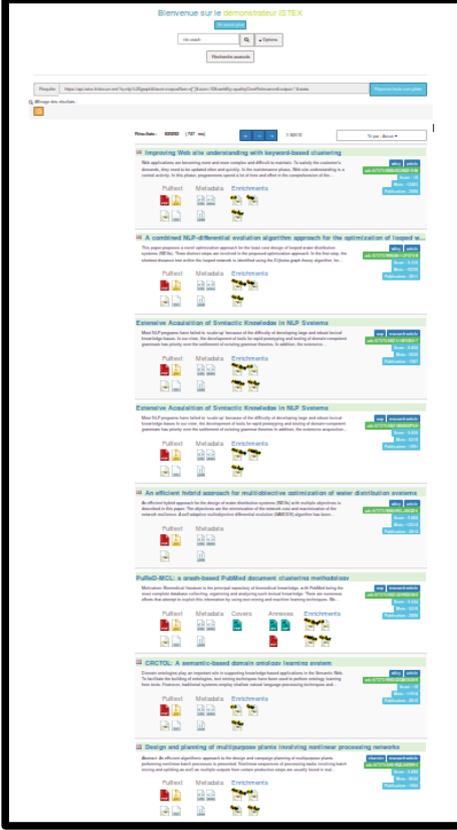
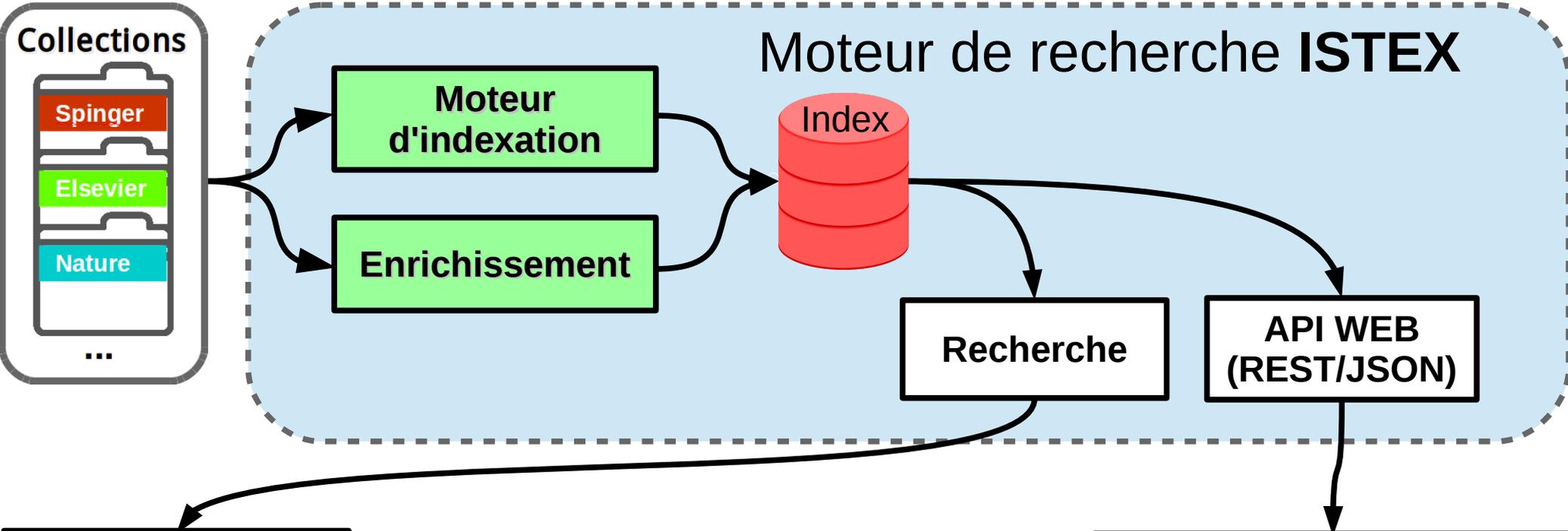
Recherche avancée

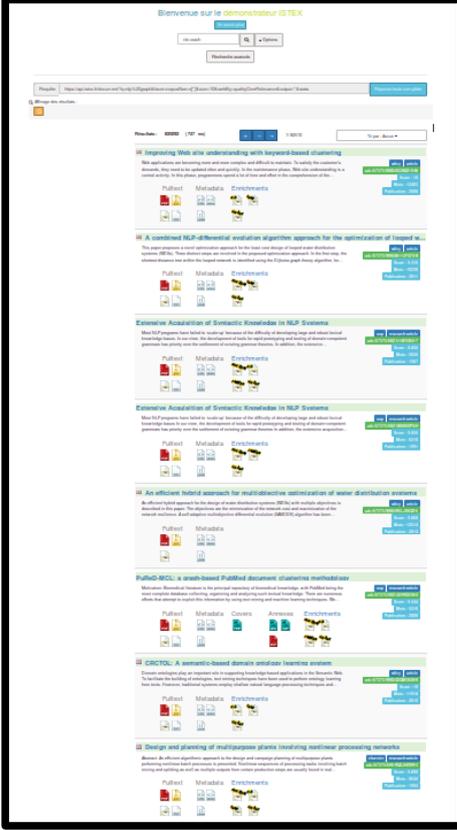
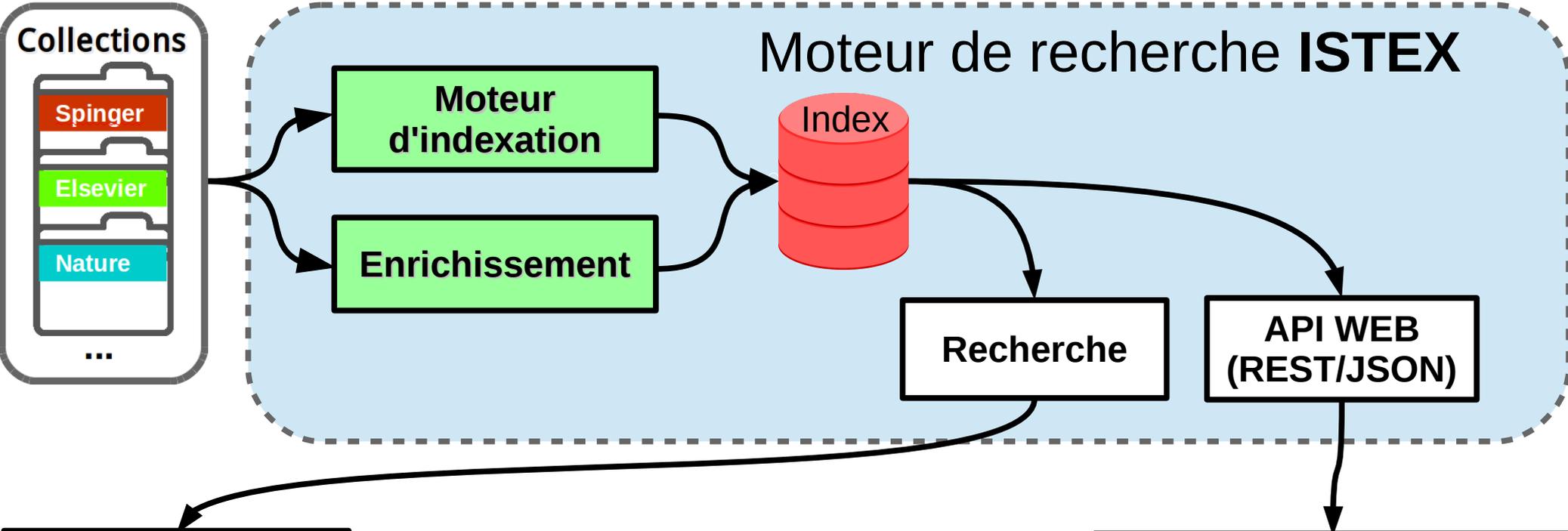






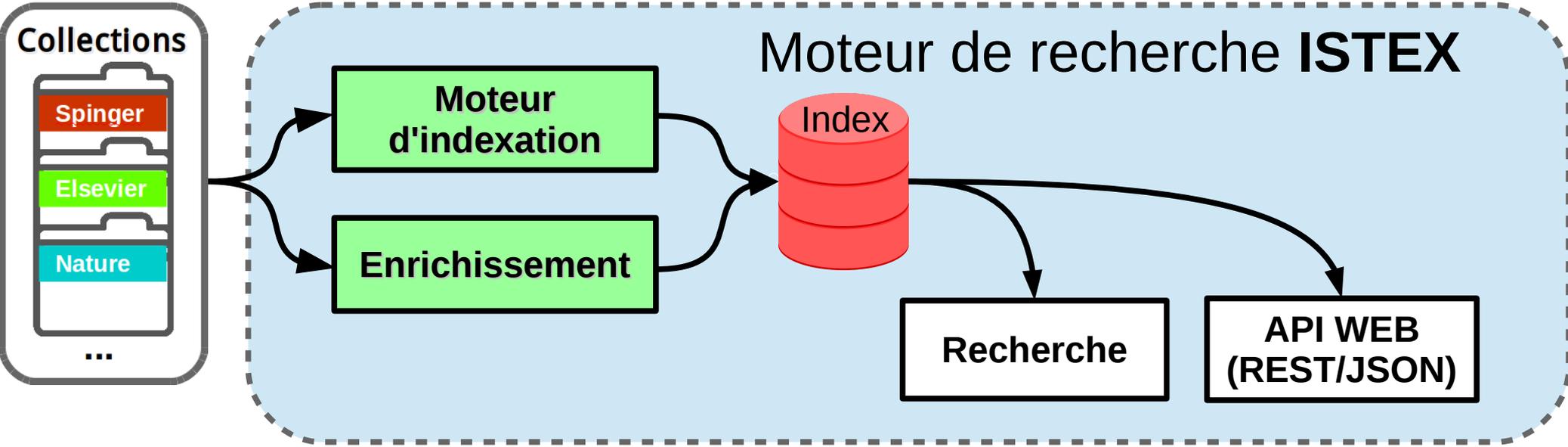




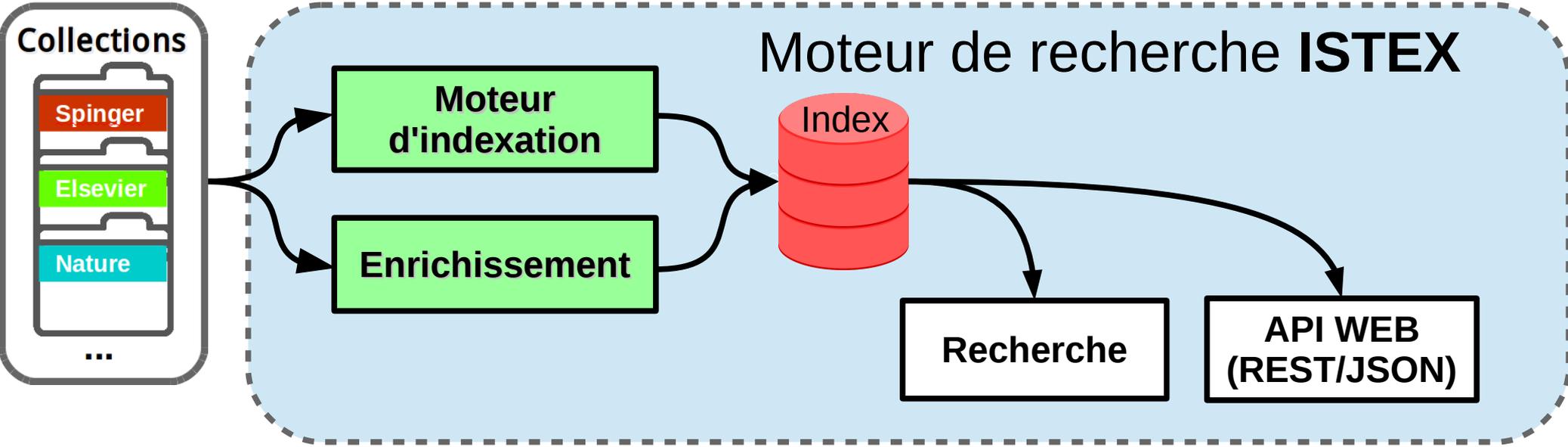


```

  hits:
    0:
      corpusName:
      keywords:
      author:
      subject:
      articleId:
      arkIstex:
      language:
      originalGenre:
      abstract:
      qualityIndicators:
      title:
      refBibs:
      genre:
      host:
      namedEntities:
      ark:
      categories:
      publicationDate:
      copyrightDate:
      enrichments:
      doi:
      id:
      score:
      fulltext:
      metadata:
      serie:
  
```



- corpusName:
- ▶ keywords:
- ▶ author:
- ▶ subject:
- ▶ articleId:
- ▶ arkIstex:
- ▶ language:
- ▶ originalGenre:
- ▶ abstract:
- ▶ qualityIndicators:
- ▶ title:
- ▶ refBibs:
- ▶ genre:
- ▶ host:
- ▶ namedEntities:
- ▶ ark:
- ▶ categories:
- publicationDate:
- copyrightDate:
- ▶ enrichments:
- ▶ doi:
- id:
- score:
- ▶ fulltext:
- ▶ metadata:
- serie:



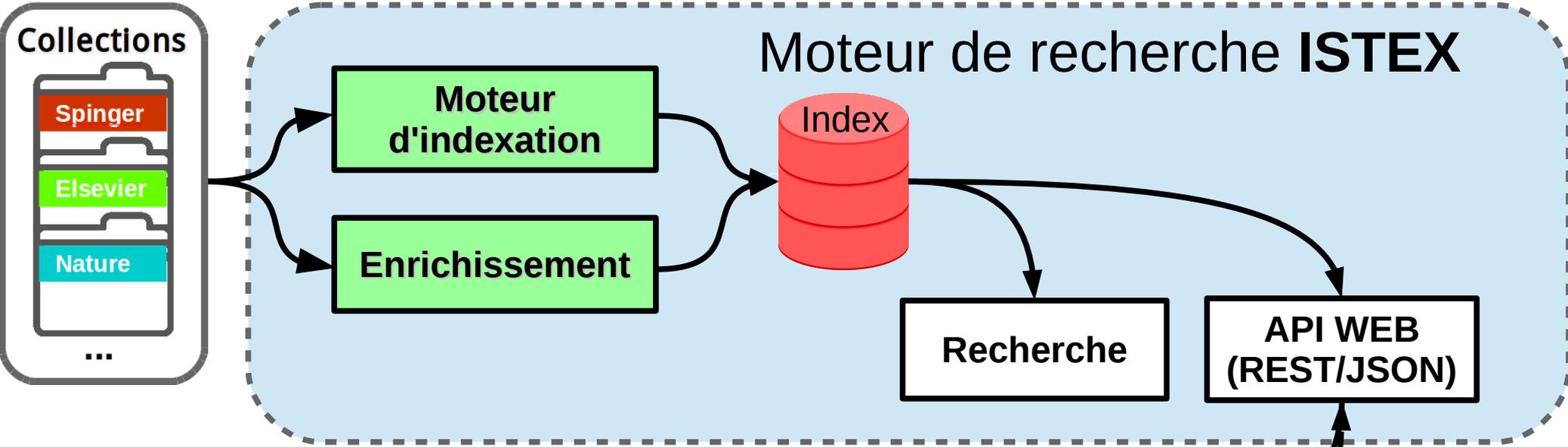
```

corpusName:
▶ keywords:
▶ author:
▶ subject:
▶ articleId:
▶ arkIstex:
▶ language:
▶ originalGenre:
▶ abstract:
▶ qualityIndicators:
▶ title:
▶ refBibs:
▶ genre:
▶ host:
▶ namedEntities:
▶ ark:
▶ categories:
publicationDate:
copyrightDate:
enrichments:
▶ doi:
id:
score:
▶ fulltext:
▶ metadata:
serie:

```

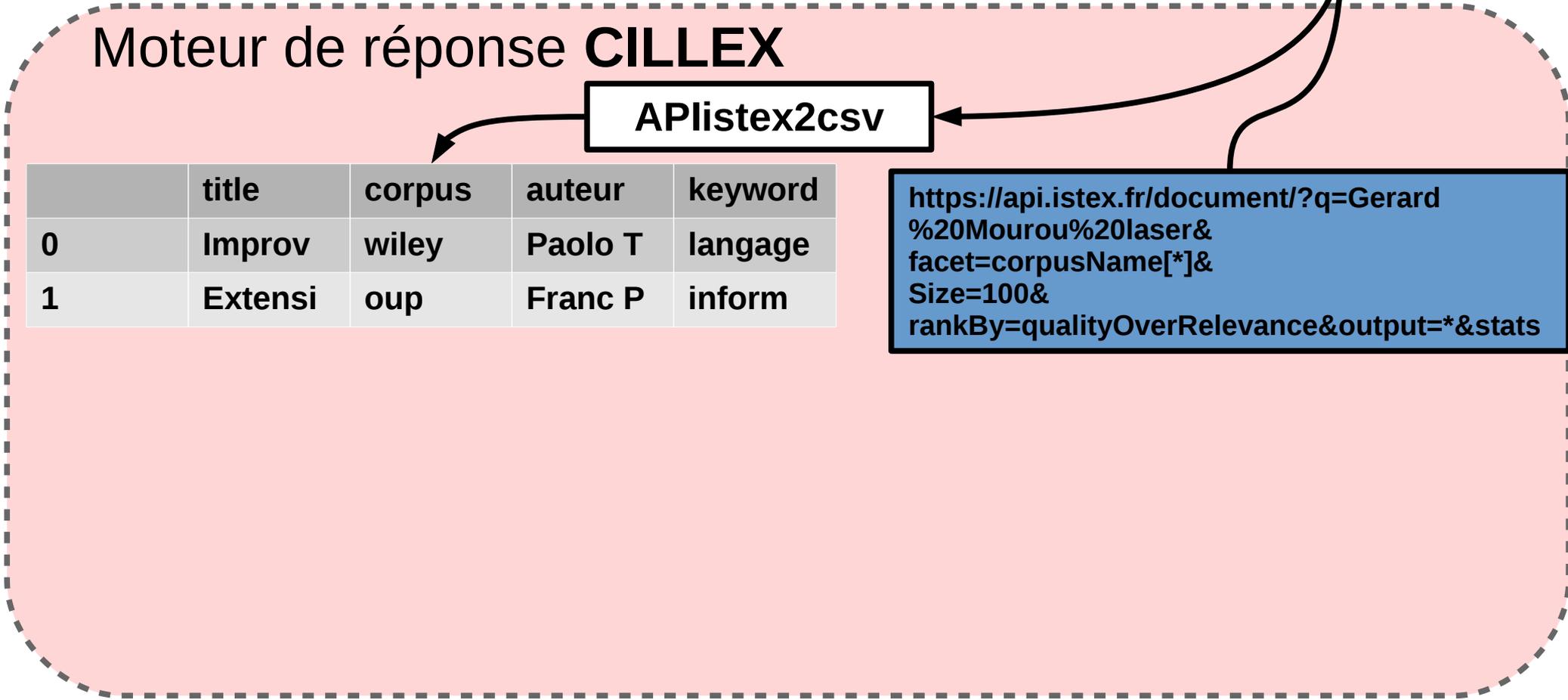
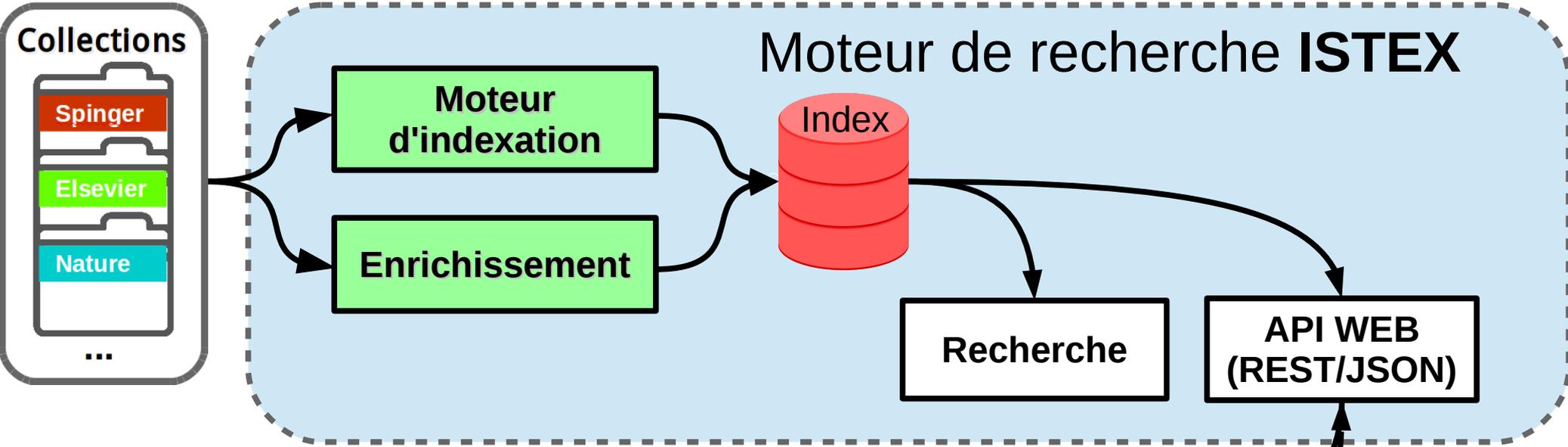




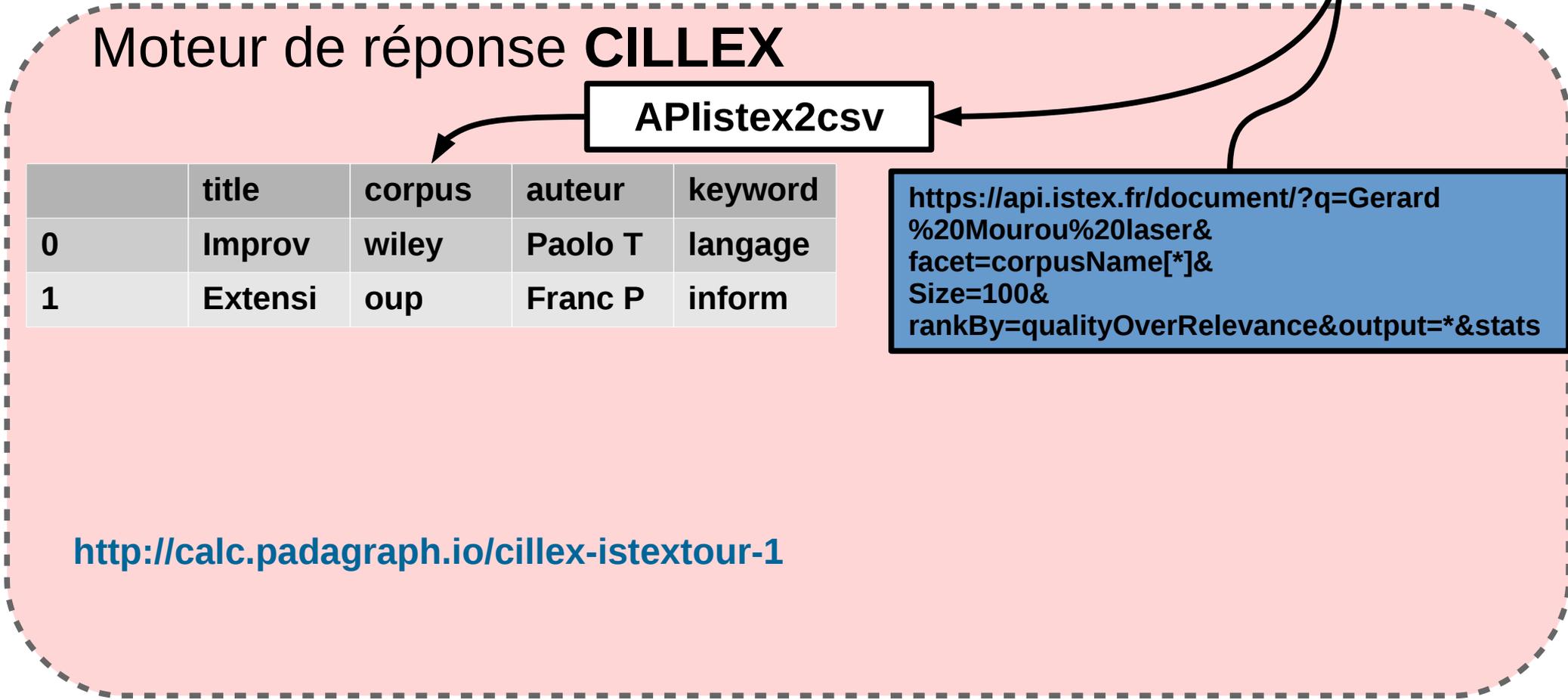
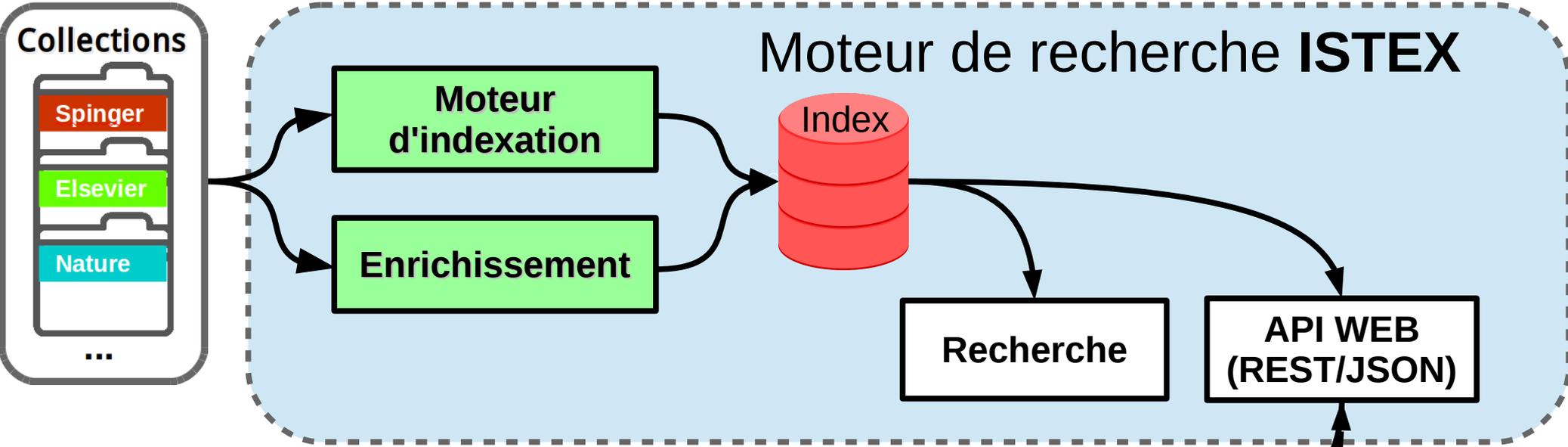


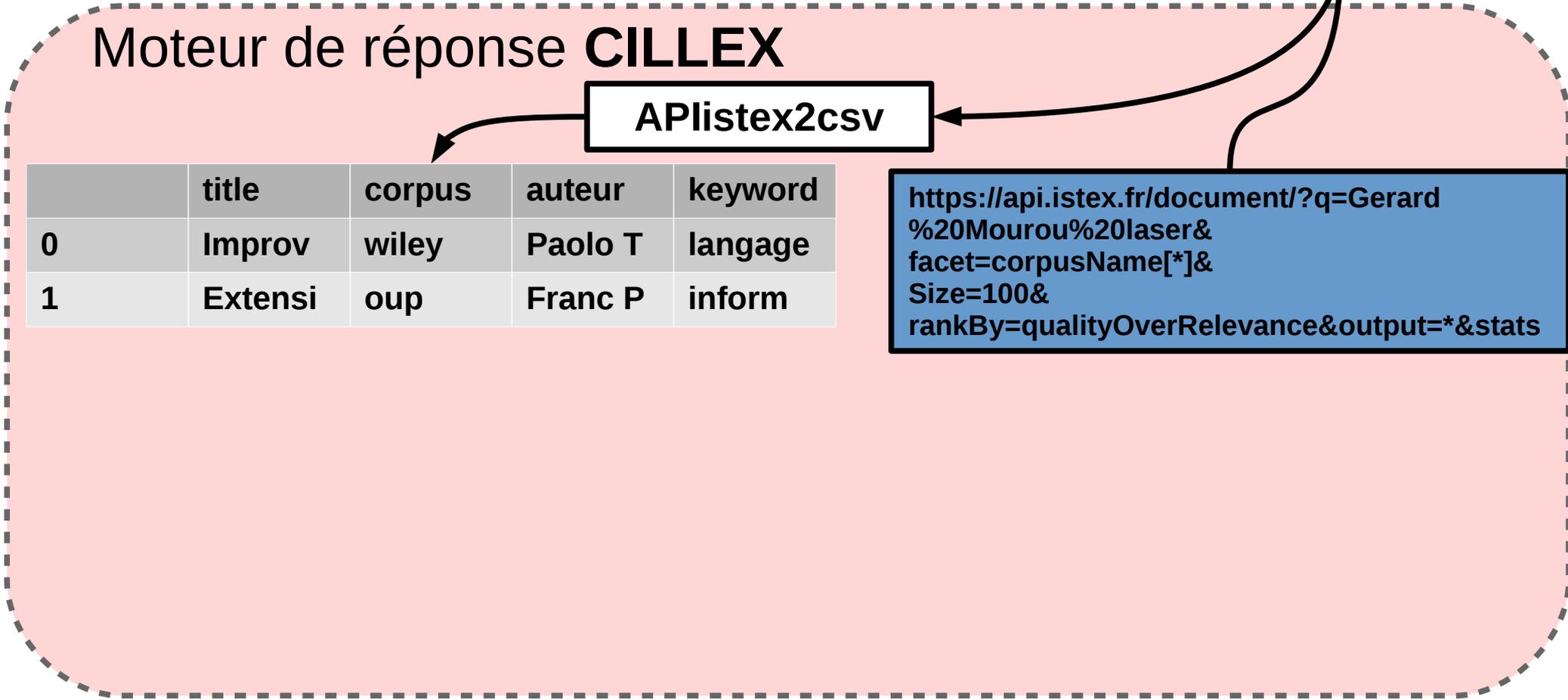
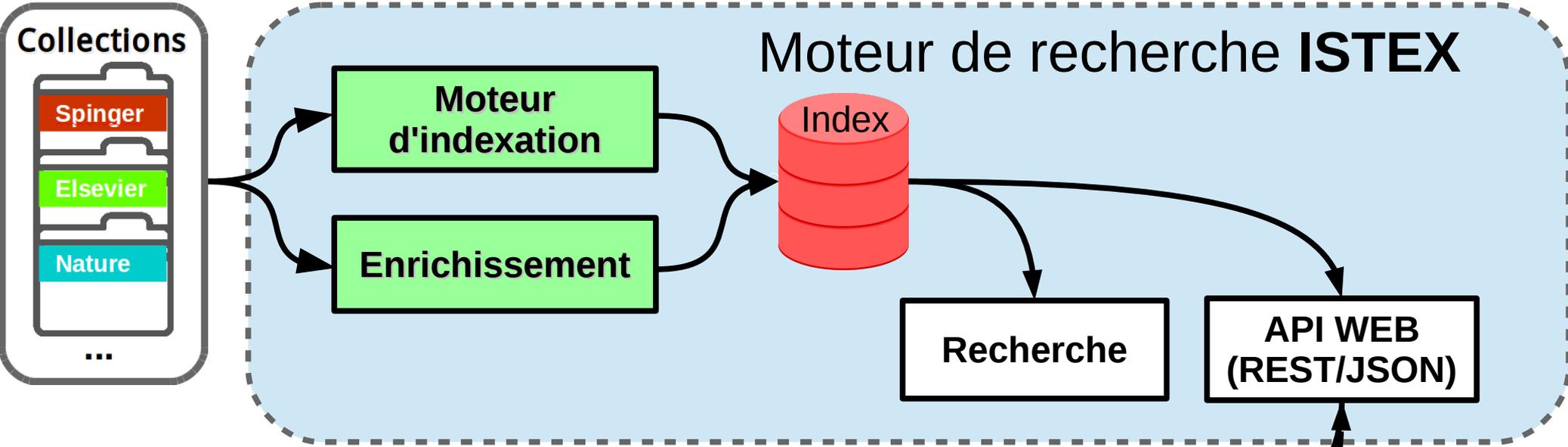
	title	corpus	auteur	keyword
0	Improv	wiley	Paolo T	langage
1	Extensi	oup	Franc P	inform

```
https://api.istex.fr/document/?q=Gerard
%20Mourou%20laser&
facet=corpusName[*]&
Size=100&
rankBy=qualityOverRelevance&output=*&stats
```

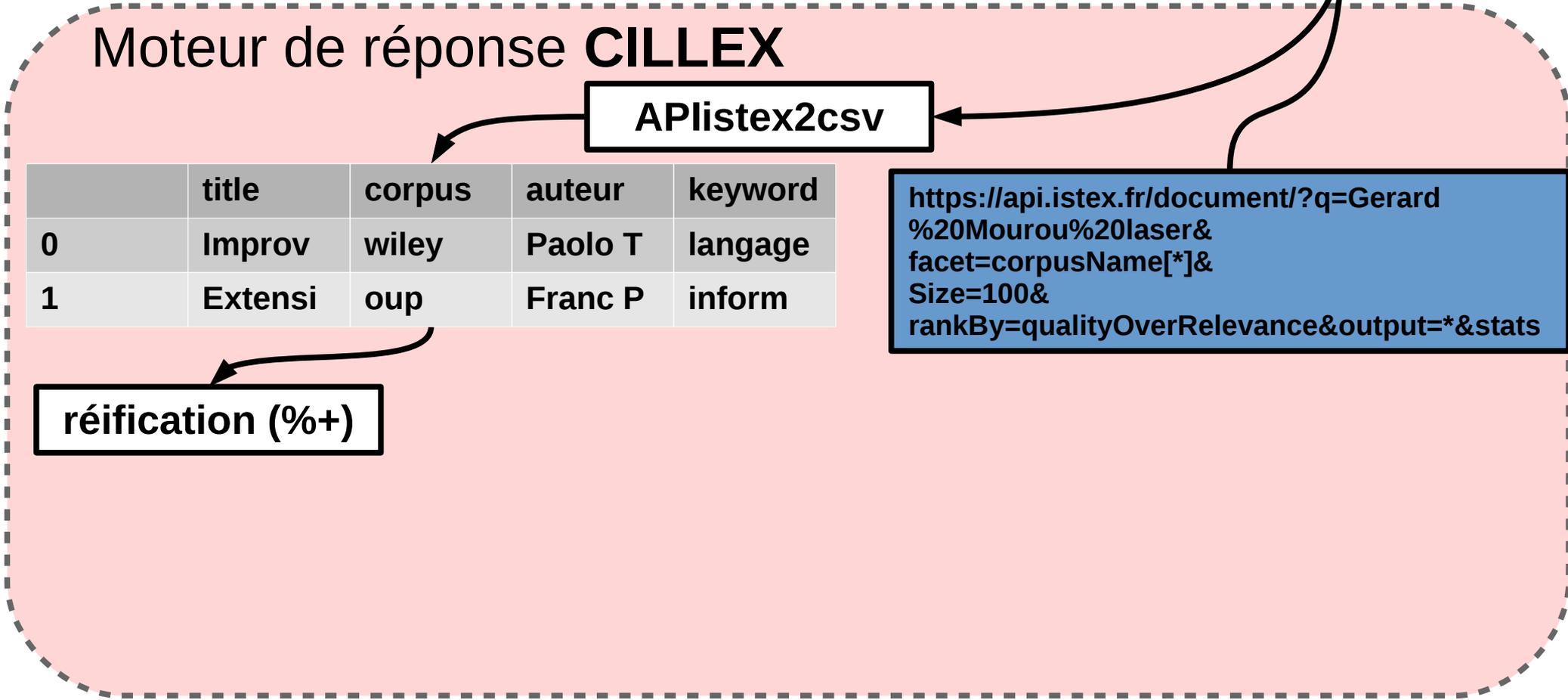
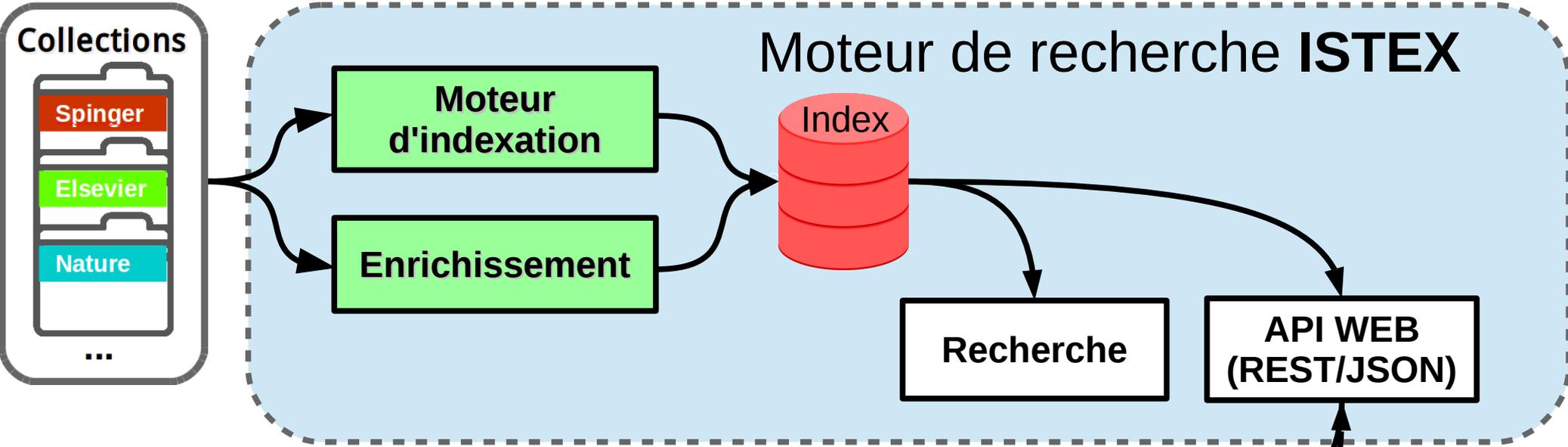


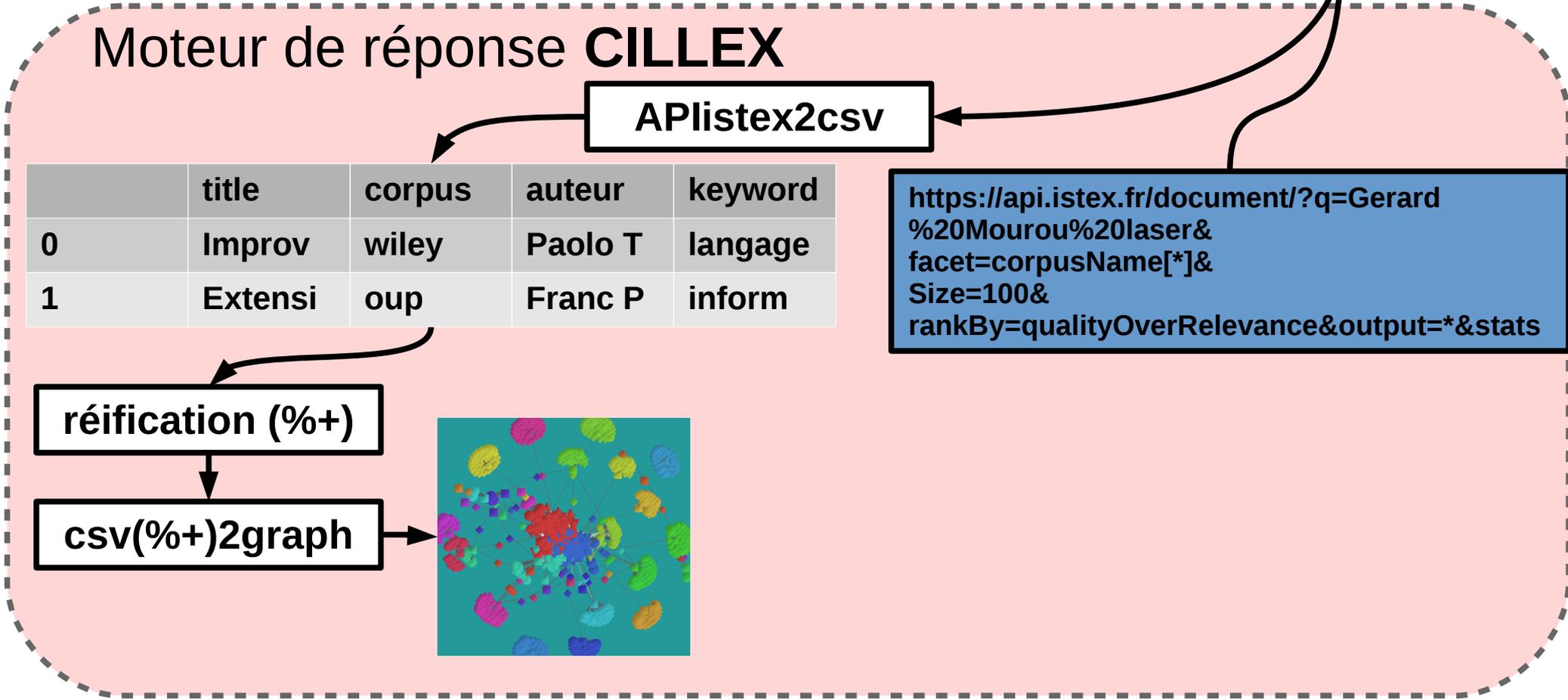
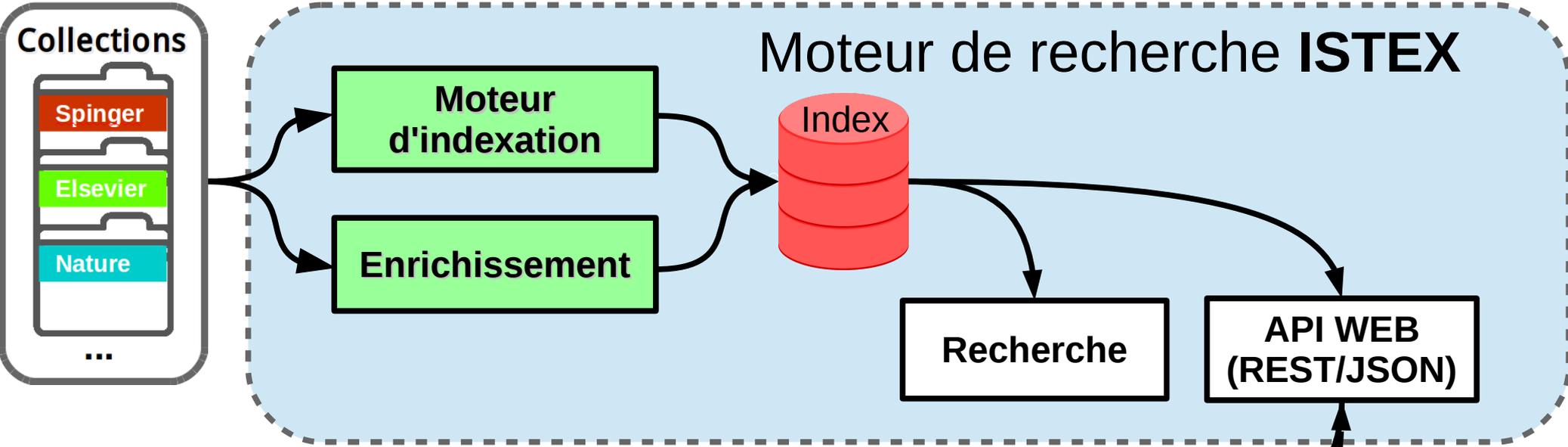
```
https://api.istex.fr/document/?q=Gerard%20Mourou%20laser&facet=corpusName[*]&Size=100&rankBy=qualityOverRelevance&output=*&stats
```

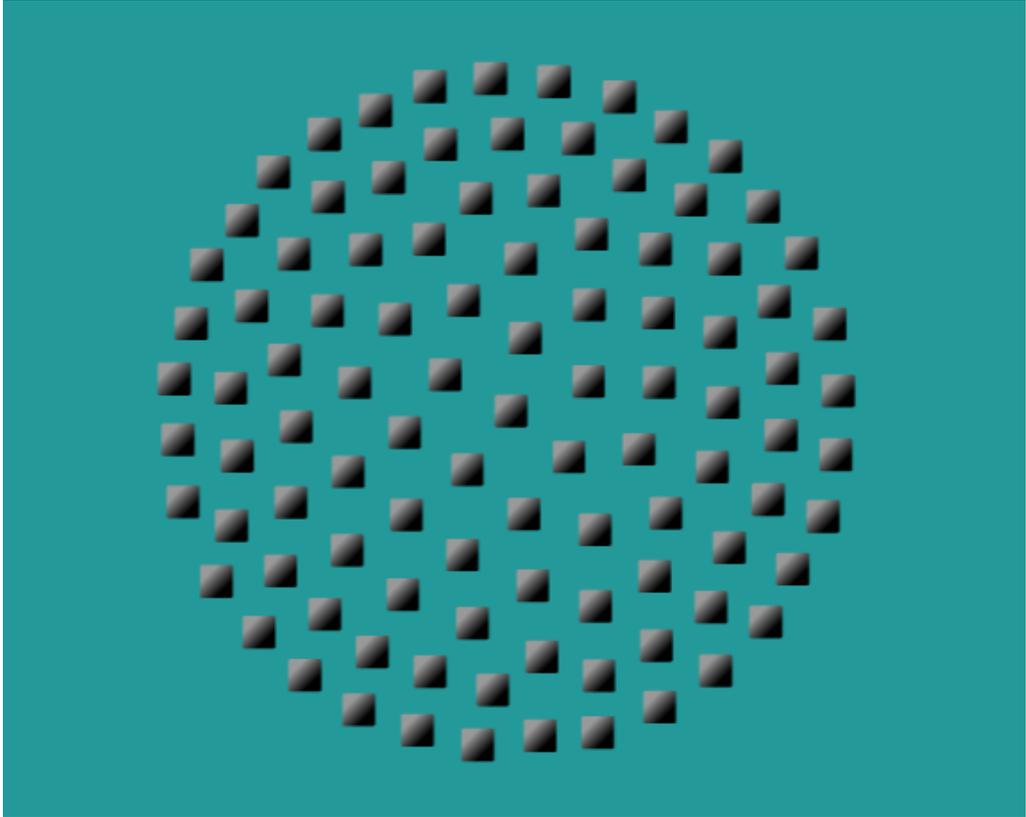


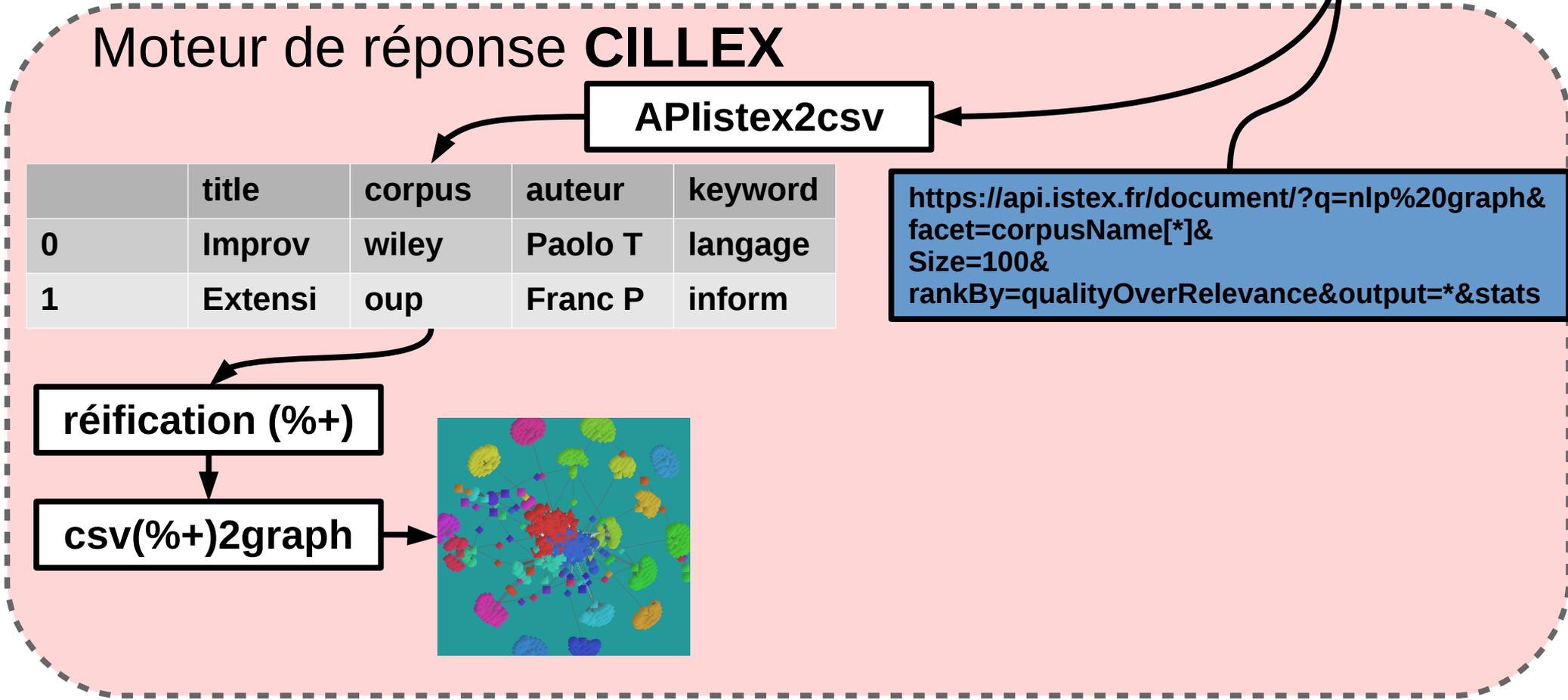
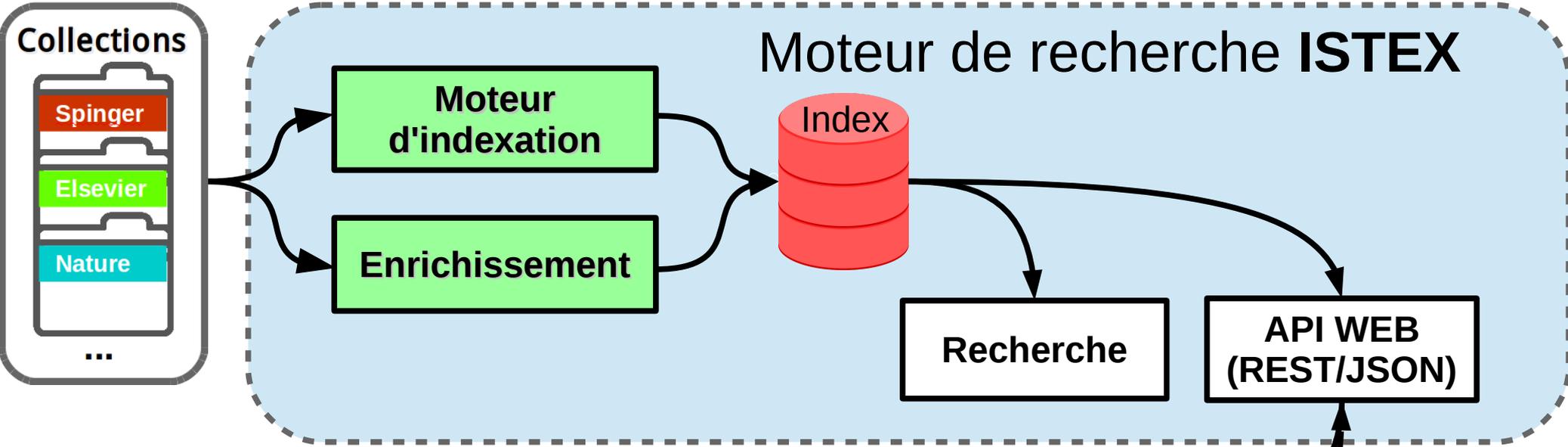


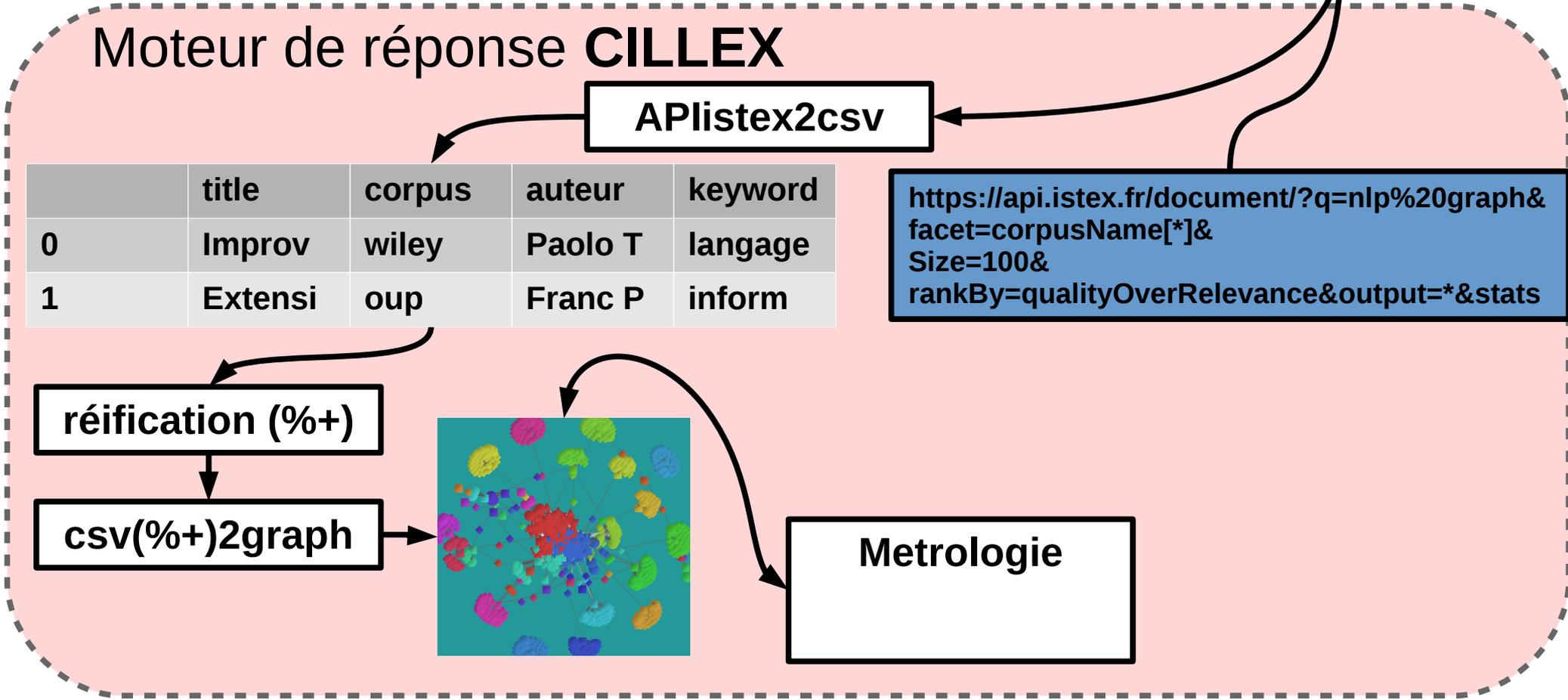
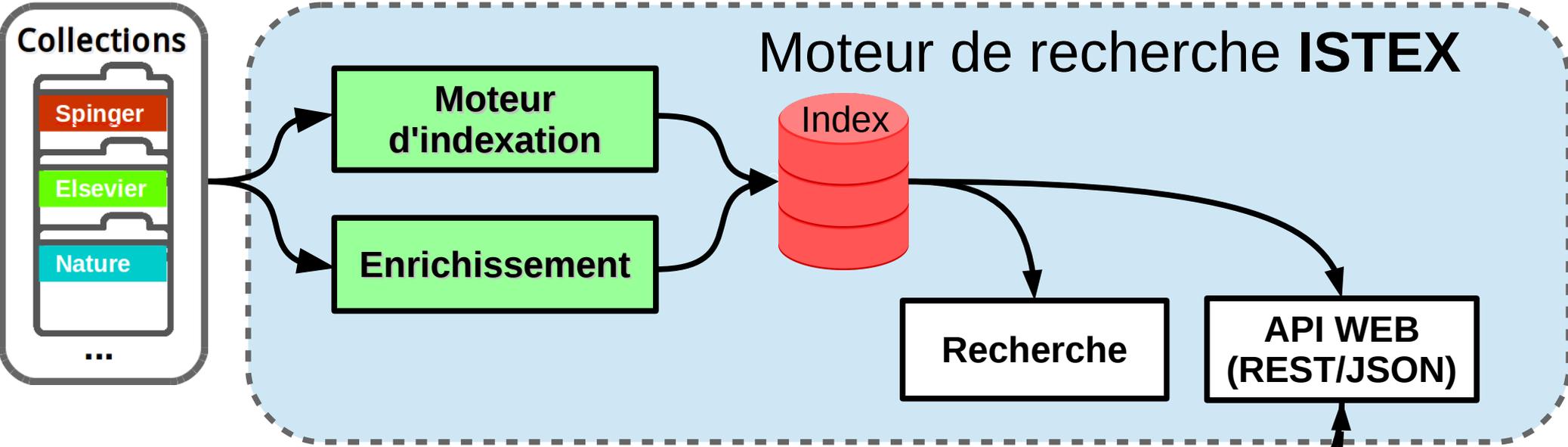
```
https://api.istex.fr/document/?q=Gerard
%20Mourou%20laser&
facet=corpusName[*]&
Size=100&
rankBy=qualityOverRelevance&output=*&stats
```

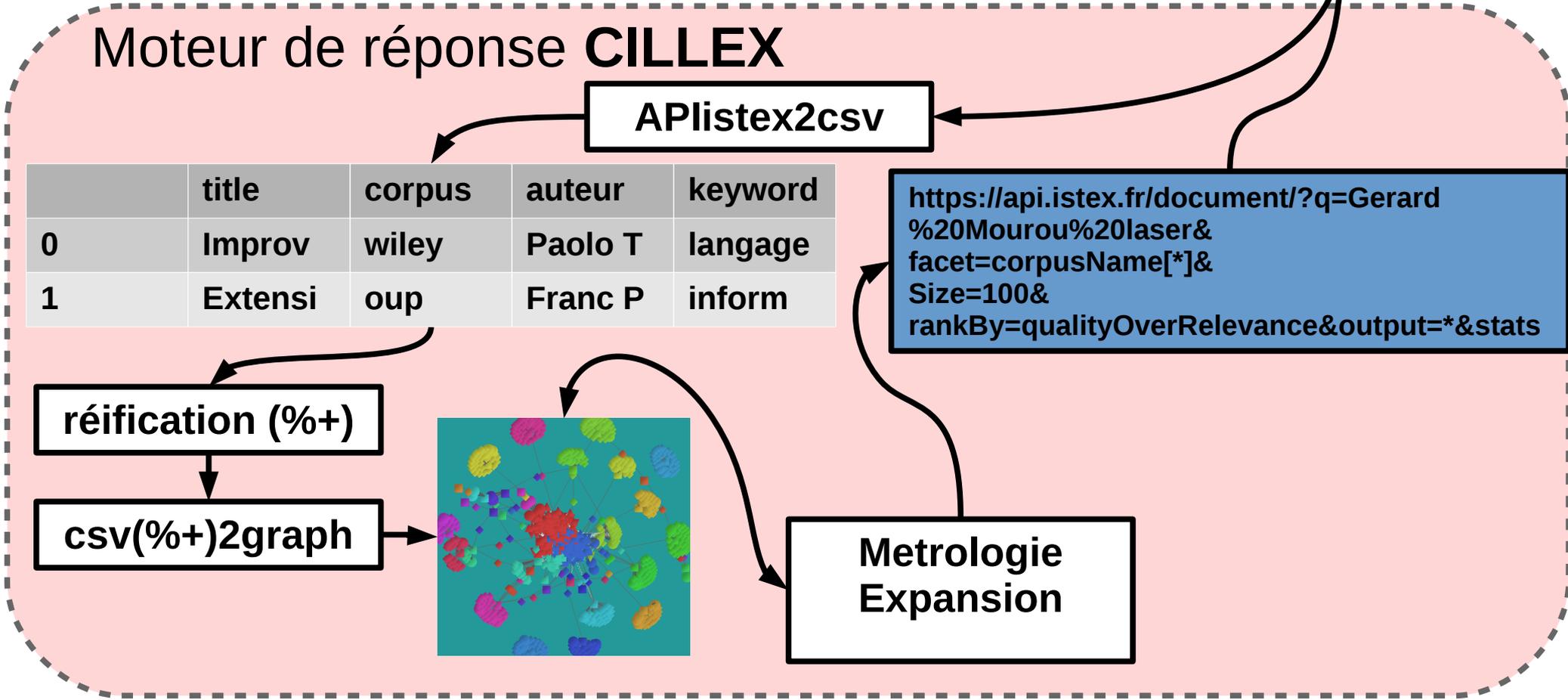
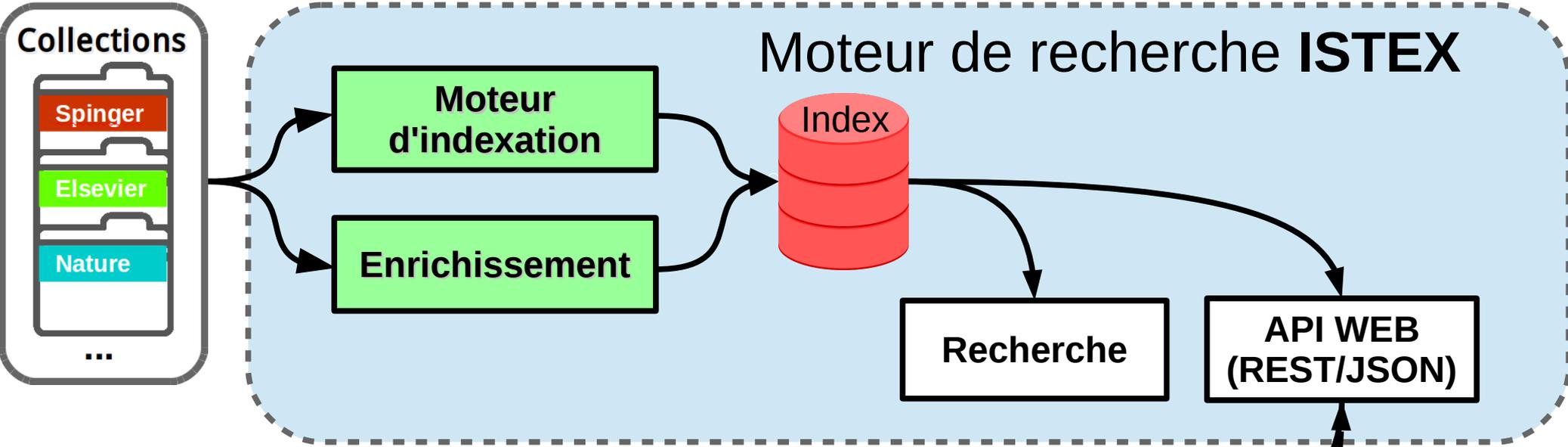


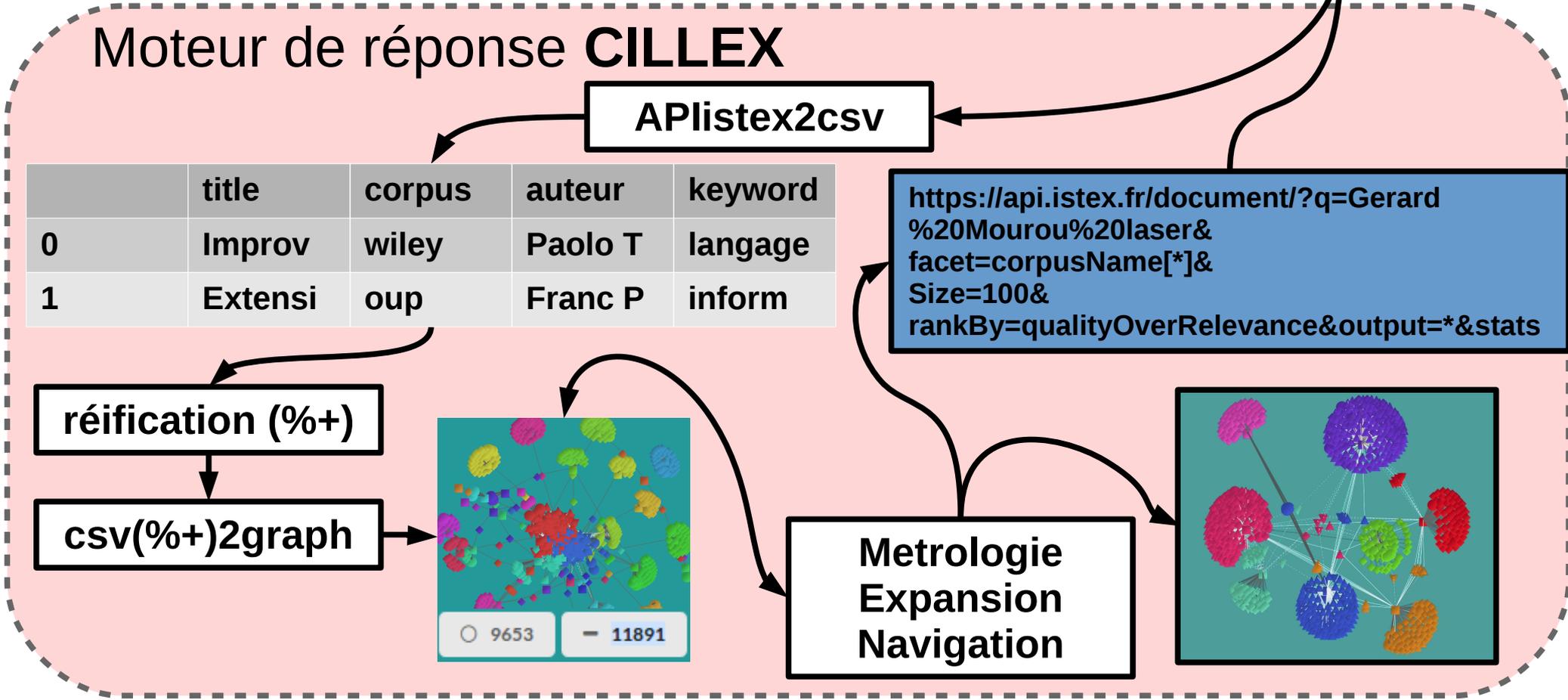
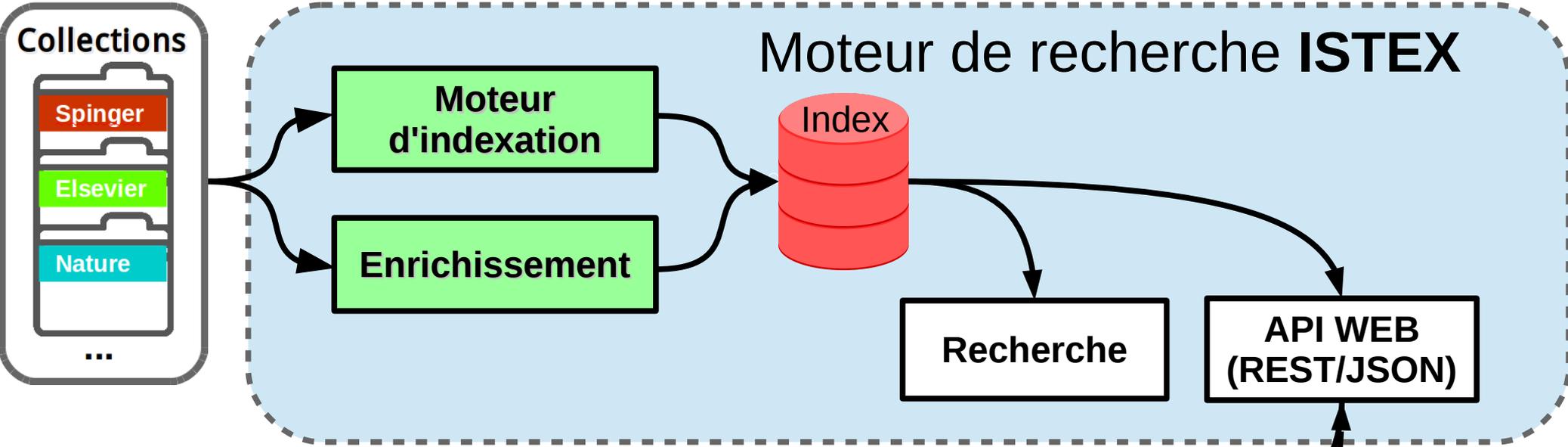




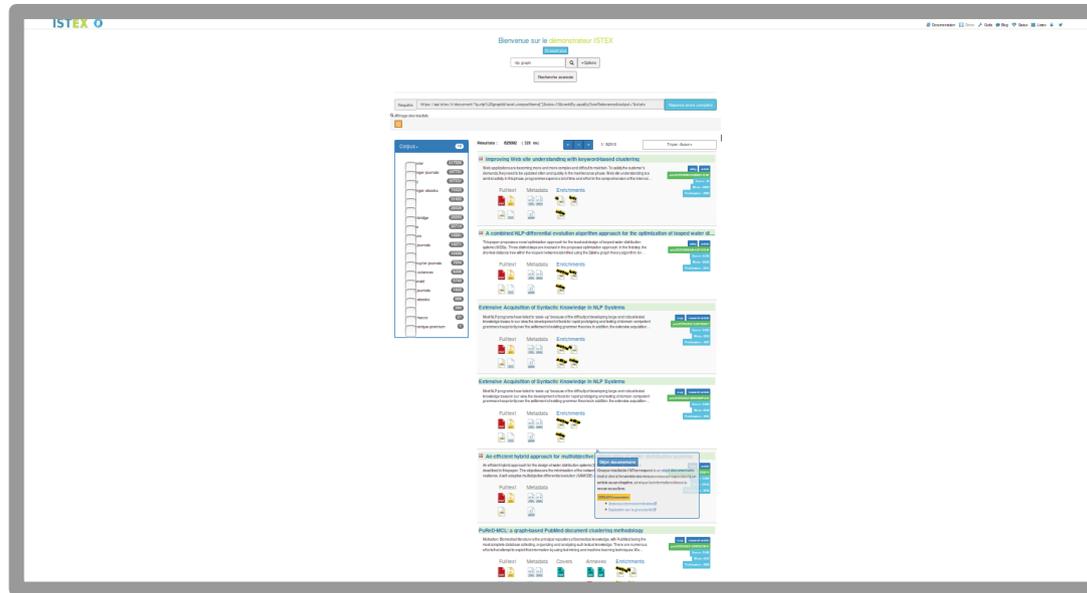




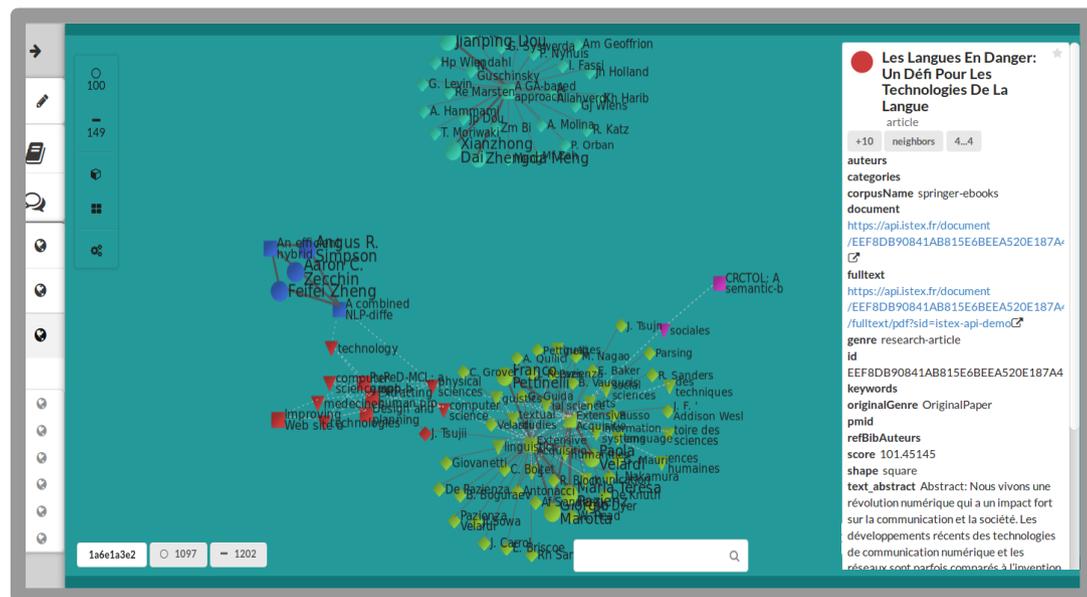




# Présentation des résultats Sous forme de liste



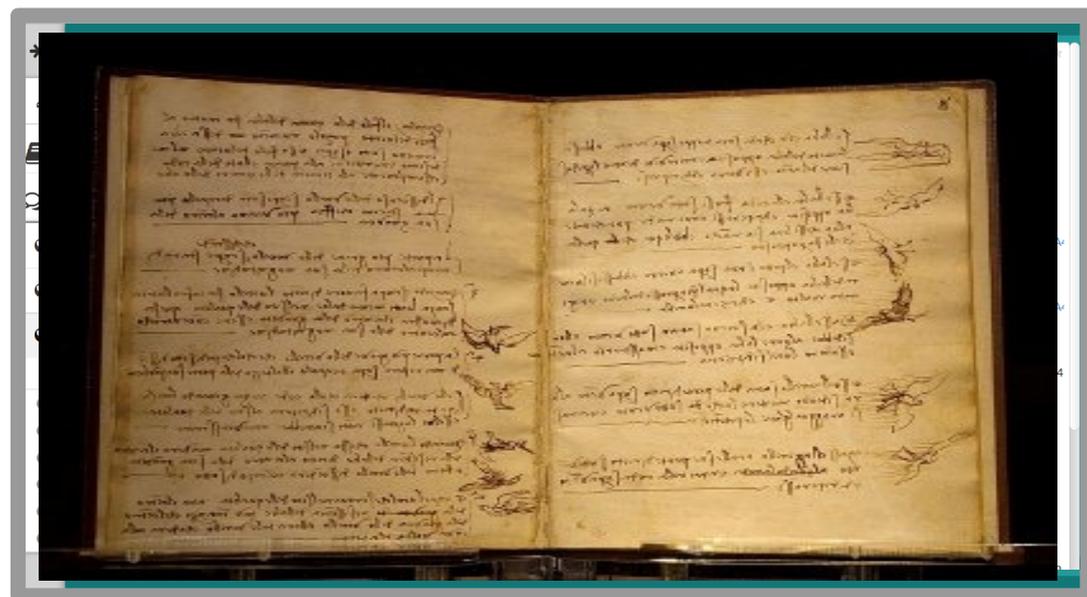
# Présentation des résultats Sous forme de carte



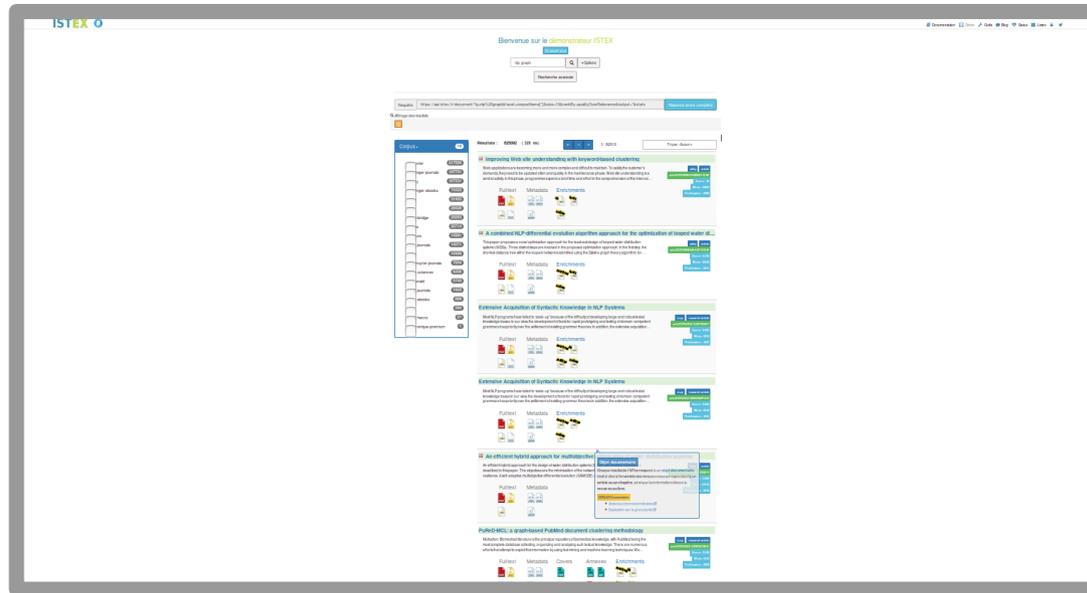
# Le Rouleau



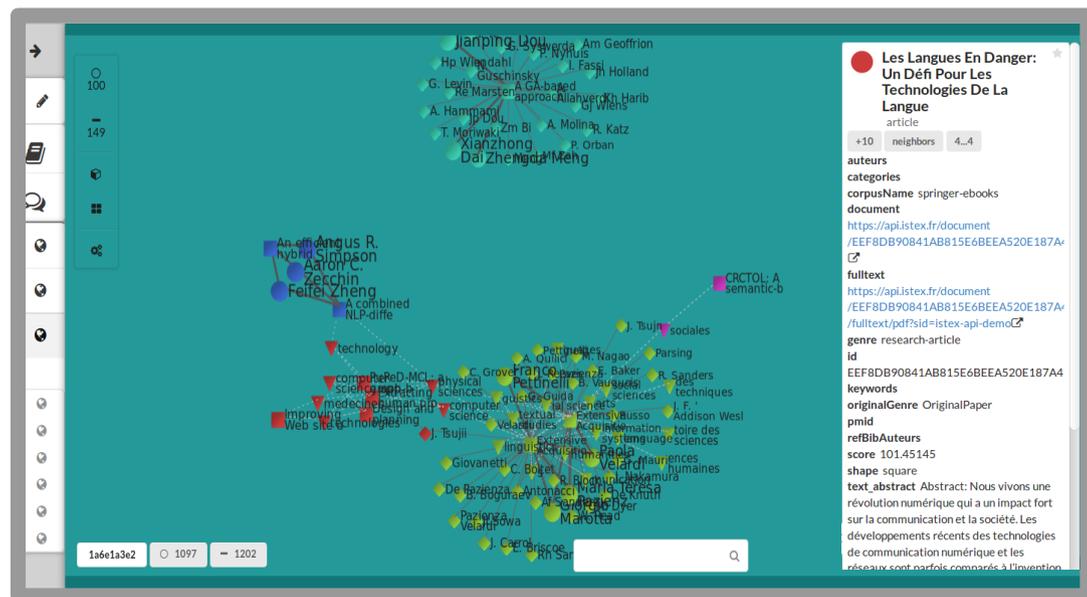
# Le Codex



# Présentation des résultats Sous forme de liste



# Présentation des résultats Sous forme de carte

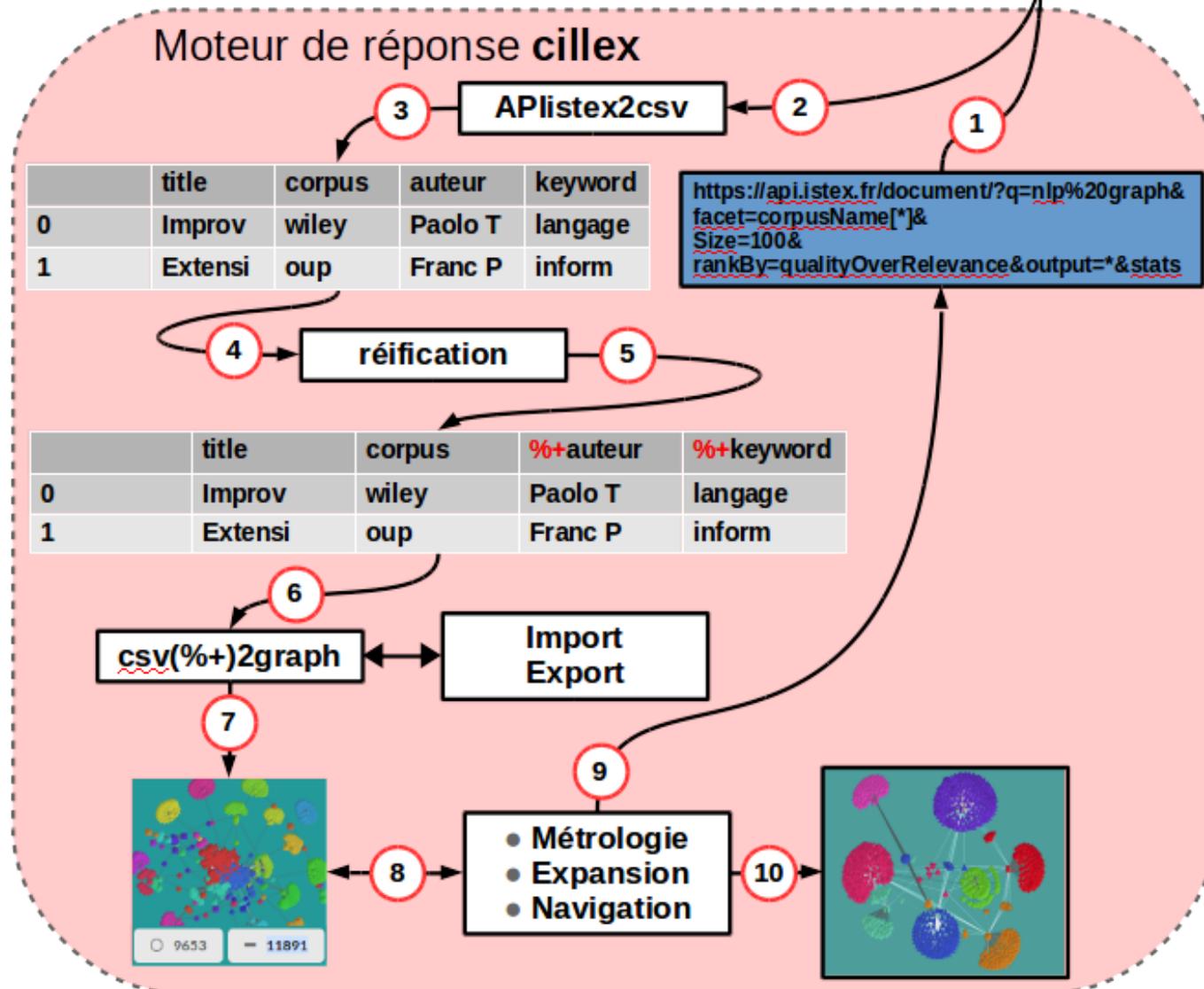
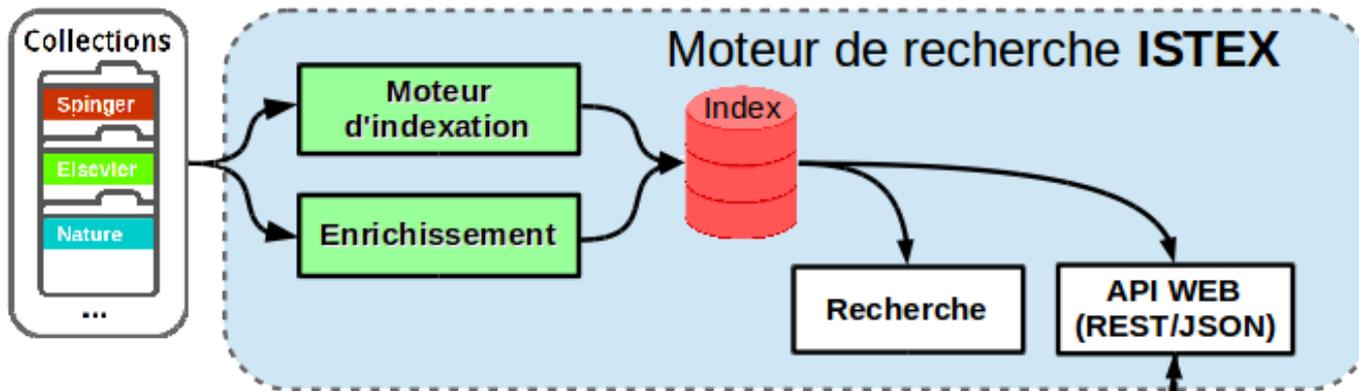


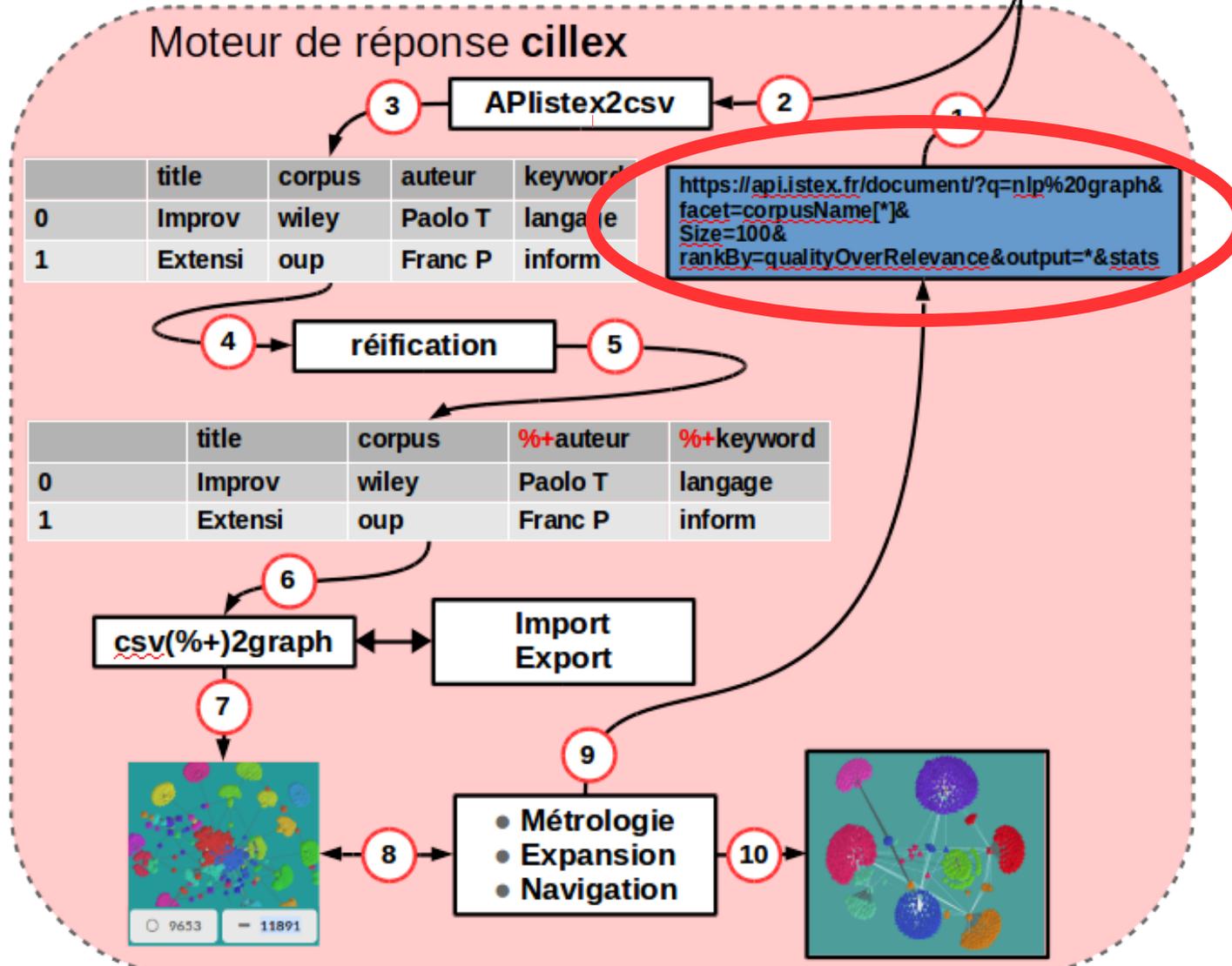
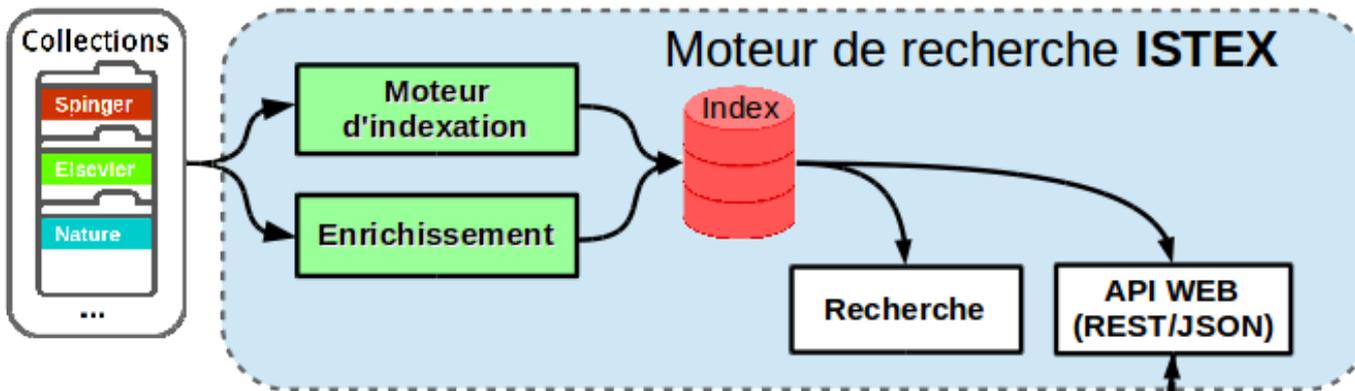


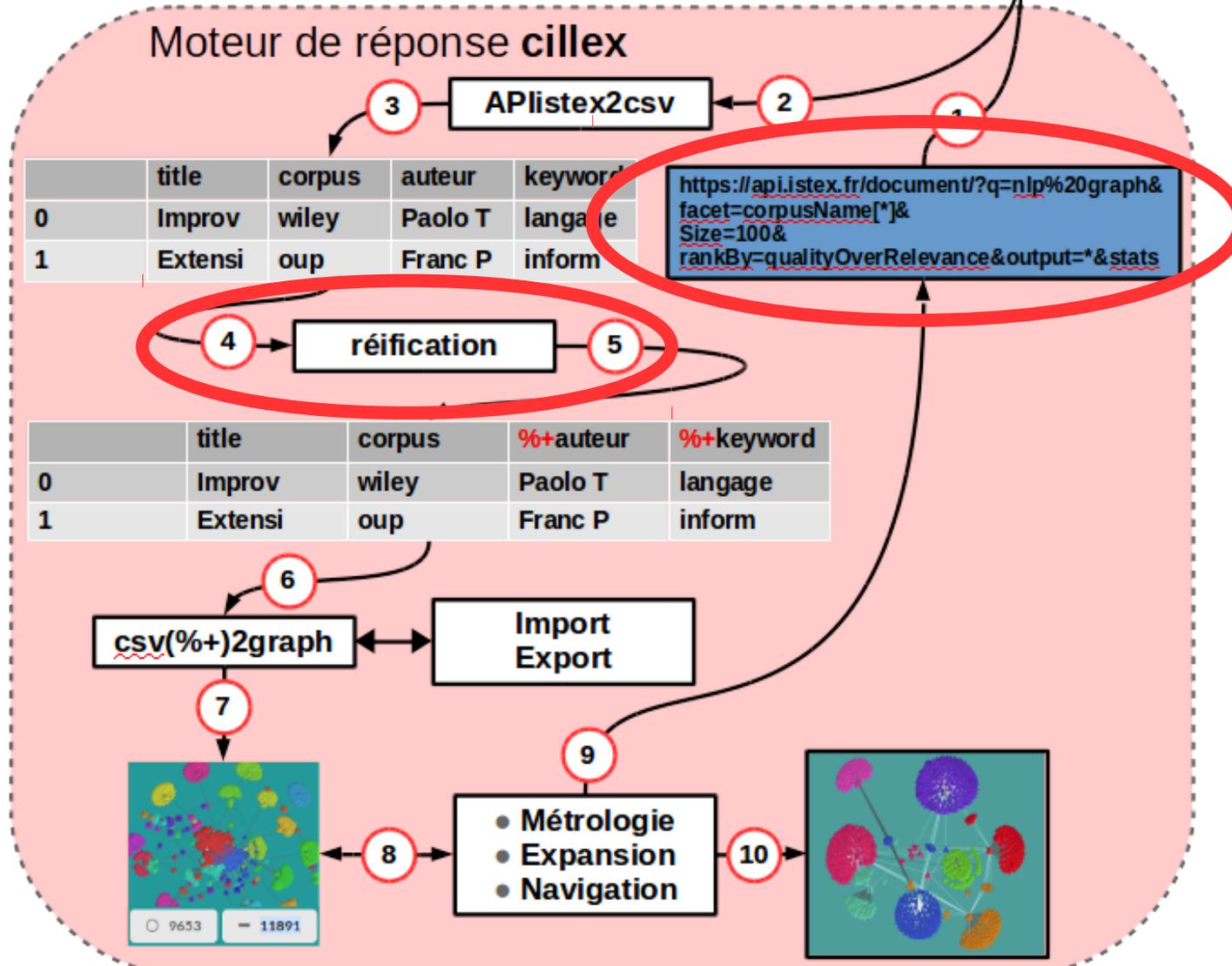
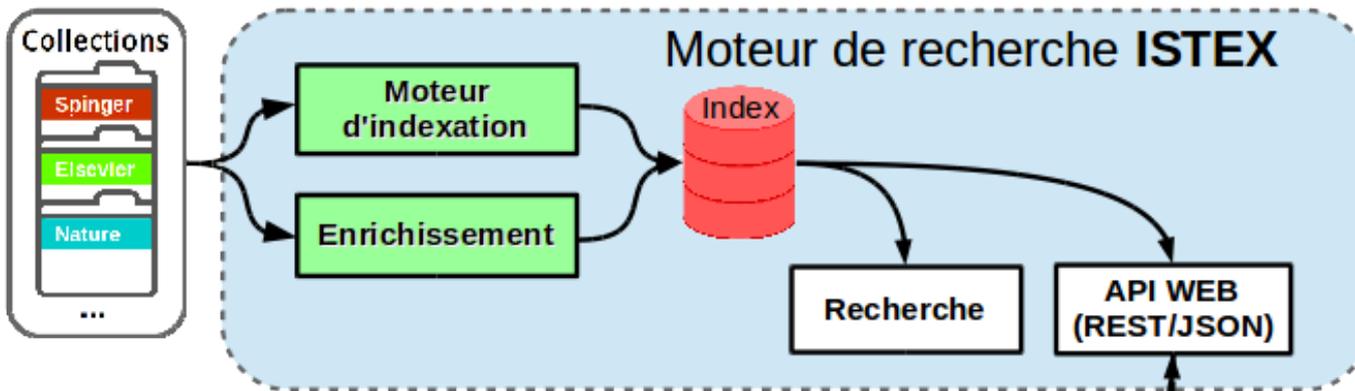


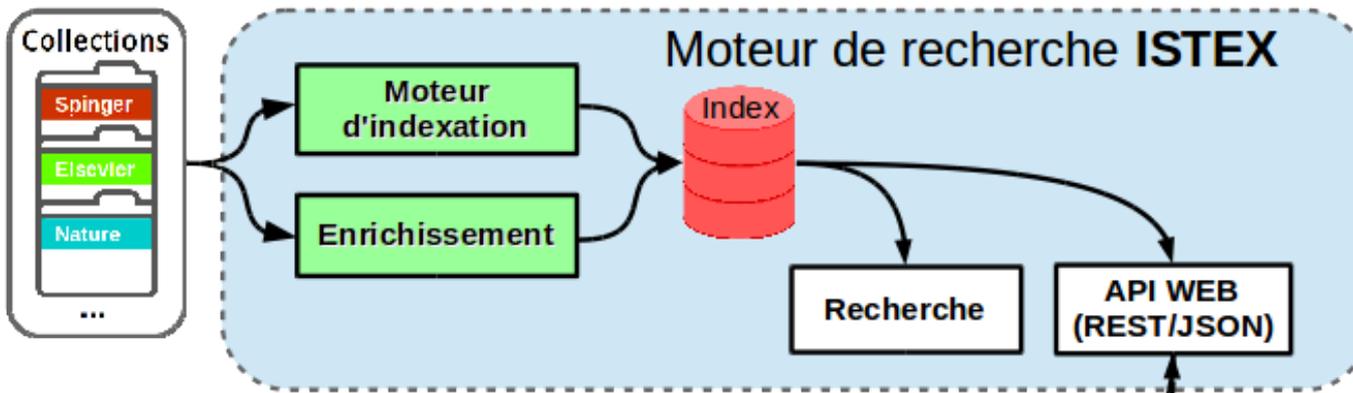
- Présentation
- Chaîne de traitement de l'information
- **Démonstration en ligne :**  
<http://cillex.padagraph.io/>



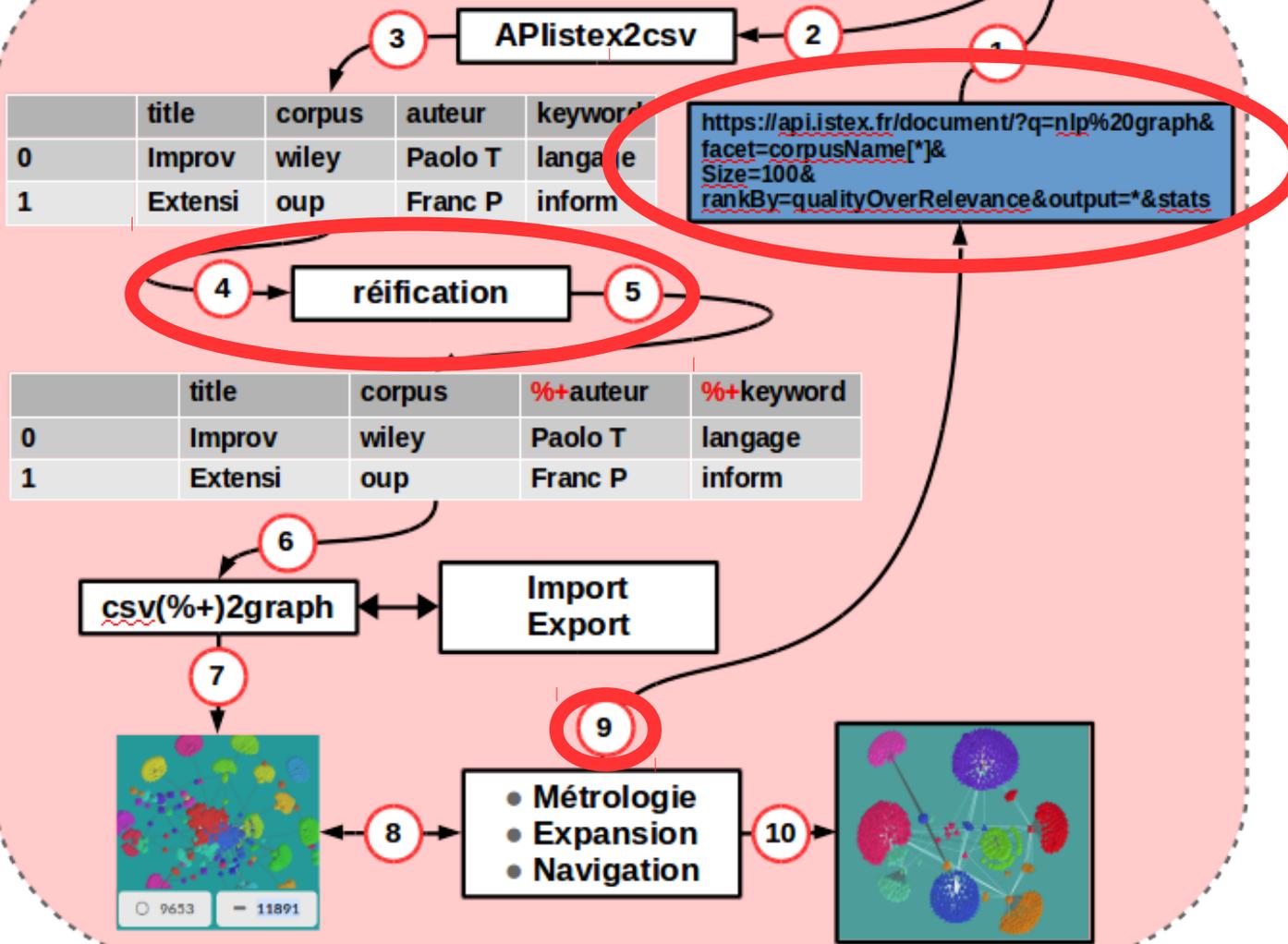


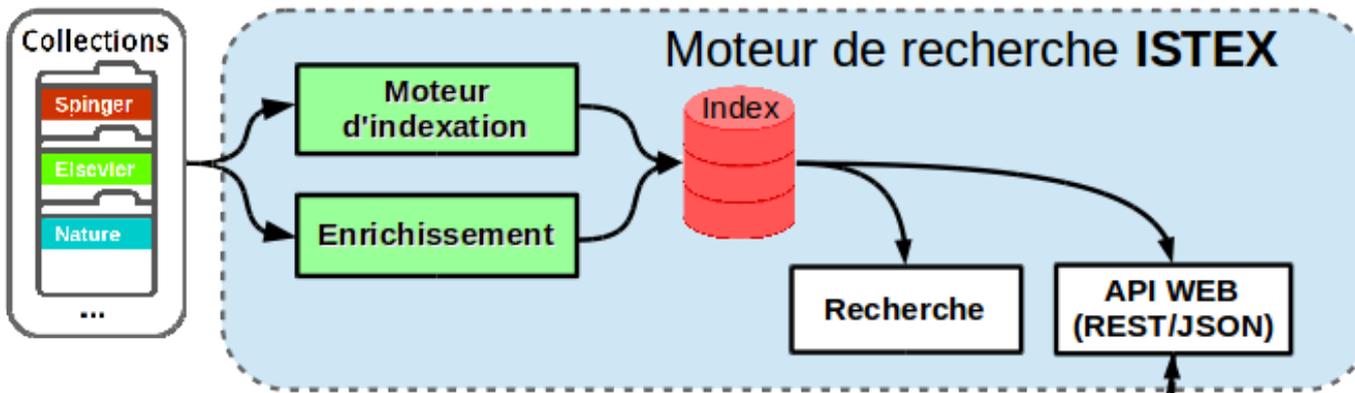




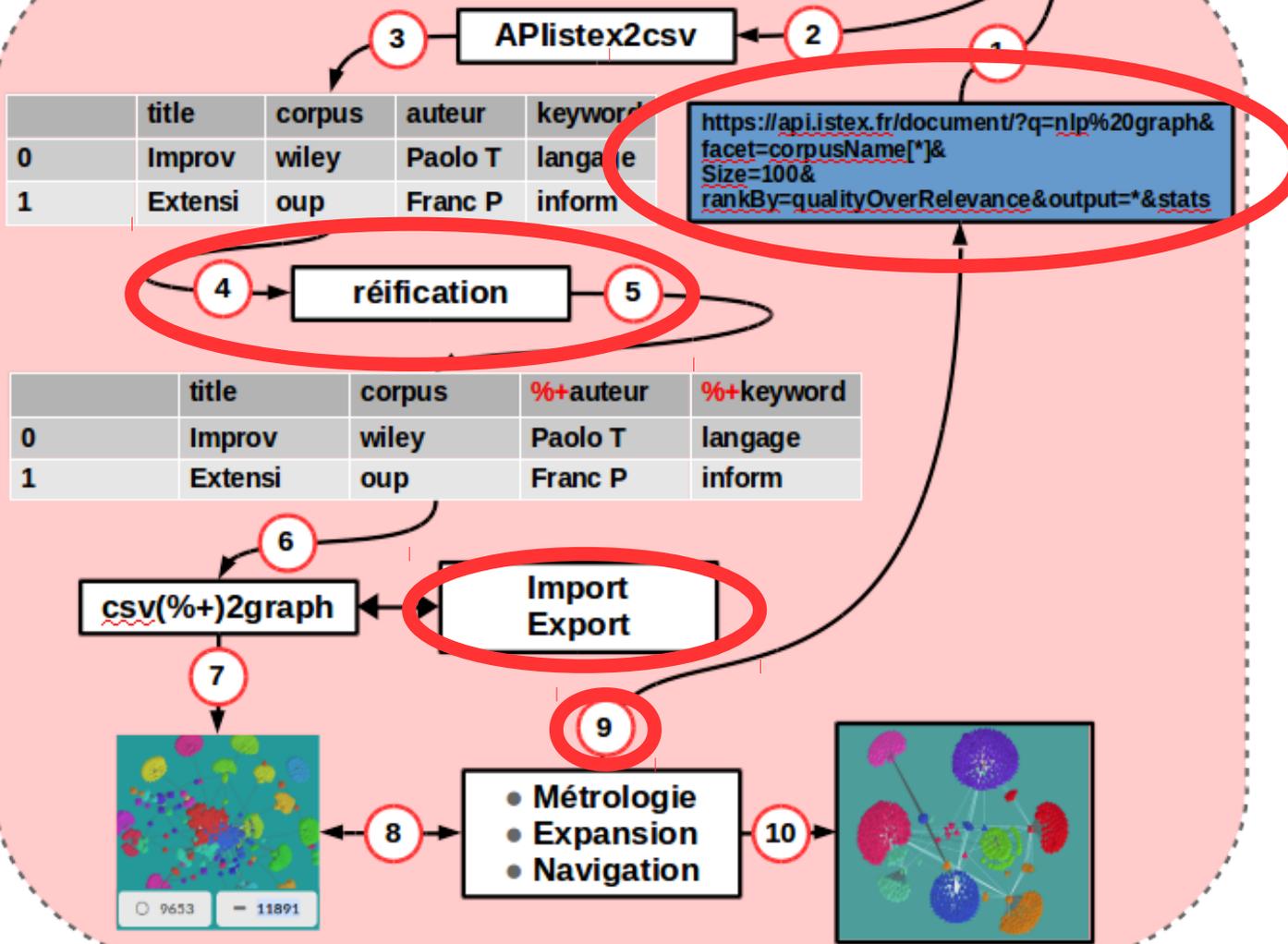


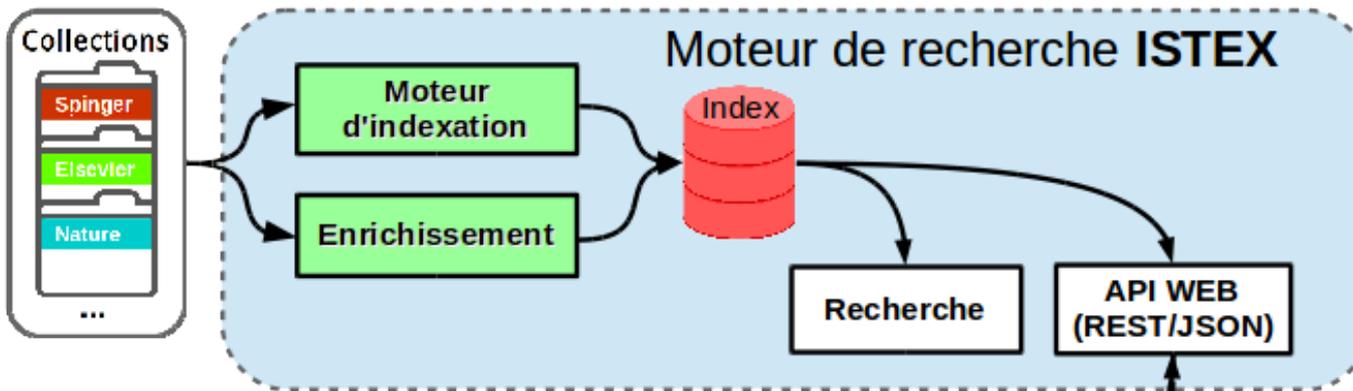
### Moteur de réponse cillex



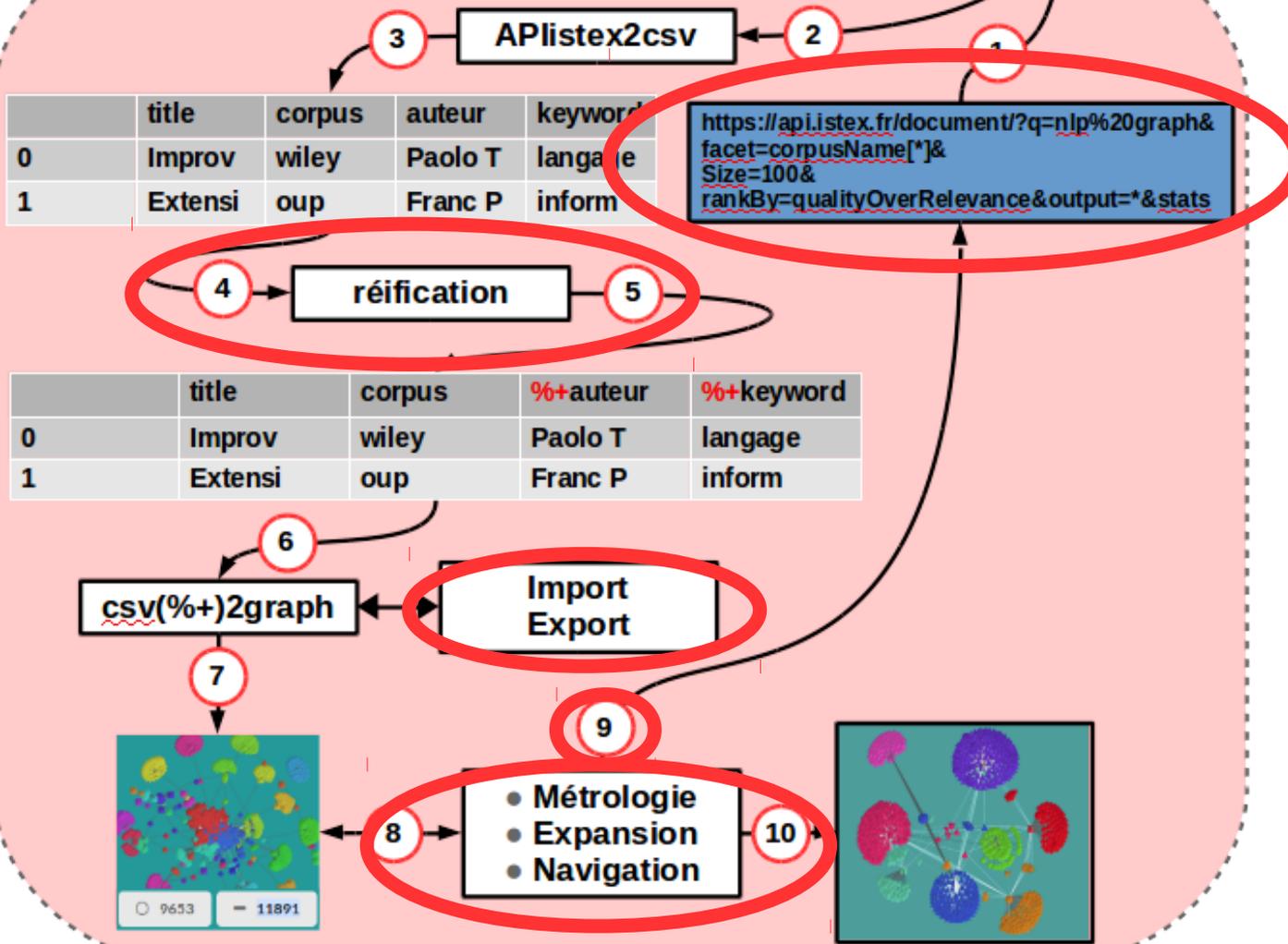


### Moteur de réponse cillex





### Moteur de réponse cillex



	title	corpus	auteur	keyword
0	Improv	wiley	Paolo T	langage
1	Extensi	oup	Franc P	inform

```
https://api.istex.fr/document/?q=nlp%20graph&facet=corpusName[*]&Size=100&rankBy=qualityOverRelevance&output=*&stats
```

	title	corpus	%+auteur	%+keyword
0	Improv	wiley	Paolo T	langage
1	Extensi	oup	Franc P	inform

cillex

CILLEX

**Cartographie** documentaire



cillex

CILLEX

**Cartographie** documentaire

  
UMR 5263  
COGNITION, LANGUES, LANGAGE, ERGONOMIE

