

Le dictionnaire des francophones, une plateforme lexicographique contributive et sémantique.

Conférence Plateforme AFIA / IC 2021

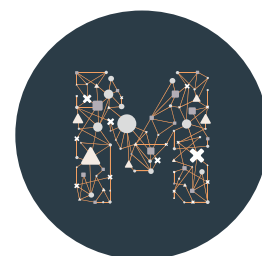
Freddy LIMPENS, Mnemotix

freddy.limpens@mnemotix.com



**MINISTÈRE
DE LA CULTURE**

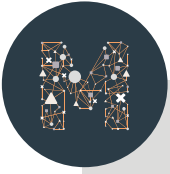
*Liberté
Égalité
Fraternité*



mnemotix

Donnons du sens à vos données

INSTITUT
INTERNATIONAL
POUR LA
FRANCOPHONIE



Le Dictionnaire des francophones (DDF)

www.dictionnairedesfrancophones.org

- **Réalisé** par Mnémotix pour le compte du **ministère de la culture**, porté par **2IF**
- **Inauguré le 16 mars**
- **Ambition:** proposer une **plateforme contributive ouverte à tous les francophones**
- **Partenaires:** CNRS/Humanum pour l'hébergement, l'AUF, et l'OIF et nombreux autres partenariats ([cf site](#))



Dictionnaire des francophones

🔍 Écrivez le terme recherché ici

France

📍 Utiliser ma position actuelle

Rechercher

—

Comprendre et partager
les mots avec plusieurs dictionnaires
du français parlé partout dans le monde.
Participer à l'enrichissement collectif

—

Enrichir le DDF 

Pourquoi est-ce un projet innovant?

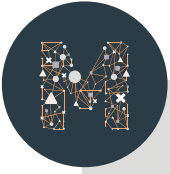
1. **Point d'accès unique** à un ensemble de ressources lexicographiques transcrites en RDF
2. **Modèle contributif basé sur l'addition** plutôt que la fusion des contributions
3. Repose sur sur une **architecture résiliente et haute disponibilité basée entièrement sur des graphes RDF** basé sur le middleware sémantique open source **Synaptix**





Les ressources lexicographiques importées

Nom	Source	Format import	Nombre mots	Nb définitions	Licence
Wiktionnaire francophone	fr.wiktionary.org/	XML, WikiML	1 830 093	2 390 494	CC BY SA
Inventaire des particularités lexicales du français en Afrique noire	Danièle Racelle-Latin (dir.), <i>Ed. Vanves (1988)</i>	PDF	6 420	7 906	CC BY SA
Belgicismes - Inventaire des particularités lexicales du français en Belgique	Willy Bal <i>et al.</i> , <i>Ed. Duculot (1994)</i>	CSV	2 150	2 443	CC BY SA ND
BDLP - Base de données lexicographiques panfrancophone	www.bdlp.org/	CSV	14 494	20 776	CC BY SA ND
ASOM - Dictionnaire des synonymes, des mots et expression des français parlés dans le monde	www.dictionnaire-synonymes-francophones.fr/	CSV	2 374	3 438	CC BY SA
DRF - Dictionnaire des régionalismes de France	Pierre Rézeau (dir.), <i>Ed. Duculot (2001)</i>	CSV	2 070	3 000	CC BY SA ND
GDT - Grand dictionnaire terminologique	gdt.oqlf.gouv.qc.ca	CSV	4 321	4 494	CC BY SA
DDF - Dictionnaire des francophones	dictionnairedesfrancophones.org	RDF	460	495	CC BY SA



Exemple : “canne”

/form/canne

- **8 aires d’usages**, réparties sur **4 continents** (code couleur)
- **19 définitions**
 - avec les **flexions** du verbe “canner”
 - avec **2 variantes d’un même sens** (canne comme accessoire de sport ou de combat)
- **3 sources :**
 - Wiktionnaire (14)
 - BDLP (4)
 - Inventaire (1)

Terme recherché : canne

[Définition]

Normandie

nom, féminin
Bidon (pour le lait).

Source : Wikit

Suisse

nom, féminin
Long manche de bois à l’extrémité plate et légèrement recourbée, faisant angle avec la hampe, servant à manipuler le puck* au hockey sur glace. (fiche originale)

Source : BDLP

Cameroun, Togo

nom, féminin
Tige de canne à sucre.

Source : Inv

Québec

nom, féminin
Boîte de conserve (de l’anglais *can*).

Source : Wikit

Louisiane

nom, féminin
Pot. (fiche originale)

Source : BDLP

Louisiane

nom, féminin
Paille (pour boire). (fiche originale)

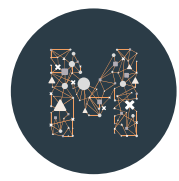
Source : BDLP

Louisiane, Madagascar, Réunion

nom, féminin
Canne à sucre. (fiche originale)

Monde francophone

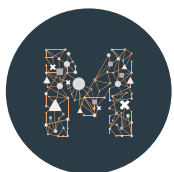
nom, féminin
Instrument long et cylindrique dont les usages diffèrent.



Modélisation

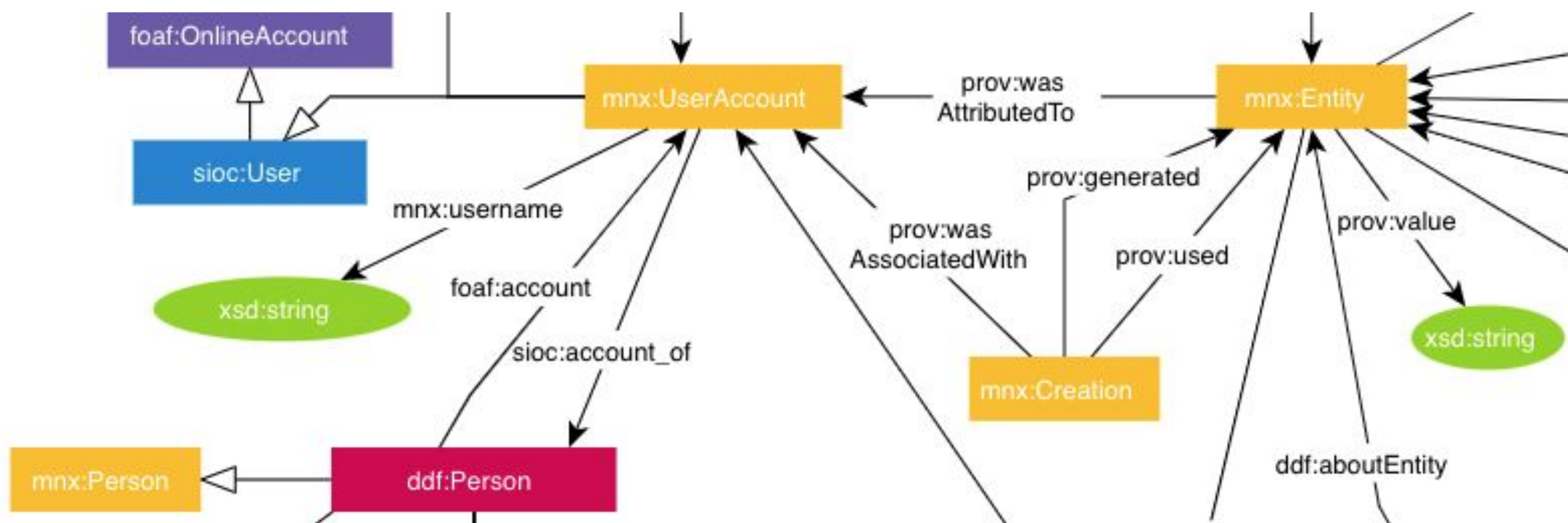
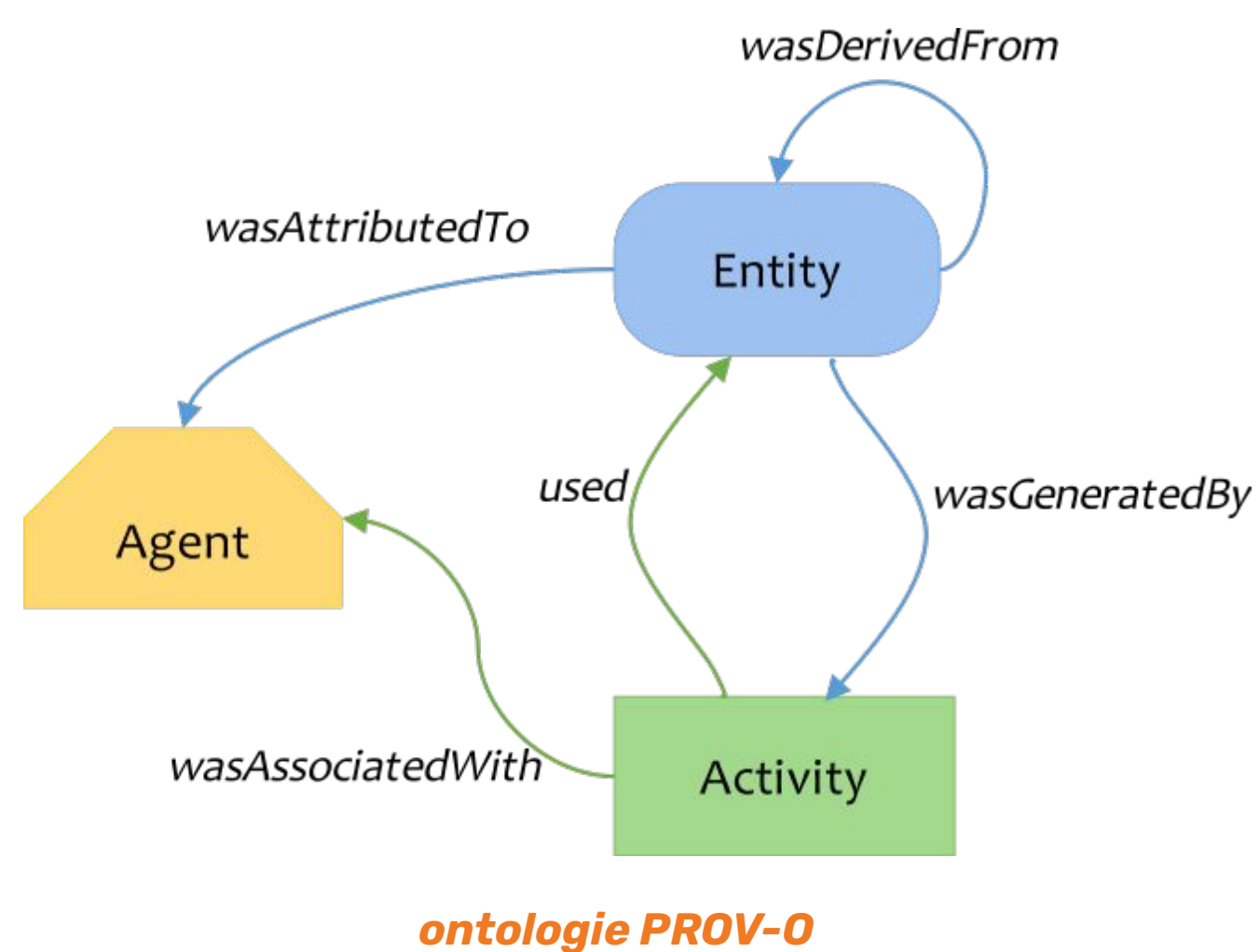
Namespace & doc.: <http://data.dictionnairedesfrancophones.org/ontology/ddf>

- **1 ontologie OWL + 22 vocabulaires SKOS**
- **Basé sur Ontolex - Lemon + Lexinfo**
pour articulation entre :
 - **formes** (mots) e.g. “*canne*”
 - **entrée lexicale**, eg. “*canne*” *nom féminin*
 - **définition** (sens), eg. “tige de canne à sucre” + **exemple**
 - **propriété** grammaticale (genre, etc.)
 - **marqueur** de sens (registre, etc)
- **Extension DDF pour :**
 - **localisation** sur **forme**, **sens**, et **exemples**
 - **autres marqueurs** de sens (glossaire, connotation, etc.)
 - **relations sémantiques** (entre mots et entre sens)
- **Étymologies**, ainsi que **notes d’usages** sur les **formes** et les **sens**, modélisées comme des **discussions** avec **SIOC**



Modèle contributif

- **Actions contribution réifiées** à l'aide d'une extension de **PROV** et son modèle **Agent-Activity-Entity**
- Actions de **vote / signalement** étendent **schema:AggregateRating**



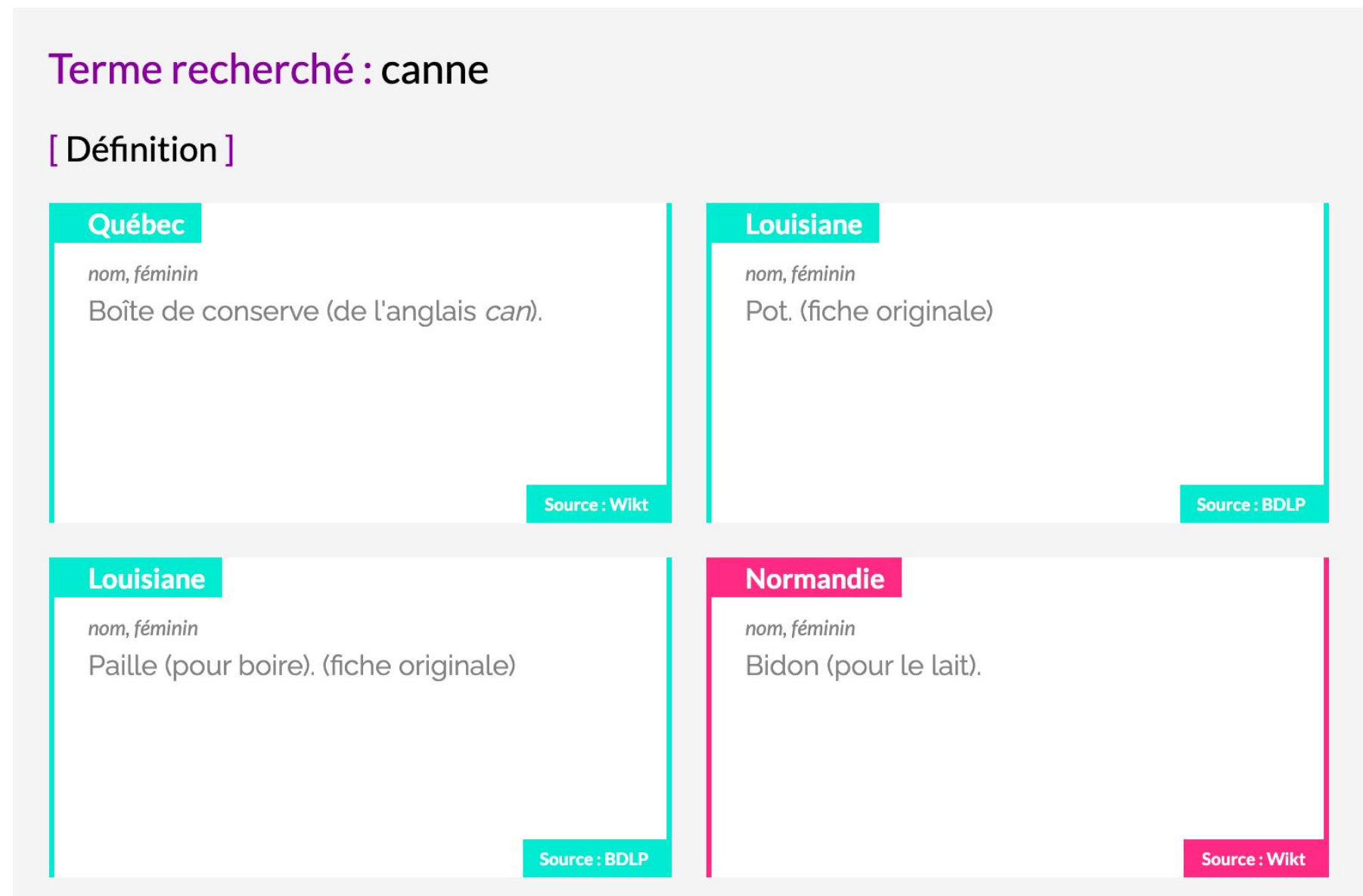
ontologie DDF, module "contribution"



Contribution additive

Contribution additive = pas de fusion des contributions

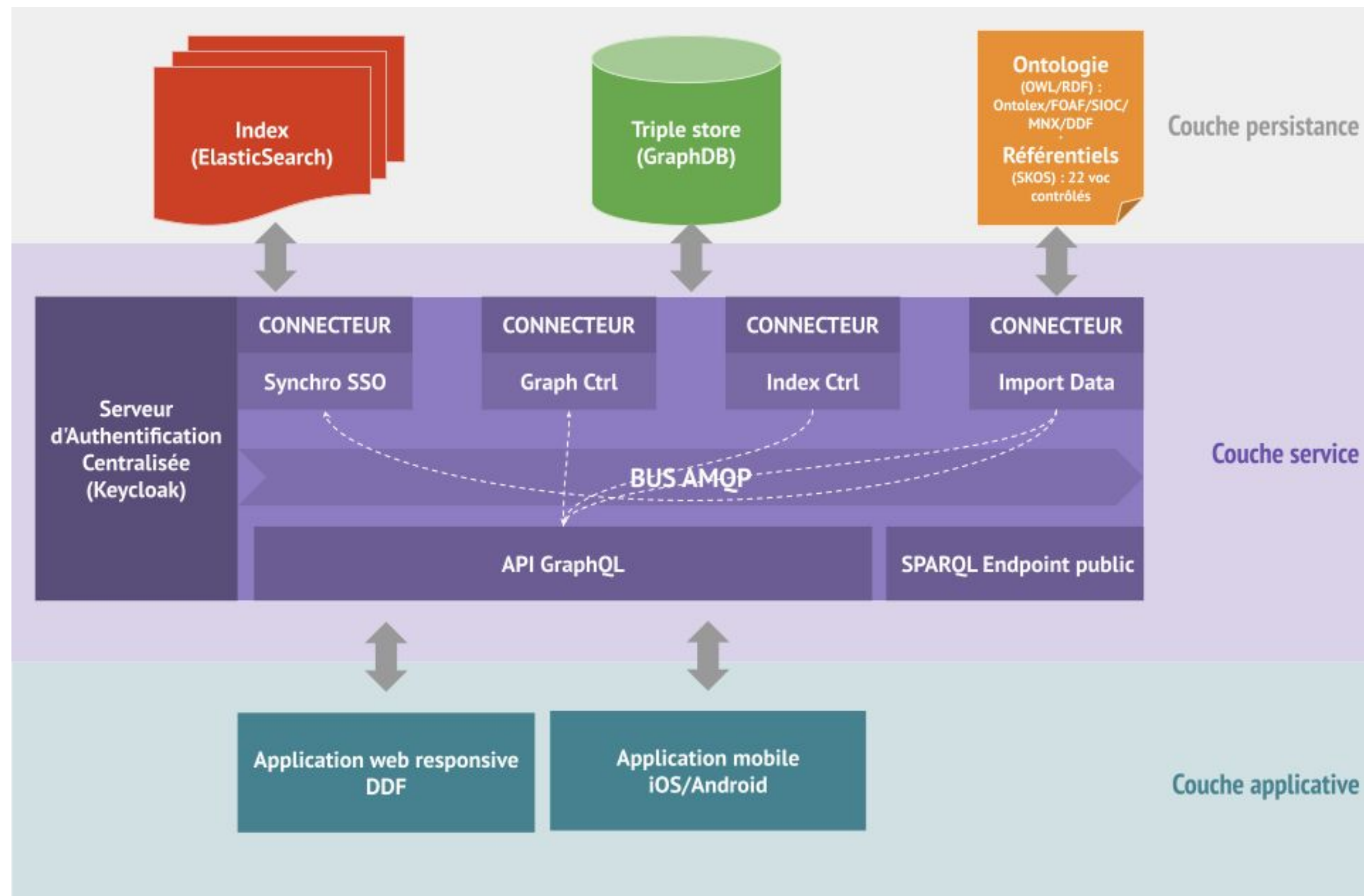
- **modération *a posteriori*** à l'aide de la communauté (signalement, vote) et d'utilisateurs cooptés au rang d'opérateurs
- **nécessité de trier les résultats de recherche** en calculant **score**:
 - ◆ basé sur la localisation de l'utilisateur
 - ◆ basé sur le nombre de vote positif et/ou de signalements

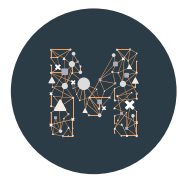




Architecture

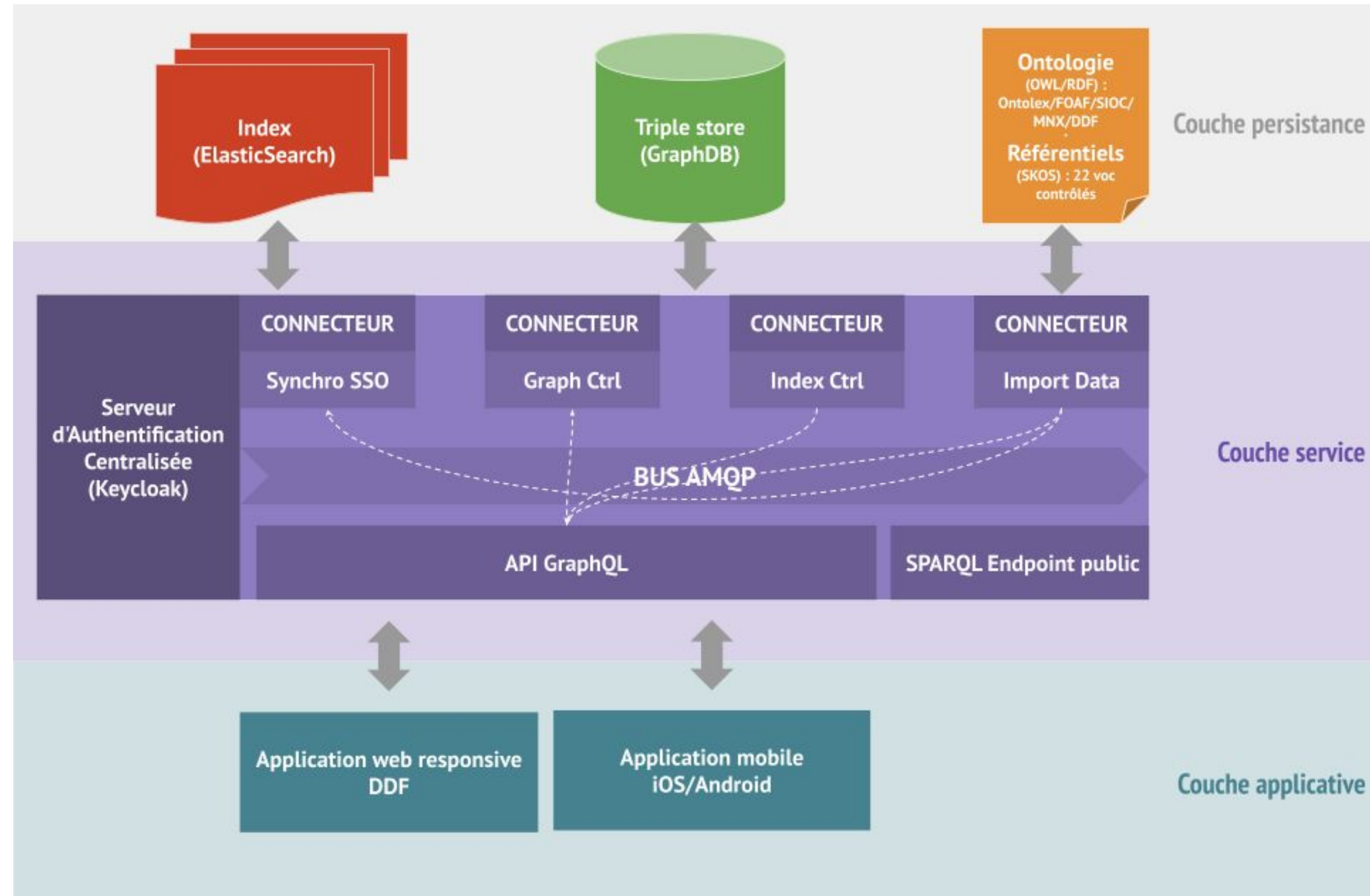
- basée sur **Synaptix** le **middleware sémantique** articulant “**micro-services**”
- Processus **communiquent par messages** via bus **AMQP**
- files de messages **consommés progressivement**
- **amortit** les **afflux ou requêtes massifs** de données

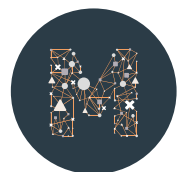




Architecture

- **triple store : graphDB**
- données sont synchronisés dans un index (ElasticSearch) pour accès rapide
- Exposition des données via API GRAPHQL
- 3 applis consomment données:
 - appli web responsive développée en React.js
 - 2 applis mobiles iOS+Android





Démo

<https://www.dictionnairedesfrancophones.org/>

“**dictionnaire des francophones**” dans un moteur de recherche



Dictionnaire des francophones

🔍 Écrivez le terme recherché ici

France

📍 Utiliser ma position actuelle

Rechercher

Comprendre et partager
les mots avec plusieurs dictionnaires
du français parlé partout dans le monde.
Participer à l'enrichissement collectif

Enrichir le DDF 



Conclusion et perspectives d'évolution

Take away message

- Application web basée sur **stack sémantique**, responsive et contributive, qui embarque **middleware open source Synaptix**
- Agrégation de multiples **ressources lexico en RDF, dont Wiktionnaire**
- Plateforme **contributive ouverte à tous les francophones**, sans agrégation des contributions grâce **stratégie de modération** et de **tri des données**

Perspectives

- Nouvelle ressource fin 2021, France Terme
- début 2022 nouvelle version avec
 - contribution complète
 - ergonomie et accessibilité améliorée
 - exploration de listes de termes
- courant 2022:
 - espaces de discussions
 - prononciations audios
 - projections cartographiques des aires d'usage



Contact :

+33 6 16 22 71 18
contact@mnemotix.com
@mnemotix.com

Siège social :

318 avenue de la Carraire
06210 MANDELIEU-LA NAPOULE

Bureaux :

Hôtel d'entreprises CréaCannes Lerins
Bâtiment 1, allée Maurice Bellonte
Parc d'Activités de l'Échangeur
06210 MANDELIEU-LA NAPOULE