



## Indexation de ressources pédagogiques intelligente et assistée

Commande : Ministère de l'éducation nationale - Bureau du soutien à l'innovation numérique et à la recherche appliquée

Réalisation : Joachim Dornbusch, Rosa-Maria Gomez de Regil, Jean Delahousse

# Contexte

Le Gestionnaire d'Accès aux Ressources Numérique (GAR - Education nationale) gère un catalogue de ressources pédagogiques numériques mises à la disposition des établissements scolaires.

Les notices descriptives des ressources pédagogiques sont normalisées selon le standard ScoLOMFR, déclinaison française du standard international LOM (Learning Object Metadata).

La description d'une ressource pédagogique conformément au standard ScoLOMFR est une tâche complexe, l'utilisateur est confronté à de nombreux champs, il doit choisir des valeurs dans des vocabulaires de grande taille.

Les informations à donner sont en général déjà présentes dans la présentation de la ressource dans le catalogue de l'éditeur.



vocabulaires skos  
schéma d'annotation

# Objectifs du projet IRPIA

Valider dans un POC la possibilité de simplifier la saisie des notices des ressources pédagogiques par les éditeurs

Tester l'aide apportée par l'IA dans les tâches de saisie d'une notice en combinant le machine learning (transformers) et les graphes de connaissance.



# Approche

Simplification du modèle ScoLOMFR de description des ressources conforme à l'alimentation du GAR

Suggestion par extraction de contenus

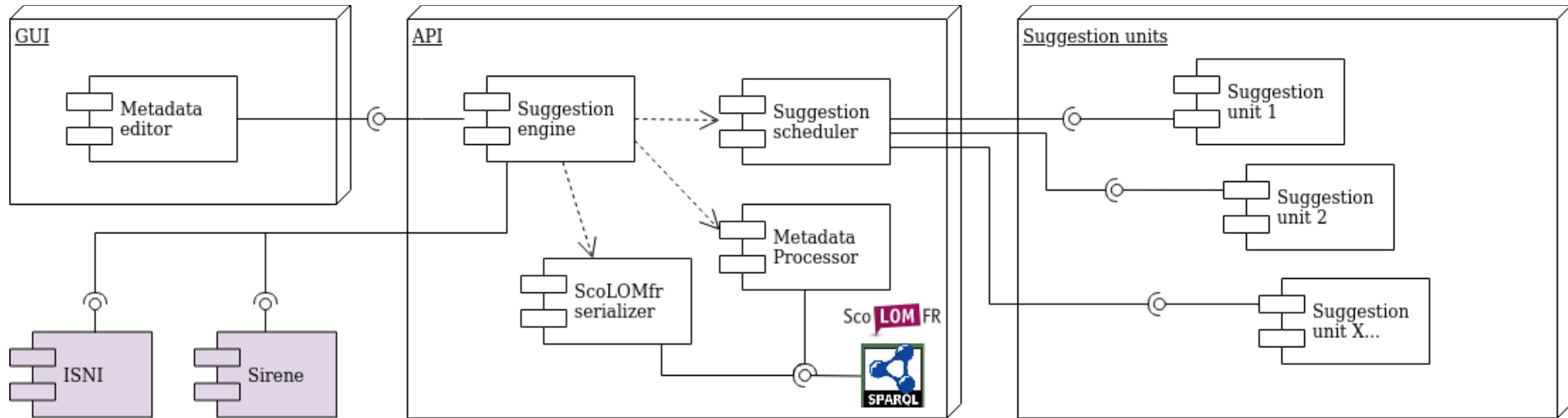
Suggestion par Machine Learning

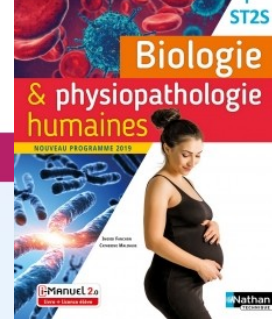
Suggestion par Inférence

Enrichissement par l'appel aux APIs Sirene et ISNI

Optimisation UX/UI

# Architecture du POC IRPIA





## Prototype d'assistant intelligent pour la description de vos ressources pédagogiques

Vous souhaitez indexer une ressource pédagogique afin que vos utilisateurs la trouvent ? Essayez l'assistant IRPIA !  
En mobilisant des technologies d'assistance mues par l'intelligence artificielle, IRPIA vous facilite la tâche.

COMMENCER



# Entraîner le modèle

Grands ensembles de notices de ressources pédagogiques décrites manuellement par les éditeurs de manuels scolaires et les professeurs dans **Édubase** et dans le **catalogue du GAR**.

Conversion de ces bases de métadonnées de ressources pédagogiques en données d'entraînement pour l'algorithmes de classification automatique.

Les prédictions du modèle sont proposées à l'utilisateur en tant que suggestions afin d'accélérer la saisie des métadonnées.

Utilisation de ChatGPT qui surclasse CamemBert

Les données d'entraînement et les modèles sont ouverts et disponibles pour tous

# Une annotation basée sur le graphe de connaissance ScoLOMFR

Canopé édite et publie un ensemble de vocabulaires contrôlés (SKOS) pour l'annotation des ressources documentaires.

Les vocabulaires utiles au projet sont chargés dans le triple store Fuseki.

Les valeurs dans la notice finale sont des URIs des vocabulaires ScoLOMFR : type de document, type de ressource éducative, domaine d'enseignement, niveau éducatif



# Enchaîner réseau de neurones et graphe de connaissance

“Mesure de la vitesse de propagation de la lumière”

Physique-Chimie

DOMAINES

- physique-chimie (enseignement spécialité)
- physique-chimie (cycle 4)

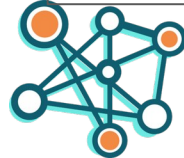
NIVEAU

Cycle 4

suggestion engine

metadata processor

metadata processor



## Optimiser l'UX/UI

“Chemin de fer” numéroté, qui incite l'utilisateur à respecter une marche à suivre, sans pour autant le contraindre : les badges numérotés restent en réalité cliquables, même lorsqu'ils sont grisés.

Cet effort d'allègement de la charge mentale et visuelle se poursuit ensuite au niveau graphique :

- l'information textuelle autour des champs est réduite et opte pour des polices de petite taille
- la charte de couleurs est limitée à deux ou trois valeurs pour induire une sensation de simplicité



# Conclusion

Une preuve de l'apport de la combinaison de technologies pour faciliter la saisie de notices de ressources pédagogiques

Une expérimentation dont les résultats sont valables dans les nombreux contextes de description de ressources

# Conclusion et perspectives

Industrialisation à venir du POC IRPIA au travers du projet **PIAGET** mené par le Ministère de l'éducation nationale et CANOPÉ qui sera mis à disposition des éditeurs pour l'alimentation du GAR

- Réflexion sur le choix du LLM le plus approprié pour une industrialisation dans le cadre d'un service public
- Meilleure exploitation du graphe de connaissance des vocabulaires ScoLOMFR

SemWeb.Pro

# Vos questions



MINISTÈRE  
DE L'ÉDUCATION  
NATIONALE  
ET DE LA JEUNESSE

*Liberté  
Égalité  
Fraternité*

Sco **LOM** FR

Nos remerciements à :

CANOPÉ

**GAR** LE GESTIONNAIRE  
D'ACCÈS  
AUX RESSOURCES  
NUMÉRIQUES