

Comment concevoir un moteur de recherche IA pour les données géographiques ?

Viet Hang Nguyen



Journée romande de la
géoinformation2025

Présentation



Viet Hang Nguyen

- Docteure en mathématiques & informatique, Grenoble, France
- Chercheuse postdoctorale à l'EPFL
- +10 ans en développement informatique
- Fondatrice & CEO de DeepAlps Sàrl



- Startup basée à St-Sulpice, VD
- Spécialisée en solutions basées sur l'IA
 - ✓ Chatbot IA (pour les communes, les événements)
 - ✓ Moteur de recherche IA
 - ✓ Solutions IA internes pour les entreprises, organisations

Système de recherche courant

Boîte de recherche par mot-clé / texte intégral + Des filtres

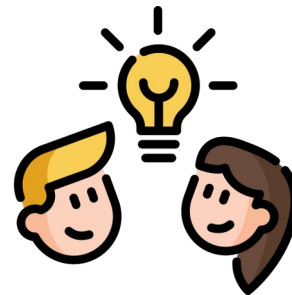
Problèmes courants

- **!** La recherche par mot-clé/texte intégral donne soit trop de résultats soit aucun résultat
- **!** Le critère qui intéresse l'utilisateur n'est pas prévu
- **!** Plus de filtres, filtres plus sophistiqués = plus compliqué à utiliser, plus compliqué à concevoir la UI
- **!!!** Le système ne permet pas d'apprendre le besoin utilisateur pour s'améliorer



Et si l'utilisateur peut simplement décrire ce qu'il/elle cherche ?

- ✓ La requête sera beaucoup plus expressive
- ✓ Le dev pourra apprendre quels critères sont importants pour les utilisateurs (et les supporter)
- ✓ L'interface utilisateur restera toujours aussi simple qu'une boîte de saisie textuelle
- ✓ Comme le dev n'est pas contraint par le design de la UI, il pourra intégrer beaucoup de filtres



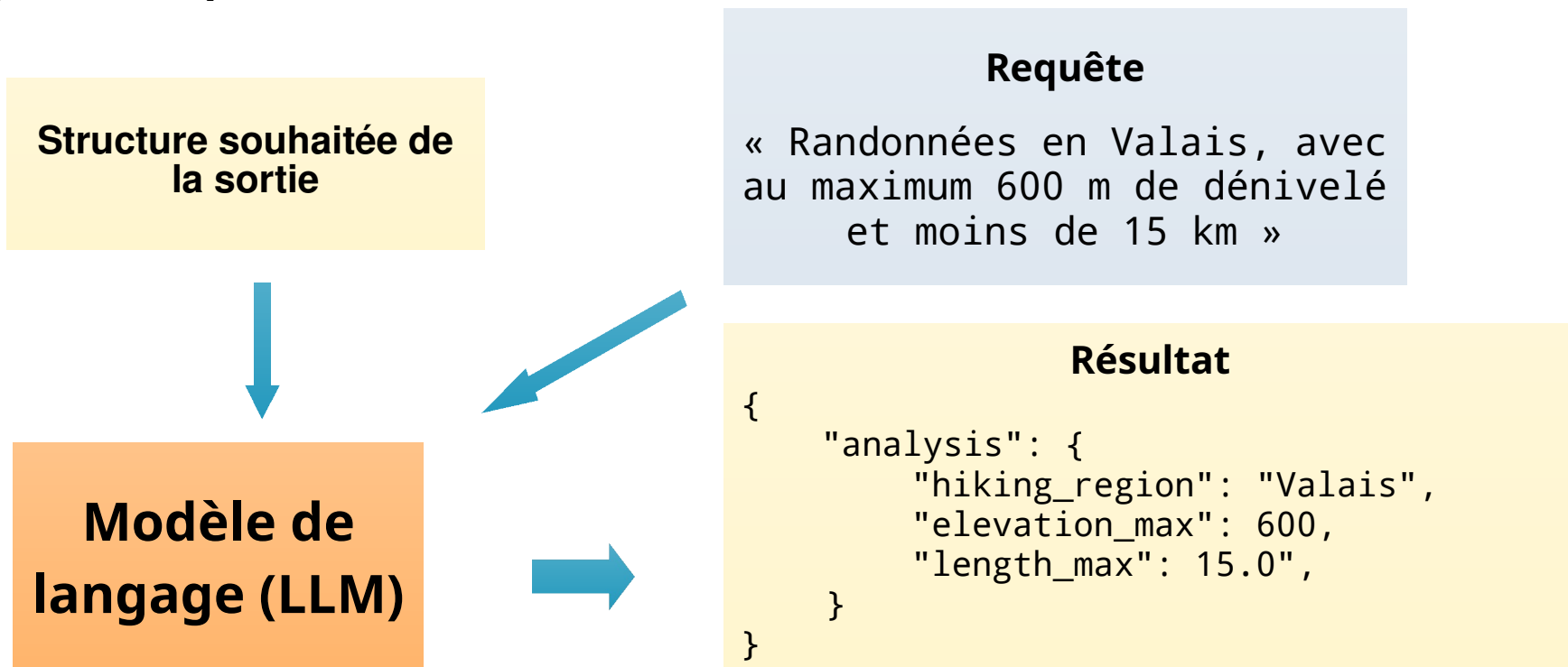
Comment ? → L'IA à l'aide

- Analyse la requête de l'utilisateur, extrait automatiquement les critères et les convertit en filtres.
- Les critères non filtrants sont utilisés dans la recherche sémantique.
- Analyse les requêtes collectées pour comprendre les critères importants, suggérer de nouveaux filtres



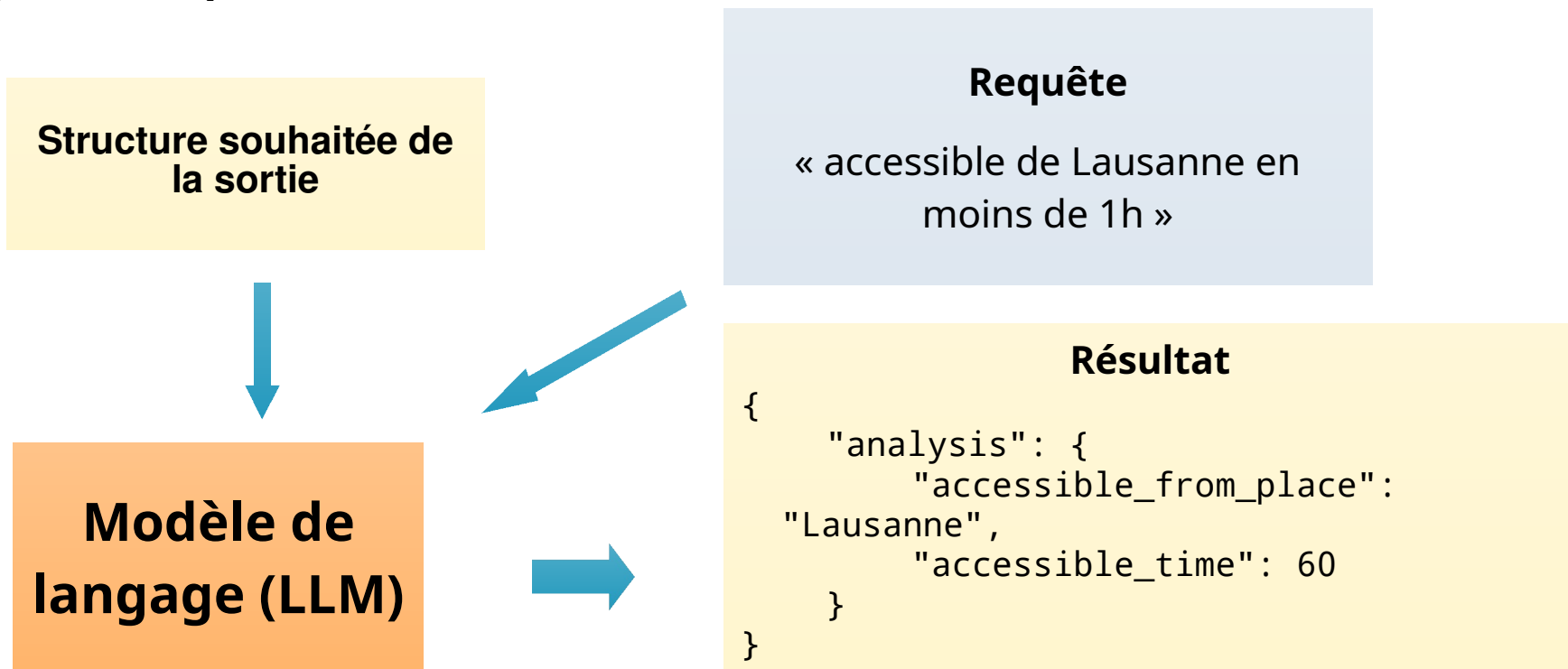
IA à l'aide

- Analyser la requête



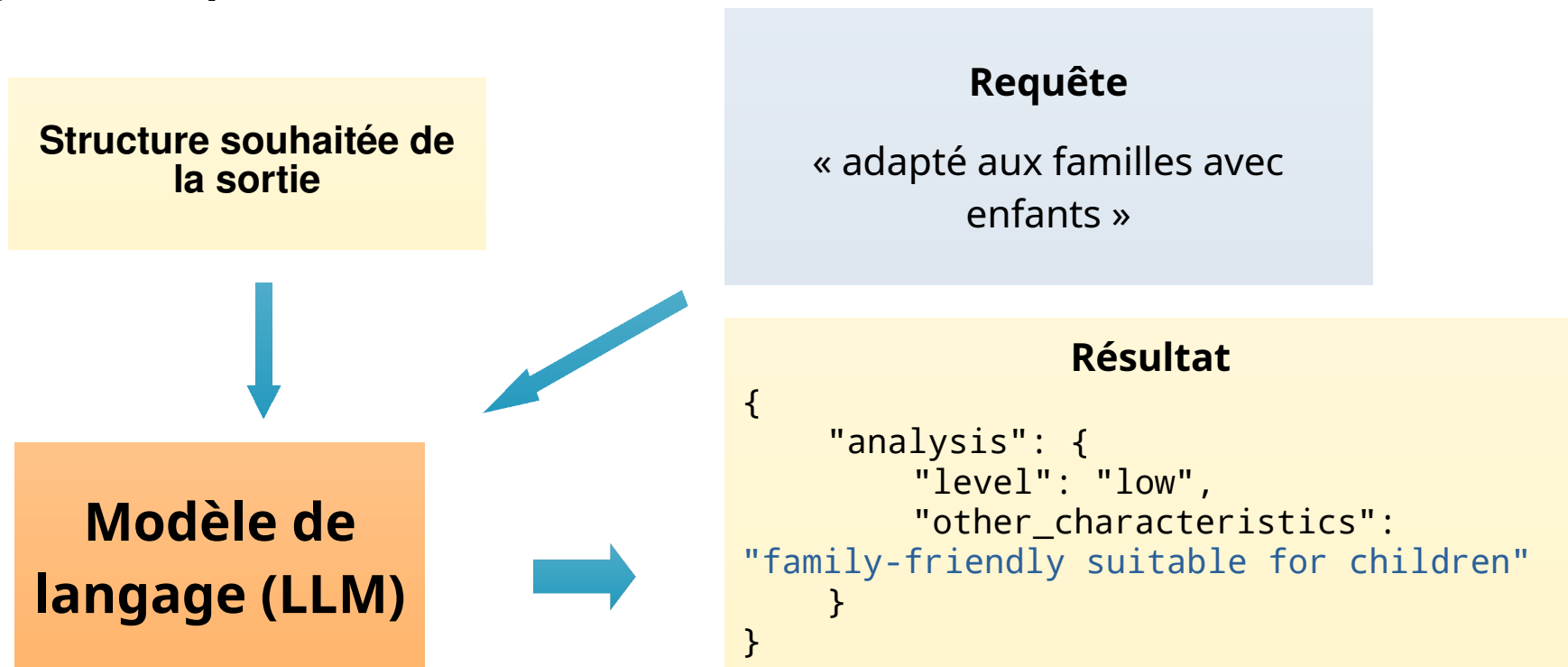
IA à l'aide

- Analyser la requête



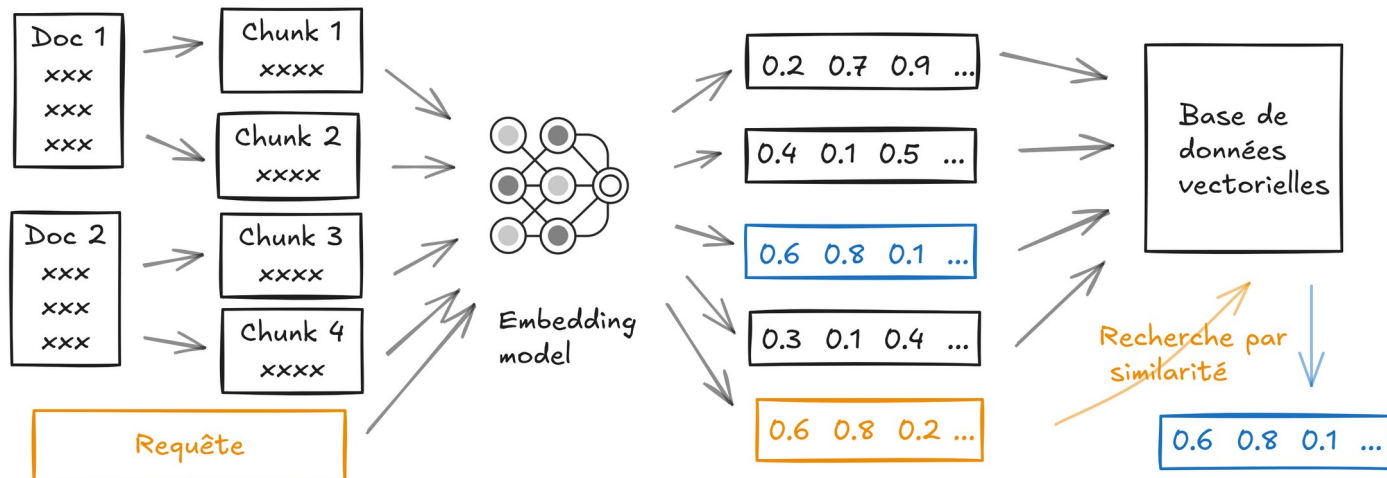
IA à l'aide

- Analyser la requête



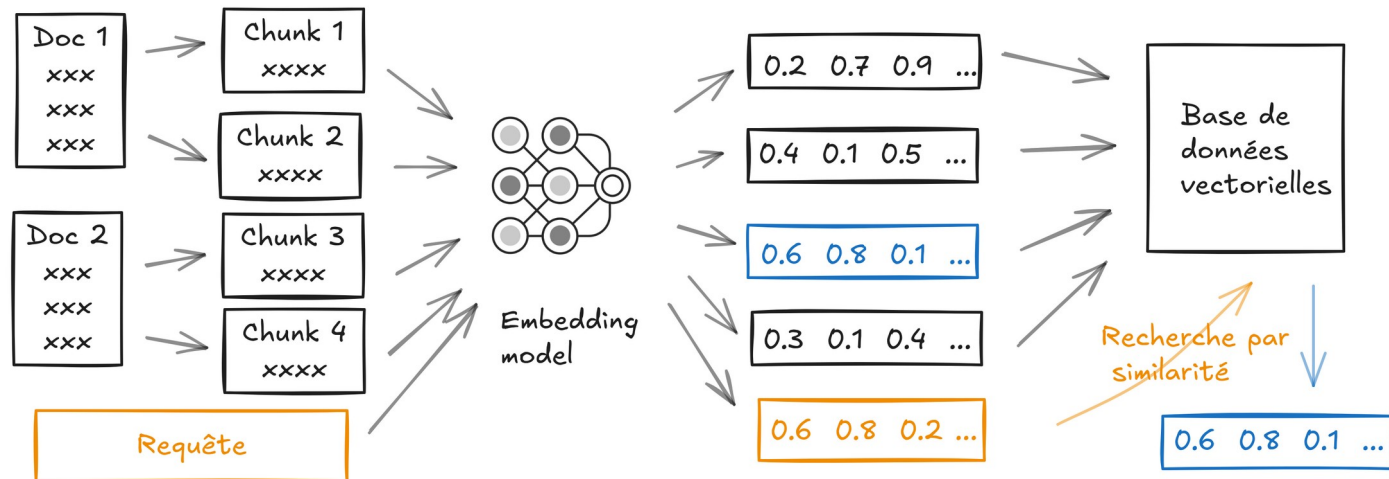
IA à l'aide

- Recherche sémantique



IA à l'aide

- Recherche sémantique



- Pour augmenter la qualité de la recherche, on "augmente" les textes avec **Named-Entity Recognition (NER)**

From Frauenkappelen[LOCATION], the path descends to Lake Wohlen[LAKE] and follows its shore and then the course of the Aare[RIVER].

Démo

Hike AI Search

randonnée dans Jura, moins de 600m d'élévation, max 15km



Exemples

randonnée dans Jura, moins de 600m d'élévation, max 15km

accessible de Lausanne en moins de 1h

adapté aux familles avec enfants

Analyse de votre requête

hiking region: Jura

max elevation: 600 m

max hiking length: 15 km

1. Des hauteurs à la rive du Doubs

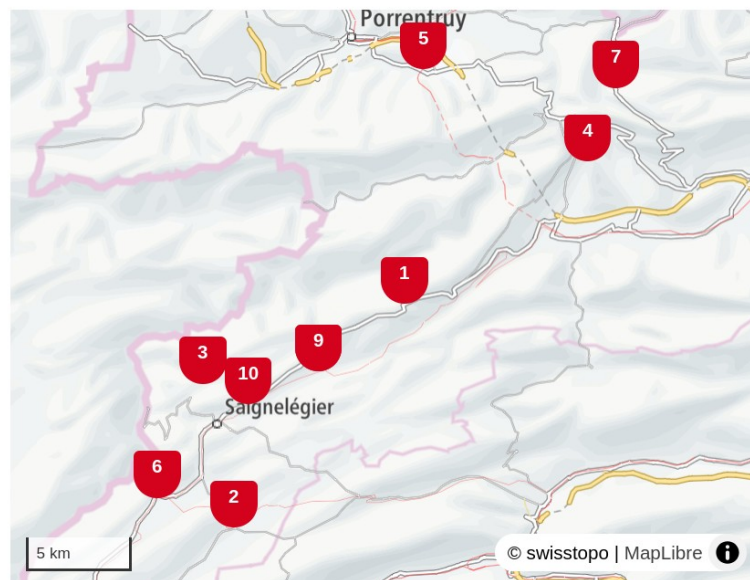
<i>Level:</i>	<i>Ascent:</i>	<i>Descent:</i>	<i>Hike time:</i>	<i>Length:</i>
medium	139m	670m	3.17h	12.4km

A la vue des grandes armoiries dessinées sur le sol, devant la fontaine du village de St-Brais, aucun doute ne subsiste: nous voilà dans le canton du Jura. Non loin du village, à gauche dans le sens de la marche, apparaissent d'impressionnants rochers dans la forêt. Dans les grottes situées en contrebas, différents objets...

Plan hike

2. Les Franches-Montagnes en hiver

<i>Level:</i>	<i>Ascent:</i>	<i>Descent:</i>	<i>Hike time:</i>	<i>Length:</i>
medium	169m	169m	1.95h	5.5km



IA à l'aide

- **Analyse des requêtes collectées** : utiliser l'apprentissage non-supervisé pour identifier les critères les plus cherchés
 - Déterminer les filtres à supporter
 - Classifier les données pour faciliter la recherche

Pour concevoir un moteur de recherche IA

- **Étape 0 : Data pre-processing** : Augmenter les données textuelles avec NER
 - **Étape 1 : Définir** un certain nombre de filtres à supporter
 - **Étape 2 : Implémenter:**
 - Frontend : boîte de saisie de texte + résultat de l'analyse
 - Backend :
 - analyser la requête en utilisant un LLM
 - appliquer la recherche par filtres
 - appliquer la recherche sémantique
 - **Étape 3** : Déployer et **collecter** les requêtes
 - **Étape 4 : Analyser** les données collectées
- 

Pour concevoir un moteur de recherche IA

- **Étape 5** : Utiliser les données collectées pour améliorer davantage le fonctionnement
 - améliorer les prompts
 - fine-tuner le modèle d'embedding
 - fine-tuner un LLM pour mieux analyser les requêtes

Choix des modèles IA

Principe : **petit, performant, open**

Opensource / open-weight
transparence



Modèles petits
rapidité, écologie



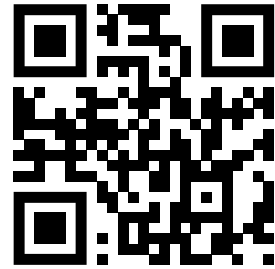
déployable sur des infrastructures locales
souveraineté des données

**Privilégier un
modèle
spécifique**
à un modèle
générique

Merci pour votre attention !

Viet Hang Nguyen

DeepAlps 
DEEP INSIGHT SOLUTIONS



Attributions

Icons :

- Question icon made by justicon from www.flaticon.com
- Idea icon made by Freepik from www.flaticon.com
- Artificial-intelligence icon made by Freepik from www.flaticon.com
- Neural network icon made by Creative_Captain from www.flaticon.com