



INNOVATIVE SOLUTIONS  
BY OPEN SOURCE EXPERTS

# Les LLM au service de la découvrabilité des données

Les recettes du succès

Rencontres Plateformes et Innovations #5



# Florent Gravin

## Directeur Technique



[florent.gravin@camptocamp.com](mailto:florent.gravin@camptocamp.com)



<https://www.linkedin.com/in/fgravin/>



<https://github.com/fgravin>



<https://twitter.com/fgravin>



<https://bsky.app/profile/fgravin.bsky.social>



**01**

Difficultés et risques  
des projets AI

**02**

Développer son  
application  
géospatiale IA

**03**

Tester et évaluer, la  
clé de la  
robustesse

**04**

Autres contraintes

# Projets IA

## Difficultés et risques



# La hype autour de l'IA

## Mais peu de résultats convaincants

- Expérience utilisateur insatisfaisante
- Hallucinations
- Résout un problème que personne n'a.



# Et dans le géospatial ?





**80% des projets échouent \***  
**et seulement 30% dépassent le stade pilote \*\***

\* Rand Corporation, \*\* Gartner



# Risques des projets basés sur l'IA



## Les principales raisons des échecs

- **Attentes exagérées (*hype vs reality check*):**
  - La "magie" de ChatGPT génère des attentes importantes, souvent irréalistes.
- **Spécifications défailiantes (fuzzy goals):**
  - Les objectifs ne sont pas précisément définis et détaillés.
  - Les méthodes d'évaluations et de mesure du succès sont défailiantes.
- **Qualité des données insuffisantes (garbage in, garbage out):**
  - Si les données ne sont pas de bonne qualité, les réponses ne le seront pas.
- **Manque d'expertise (talent tangle):**
  - L'IA est de l'ingénierie, il faut des gens qui comprennent ce qu'ils font.



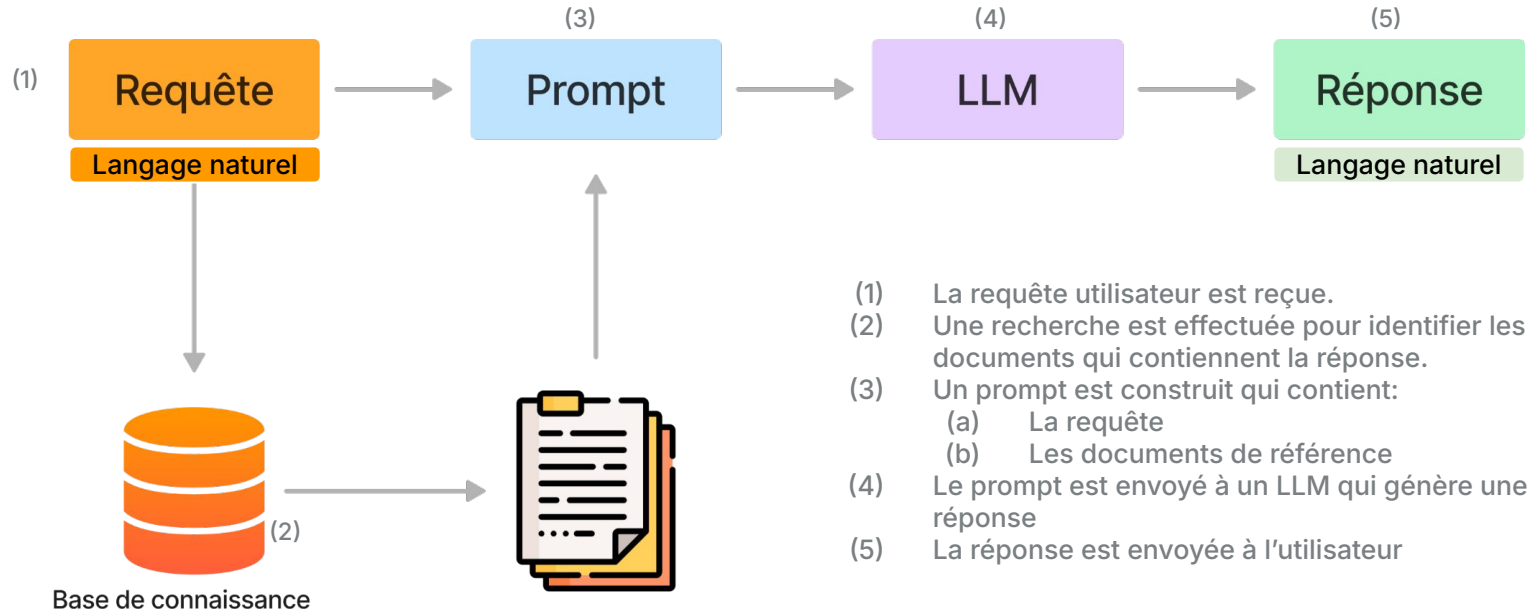
# Applications IA

Développez votre système



# Application Retrieval-Augmented Generation

## Architecture RAG



- (1) La requête utilisateur est reçue.
- (2) Une recherche est effectuée pour identifier les documents qui contiennent la réponse.
- (3) Un prompt est construit qui contient:
  - (a) La requête
  - (b) Les documents de référence
- (4) Le prompt est envoyé à un LLM qui génère une réponse
- (5) La réponse est envoyée à l'utilisateur



**Dans une application RAG, la qualité de la réponse dépend de la qualité de la recherche.**



## Agents

Sélectionnez l'agent utilisé pour la recherche.

### Recherche sémantique simple

Une recherche sémantique sans transformation de la requête



### Recherche sémantique avec reformulation

Une recherche sémantique avec extraction des termes par un LLM

### Recherche sémantique et géographique

Une recherche sémantique avec un filtre géographique

### Réordonne les résultats

Recherche sémantique et géographique avec tri des résultats

### Recherche Geocat

Recherche lexicale avec filtre géographique de Geocat

### Recherche hybride

Recherche hybride ( sémantique / geocat.ch )

# Bienvenue sur la Plateforme Nationale de Données Géographiques

## Comment puis-je vous aider ?

Saisissez votre requête ici



Saisissez votre requête ou laissez-vous guider avec ces suggestions :

Montre-moi les possibilités de randonnée en hiver dans la région du Col des Mosses.

 COMMENCEZ AVEC CETTE REQUÊTE

Montre-moi les sentiers de randonnée dans la région de Innertkirchen

 COMMENCEZ AVEC CETTE REQUÊTE

Quel Canton a des données sur le revêtement des routes?

 COMMENCEZ AVEC CETTE REQUÊTE

Tous les jeux de données sur les toitures et les objets qu'il y a sur les toitures à Genève.

 COMMENCEZ AVEC CETTE REQUÊTE



**La principale cause ? Les données**

# Les 5 grands défis liés aux données



- **Disponibilité** : données absentes, silotées, ou inaccessibles.
- **Qualité** : données incomplètes, bruitées ou mal étiquetées.
- **Gouvernance** : absence de cadre juridique, sécurité et conformité floues.
- **Format** : données non exploitables pour l'IA (format non adapté).
- **Utilisabilité** : les systèmes métiers ne savent pas consommer les résultats IA.

# Qualité des données



Surtout la **Qualité des métadonnées !**

Les IA comme assistant de catalogage

- Génération
- Enrichissement

[Prototype Ticino](#)



# JUMEAU NUMÉRIQUE DE LA FRANCE ET DES TERRITOIRES

ANNEXE TECHNIQUE  
PROGRAMME DE TRAVAIL

JANVIER 2025

IGN CHANGER D'ECHELLE

Cerema

Inria

spatial

camptocamp

igo geofit

SIRADEL ENGIE

Geometry Factory

LuxCarta

cnes

cap-digital

Université Gustave Eiffel

ifp Energies nouvelles

© GEOFIT-IGO



# Aller plus loin

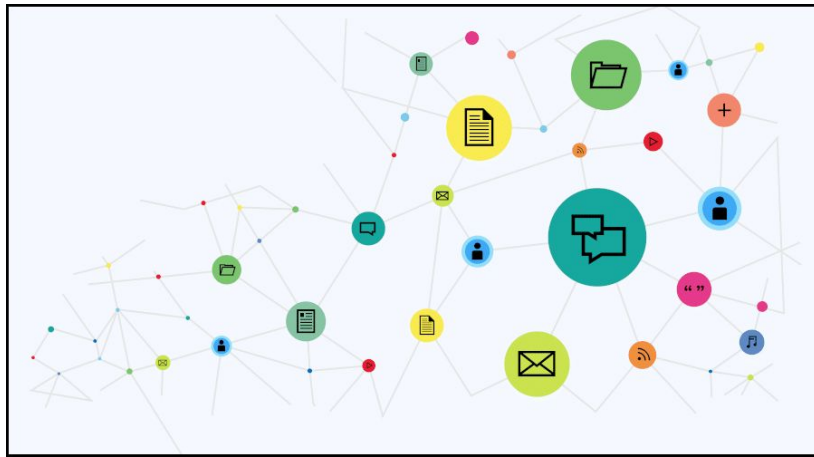
Anticiper les usages de demain



**Ontologie** : Lien entre données et connaissances

⇒ **Contextualiser les informations**

## Graphe de connaissance



- Concepts
- Données
- Relations



# Consolider votre expertise IA

# Ne pas réinventer la roue



- Veille
- Bonnes pratiques
- Communautés
- Framework

**Anticiper la prochaine itération**



# Restez connectés



- RAG / KAG
- Modèles multimodaux
- Agentic AI
- Model Context Protocol (MCP)

# Applications IA

Evaluez votre système





**Sémantique, lexicale ou géographique, la recherche doit être testée !**

- On veut tester le R (le retrieval) du **RAG**.
- Pour une question donnée, quels documents faut-il consulter ?
- Un jeu de tests est constitué pour mesurer la performance du "*retrival*". Chaque élément est composé d'une question et d'une liste de documents qui permettent d'y répondre.

**[question, [documents pertinents]]**

# Jeux de test



 Semantic search geocat.ch  ID

Latest ▾

Add Auto-Evaluator

+ Experiment



Experiments Pairwise Experiments **Examples** Few-shot search

≡ Filters

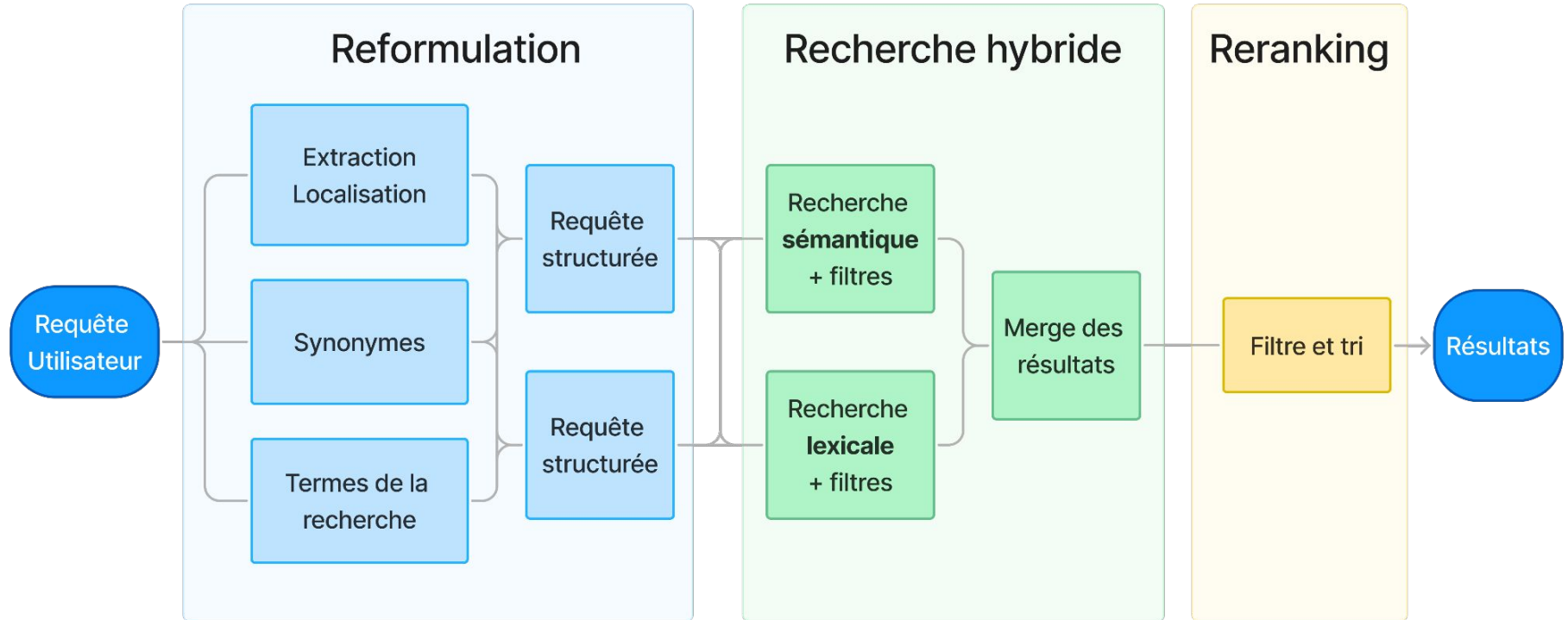
Select split ▾

+ Example

<input type="checkbox"/> Inputs	Reference Outputs
<input type="checkbox"/> Affiche moi les routes nationales.	{"to_avoid":["fe6115f2-cff8-4deb-8849-53e3d501b747","d85abbca-b090-4512-k
<input type="checkbox"/> Je veux voir les parks à vélo de Neuchâtel	["16d16fa3-a416-4e8b-99fc-69c7267f134d"]
<input type="checkbox"/> Je veux voir les parks à vélo de Lausanne	["eb867568-c511-4304-b902-1bbfbd0d8806"]
<input type="checkbox"/> How does the geographical diversity of Switzerland impact its agricultu...	["6e4544bc-55f4-4aef-b36a-e36fe376eaf8","73752124-516a-5c19-aff8-2d0d119
<input type="checkbox"/> What are the major rivers in Switzerland?	["edff6d33-ba8c-441e-b8bd-7ab4969d4440","urn:isogeo:metadata:uuid:bc6264
<input type="checkbox"/> Is there a map with the lakes?	{"to_avoid":["c7db9bf6-4b08-49c6-a759-f3bdefef3121","cfe43baa-75d9-4c37-e
<input type="checkbox"/> Come si chiama il lago di Lugano?	{"to_avoid":["6504bf41-0e48-402a-bd30-5e80f931ae76"],"retrieved_layers":["ak
<input type="checkbox"/> Quelle est la population de Bern?	["4bfbbf20-d90e-4131-8fe2-4c454ad45c16","d89e249e-ab40-4c9a-a8ef-24619

# Recherche de Métadonnée

## Phase de récupération





# Evaluations



Semantic search geocat.ch [ID](#)

All Versions ▾

Add Auto-Evaluator

+ Experiment



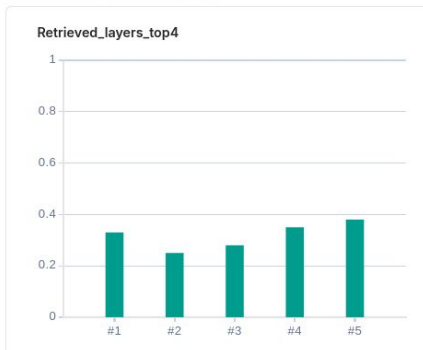
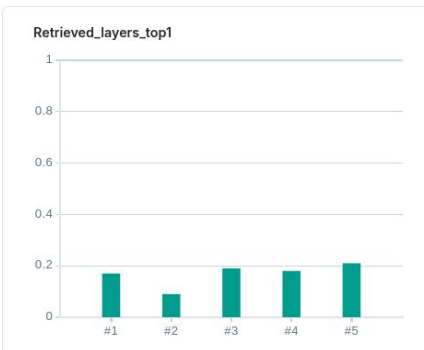
Experiments [Pairwise Experiments](#) [Examples](#) [Few-shot search](#)

Q Search by name...

Charts

x-axis Experiment ▾

Filters



<input type="checkbox"/>	Experiment ↑↓	Splits	Repetitions	Retrieved_layers_relative ↑↓	Retrieved_layers_top1 ↑↓	Retrieved_layers_top4 ↑↓	Created At ↓	P50 Latency ↑↓	P90
#5	4 - Agent hybrid-1825...	base	1	0.00	0.21	0.38	11/26/2024, 3:00:...	2.11s	
#4	3 - Agent geoCat-20a...	base	1	0.00	0.18	0.35	11/26/2024, 2:58:...	0.36s	
#3	2 - Agent similarity + ...	base	1	0.00	0.19	0.28	11/26/2024, 2:57:...	1.94s	
#2	1 - Agent similarity + e...	base	1	0.00	0.09	0.25	11/26/2024, 2:55:...	1.67s	
#1	0 - Agent similarity-f0...	base	1	0.00	0.17	0.33	11/26/2024, 2:54:...	0.65s	

# Tester la generation



Les tests logiciels standard, boolean, ne s'applique pas bien au texte

- On veut tester le G (la génération de texte) du RAG.
- Est-ce que la réponse générée correspond à la réponse attendue ?
- Comment les comparer? (exact match pas pertinent pour du texte)
- Un LLM est utilisé pour juger la réponse et/ou comparer la réponse générée et la réponse attendue. ("*LLM as a judge*").



# Monitorer et debugger un chatbot RAG

# Monitorer une application basée sur les LLM



“Ce qui ne se mesure pas ne s'améliore pas.”

- Quels documents ont été trouvés et fournis au LLM?
- Quel prompt exact a été envoyé au LLM?
  - Puis-je obtenir une meilleure réponse en modifiant le prompt?
  - Est-ce mieux avec un autre modèle?
- Combien coûte la génération d'une réponse?
  - En euro?
  - Et en temps (latence)?

# Traçabilité



TRACE



☰ Waterfall Most relevant ↕

/semsearch-with-geo-filter-and-rerank

🕒 5.53s 🗨️ 2,127

Agent3-SemanticSearchWithGeo... 5.52s ▾

Agent3-SemanticSearchWithG... 5.52s ▾

extract params from query 1.24s ▾

ChatOpenAI gpt-4o-mini 1.13s

PydanticToolsParser 0.00s

get geo filter 1.50s

semantic search with geo filter 1.33s

rerank 1.43s ▾

ChatOpenAI gpt-4o-mini 1.18s

PydanticToolsParser 0.00s

top N layers 0.00s

format\_output 0.00s

🔔 Some runs have been hidden. [Show 11 hidden runs](#)



## extract params from query



ID

Playground

Add to ▾



Run Feedback Metadata



Input ▾

▾ User Query

Je voudrais nager près de Vevey

Output ▾

▾ Location Name

Vevey

▾ Location Scale

local

▾ Query Lang

fre

▾ Similarity Search Query

nager

START TIME

12/09/2024, 02:23:56 PM

END TIME

12/09/2024, 02:23:57 PM

TIME TO FIRST TOKEN

STATUS

Success

TOTAL TOKENS

230 tokens /  
\$0.00004485

LATENCY

🕒 1.24s

TYPE

Chain

TAGS

graph:step:1



# Amélioration continue

# Comment améliorer un chatbot au cours de temps ?



## Collecter et utiliser les retours des utilisateurs

- Le meilleur jeu de test est celui constitué des retours utilisateurs
  - Tests synthétiques imparfaits
  - Tests générés par des humains laborieux
  - Les retours d'utilisateur, une fois nettoyés, sont les meilleurs tests et permettent d'améliorer l'application.
- Collecter et utiliser les retours utilisateur
  - Intégrer un élément de retour d'information dans la UI
  - Collecter les retours dans un service du backend
  - Nettoyer et créer un jeu de test



# IA générative

## Les autres contraintes





# Nos réponses aux enjeux de l'IA



## Souveraineté :

- Utiliser des modèles open source
- Contrôler leur hébergement

## Eco-conception :

- Rationaliser et mutualiser les infrastructures.
- Optimiser la taille du modèle



# Les points à retenir



- Visez des solutions opérationnelles
- Travaillez sur votre base de connaissance
- Appuyez-vous sur de la compétence éprouvée
- Ce qui n'est pas mesuré ne s'améliore pas
- Utilisez des offres d'Infra as a service pour l'inférence des modèles

Merci pour votre écoute



**Florent Gravin**  
[florent.gravin@camptocamp.com](mailto:florent.gravin@camptocamp.com)

# camptocamp<sup>®</sup>

INNOVATIVE SOLUTIONS  
BY OPEN SOURCE EXPERTS