



HIBERNATE SEARCH 7

(et 6)

Nouveautés et cas pratiques
de la recherche full-text et vectorielle Java



Yoann Rodière

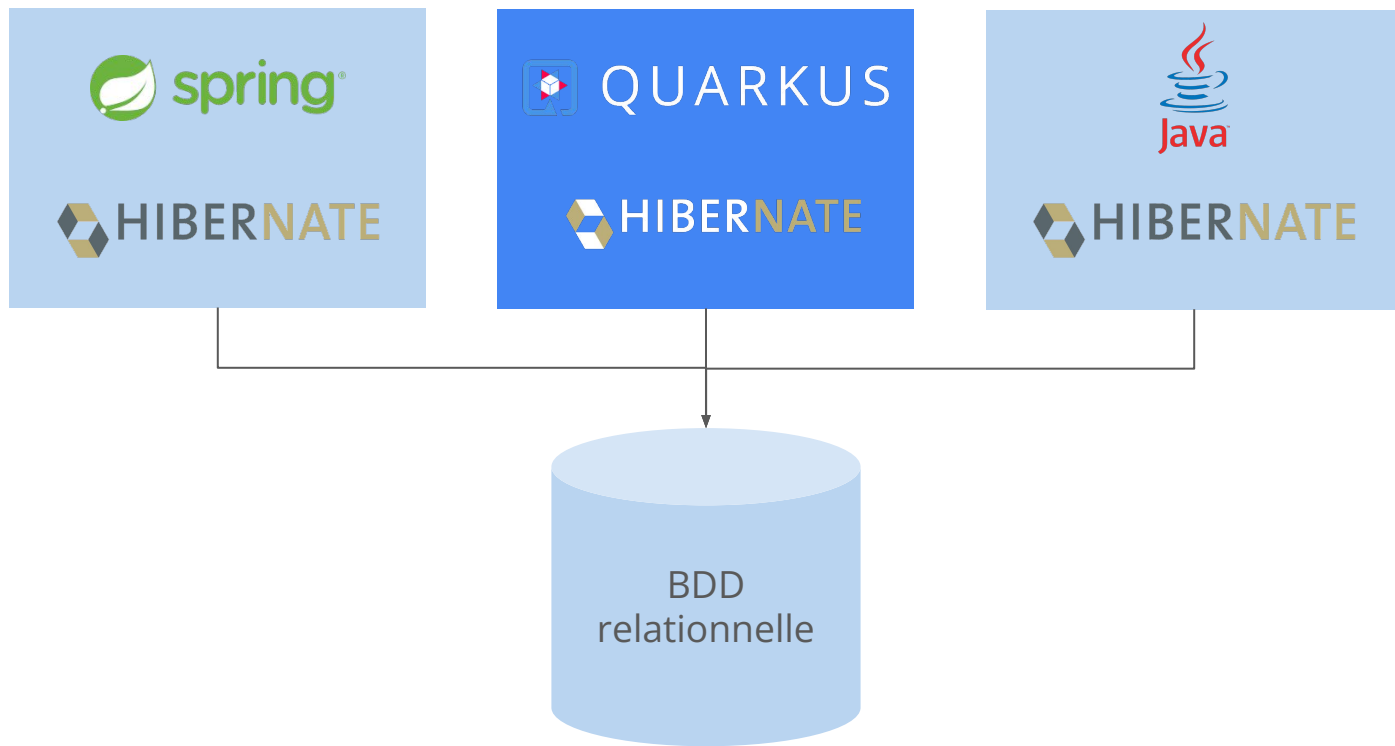
Red Hat - Hibernate - Quarkus



Katia Aresti

Red Hat - Infinispan - Quarkus

Prenons une application Java...



Dossier de formation



<input type="text" value="Numéro (dossier, archivage...)"/>	<input type="text" value="Secteur d'activité"/>	<input type="text" value="Date de modification"/>	<input type="text" value="Date de modification"/>
---	---	---	---

Session

<input type="text" value="Organisme de formation (nom ou numéro DA)"/>	<input type="text" value="Type de modalité"/>	<input type="text" value="Modalité d'exécution"/>	
<input type="text" value="Intitulé"/>	<input type="text" value="Nature"/>	<input type="text" value="Thème"/>	<input type="text" value="Modalité particulière"/>
<input type="text" value="Département"/>	<input type="text" value="Région"/>	<input type="text" value="Ville"/>	<input type="text" value="Lieu"/>
<input type="text" value="Année de début"/>	<input type="text" value="Date de début"/>	<input type="text" value="Date de début"/>	

 Filtrer

952382 dossiers de formations trouvés

N° dossier	N° archivage	Intitulé	Situation	Stagiaires
M1337001	-	INTITULE STAGE FORMATION	 Agrément	Stagiaire Nom Prénom  
M1337001	-	INTITULE STAGE FORMATION	 Agrément	Stagiaire Nom Prénom  

Recherche de texte : SQL et ILIKE ?

ILIKE '%thé%'

'kinési**th**érapie'



ILIKE '%pommes**s**%'

'pomme'



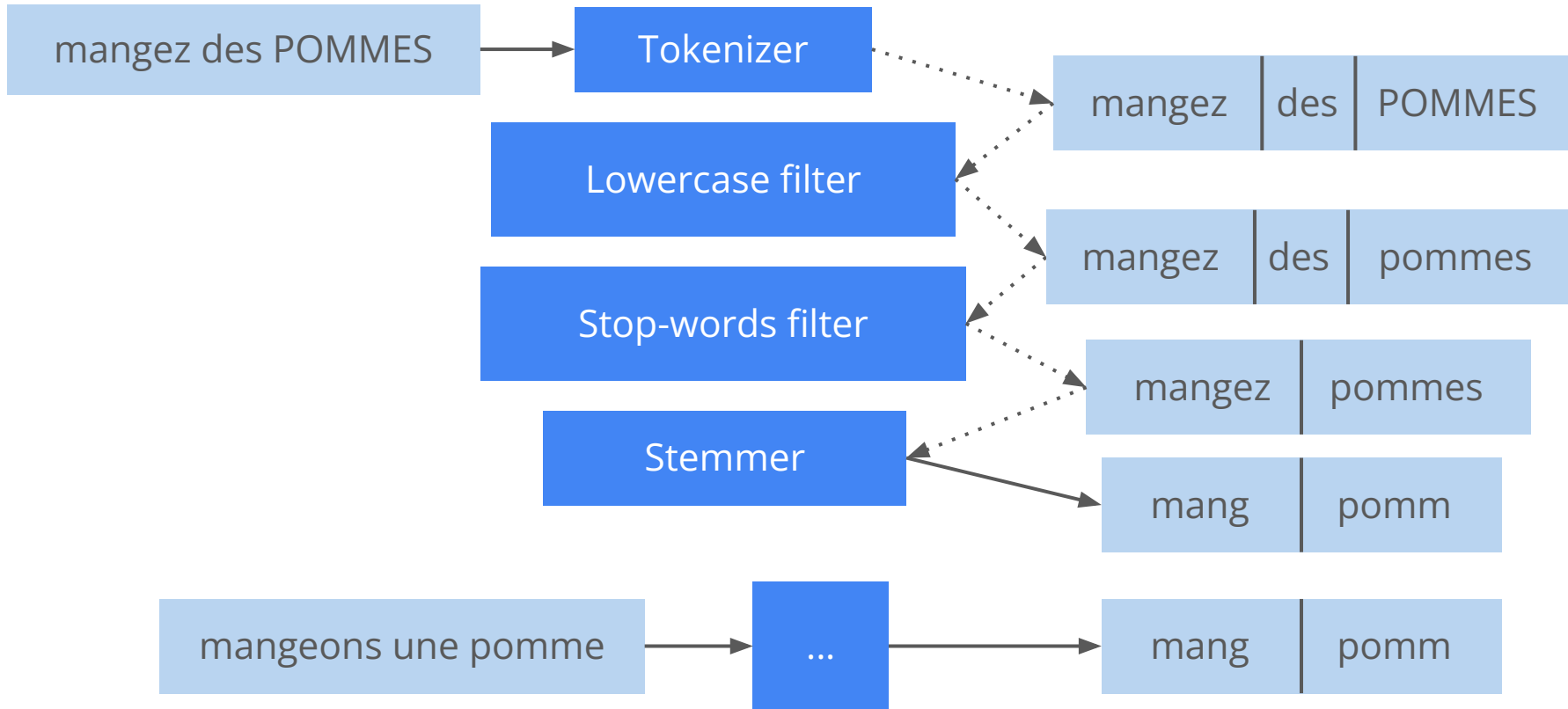
Perf



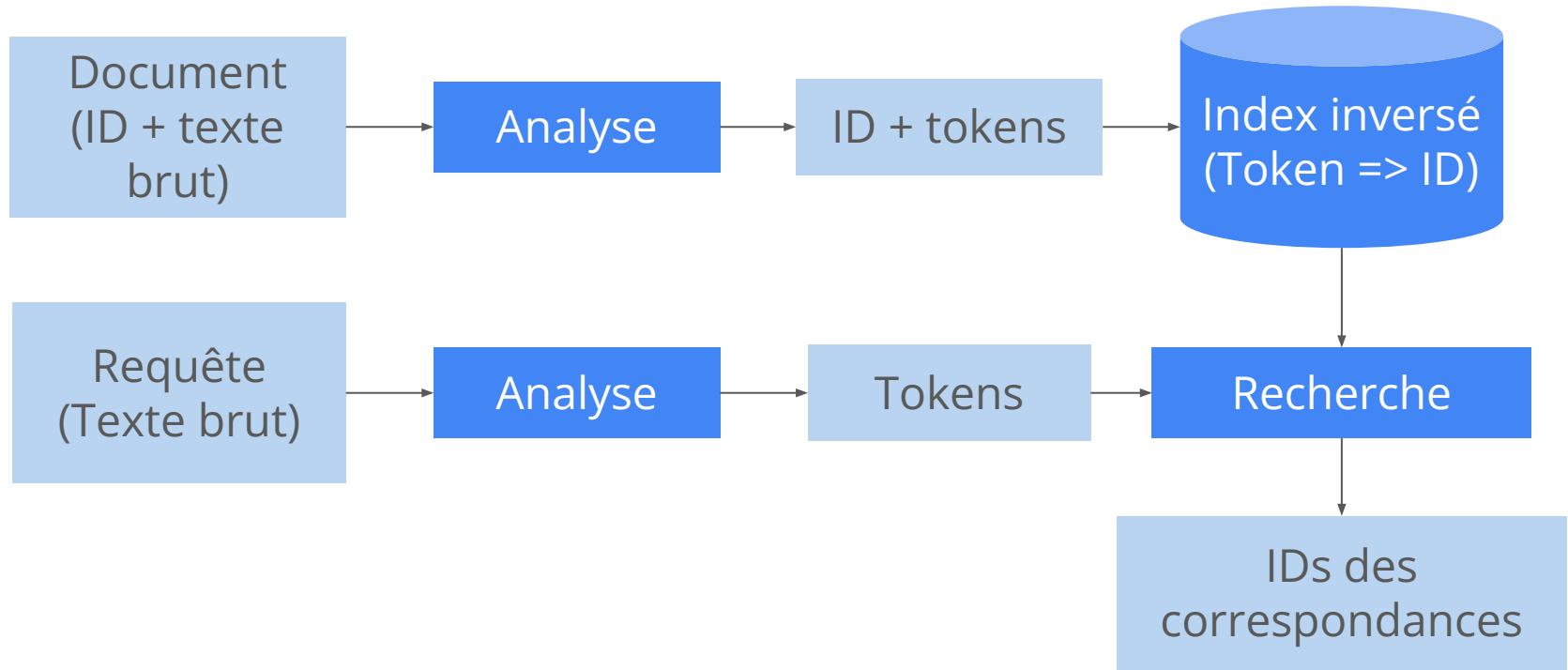
Tri



Analyse



Recherche full-text



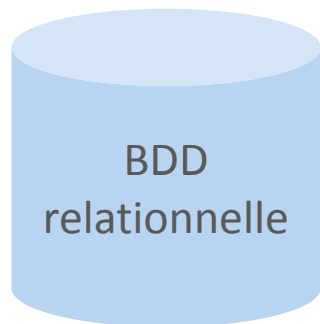


- Fonctionnalités dépendent de la BDD
- Syntaxe dépend de la BDD
- Charge supplémentaire... ?

APACHE
LUCENETM



Comment transférer les données ?

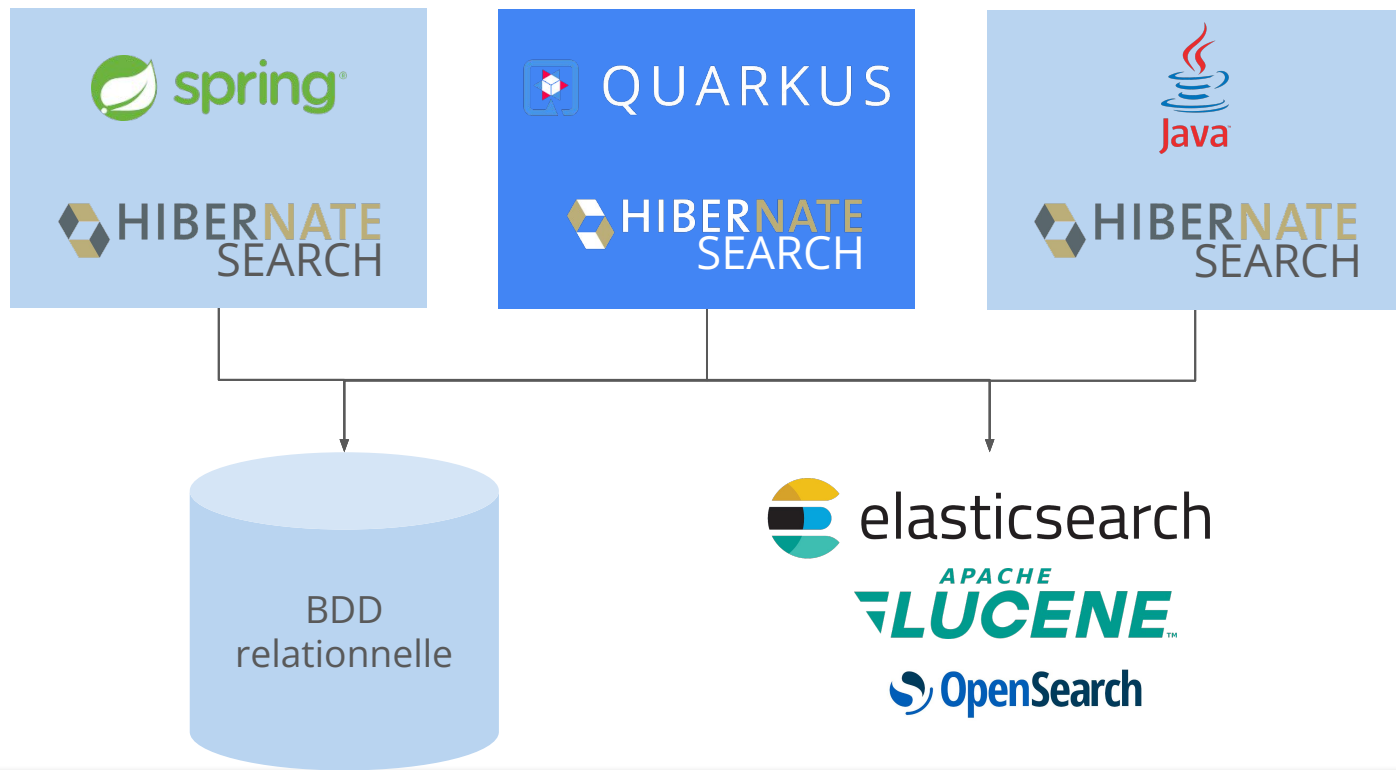


The logo consists of three interlocking hexagons in shades of gold and white.

HIBERNATE SEARCH

ORM. Lucene. Elasticsearch.
Integrated.

Quarkus... mais pas seulement

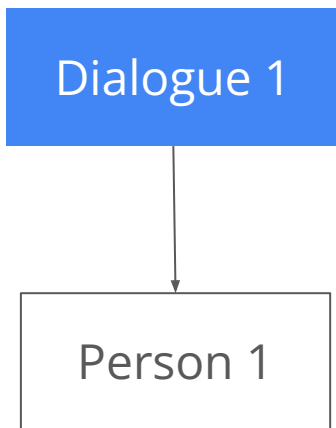


Démo !

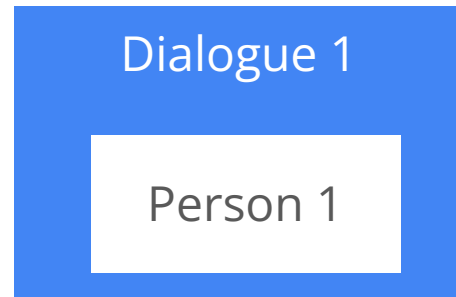
<https://github.com/yrodiere/rivieradev-2024-hibernate-search-demo>

Dénormalisation – cas trivial

HIBERNATE Entités

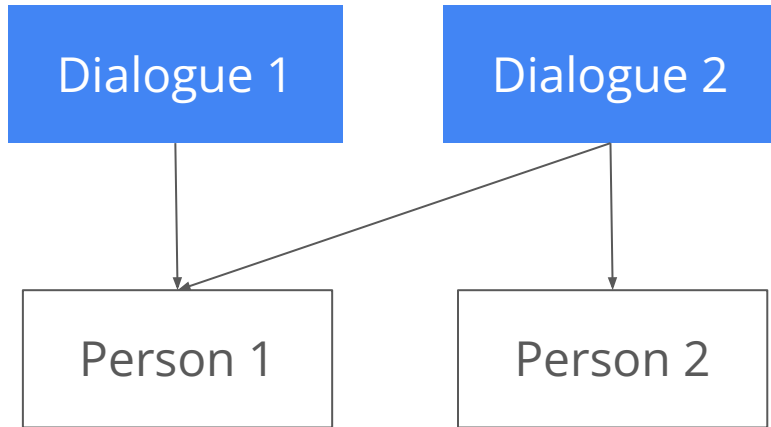


APACHE LUCENE Document

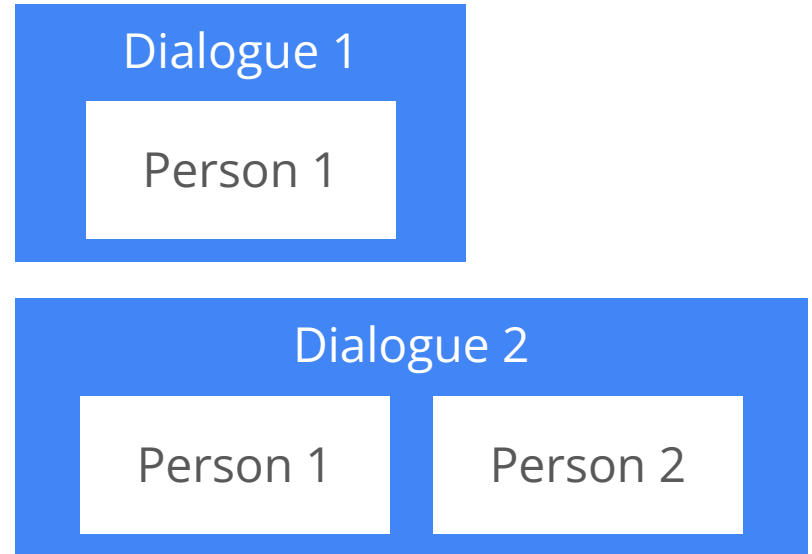


Dénormalisation – cas plus commun

HIBERNATE Entités

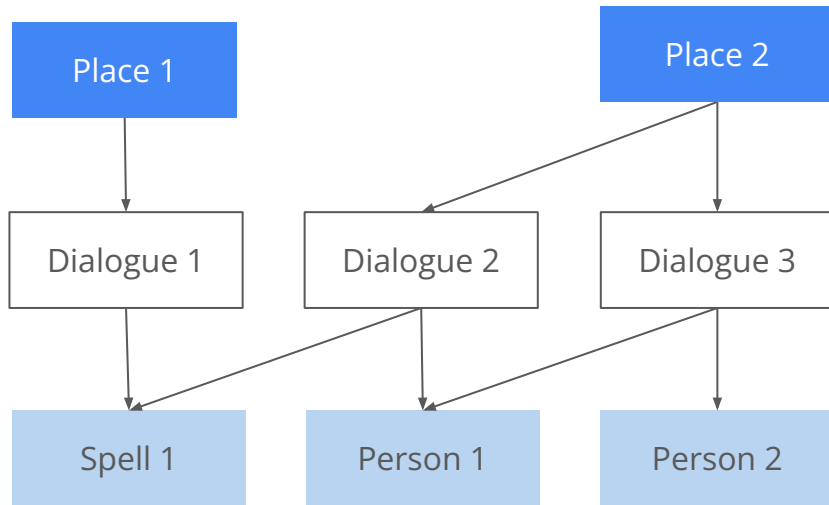


APACHE LUCENE Document

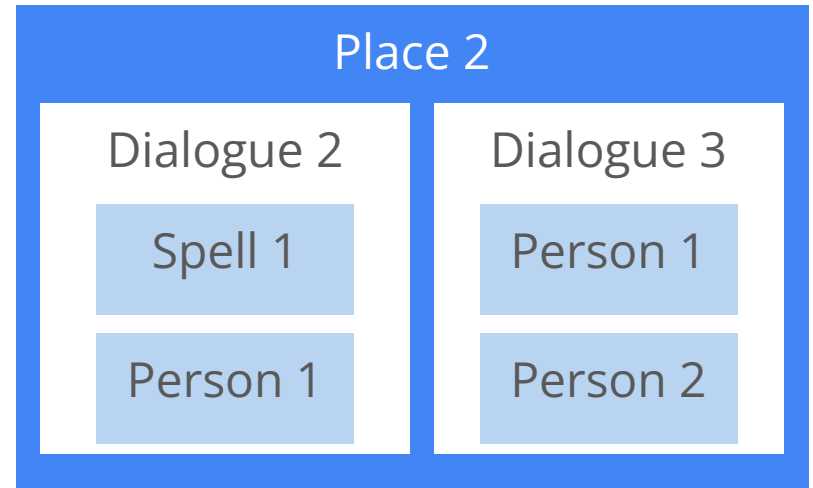


Dénormalisation – cas complexe

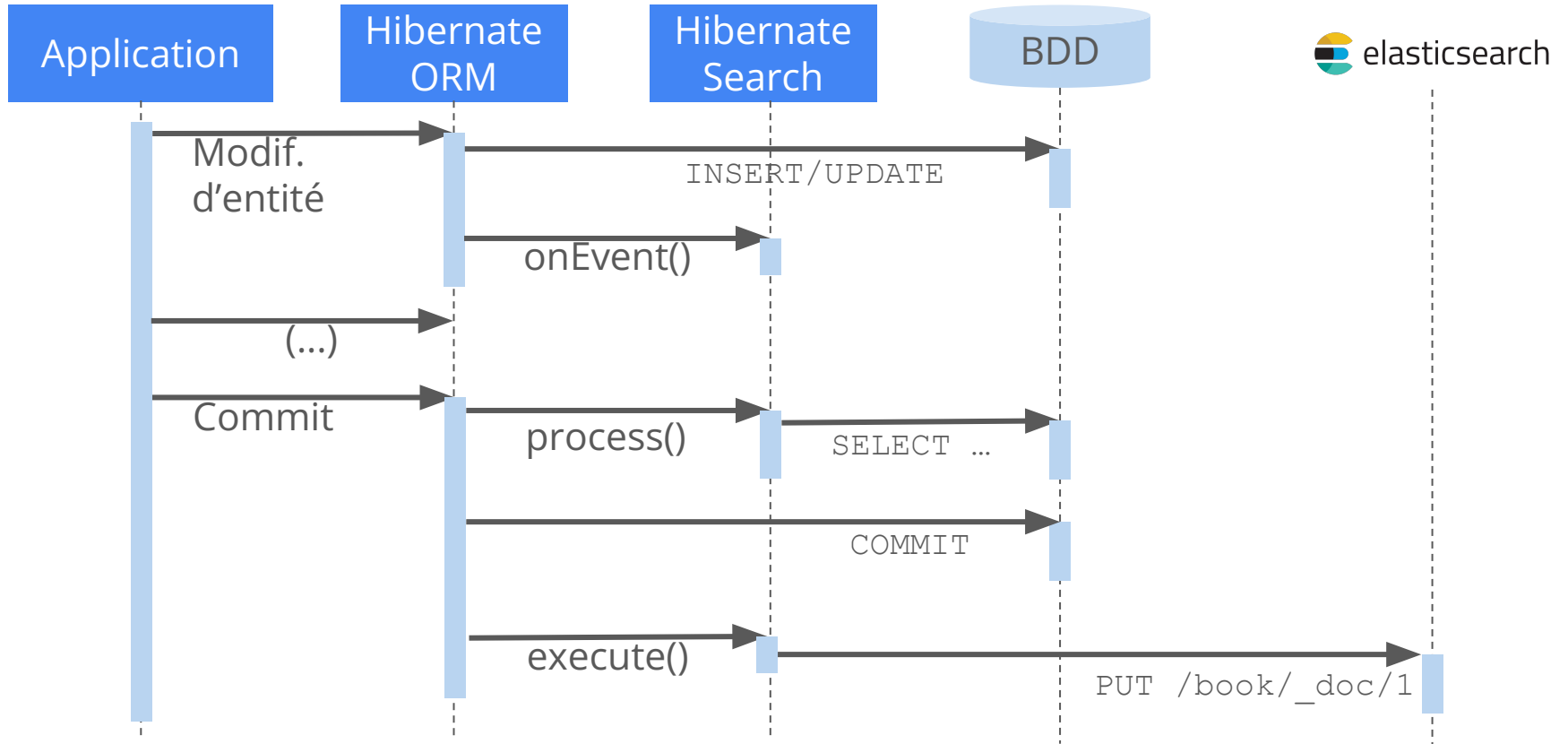
HIBERNATE Entités



APACHE LUCENE Document



Indexation



The logo consists of three interlocking hexagons in shades of white, yellow, and blue.

HIBERNATE SEARCH

Version 6+
2020

Refonte de l'API



Refonte de l'API



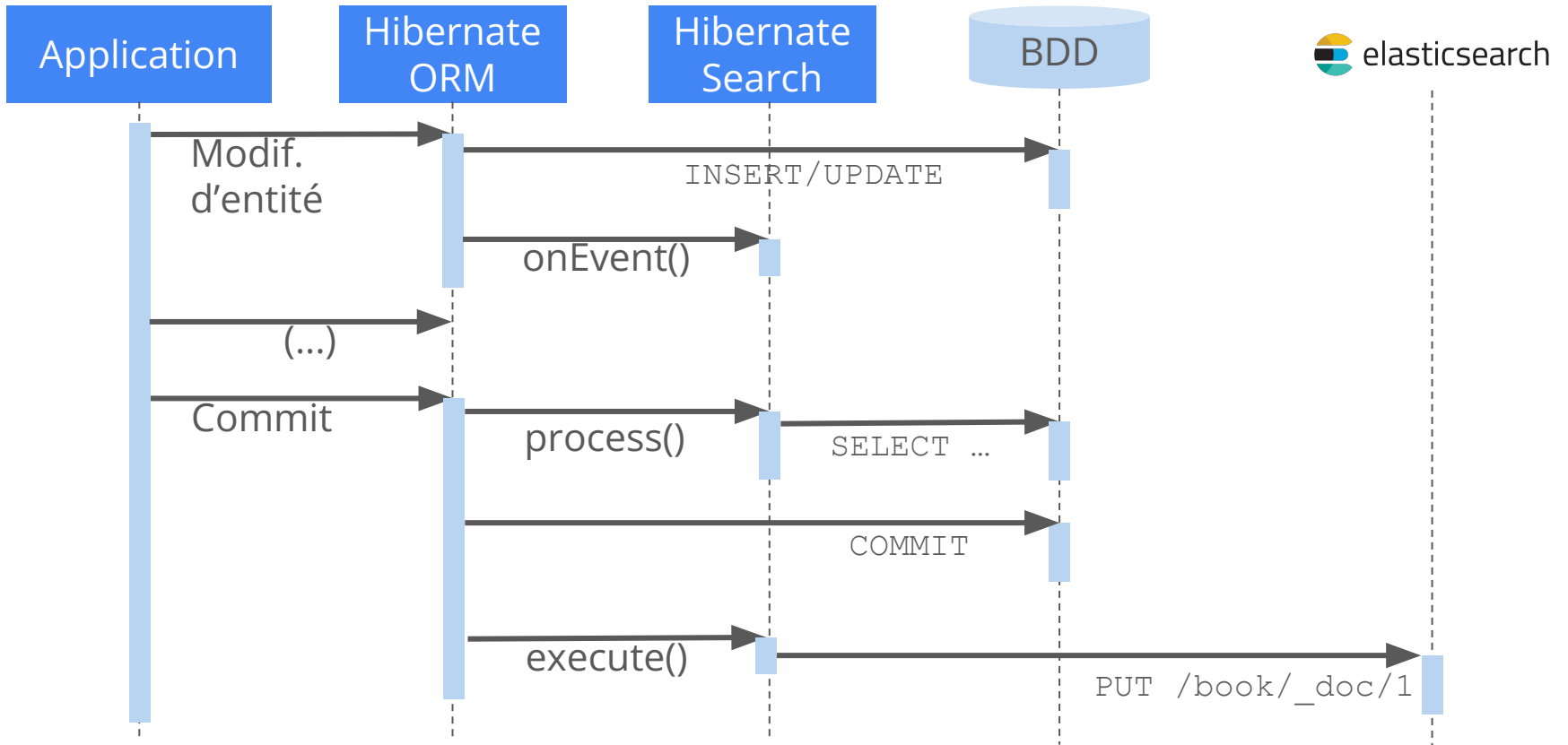
GÉNÉRIQUE

CONCIS

CUSTOMISABLE

ÉPROUVÉ

Et si ça se passe mal... ?

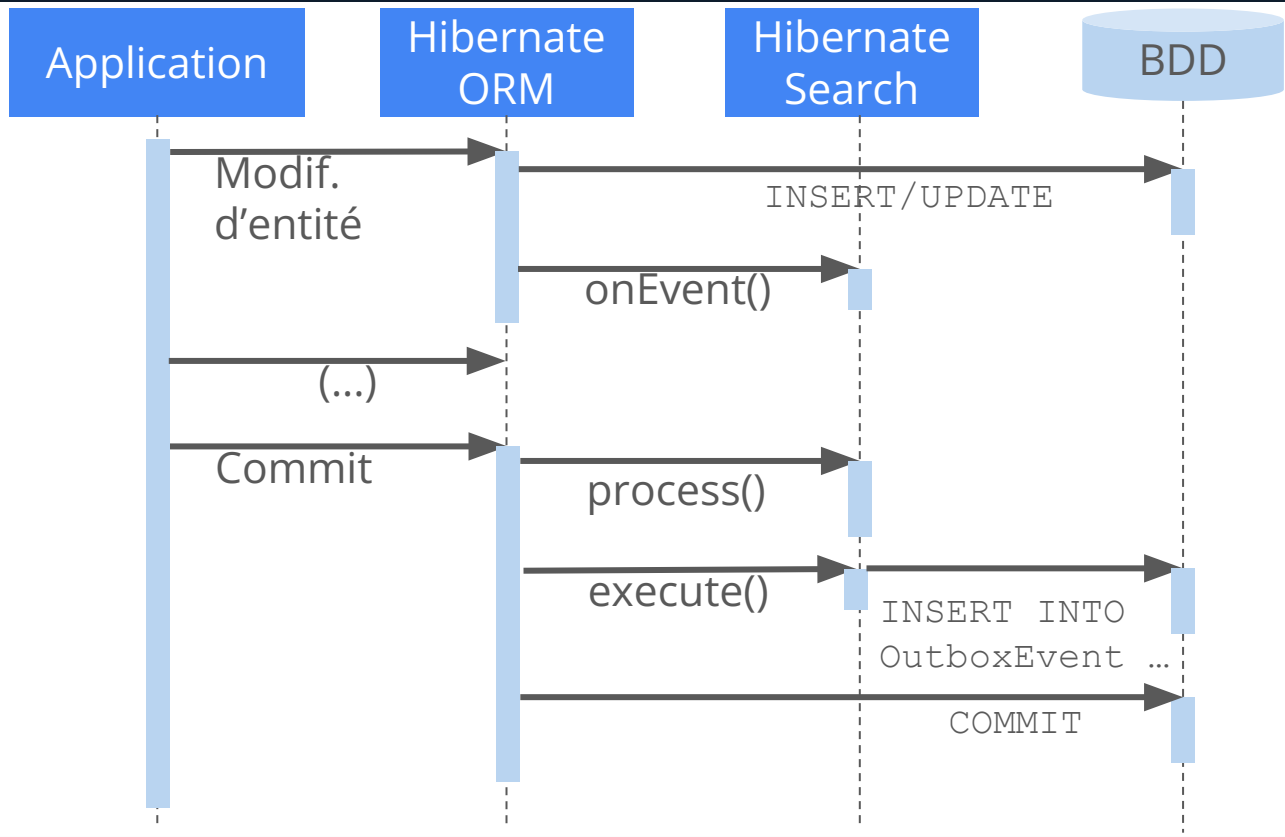


Transactional outbox

https://docs.jboss.org/hibernate/stable/search/reference/en-US/html_single/#coordination-outbox-polling



Transactional outbox 1 - application

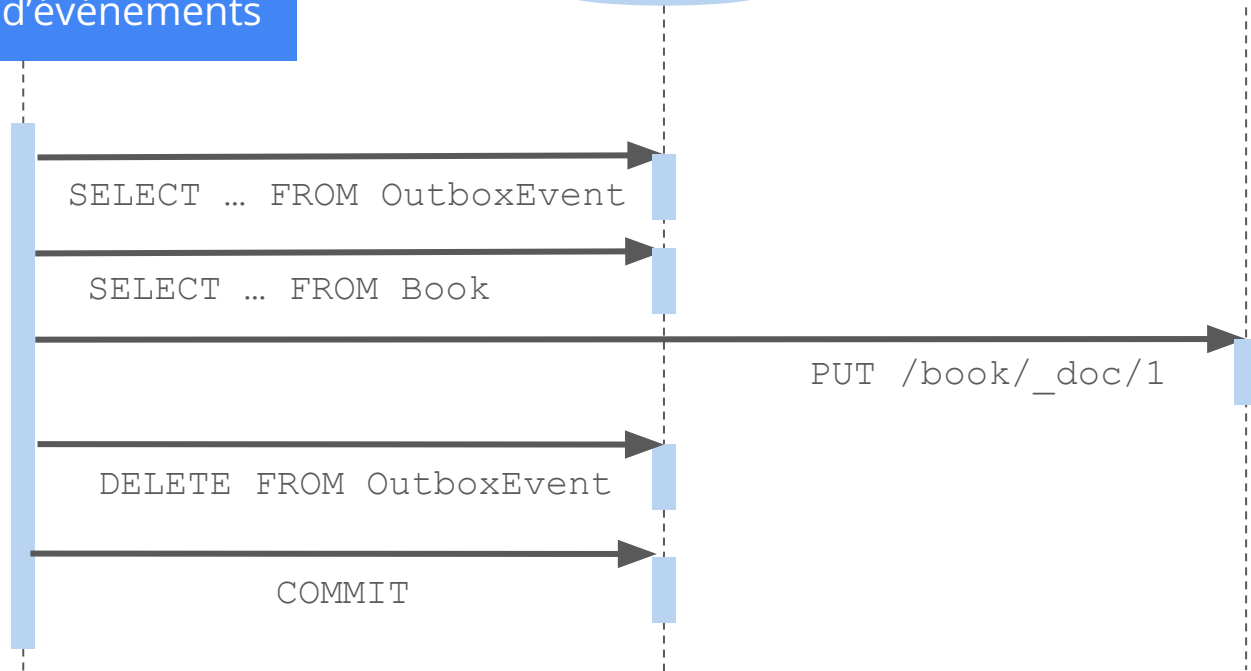


Transactional outbox 2 - processeur d'événements

Hibernate Search:
Processeur d'événements

BDD

 elasticsearch



Plus
puissant...

Transactionnel

Perte d'événement impossible

Asynchrone

Latence minimale

Partitionné

Pas de conflits d'indexation
Parallélisme

... et tout
aussi
simple!

Infrastructure :



Configuration :

```
<dependency>  
  <groupId>io.quarkus</groupId>  
  <artifactId>  
    quarkus-hibernate-search-orm-outbox-polling  
  </artifactId>  
</dependency>
```

```
quarkus.hibernate-search-orm.coordination.strategy  
=outbox-polling
```

Standalone Pojo Mapper

https://docs.jboss.org/hibernate/stable/search/getting-started/standalone/en-US/html_single/index.html#mapper-pojo-standalone-getting-started

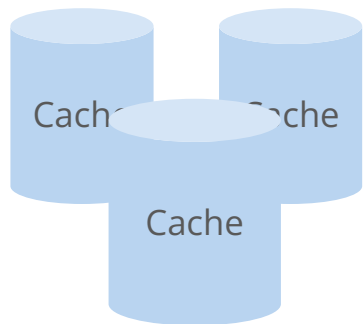


Standalone Pojo Mapper





Infinispan



Hibernate Search +
intégration custom

APACHE
LUCENE™



Infinispan

LangChain Embedding Store



Infinispan



Lanchain 4J



LangChain

The logo consists of three interlocking hexagons in shades of white, yellow, and blue.

HIBERNATE SEARCH

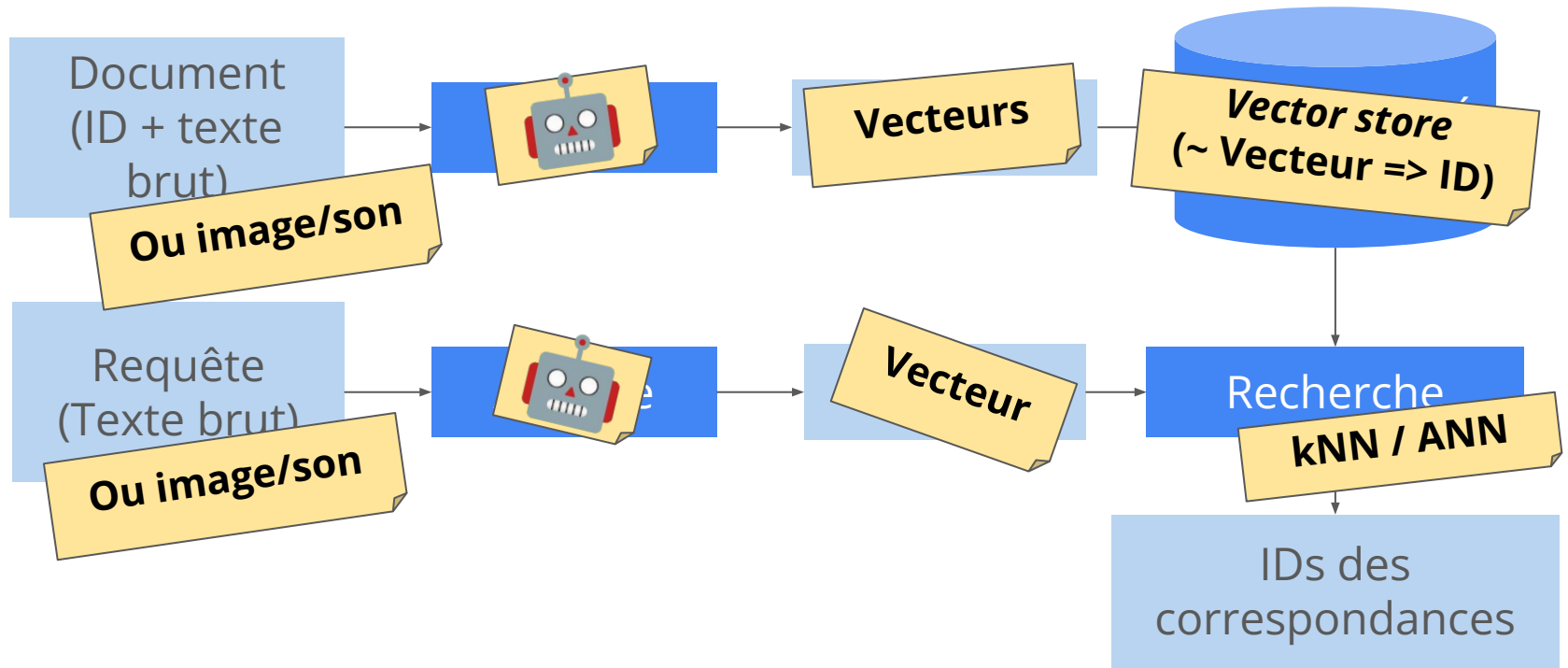
Version 7
2023 / 2024

Vecteurs et *embeddings*



```
[ -0.7754,  
  -0.7972,  
  -0.5784,  
    ...  
  -0.4750,  
  -0.4254 ]
```

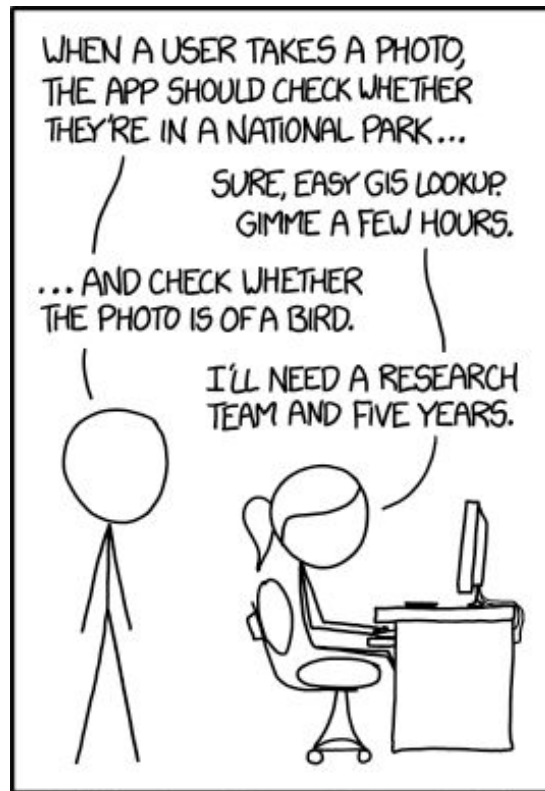

Recherche vectorielle



Démo !

<https://github.com/yrodiere/rivieradev-2024-hibernate-search-demo>

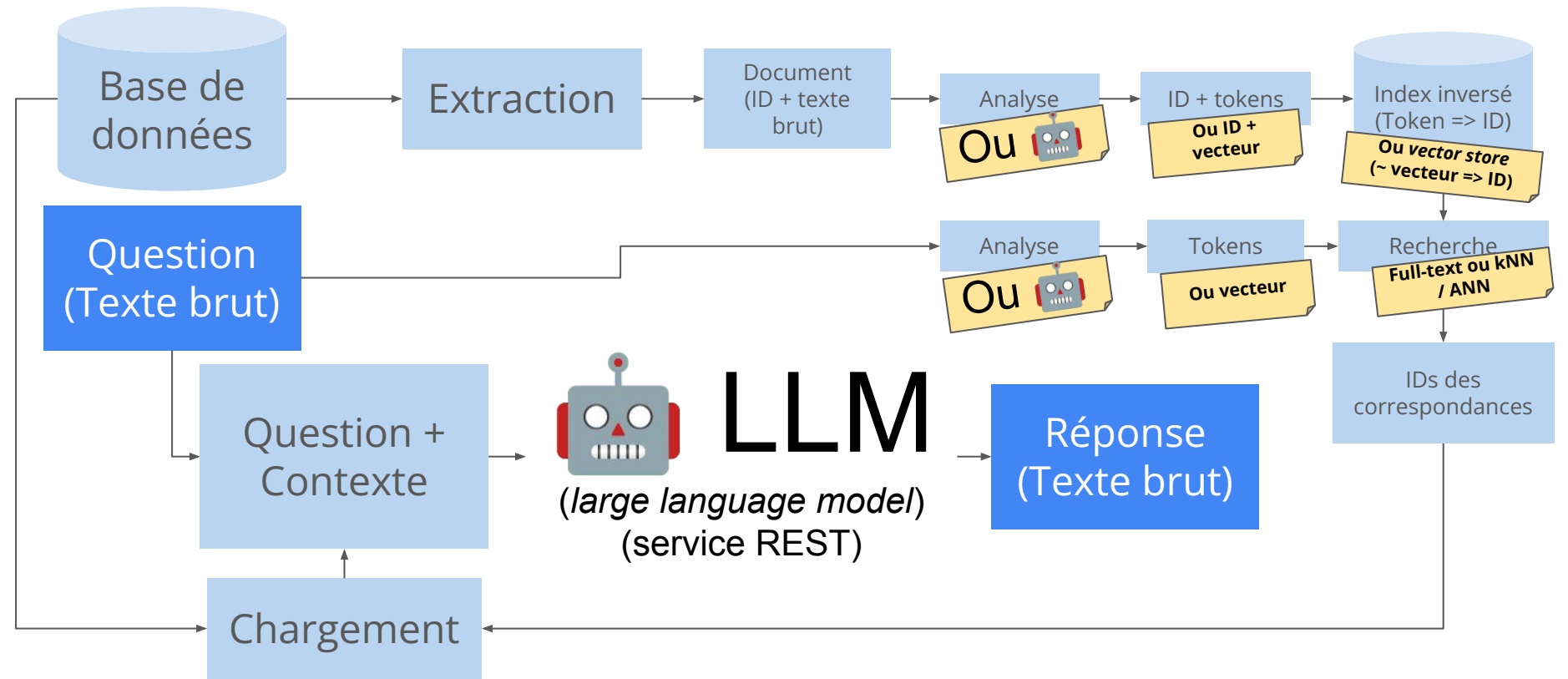
Créer le bon modèle



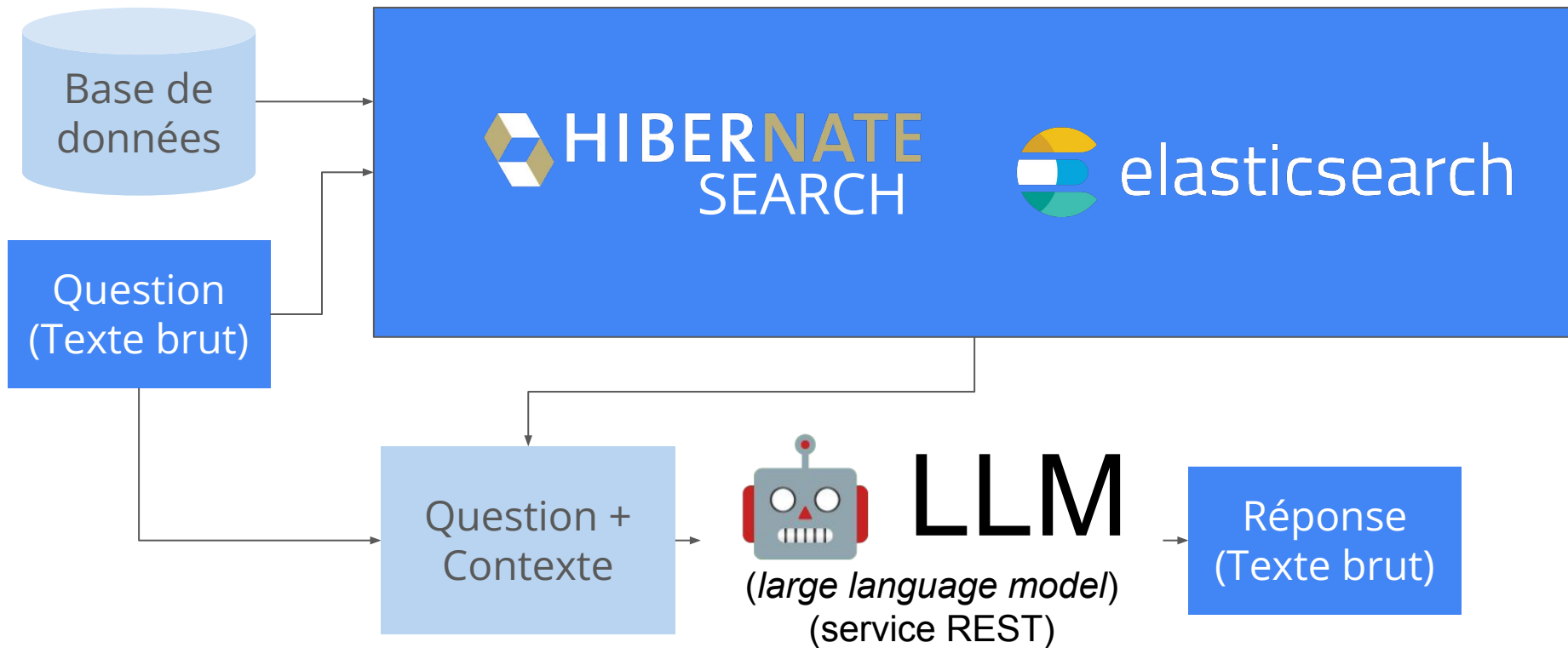
<https://xkcd.com/1425/>

This work is licensed under a [Creative Commons Attribution-NonCommercial 2.5 License](#).

Retrieval Augmented Generation (RAG)



Retrieval Augmented Generation (RAG)



Podman Desktop & AI Lab

The image shows two overlapping screenshots of the Podman Desktop AI Lab interface. The top screenshot displays the 'Models' section with a table of installed models:

NAME	SIZE	AGE	ACTIONS
TheBloke/Mistral-7B-Instruct-v0.1-GGUF Hugging Face - Apache-2.0	4.37 GB	5 days	RAM usage: 4.07 GiB CPU
NousResearch/Hermes-2-Pro-Mistral-7B-GGUF Hugging Face - Apache-2.0	4.37 GB	6 days	RAM usage: 4.07 GiB CPU
ibm/merlinite-7b-gguf Hugging Face - Apache-2.0			RAM usage: 4.07 GiB CPU
TheBloke/Mistral-7B-codealpaca-lora-GGUF Hugging Face - Apache-2.0			RAM usage: 4.07 GiB CPU

The bottom screenshot shows a 'Playgrounds' view for 'demo-redhat' with a chat interface. The chat contains the prompt 'Can you tell me the history of Red Hat?' and the assistant's response: 'Red Hat is a software company that was founded in 1993 by Marc Ewing and Robert Young. The company's initial focus was on providing Linux-based solutions for enterprise customers, but it has since expanded to offer a wide range of open-source software products and services. In its early years, Red Hat struggled to gain traction in the market, as many businesses were hesitant to adopt open-source technology. However, the company persevered and in 1995, it released its first version of the Red Hat Enterprise Linux operating system. This release marked a turning point for the company, as it helped to establish Red Hat as a leader in the open-source software space.'



Crafting a Local RAG application with Quarkus

Clément Escoffier

<https://quarkus.io/blog/granite-rag/>

Merci !



<https://quarkus.io/>



<https://hibernate.org/search/>



<https://quarkus.io/guides/#q=hibernate+search>