



HIBERNATE SEARCH 7

(et 6)

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

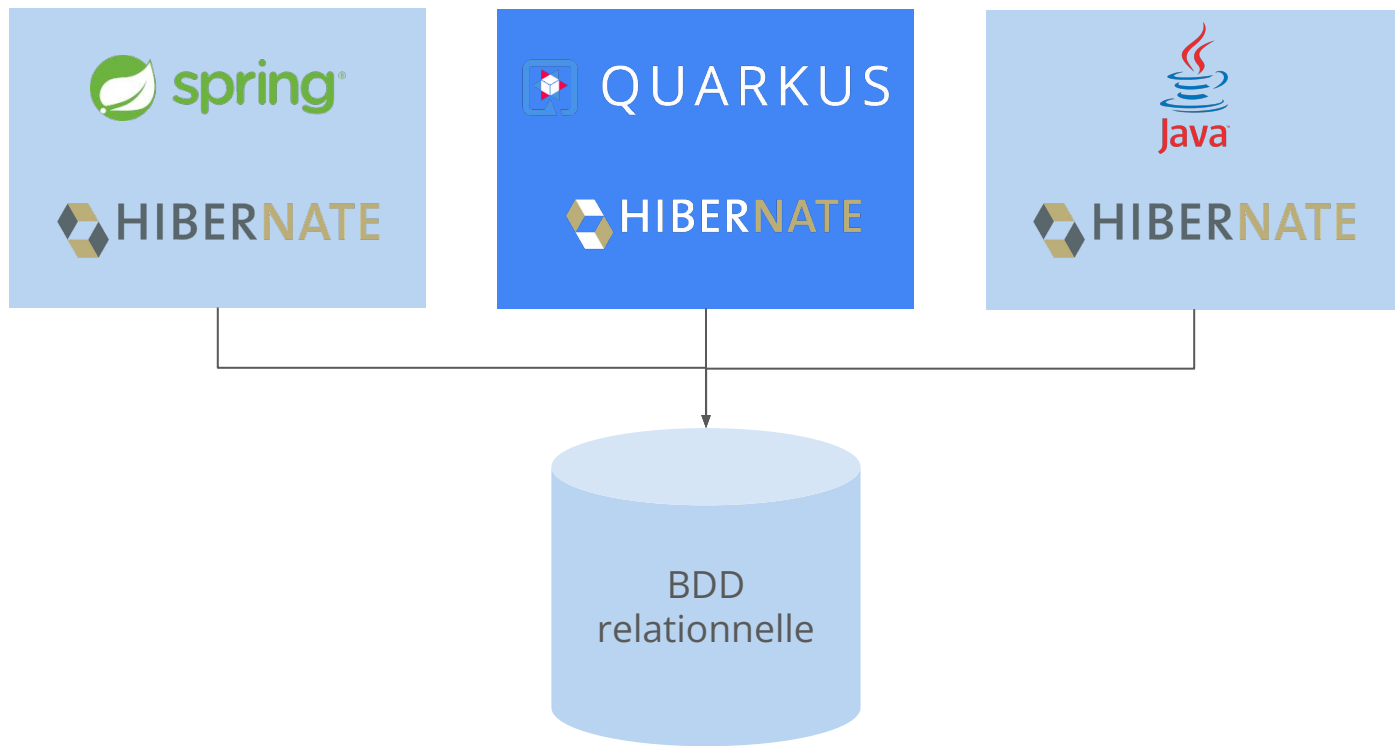


Yoann Rodière
Red Hat - Hibernate - Quarkus



Katia Aresti
Red Hat - Infinispan - Quarkus

Prenons une application Java...



Dossier de formation



Session



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 : ILIKE ?

ILIKE '%thé%'

'kinési**th**érapie'



ILIKE '%pommes**s**%'

'pomme'



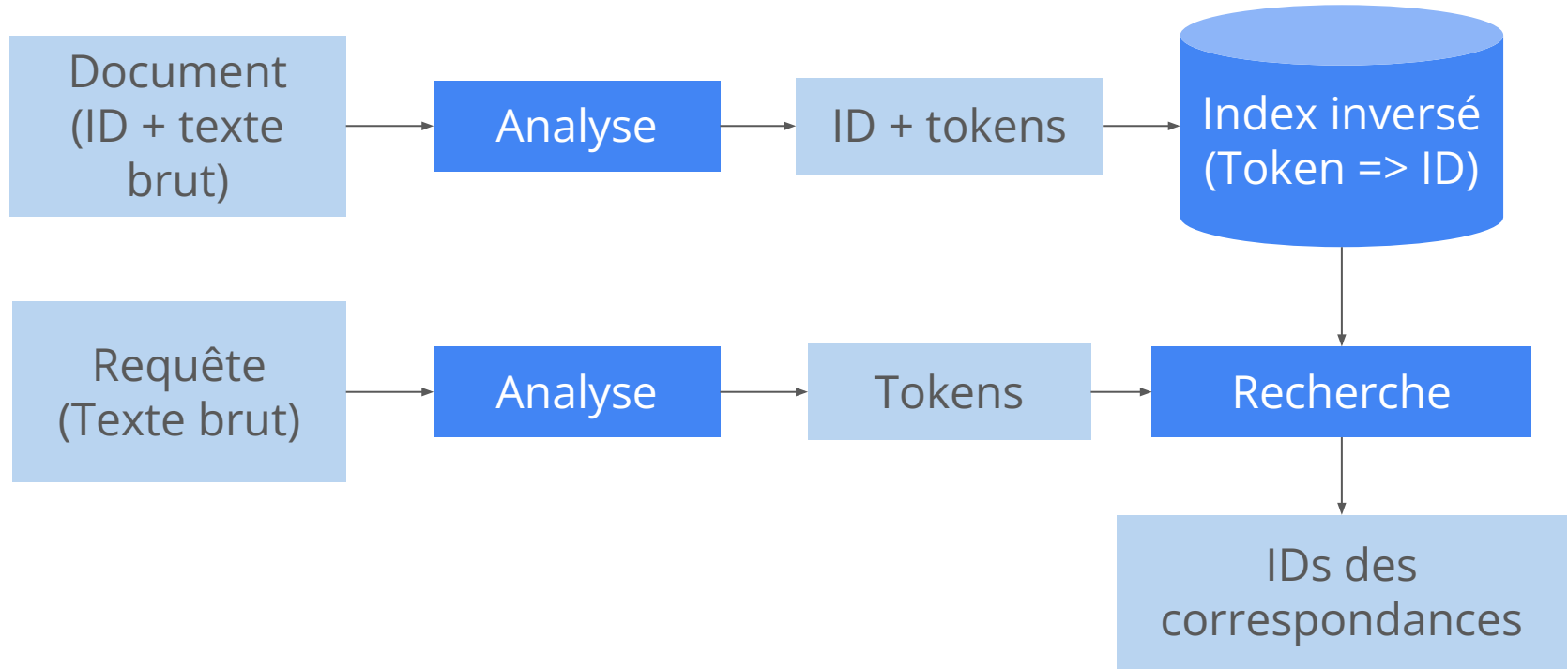
Perf



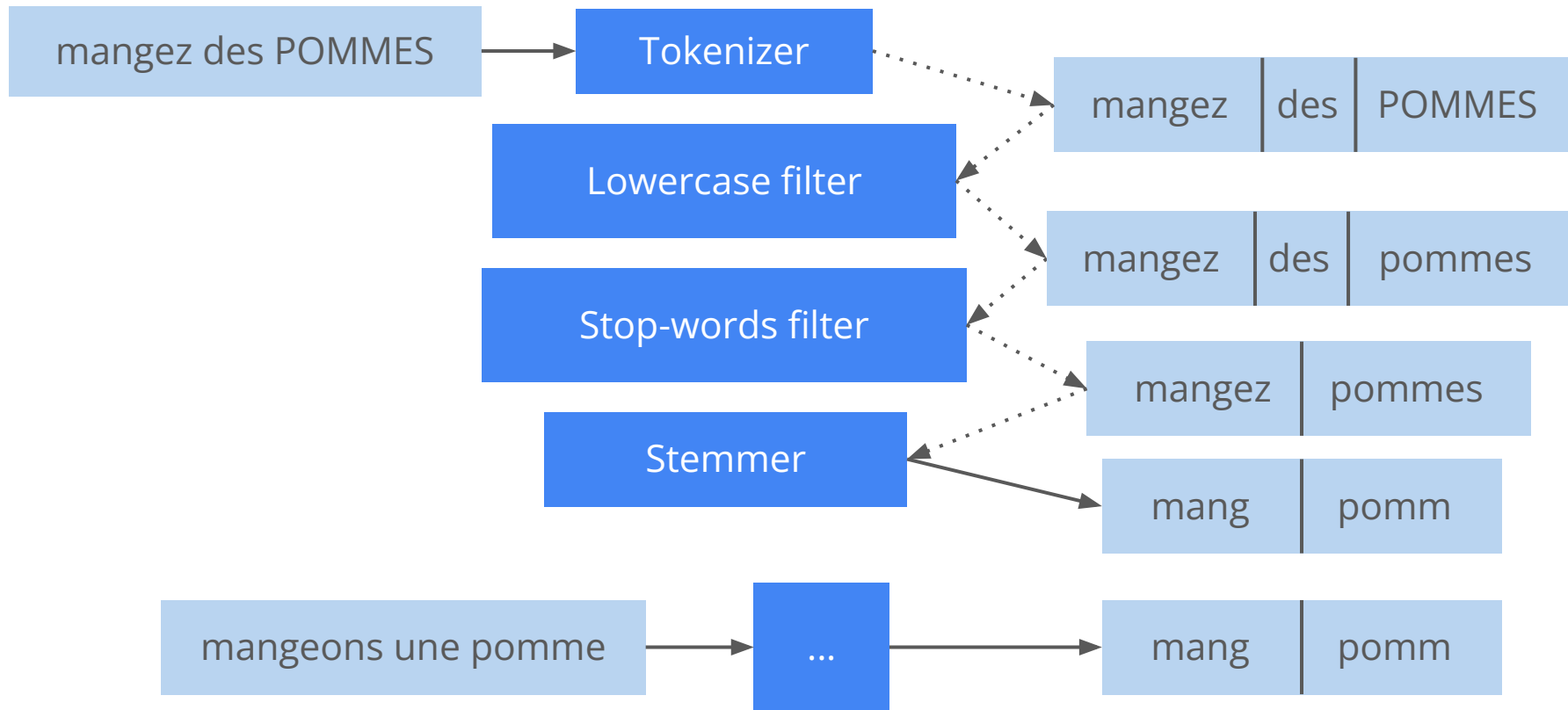
Tri

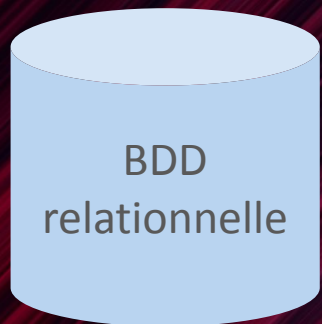


Recherche full-text



Analyse



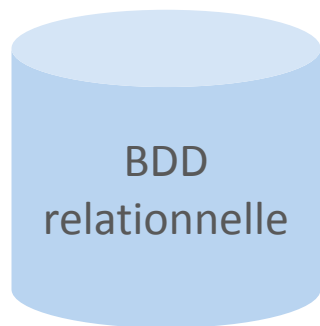


- Syntaxe dépend de la BDD
- Fonctionnalités dépendent de la BDD
- Charge supplémentaire... ?

APACHE
LUCENETM



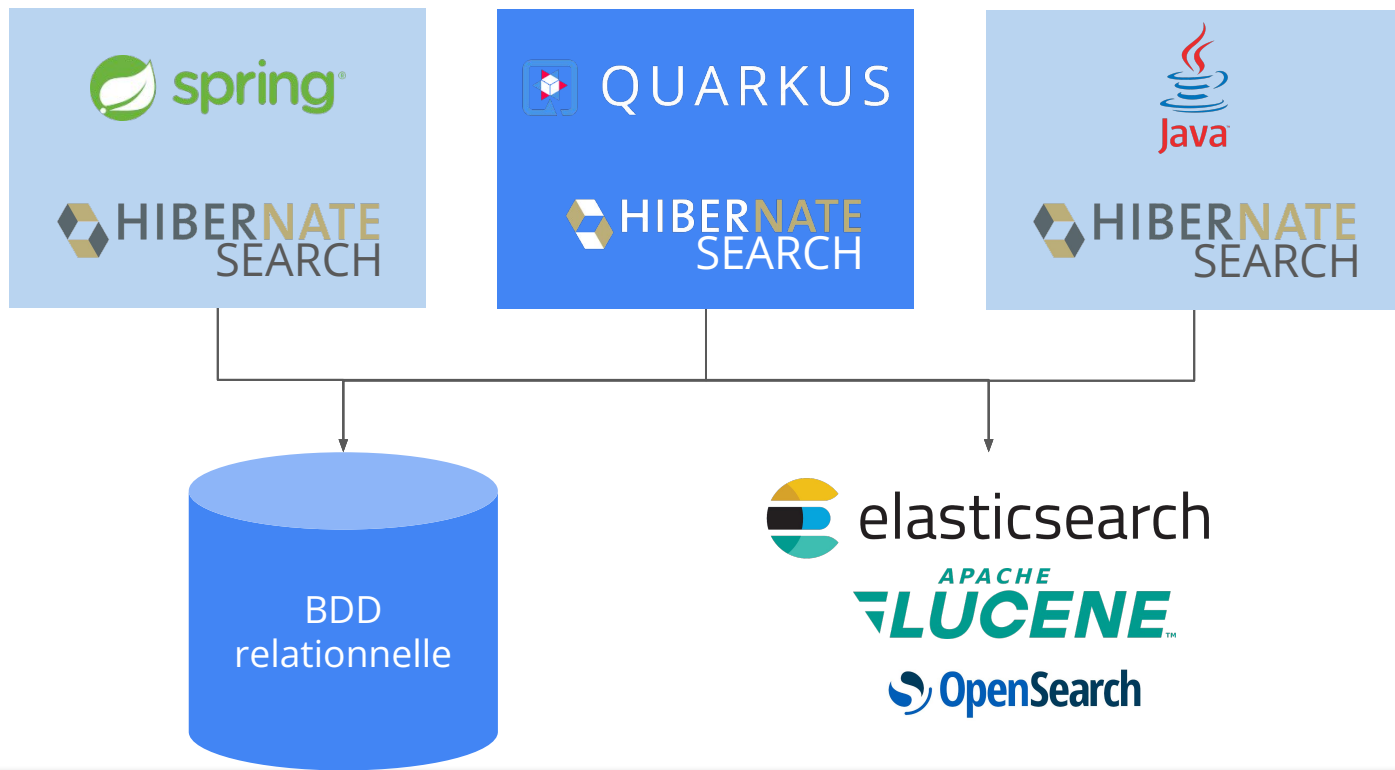
Comment transférer les données ?





ORM. Lucene. Elasticsearch.
Integrated.

Toujours Quarkus... mais pas seulement

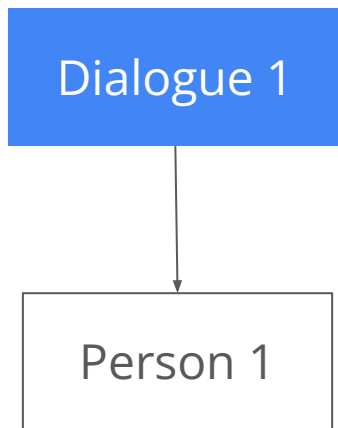


Démo !

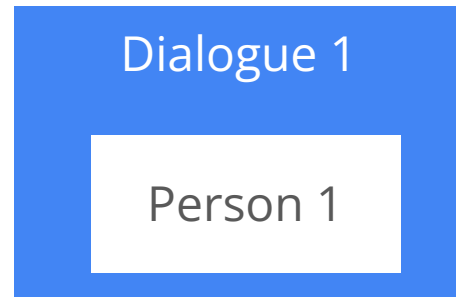
<https://github.com/yrodiere/rivieradev-2024-hibernate-search-demo>

Dénormalisation – cas trivial

 **HIBERNATE** Entités

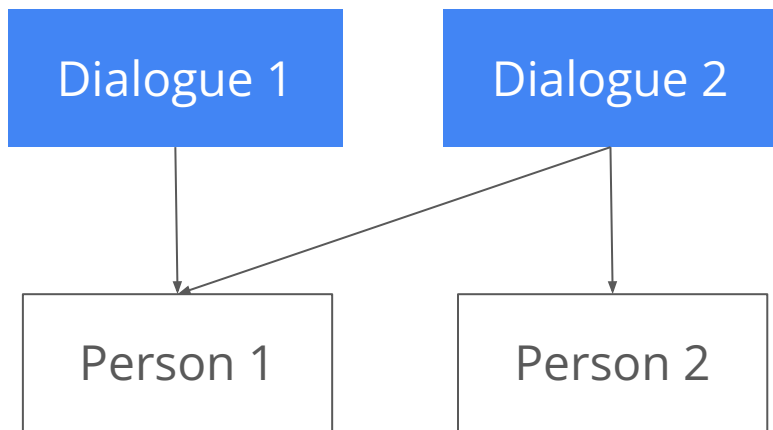


 **APACHE LUCENE** Document

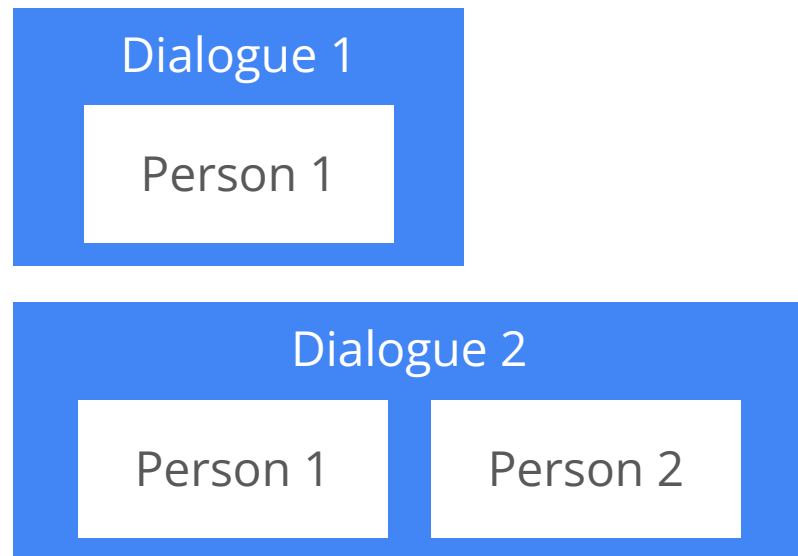


Dénormalisation – cas plus commun

HIBERNATE Entités

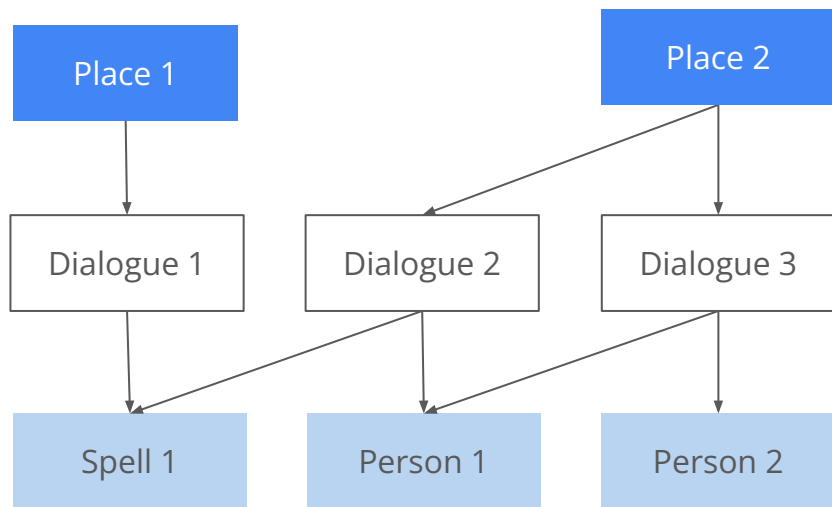


APACHE LUCENE Document

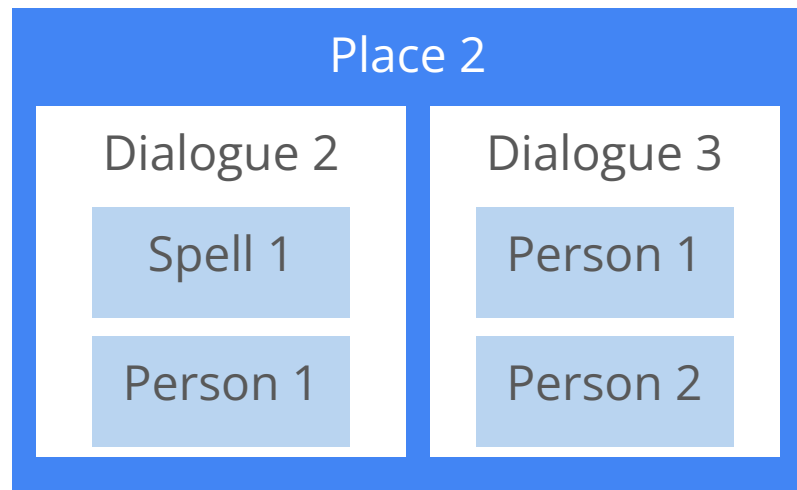


Dénormalisation – cas complexe

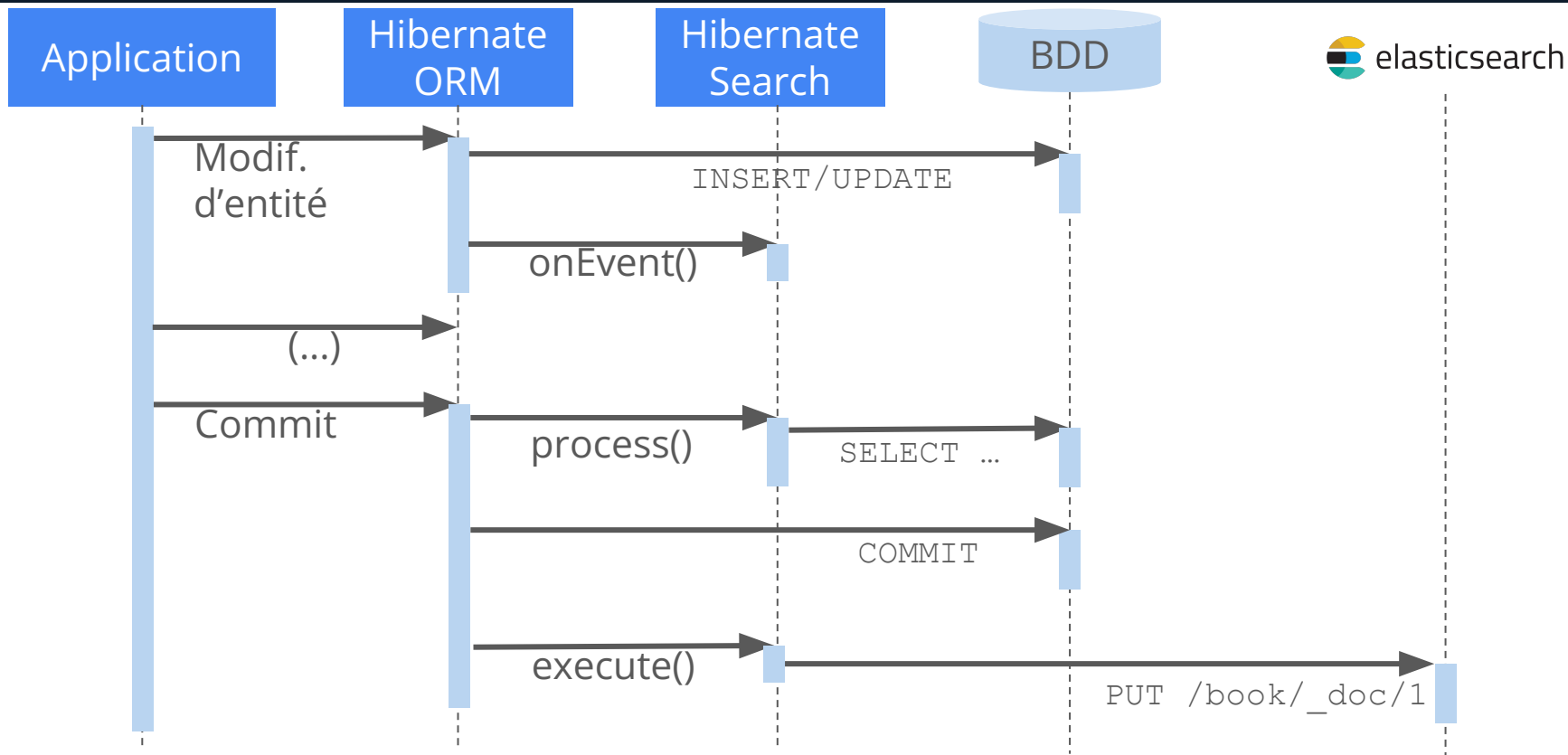
HIBERNATE Entités



APACHE LUCENE Document



Indexation



The logo consists of a stylized hexagon made of three smaller hexagons in white and gold.

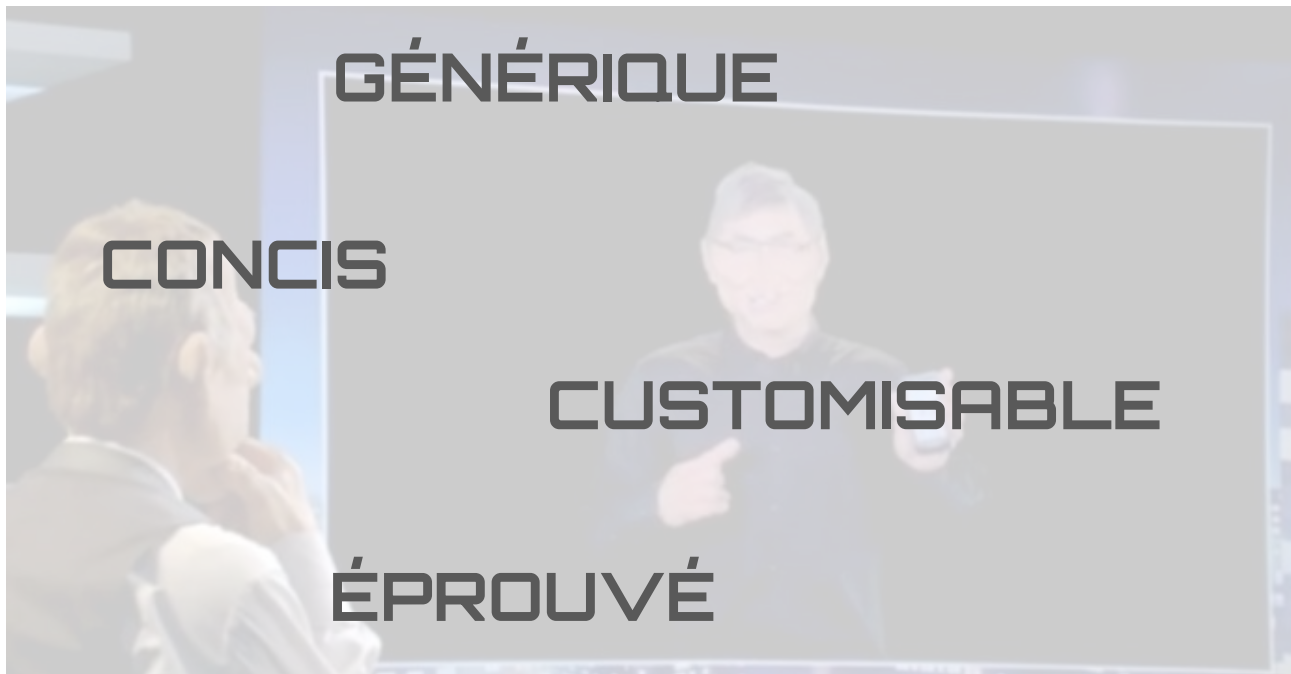
HIBERNATE SEARCH

Version 6+
2020

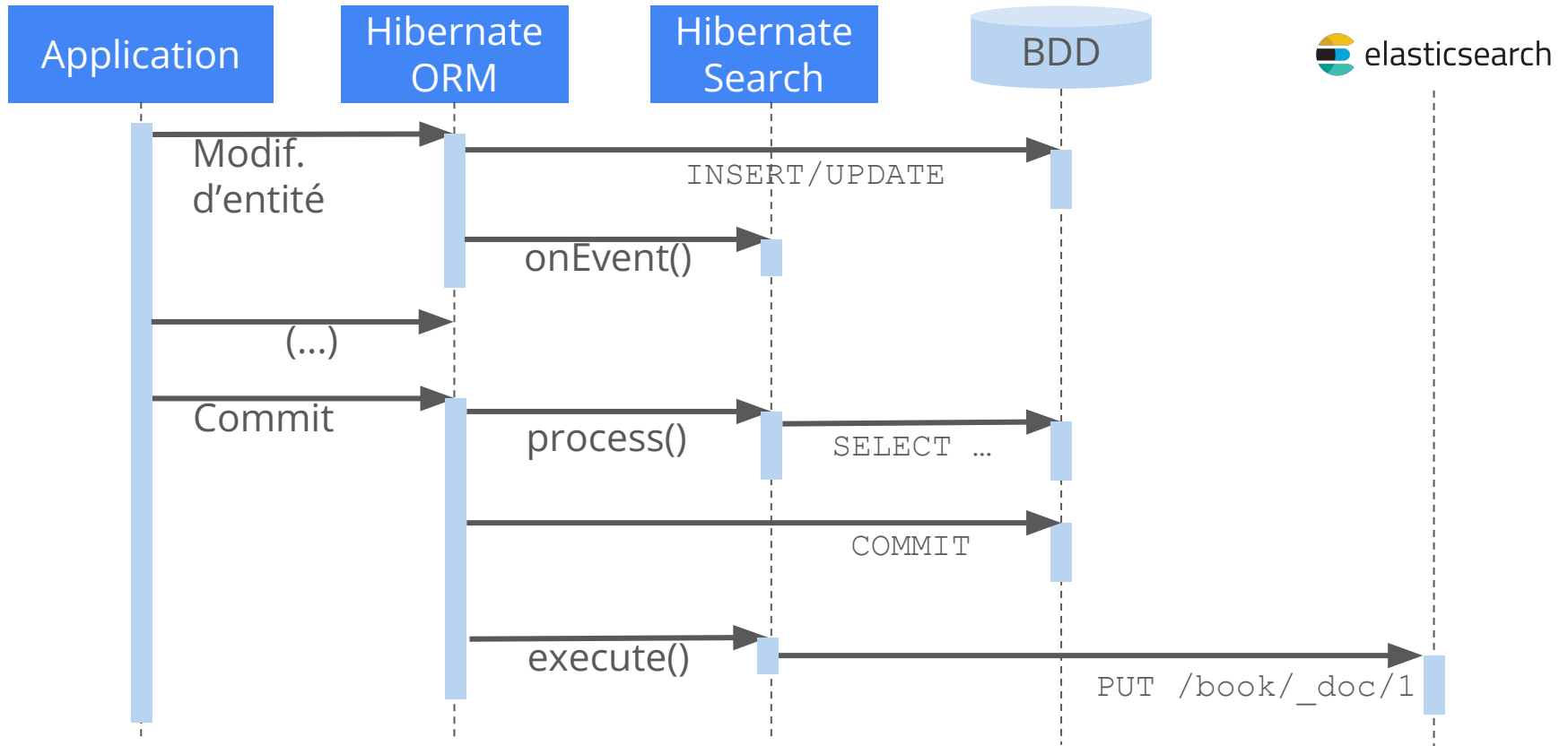
Refonte de l'API



Refonte de l'API



Et si ça se passe mal... ?

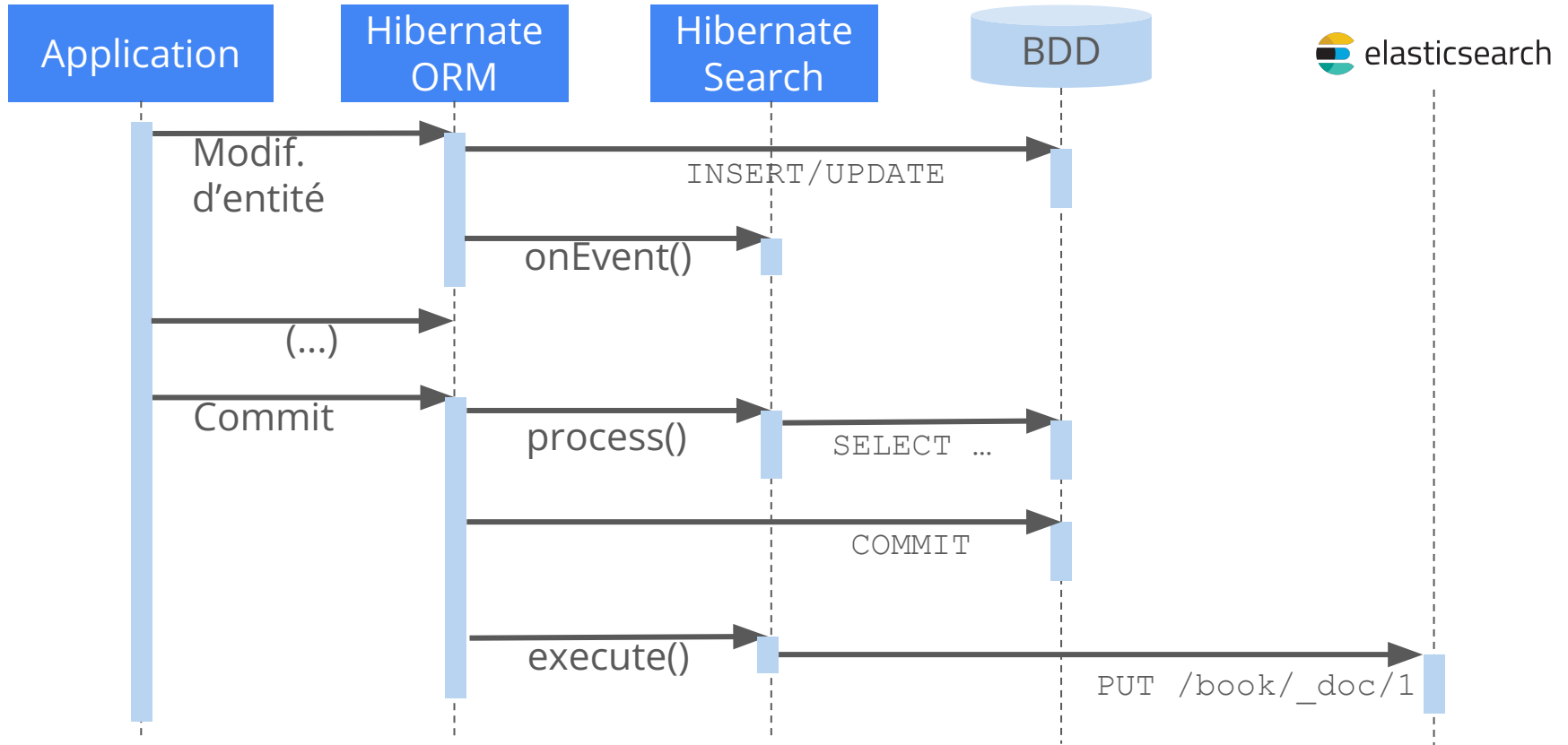


Transactional outbox

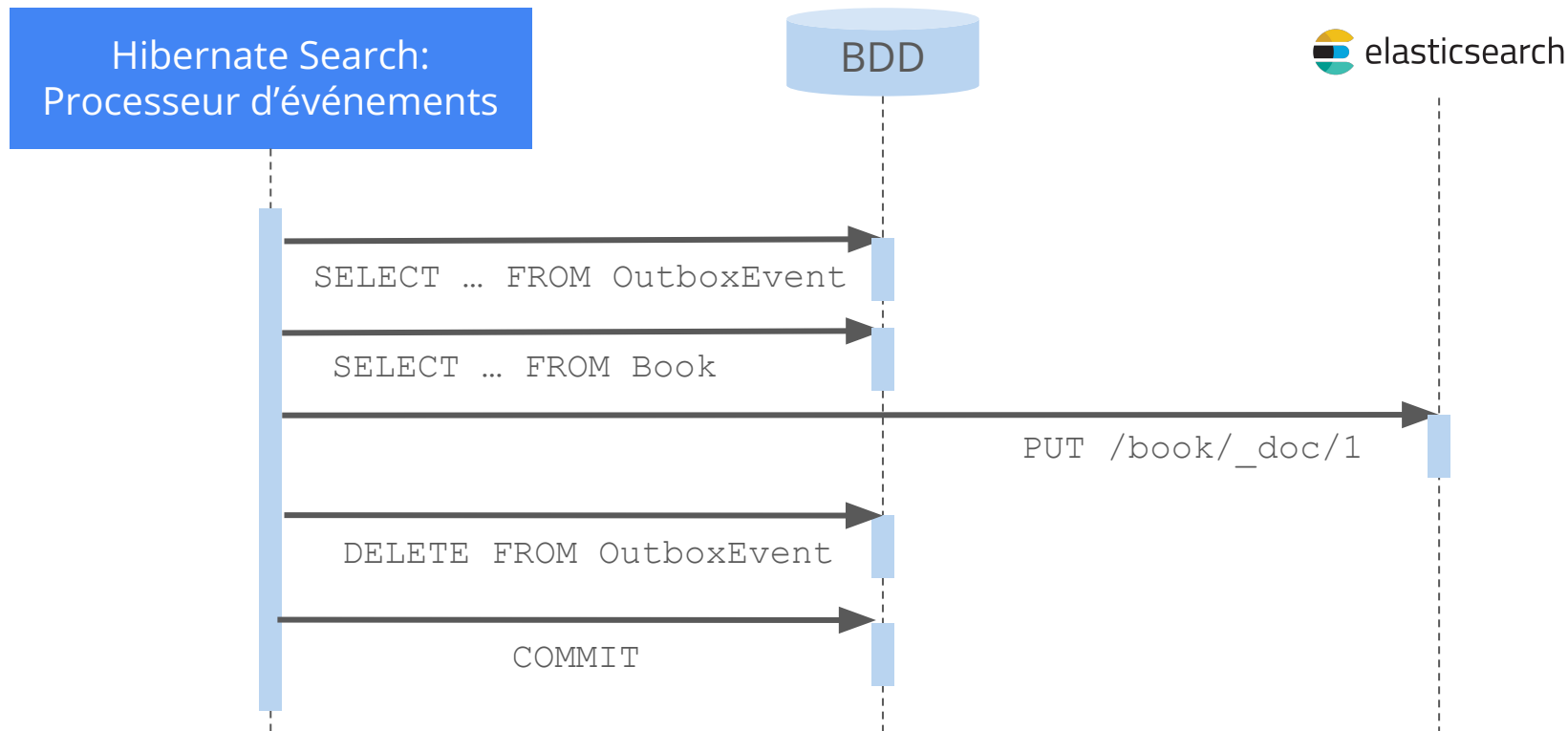
https://docs.jboss.org/hibernate/stable/search/reference/en-US/html_single/#coordination-outbox-polling



Transactional outbox 1 - application



Transactional outbox 2 - processeur d'événements



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
```

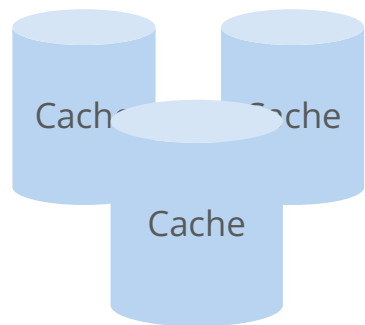
Standalone Pojo Mapper

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



Standalone Pojo Mapper





Hibernate Search +
intégration custom

APACHE
LUCENETM



Infinispan

LangChain Embedding Store



Infinispan



Lanchain 4J

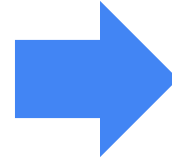


LangChain



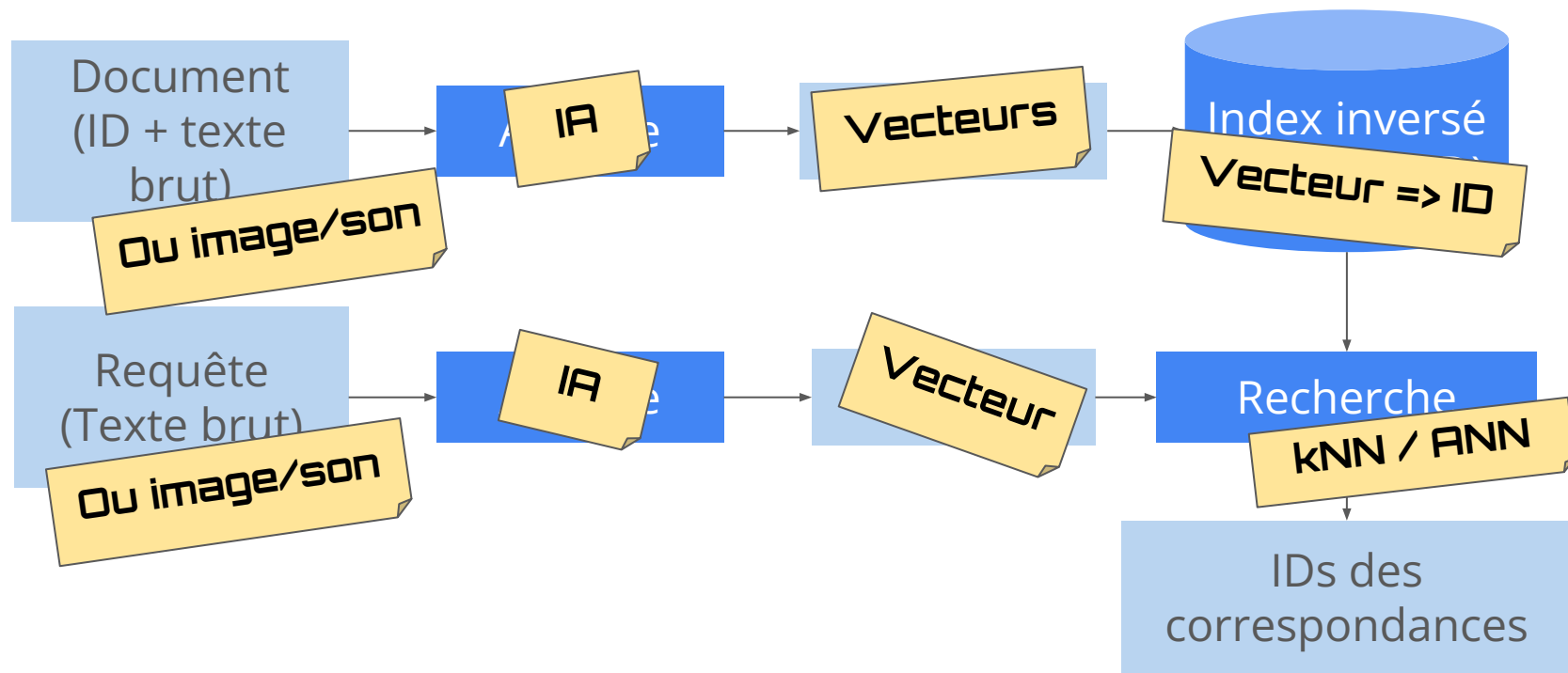
Version 7
2023 / 2024

Vecteurs et *embeddings*



$[-0.7754,$
 $-0.7972,$
 $-0.5784,$
 \dots
 $-0.4750,$
 $-0.4254]$

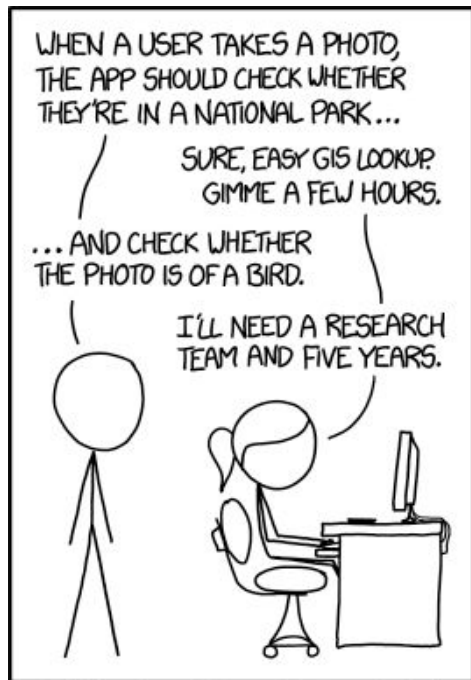
Recherche vectorielle



Démo !

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

Créer le bon modèle



STAY TUNED

Merci !



<https://quarkus.io/>



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



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