

NoSQL

Module: 165

Utiliser des bases de données
NoSQL

Base de données NoSQL

*Les indexes, recherche textuelle
et géospatiale*

MongoDB

Les index

MongoDB

Introduction

Les index peuvent être considérés comme des raccourcis pour accéder à vos données, de sorte qu'il n'est pas nécessaire d'analyser l'intégralité de la base de données pour trouver ce que vous recherchez.



MongoDB

À quoi servent les index ?

Basse de données

Collection

```
{  
  _id: ObjectId('798'),  
  model: 'Twingo',  
  prix: 4500  
}
```

```
{  
  _id: ObjectId('458'),  
  model: '206',  
  prix: 3800  
}
```

```
{  
  _id: ObjectId('324'),  
  model: 'Clio',  
  prix: 14000  
}
```



Requête de recherche **SANS** indexes

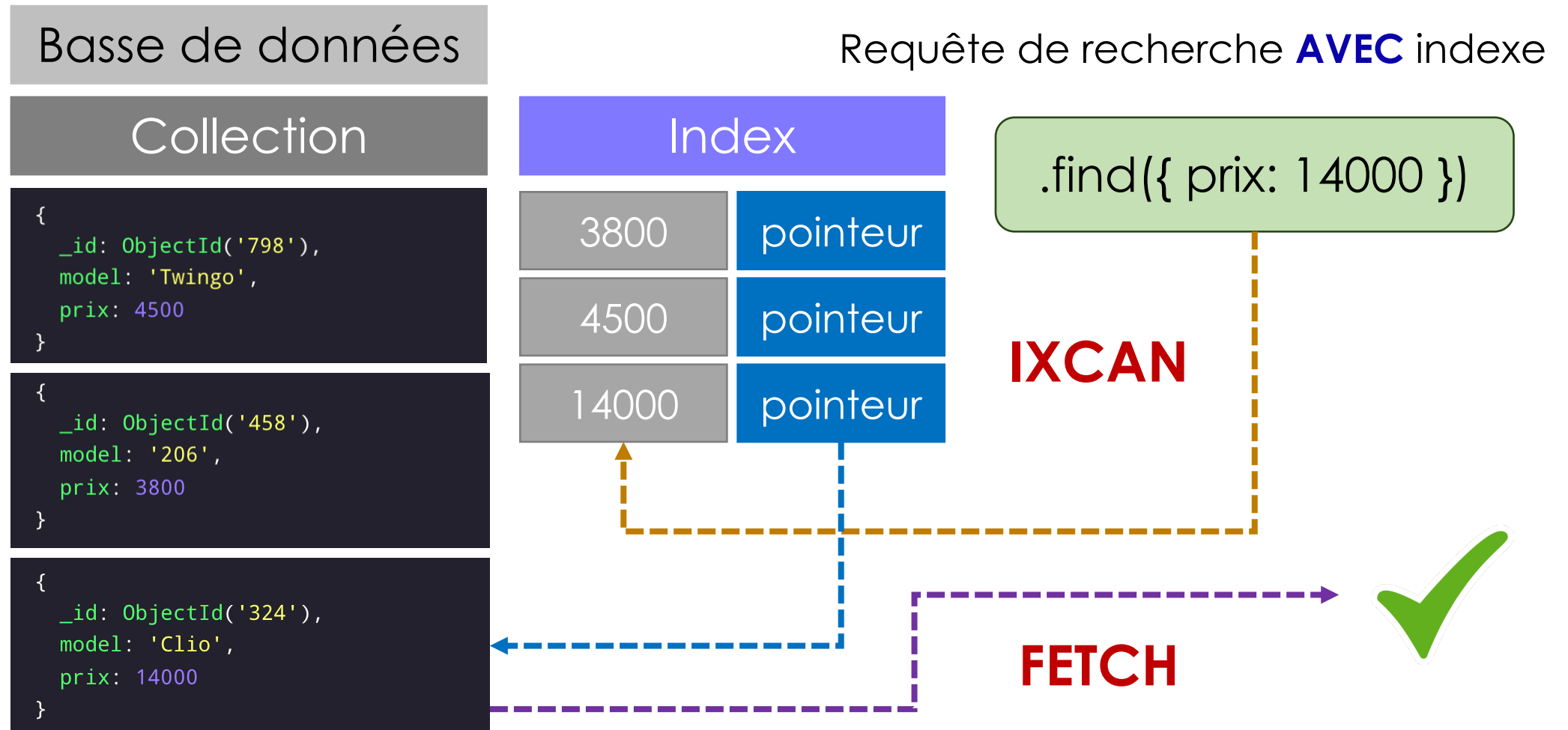
```
.find({ prix: 14000 })
```

Scan de la collection

COLLSCAN

MongoDB

À quoi servent les index ?



MongoDB

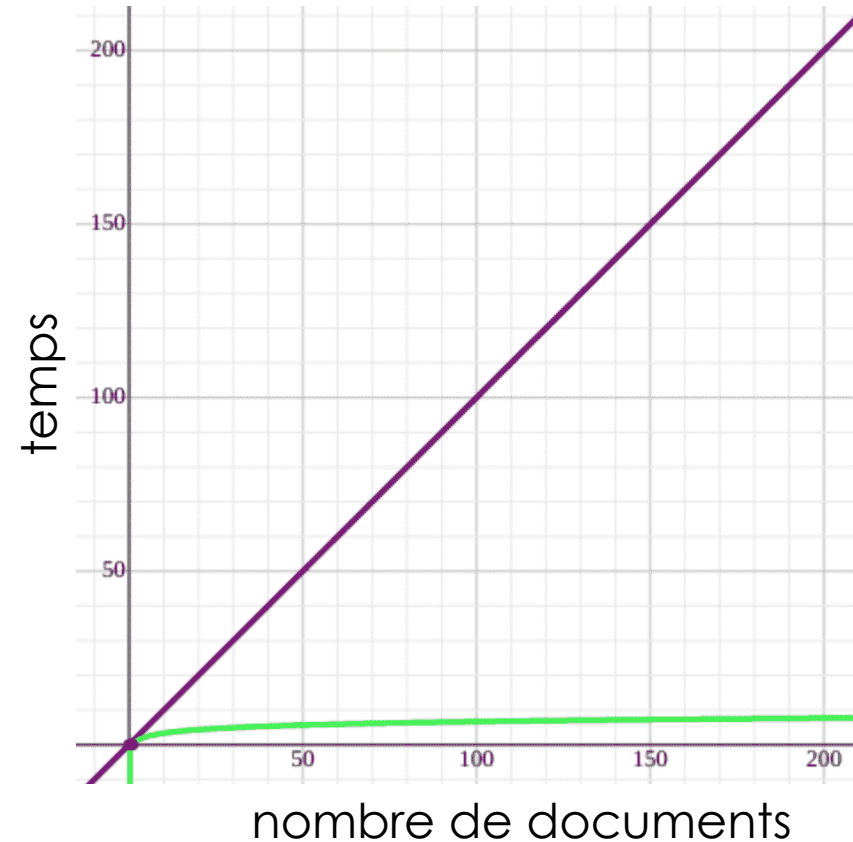
Les indexes : *Pros* et *Cons*

Cons

- Ralentit l'écriture
- Prend de l'espace disque
- Prend de la RAM

Pros

- Accélère les recherches



1. En **vert**, nombre de documents à examiner pour une recherche avec un index.
2. En **violet**, nombre de documents à examiner pour une recherche sans index.

MongoDB

Quand créer un index ?

Quand une requête fréquente devient lente



Nombre de document examinés >> Nombre de document retournés

```
db.collection.explain('executionStats').find({})
```


MongoDB

Visualiser les index

Vous pouvez le visualiser facilement dans **Compass**, en cliquant sur n'importe quelle collection, puis en cliquant sur l'onglet **Indexes** :

sample_mflix.users 185 DOCUMENTS 2 INDEXES

Documents Aggregations Schema **Indexes** Validation

Create Index Refresh VIEWING **INDEXES** SEARCH INDEXES

Name and Definition	Type	Size	Usage	Properties
> _id_	REGULAR ⓘ	24.6 KB	Usage data unavailable	UNIQUE ⓘ
> email_1	REGULAR ⓘ	24.6 KB	Usage data unavailable	UNIQUE ⓘ

MongoDB

Afficher les indexes

```
db.myColl.getIndexes()
```

MongoDB

Supprimer des indexes

Supprimer un indexe spécifique:

```
db.myColl.dropIndex( "indexName" )
```

```
db.myColl.dropIndex( "key": -1 )
```

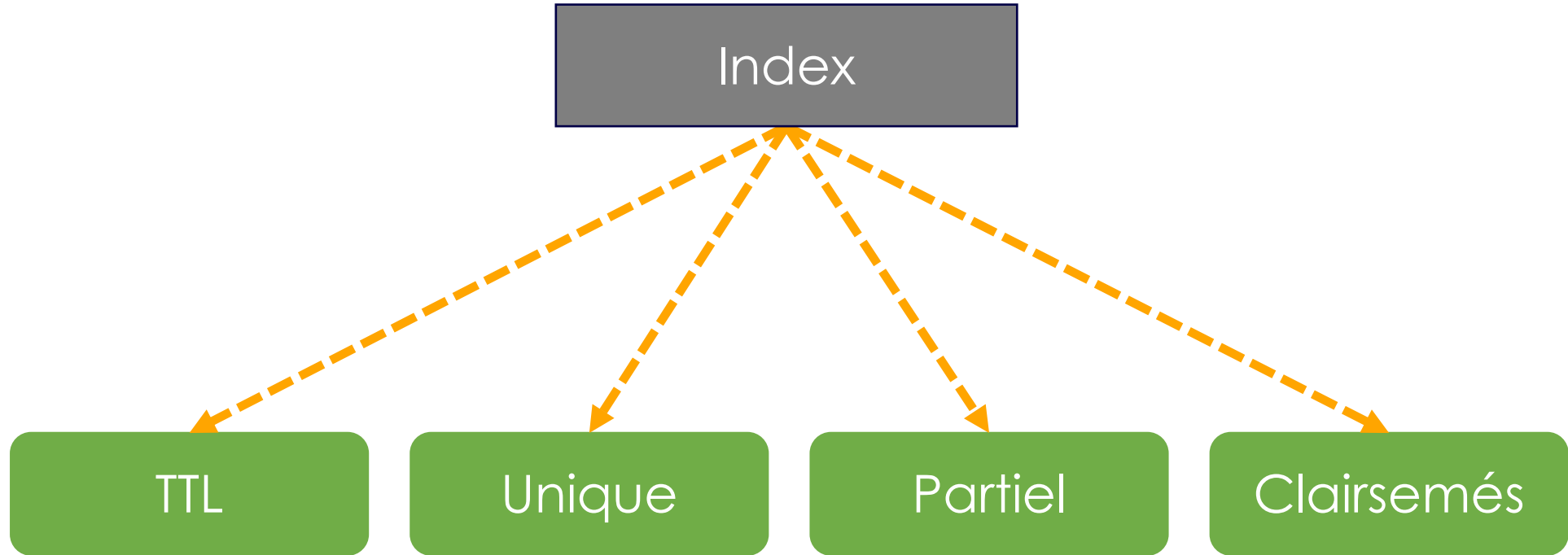
Supprimer tous les indexes ou des indexes spécifiques:

```
db.myColl.dropIndexes()
```

```
db.myColl.dropIndexes(["indexNameA", "indexNameB"])
```

MongoDB

Les types d'index



MongoDB

La recherche textuelle

MongoDB

La recherche textuelle

Comment apprendre à bien jouer de la guitare ?



Collection: Articles

```
{  
  article: 'Guide sur l\'apprentissage de la guitare',  
  description: 'Vous souhaitez bien jouer de la guitare voilà comment faire.'  
}
```

```
  article: 'Un gros plan sur les oiseaux migrateurs',  
  description: 'Que savez-vous réellement des oiseaux migrateurs ?'  
}
```

?

MongoDB

La recherche textuelle

```
{  
  article: 'Un gros plan sur les oiseaux migrateurs',  
  description: 'Que savez-vous réellement des oiseaux migrateurs ?'  
}
```

Index TEXTUEL

article & description

Texte

pointeur

pointeur



MongoDB

La recherche textuelle

Algorithmes de NLP (Natural Language Processing)

```
article: 'Un gros plan sur les oiseaux migrants',  
description: 'Que savez-vous réellement des oiseaux migrants ?'  
}
```

Étape 1 => Transformation

que

savez

vous

réellement

des

oiseaux

migrateur

?

Étape 2 => Stop-words

savez

réellement

oiseaux

migrateur

Étape 3 => Stemming

sav

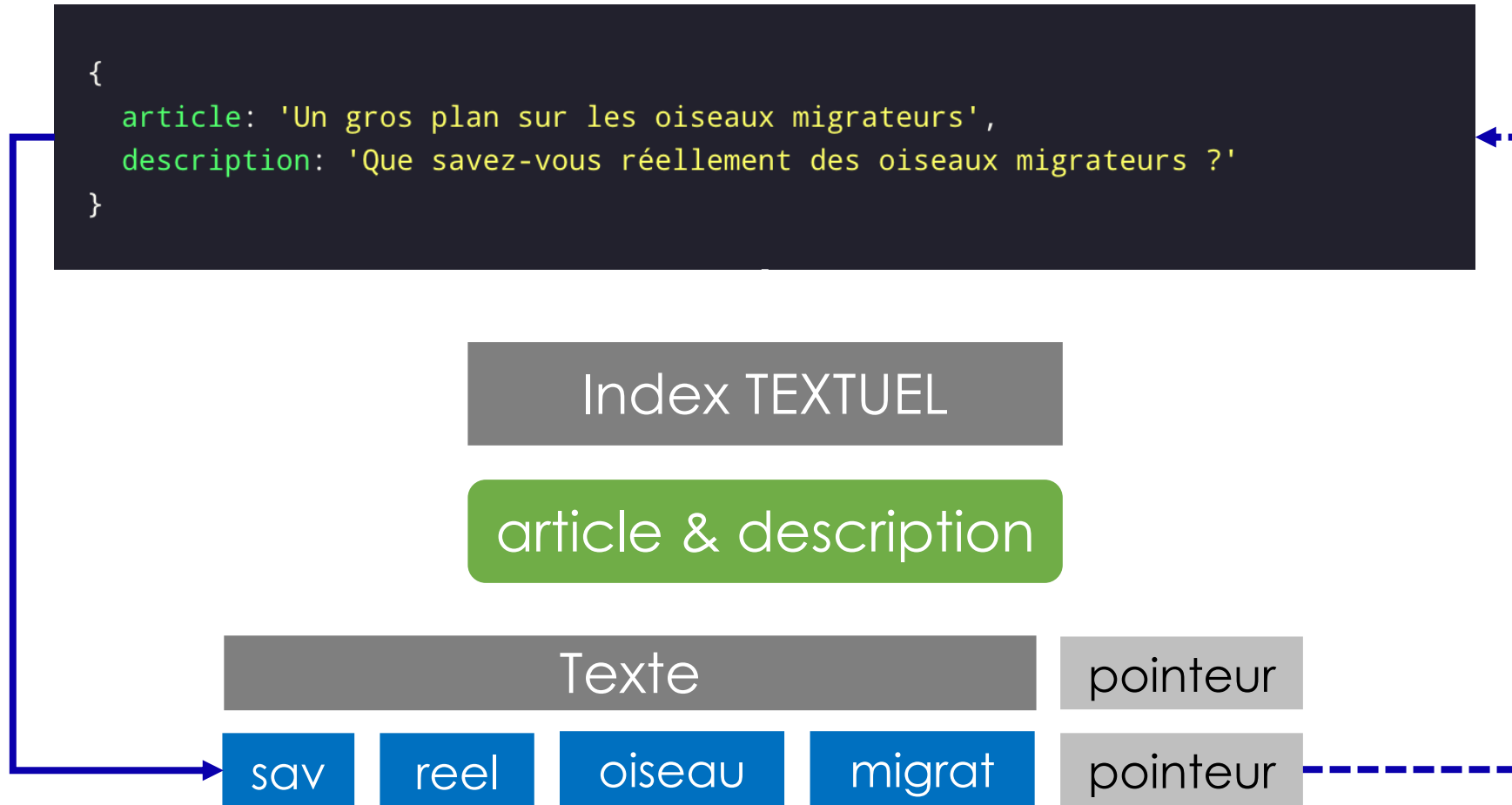
reel

oiseau

migrat

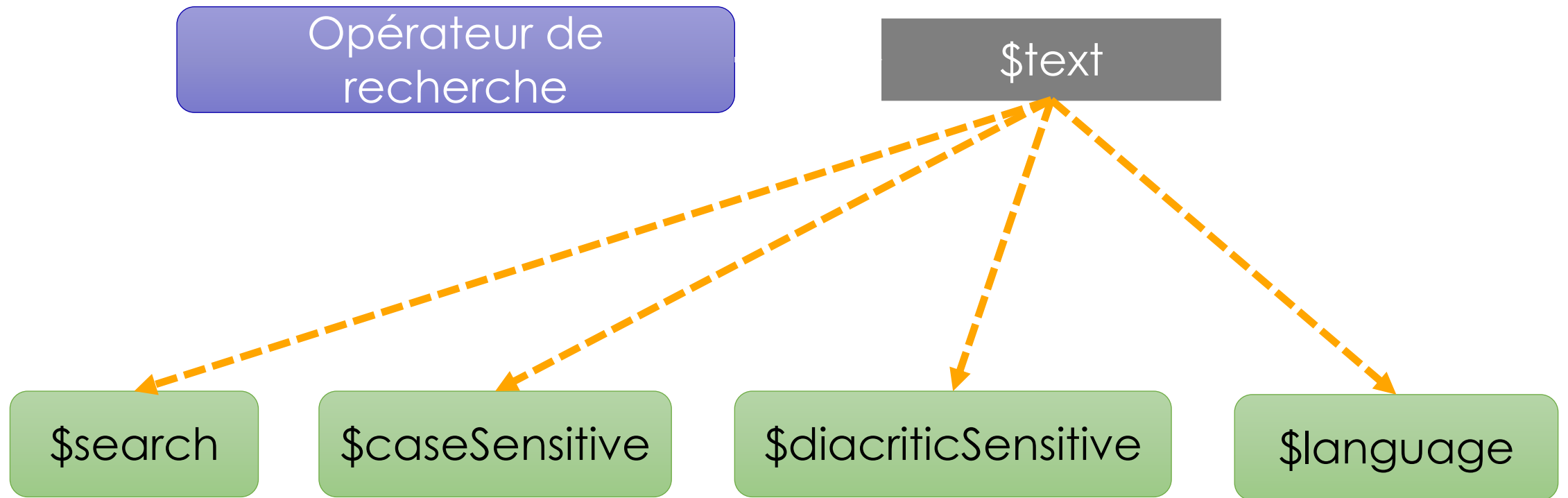
MongoDB

La recherche textuelle



MongoDB

La recherche textuelle



MongoDB

La recherche textuelle

Création d'un index texte spécifique:

```
db.articles.createIndex(  
  { content : "text" },  
  { default_language: "french" }  
);
```

Par défaut, la langue utilisée par les index textuels par MongoDB est l'anglais. Il est donc important de spécifier la langue.

MongoDB

La recherche textuelle

Une fois vos index textuels en place, vous allez pouvoir utiliser l'opérateur **\$text** pour rechercher parmi ces index.

```
db.articles.find( { $text: { $search: "une abeille bourdonne" } } )
```

Si vous ne voulez que les documents qui contiennent obligatoirement les deux termes, c'est-à-dire une recherche ET. Il faut utiliser **\ "termes exacts" **.

Si à l'inverse, vous voulez exclure un terme d'une recherche, il suffit d'utiliser « - ». Exemple **"abeille bourdonne -frelon"**.

MongoDB

La recherche textuelle

L'algorithme de recherche assigne un score pour chaque document qui contient les termes recherchés dans les champs textuels indexés. En utilisant l'opérateur **\$meta**, l'on va pouvoir afficher et aussi trier selon le score.

```
db.articles.find(  
  { $text: { $search: "Comment bien jouer de la guitare ?" } },  
  { score: { $meta: 'textScore' } }  
) .sort( { score: { $meta: "textScore" } } );
```

Pour afficher le score de recherche ajouter : **score: { \$meta: 'textScore' }**.

Afin vous pouvez trier directement par score de recherche avec:
.sort({ score: { \$meta: 'textScore' } }).

MongoDB

La recherche géospatiale

MongoDB

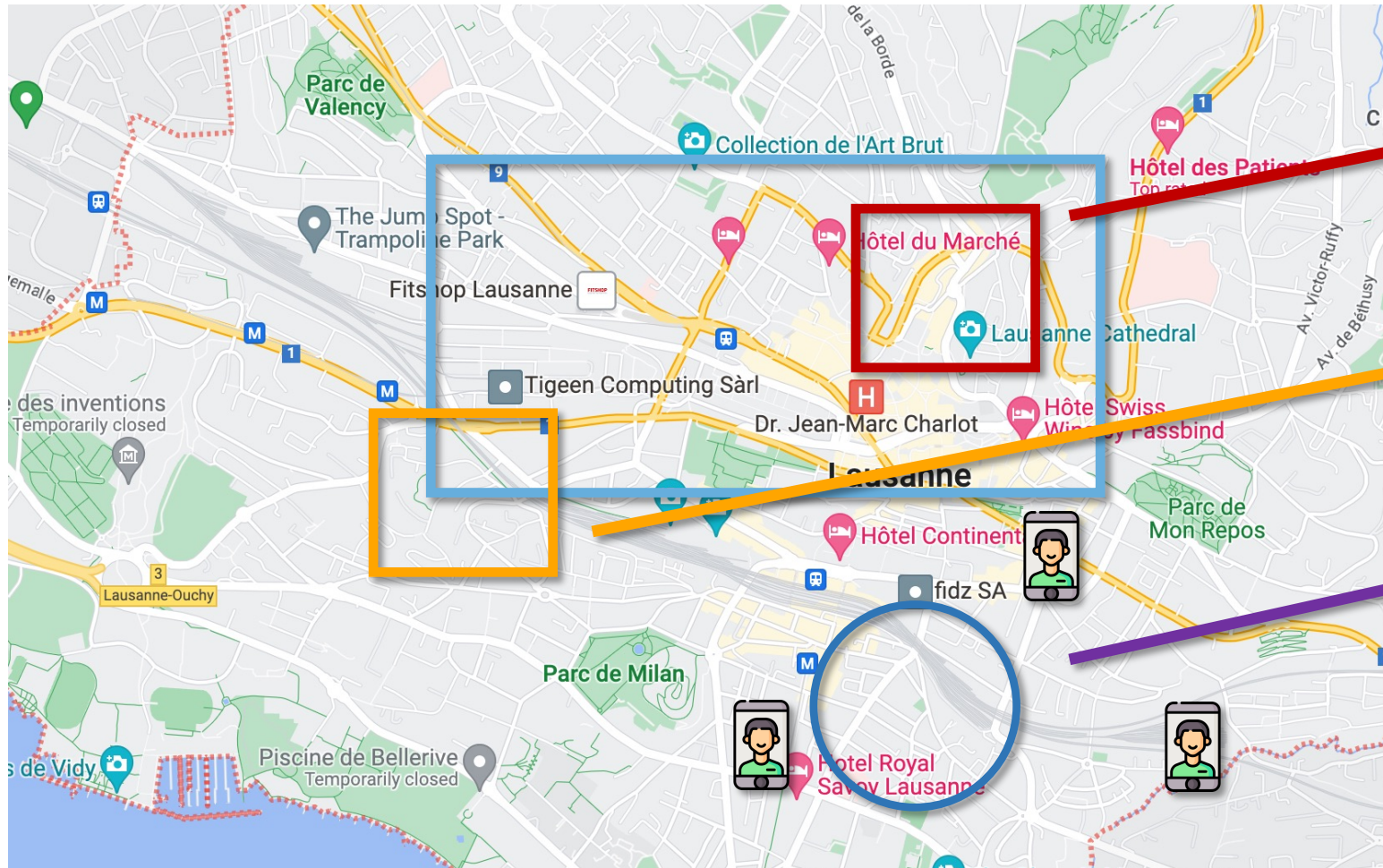
La recherche géospatiale

MongoDB offre aussi des indexes et des opérateurs spécifiques pour des recherche géospatiale. Les index **2dsphere** permettent d'effectuer des requêtes géospatiales sur des documents contenant des coordonnées au format **GeoJSON**.

- **\$geoIntersects** : sélectionne les documents dont les données géospatiales intersectent l'objet GeoJSON spécifié.
- **\$geoWithin** : sélectionne les documents dont les données géospatiales sont entièrement contenues par l'objet GeoJSON spécifié.
- **\$near / \$nearSphere** : sélectionne les documents dont les données géospatiales sont proches de l'objet GeoJSON spécifié.

MongoDB

La recherche géospatiale



\$geoWithin

\$geoIntersects

\$near
\$nearSphere