

A story of auto-completions

MICES US 2022

@M_I_C_E_S

Lucian Precup and Radu Pop

@lucianprecup

@search_with_all

@a2lean

[adelean]

- Experts in Search Technologies
- Integrators of Elasticsearch and Solr
- Consulting, Services and Training providers
- Developers of a2 – your Search Engine Technology for E-Commerce and Enterprise Search



adelean
EXTRACT TRANSFORM SEARCH

[all.site]



Wikipedia and Wikidata -
knowledge databases →
powered by the community



Github - code and coders'
social network → powered by
the community

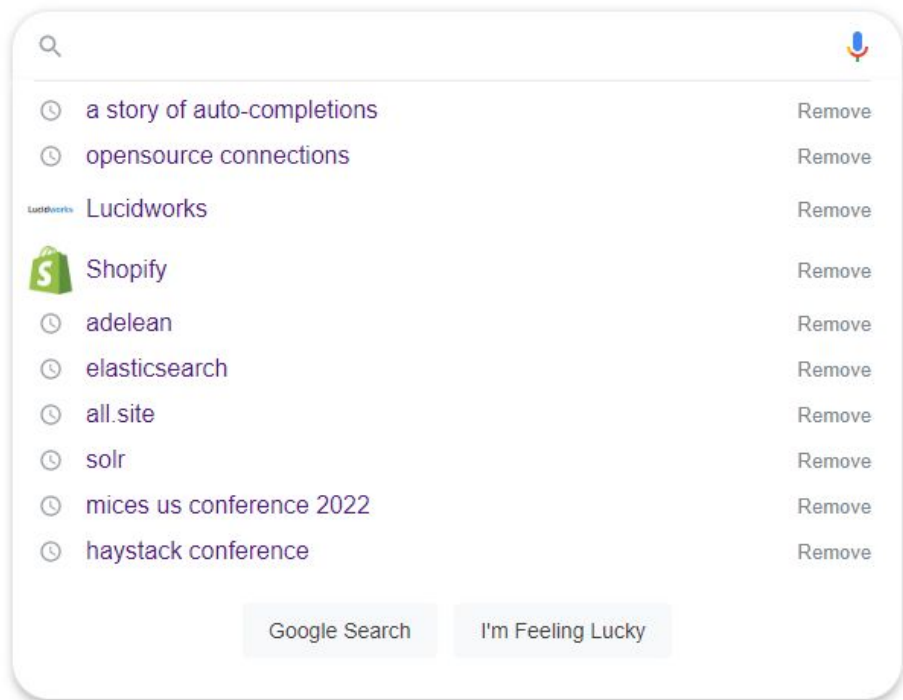


Reddit - the frontpage of the
Internet → powered by the
community

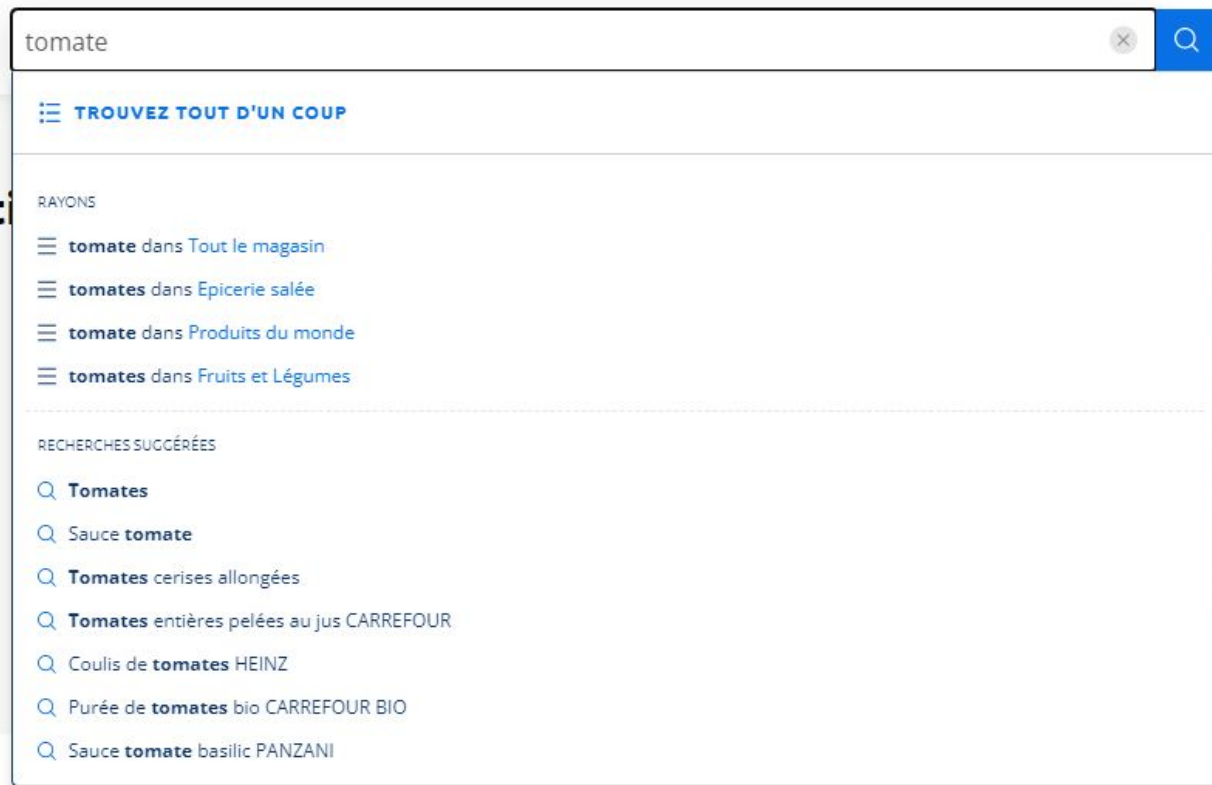


a// - your collaborative search engine
→ powered by the community

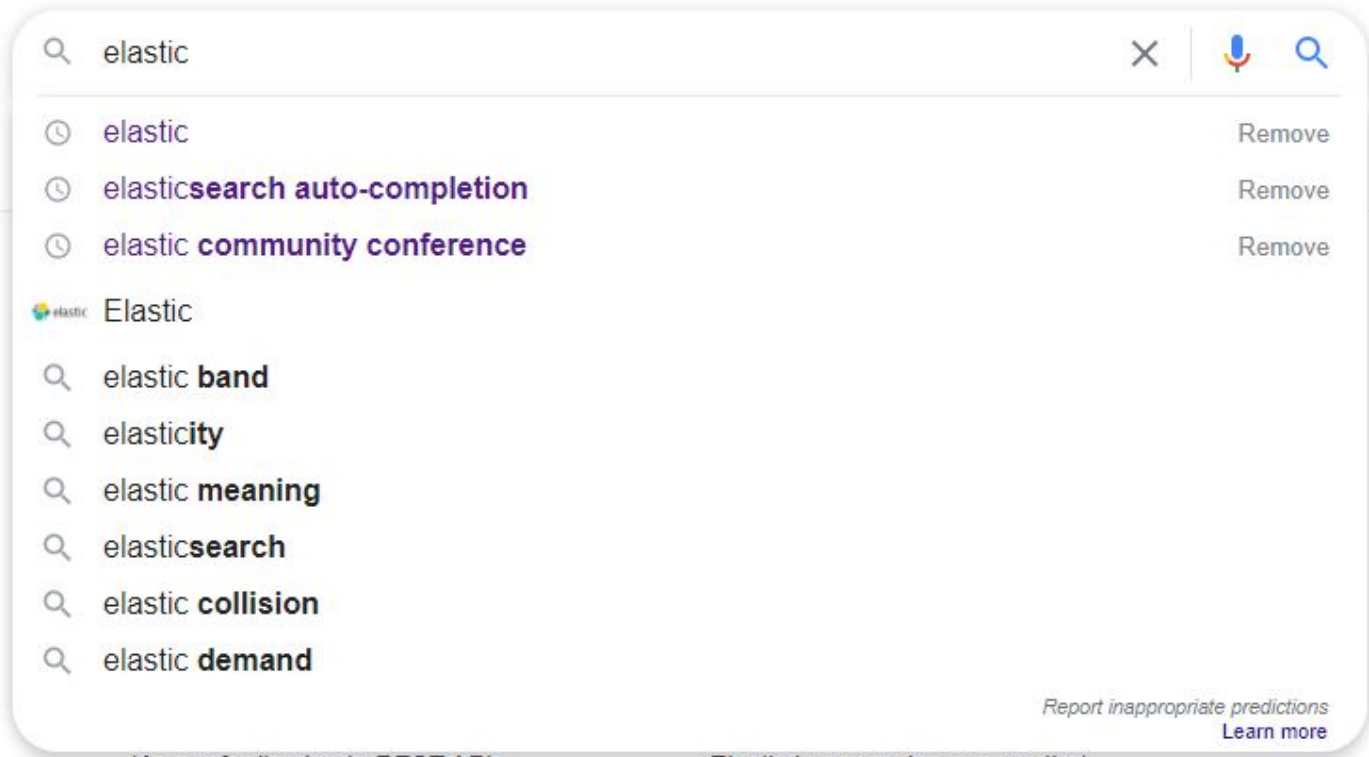
Auto-completion?



Auto-complétion?



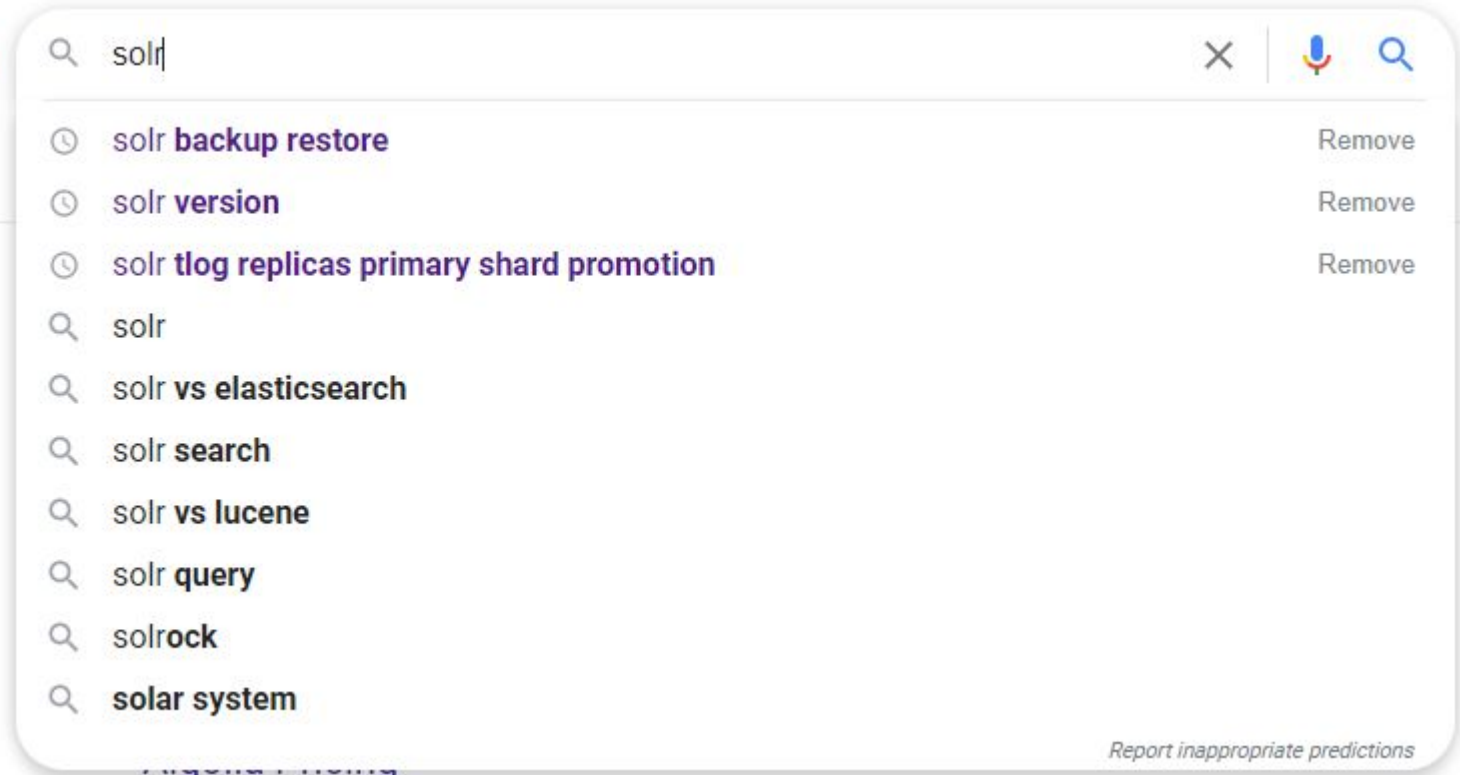
Auto-completion?



Known for its simple REST APIs,

Elastic is a search company that

Auto-completion?



Auto-completion?



elasticsearch auto-completion

elastic community conference

elasticsearch

elastic for sewing

elastic software

elastic stack

elastic company

elastic login

elastic in economics

elastic material

elastic band

kibana

Remove

Remove

Remove

People also ask

What is elastic use?

What is elastic tool?

What is elastic server?

People also search for

Michaels

elastic

JOANN

Michaels

Elastic NV

Jo-Ann Stores

Elastic Heart

Song by Diplo, Sia, and The...

Report inappropriate predictions

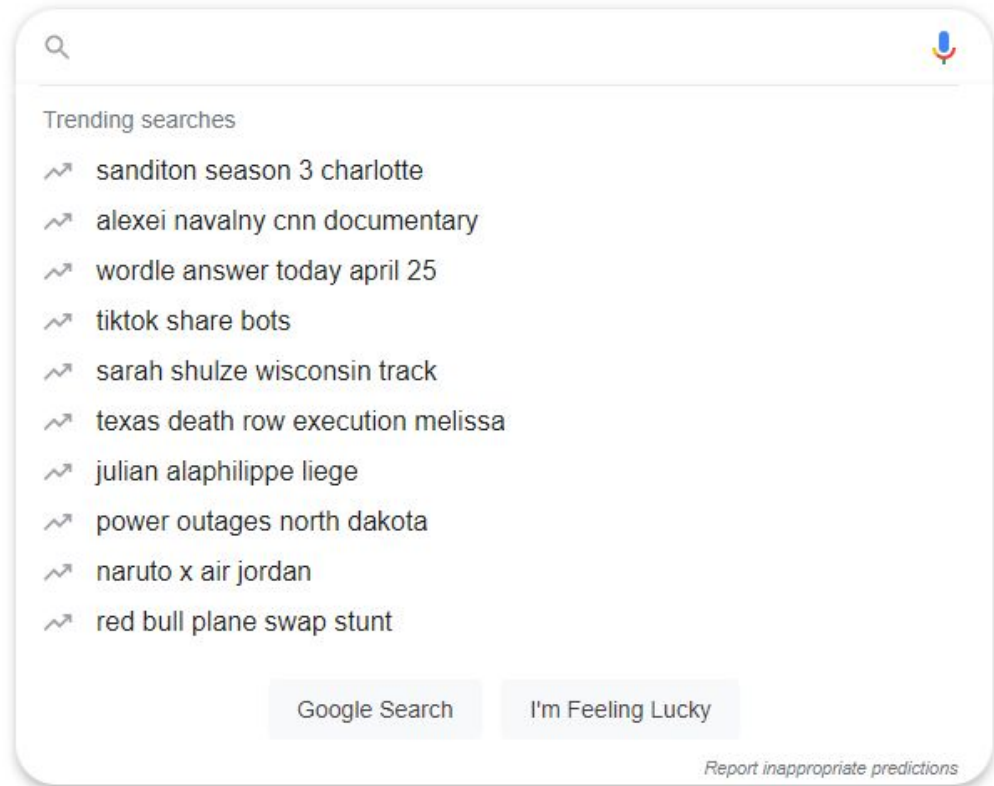
Learn more

Careers - Meet our teams - Board

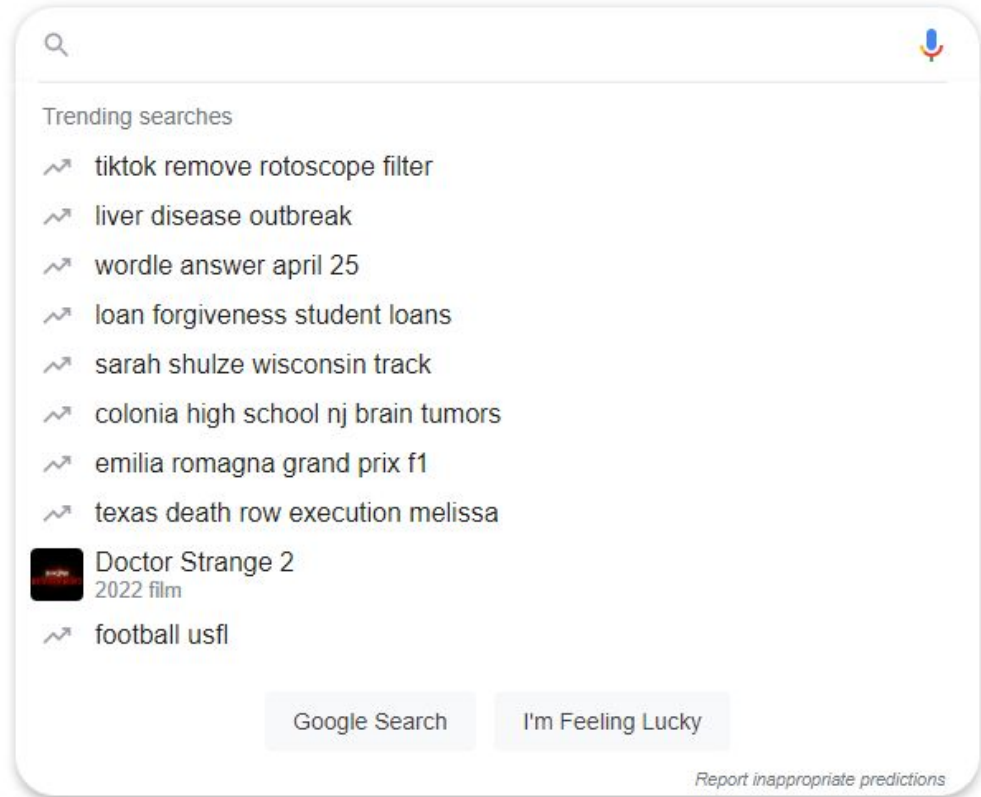
Download Kibana - Download

observability and analytics use cases [Wikipedia](#)

Auto-completion?



Auto-completion?



Auto-complétion?

29 433

pica

Réponses dans les filtres

COLLECTION [1 / 1]

Frac Picardie [1 307]

NOM ARTISTE [3 / 3]

PICA Amalia [2]

PICARD Albert Julien Theresius [1]

PICARD LII (BENEDICK LIII Elisabeth, dit) [1]

Réponses complémentaires par le texte intégral dans les rubriques non filtres

Mots dans la rubrique "Infos image"

picardie

picasso

Mots dans la rubrique "Infos ensemble"

picardie

Mots dans la rubrique "Titres"

picasso

picabia

picardie

picart

Mots dans la rubrique "Collaborateurs"

picardie

picard

Mots dans la rubrique "Stade de la création"

picardie

Christian LHOPITA

Sans titre

1984

Inv.: 85.062

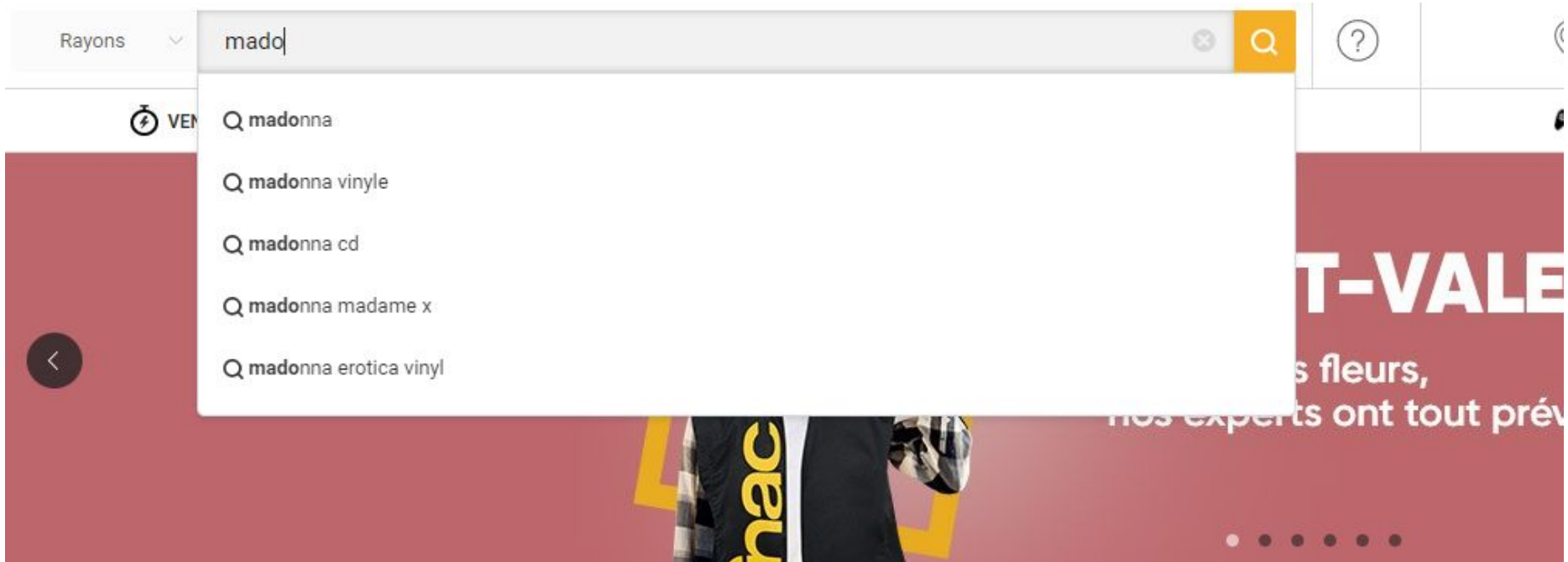
IAC, Villeurbanne/Rhône-

Teresa MURAJ

François BO

11

Auto-completion?



Auto-completion?



Documentation



Elasticsearch



7.17



auto



Manage **auto**-follow patterns | Elasticsearch Guide [7.17]

DOC — To replicate time series indices, you configure an **auto**-follow pattern so that each new index in

Pause **auto**-follow pattern API | Elasticsearch Guide [7.17]

DOC — Pauses an **auto**-follow pattern. Request `POST /_ccr/auto_follow/<auto_follow_pattern_name>/pause`

Cross-cluster replication APIs | Elasticsearch Guide [7.17]

DOC — follower indices Get info about follower indices **Auto**-follow Create **auto**-follow pattern Delete **auto**-follow

Resume **auto**-follow pattern API | Elasticsearch Guide [7.17]

DOC — Resumes an **auto**-follow pattern. Request `POST /_ccr/auto_follow/<auto_follow_pattern_name>/resume`

Get **auto**-follow pattern API | Elasticsearch Guide [7.17]

DOC — Get **auto**-follow patterns. Request `GET /_ccr/auto_follow/` `GET /_ccr/auto_follow`

Auto-completion?



Documentation



Elasticsearch



7.17



auto

Manage **auto**-follow patterns | Elasticsearch Guide [7.17]

doc — To replicate time series indices, you configure an **auto**-follow pattern so that each new index in

Pause **auto**-follow pattern API | Elasticsearch Guide [7.17]

doc — Pauses an **auto**-follow pattern. Request POST `/_ccr/auto_follow/<auto_follow_pattern_name>/pause`

Cross-cluster replication APIs | Elasticsearch Guide [7.17]

doc — follower indices Get info about follower indices **Auto**-follow Create **auto**-follow pattern Delete **auto**-follow

Resume **auto**-follow pattern API | Elasticsearch Guide [7.17]

doc — Resumes an **auto**-follow pattern. Request POST `/_ccr/auto_follow/<auto_follow_pattern_name>/resume`

Get **auto**-follow pattern API | Elasticsearch Guide [7.17]

doc — Get **auto**-follow patterns. Request GET `/_ccr/auto_follow/` GET `/_ccr/auto_follow`

Apache Solr Reference Guide

auto

[Migrating Rule-Based Replica Rules to Autoscaling Policies](#)

[Autoscaling API](#)

[SolrCloud Autoscaling Automatically Adding Replicas](#)

[SolrCloud Autoscaling Fault Tolerance](#)

[SolrCloud Autoscaling Listeners](#)

[Overview of SolrCloud Autoscaling](#)


[Autoscaling Policy and Preferences](#)


[SolrCloud Autoscaling Trigger Actions](#)

[SolrCloud Autoscaling Triggers](#)


[SolrCloud Autoscaling](#)

Auto-complétion?





PAPIERS - CITOYEN
État-civil, Passeport, ...
Papiers à conserver
d'identité...



TRANSPORT
Carte grise, Permis de
Contrôle techniq
Infractions...

passport
passage
passe
passif
passager

ACCÈS DIRECTS

Fiches pratiques

- > Passeport
- > Passage rapide aux frontières : comment utiliser un sas Parafe ?
- > Primaire et secondaire : comment s'effectue le passage du privé au public ?

Services en ligne et formulaires

- > Déclaration de perte de carte d'identité ou de passeport - 14011*02
- > Demande de réquisition de passage - article L115 du Code des pensions militaires d'invalidité et des victimes de guerre - 14611*02
- > Demande de remboursement d'un timbre électronique Passeport - 15416*01 - 2846-SD

Actualités

- > Passe sanitaire - Tout savoir sur le passe vaccinal
- > Rappel vaccination Covid-19 - Passe vaccinal valide : quels changements à partir du 15 février ?
- > Passe vaccinal - Pass+ : la nouvelle fonctionnalité de « TousAntiCovid » pour le passe vaccinal dérogatoire

COMMENT FAIRE SI...

SERVICES EN LIGNE

QUESTIONS - RÉPONSES

Auto-complétion?

Jeux de données⁽¹¹⁾Rechercher dans les réutilisations

Critères de recherche

Organisations




Mots clés

Licences

Formats

...


Venez tester notre nouvelle recherche de jeux de données.

 <p>Parcoursup 2020 - vœux de poursuite d'études et de réorientation dans l'enseignement supérieur et réponses des établissements</p> <p>Ce jeu de données présente les vœux de poursuite d'études et de réorientation dans l'enseignement supérieur ainsi que les propositions des établissements pour chaque formation – hors apprentissage – à la fin du processus d'affectation de la plateforme Parcoursup pour la session 2020 (du 22 janvier ...</p>	<p>5 ressources</p> <p>0 réutilisations</p> <p>0 favoris</p>
 <p>Parcoursup : propositions d'admission dans l'enseignement supérieur des élèves de terminale diplômés du baccalauréat général selon leurs enseignements de spécialité</p> <p>Les candidats de Parcoursup étudiés dans ces données sont les élèves de terminale générale inscrits sur Parcoursup 2021 ayant confirmé au moins un vœu en phase principale et diplômés du baccalauréat général 2021 en France (y compris DROM, COM et CNED, et hors établissements AEFE à l'étranger). Ce s...</p>	<p>2 ressources</p> <p>0 réutilisations</p> <p>0 favoris</p>
 <p>Accès à l'enseignement supérieur : premier bilan de la loi orientation et réussite des étudiants</p> <p>La loi relative à l'orientation et à la réussite des étudiants de 2018 vise à améliorer les trois étapes fondamentales du parcours des jeunes gens se destinant à l'enseignement supérieur : accompagnement et</p>	<p>1 ressources</p> <p>0 réutilisations</p> <p>0 favoris</p>


Auto-complétion?

✕ 🔍


[Devenir client](#)

 [SE CONNECTER](#)
[Créer un compte](#)


PRODUITS




Muffin chocolat pépites de chocolat 40 g
Référence : 194046




Muffin tulipe chocolat, coeur chocolat 130 g
Référence : 159676



Muffin parfum chocolat avec morceaux de chocolat noir 100 g
Référence : 249250




Crème glacée chocolat avec morceaux de chocolat 2,4 L / 1,363 kg PassionFroid
Référence : 181591




Crème glacée chocolat avec morceaux de chocolat 5 L / 2,84 kg PassionFroid
Référence : 60187

[See all products for "choco"](#) 🔍


RECETTES




Brownie chocolat-menthe sur lit de chocolat




Le choco-orange



Le cube choco



Le tout chocolat








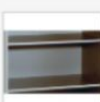
Trio de chocolats

[See all recipes for "choco"](#) 🔍







Auto-completion?

With love. armoire

Produits 1

-  **armoire** à rideaux en bois
A129088
529,00 € HT
-  **armoire** haute portes battantes
A106922
435,00 € HT
-  **armoire** haute avec réhausse
A141375
529,00 € HT
-  Tablettes supplémentaires pour **armoire**
A745825
52,90 € HT
-  **armoire** haute avec réhausse
A755571
529,00 € HT
-  Tablettes supplémentaires pour **armoire**
A745825
52,90 € HT





Catégories 2

-  **Armoire** bois
dans Armoire
-  **Armoire** métal
dans Armoire
-  **Armoire** à compartimentage
dans Armoire d'atelier
-  **Armoire** à portes battantes
dans Armoire d'atelier
-  **Armoire** à portes coulissantes
dans Armoire d'atelier
-  **Armoire** à portes transparentes
dans Armoire d'atelier

Mots clés

- armoire**
- Armoire** et rangement
- armoire** en bois 3
- armoire** à double portes
- armoire** à battantes

Marque

-  3M
-  Abus
-  Belkin
-  Matador

Devis gratuit sous 24 h

Auto-completion? Why?

- Providing meaningful results while typing
- Improve the query precision
- Guide the user
- Improve the query response time and responsiveness
- Improve the user experience



Everything around auto-completion

User Experience

Auto-completion 1: bring up “final” documents

User Interface

Auto-completion 2: bringing up “facets”

Technical details

Auto-completion 3: keywords and key phrases

Queries and Configurations

Auto-completion 4: when there is no input

Data Models

Auto-completion 5: filtered suggestions

Tips and Tricks

Return on Experience

A story of auto-completions at #ElasticCC



<https://community-conference.elastic.co/session/309778>

February 11, 2022

15:00 — 15:25 (EUROPE/PARIS)

14:00 — 14:25 (UTC)

Auto-complétion et auto-completions avec Elasticsearch

Il y a auto-complétion et auto-complétion. On dénombre au moins cinq types et implémentations différentes. Pour la plupart, un dictionnaire et index dédiés doivent être construits. Dans cette présentation, nous allons démontrer les possibilités d'Elasticsearch sur plusieurs cas d'utilisation de cette recherche avant la recherche, que nous appelons auto-complétion. Nous allons présenter les suggestions de termes à rechercher, des suggestions filtrées, des propositions de catégories, des propositions de résultats finaux, des propositions de noms propres comme les marques. Le tout avec le code Elasticsearch (requête, mapping et indexation) qui opère la magie du service.

Full session

Advanced

Enterprise search

Cloud



Lucian Precup

CTO | all.site



elastic | Community
Conference

Auto-complétion et auto-completions avec Elasticsearch

—
Lucian Precup

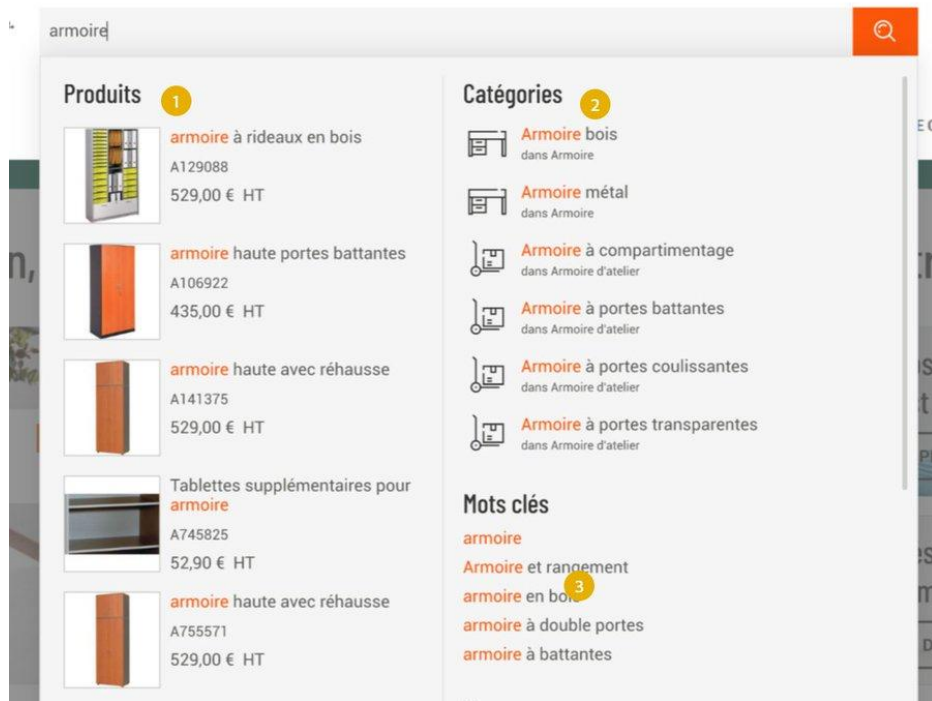


Auto-completion: Don'ts!

- Check the “activate auto-completion” option
- Use the wildcard query
 - `q=elastic*` \Leftrightarrow LIKE “elastic%”
- [Elasticsearch] Use phrase prefix
 - Or if you do, at least, activate `index_prefixes`
- Auto-completion vs. “did you mean...”
- [Elasticsearch] Think that the completion suggester is just a configuration to set up
- [Elasticsearch] Think that the `search_as_you_type` data type solves everything
- Handle the highlighting in Javascript
- Use facets or aggregations



Auto-complétion 1: bring up “final” documents



Solution : edge n-gram

edge_ngram

```
1 {  
2   "settings": {  
3     "analysis": {  
4       "analyzer": {  
5         "edge_ngram": {  
6           "tokenizer": "standard",  
7           "filter": [  
8             "my_edge_ngram"  
9           ]  
10        }  
11      },  
12      "filter": {  
13        "my_edge_ngram": {  
14          "type": "edge_ngram",  
15          "min_gram": 1,  
16          "max_gram": 19  
17        }  
18      }  
19    }  
20  },  
21  "mappings": {  
22    "properties": {  
23      "session": {  
24        "type": "text",  
25        "analyzer": "french",  
26        "fields": {  
27          "edge_ngram": {  
28            "type": "text",  
29            "analyzer": "edge_ngram"  
30          }  
31        }  
32      }  
33    }  
34  }  
35 }
```

```
1 {  
2   "query": {  
3     "multi_match": {  
4       "query": "mot",  
5       "fields": [  
6         "session",  
7         "session.edge_ngram"  
8       ],  
9       "operator": "and"  
10    }  
11  }  
12 }
```


Auto-completion with nGram

Index time

Input documents	Id	Name
<input type="checkbox"/>	1	Céline
	2	Celia
Ascii folding <input type="checkbox"/>	Celine, Celia	
Lowercase <input type="checkbox"/>	celine, celia	
EdgeNGram <input type="checkbox"/>	ce cel celi celin celine ce cel celi celia	

Query time

Name	<input type="checkbox"/> Search term
Célin	
Celin <input type="checkbox"/> Ascii folding	
celin <input type="checkbox"/> Lowercase	
ce cel celi celin <input type="checkbox"/> EdgeNGram	

Index

Key	Document id
ce	1, 2
cel	1, 2
celi	1, 2
celin	1
celine	1
celia	2



Auto-completion with nGram [search analyzer]

Index time

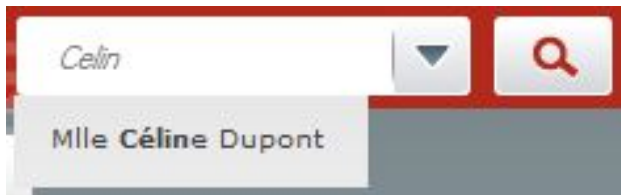
	Id	Name
Input documents <input type="checkbox"/>	1	Céline
	2	Celia
Ascii folding <input type="checkbox"/>	Celine, Celia	
Lowercase <input type="checkbox"/>	celine, celia	
EdgeNGram <input type="checkbox"/>	ce cel celi celin celine ce cel celi celia	

Query time

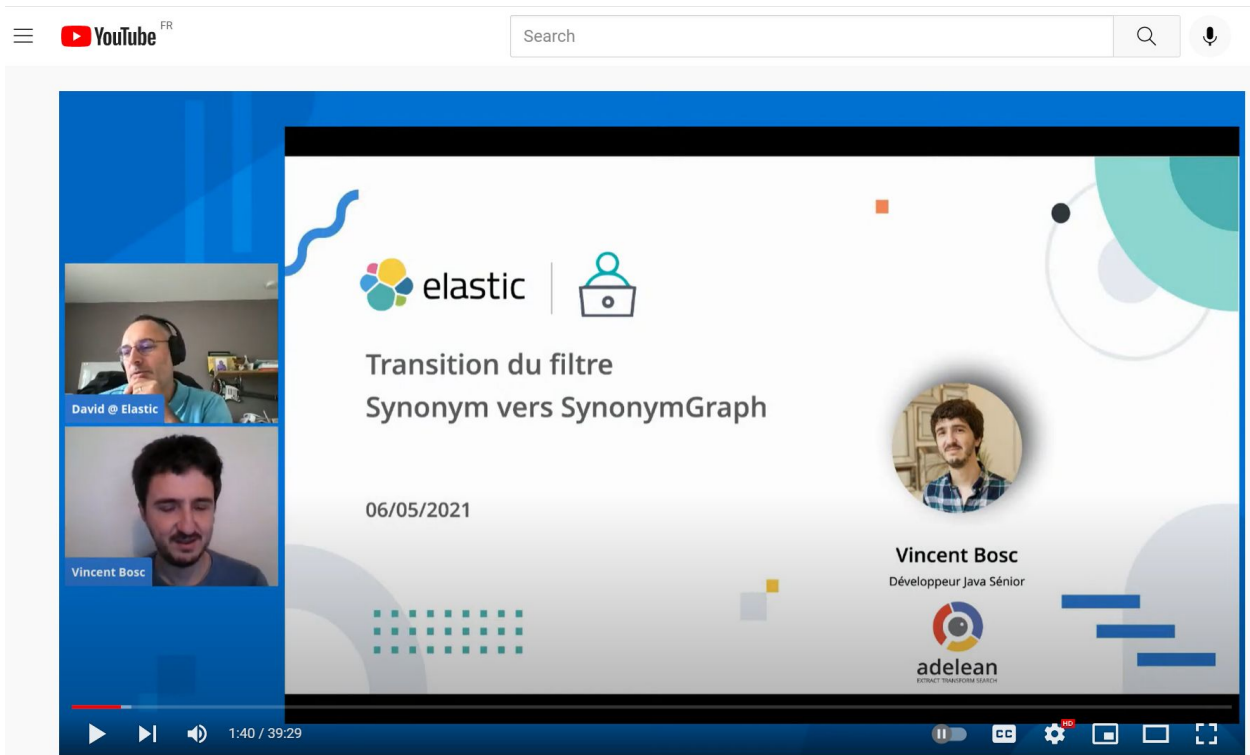
Name	<input type="checkbox"/> Search term
Célin	<input type="checkbox"/> Ascii folding
celin	<input type="checkbox"/> Lowercase

Index

Key	Document id
ce	1, 2
cel	1, 2
celi	1, 2
celin	1
celine	1
celia	2

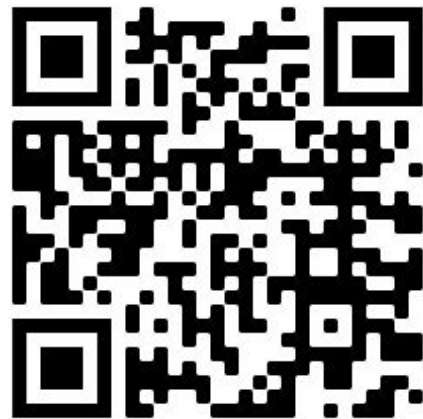


[Index & search analyzer]



Meetup ElasticFR #61 - Transition du filtre Synonym vers SynonymGraph

<https://www.youtube.com/watch?v=DcjmhkeQt-I>



[The search_as_you_type data type]

```
1 {  
2   "mappings": {  
3     "properties": {  
4       "my_field": {  
5         "type": "search_as_you_type"  
6       }  
7     }  
8   }  
9 }
```

```
1 {  
2   "query": {  
3     "multi_match": {  
4       "query": "brown f",  
5       "type": "bool_prefix",  
6       "fields": [  
7         "my_field",  
8         "my_field._2gram",  
9         "my_field._3gram"  
10      ]  
11     }  
12   }  
13 }
```

Bring up “final” documents – User Experience ?

SonoVente.com
SPECIALISTE 100% FRANÇAIS
















Search by brands

Free Shipping from 199€*

Cart My account +33 1 80 38 38 38 Language English

mixer alesis mixer behringer milus 206la mipro midas microphone home studio fbt milus numark mixtrack platinum

5,316 products

 MARSHALL MS4 Electric Guitar Combo Offers In stock 33.33€	 MARSHALL MS2 Electric Guitar Combo Offers In stock 25€	 MOTU MIDI EXPRESS XT Midi Interface Offers In stock 354.17€	 MOTU MICRO EXPRESS 2 USB Midi Interface In stock 270.83€	 MARSHALL MS-2R Electric Guitar Combo Offers In stock 25€
 VOX AMPLUG 2 CLEAN Electric Guitar Combo Offers In stock 30€	 AKG DSS1 DRUM SET SESSION Microphones Pack Offers In stock 235€	 BLACKSTAR FLY3 Electric Guitar Combo BEST-OF In stock 56.67€	 VOX AMPLUG CLASSIC ROCK V2 Electric Guitar Combo Offers In stock 29.17€	 VOX AMPLUG AC30 V2 Electric Guitar Combo Offers In stock 30€
				

The user thinks

“microphones”

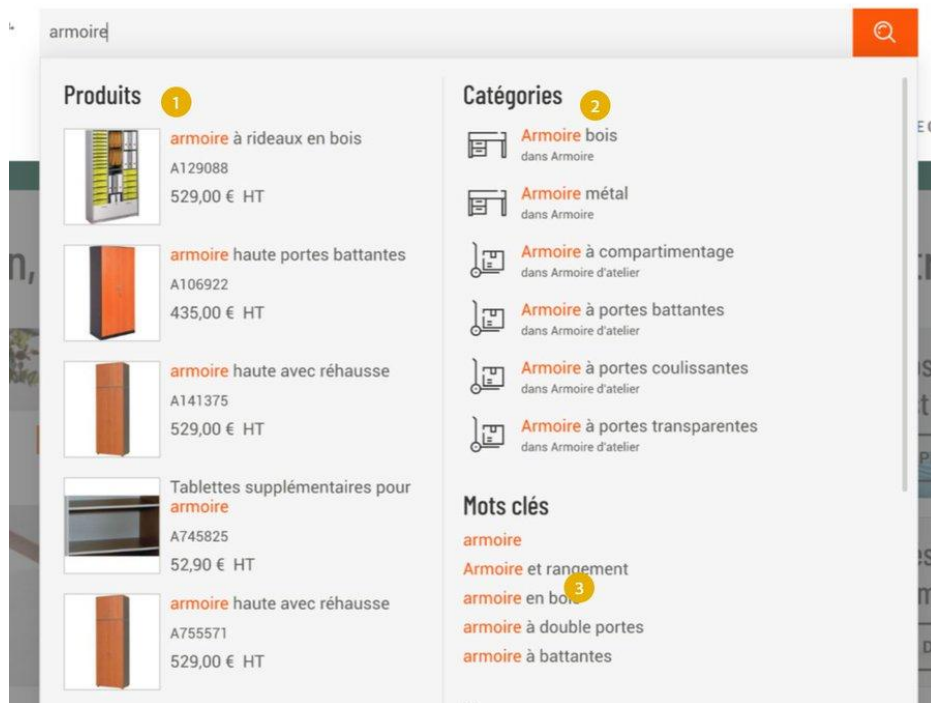
but instead, they get

“mixers” and “midi interfaces” ?

[Handle the highlighting]

```
1 {  
2   "query" : {  
3     "multi_match": {  
4       "query": "elasticsearch",  
5       "fields": [  
6         "body", "blog.title", "blog.author", "blog.comment"  
7       ]  
8     },  
9     "match": { "body": "elasticsearch" }  
10  },  
11  "highlight" : {  
12    "number_of_fragments" : 3,  
13    "fragment_size" : 150,  
14    "fields" : {  
15      "body" : { "pre_tags" : ["<em>"], "post_tags" : ["</em>"] },  
16      "blog.title" : { "number_of_fragments" : 0 },  
17      "blog.author" : { "number_of_fragments" : 0 },  
18      "blog.comment" : { "number_of_fragments" : 5, "order" : "score" }  
19    }  
20  }  
21 }
```

Auto-completion 2: bringing up “facets”

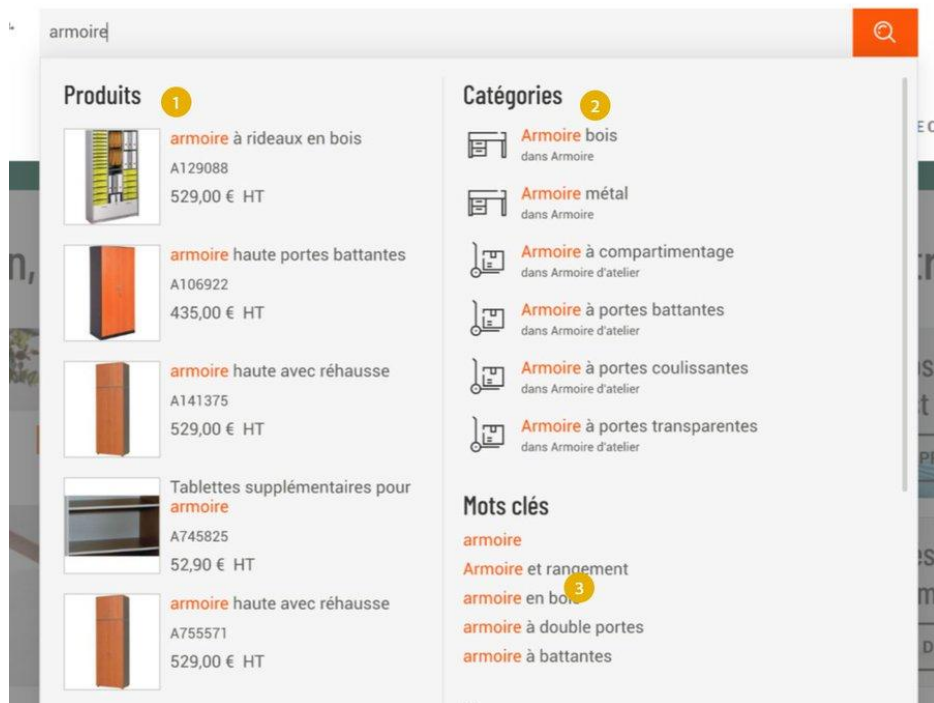


Solution :

Separate index holding only the needed fields (categories, brands)

Classical edge n-gram on the new index

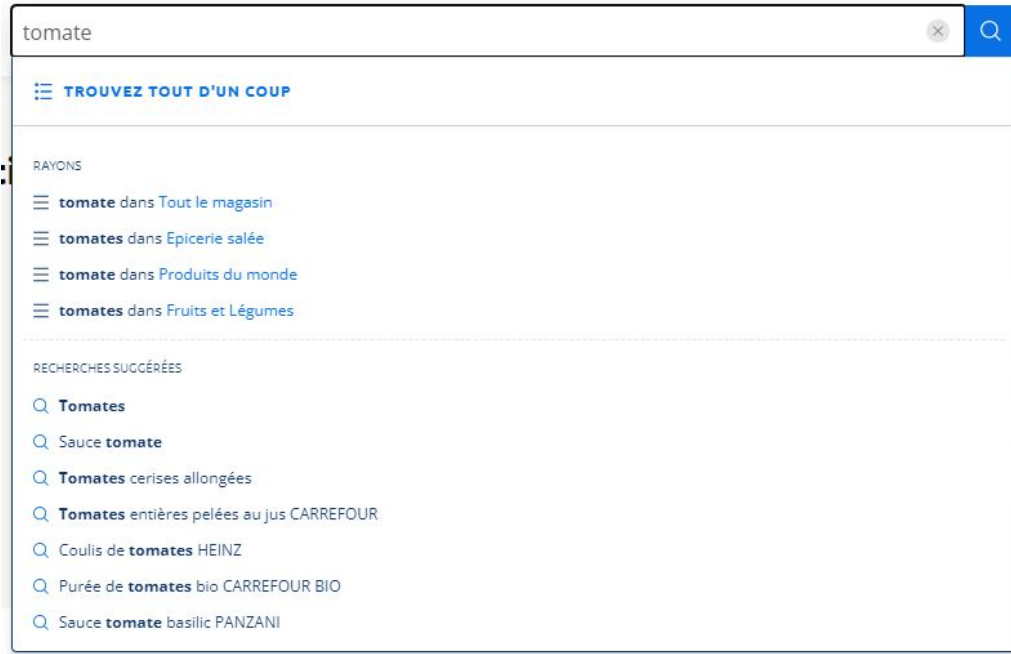
Auto-completion 3: keywords and key phrases



Solution : separate index with keywords and keyphrases

Problem : how to generate relevant keywords and keyphrases? Especially when the input data is unstructured.

Auto-completion 3: keywords and key phrases



Auto-completion 3: keywords and key phrases

Extract keywords and key phrases from a non-structured text :

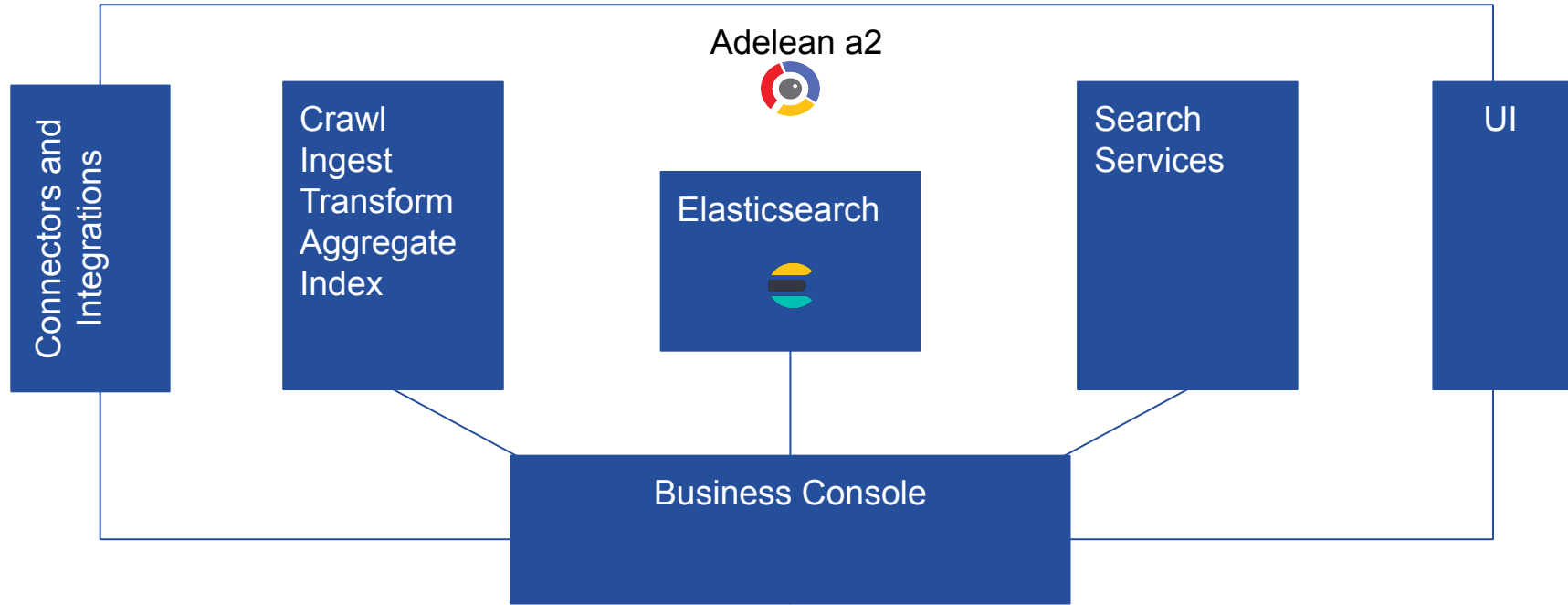
- NLP, Machine Learning
- Shingles
- Custom code and usage of analyzers to normalize the data

A community search engine to organize the Internet?

Shingle 2 \Rightarrow A community, community search, search engine, engine to, to organize, organize the, the Internet

Shingle 3 \Rightarrow A community search, community search engine, search engine *to*, engine *to* organize, *to* organize *the*, organize the Internet

[Completion suggerer]



[Completion suggester]

Mapping

```
1 {
2   "mappings": {
3     "properties": {
4       "suggest": {
5         "type": "completion"
6       },
7       "name": {
8         "type": "text"
9       },
10      "firstname": {
11        "type": "text"
12      }
13    }
14  }
15 }
```

Document

```
1 {
2   "suggest": [
3     {
4       "input": "Jean",
5       "weight": 3
6     },
7     {
8       "input": "Dujardin",
9       "weight": 10
10    }
11  ],
12  "firstname": "Jean",
13  "name": "Dujardin"
14 }
```

Query

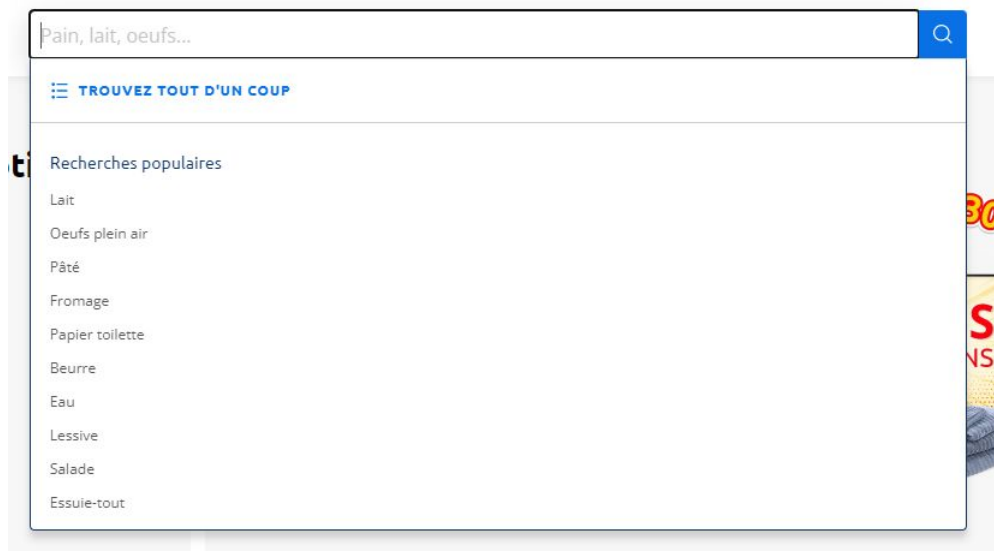
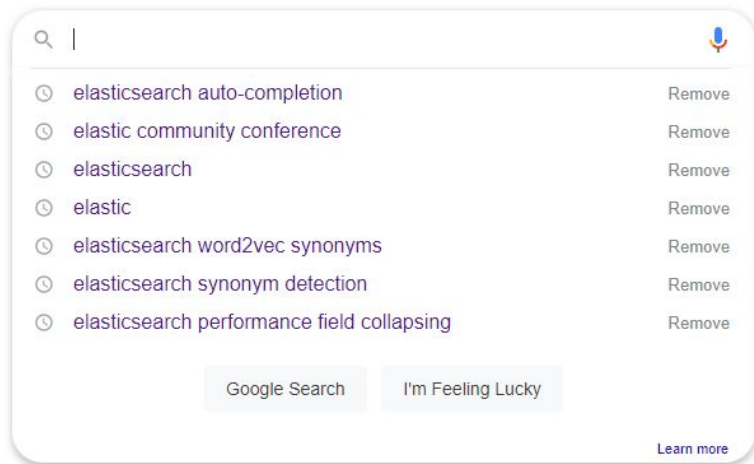
```
1 {
2   "suggest": {
3     "name-suggest": {
4       "prefix": "du",
5       "completion": {
6         "field": "suggest",
7         "skip_duplicates": true
8       }
9     }
10  }
11 }
```

[_msearch – the multi-search API]

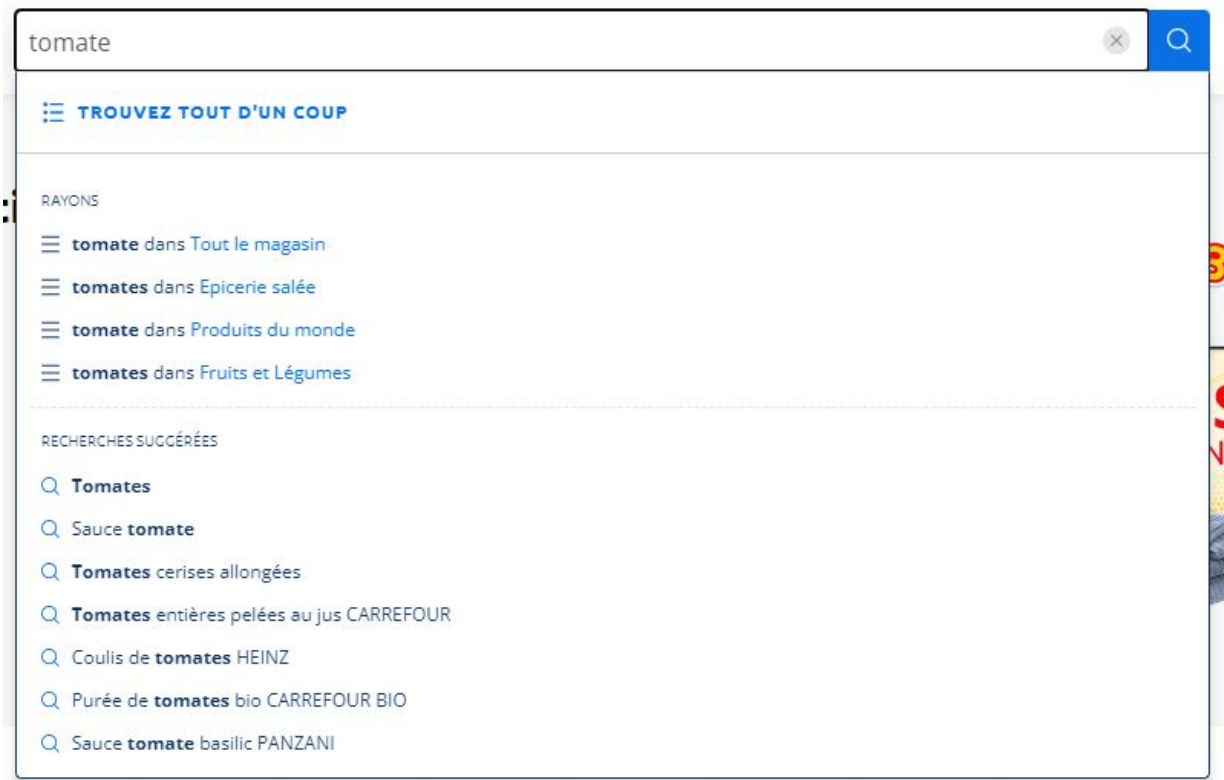
```
1 GET _msearch
2 {"index":"product_auto_complete"}
3 {"query":{"match":{"title":"tomate"}}}
4 {"index":"product_categories"}
5 {"query":{"match":{"title":"tomate"}}}
6 {"index":"keywords"}
7 {"query":{"match":{"title":"tomate"}}}
8 {"index":"brands"}
9 {"query":{"match":{"title":"tomate"}}}
```

Auto-completion 4: when there is no input

Starting with the first (second, third, more) letter?



Auto-complétion 5: filtered suggestions






Filtered suggestions




HAYSTACK[Home](#) [Schedule](#) [Tickets](#)

Relevant Facets

Lucian Precup and Radu Pop • [Back to Haystack Europe 2019](#)



 **Relevant Facets by Lucian Precup and ...**  




// Structure of the talk

- Examples of facet implementations
- Challenges with facets and possible solutions
- Challenges with facets in general and how facets can help
- Technical implementation examples are with Elasticsearch
- We are addressing less the "graphical" display of facets and more the technical issues with their relevancy

@a2lean #haystackconf

Watch on  **YouTube** **HAYSTACK**

Auto-completion 5: filtered suggestions

Products index

```
1  {  
2    "title": "Red tomatoes",  
3    "category": "Tomates",  
4    "department": "Fruits et Légumes",  
5    "price": "3"  
6  }  
7  
8  {  
9    "title": "Yellow tomatoes",  
10   "category": "Tomates",  
11   "department": "Bio et Ecologie",  
12   "price": "30"  
13 }
```

Suggestions index

```
1  {  
2    "text": "Tomates",  
3    "context": {  
4      "department": ["Fruits et Légumes"]  
5    }  
6  }  
7  
8  {  
9    "term": "Tomates",  
10   "context": {  
11     "department": ["Bio et Ecologie"]  
12   }  
13 }
```

Use the `_update` API and
increase the number of occurrences

Auto-completion 5: filtered suggestions

Suggestions dictionary

```
1 {  
2   "text": "Tomates",  
3   "context": {  
4     "product_categories_name": [  
5       "Fruits et Légumes",  
6       "Bio et Ecologie"  
7     ]  
8   },  
9   "occurrences": 81,  
10  "popularity": 16  
11 }
```

Query

```
1 {  
2   "query": {  
3     "multi_match": {  
4       "query": "tom",  
5       "fields": ["text.smartComp"]  
6     }  
7   },  
8   "sort": [  
9     {  
10      "occurrences": {  
11        "order": "desc"  
12      }  
13    }  
14  ]  
15 }
```

Auto-completion 5: filtered suggestions

The result

tom|

tom dans Tout le magasin

tomate dans Bio et Ecologie

tomate dans Fruits et Légumes

1

[Spellchecking and “did you mean ...”]

Beware of the mapping

The max_errors parameter

Terms suggester vs.
Phrase suggester

```
1  {  
2  {  
3    "suggest": {  
4      "text" : "tomtate onde",  
5      "simple_phrase" : {  
6        "phrase" : {  
7          "field" : "title",  
8          "size" : 1,  
9          "max_errors": 2  
10       }  
11     }  
12  }  
}
```

<https://www.elastic.co/guide/en/elasticsearch/reference/7.17/search-suggesters.html#phrase-suggester>

Thank you for your attention

Questions / Feedback / More ...

@lucianprecup

@search_with_all

@a2lean

info@adelean.com

<http://all.site>

<http://www.adelean.com>

<http://www.linkedin.com/company/adelean>

<http://www.meetup.com/fr-FR/search-and-data>

