

An AI Assistant in the life of a Search Engine Administrator

Lucian Precup

Maëilly Dubois

2024-06-12

MIX-CAMP E-COMMERCE SEARCH #MICES2024

@a2lean

Who are we?

🔍 Adelean

- 🔍 Experts in **search** technologies
- 🔍 Integrators of **Elasticsearch**, **OpenSearch** and **Solr**
- 🔍 **Consulting** and **Training** providers
- 🔍 Developers of **a2** E-Commerce and Enterprise Search solution
- 🔍 Developers of **all.site** - your **Collaborative** Search Engine

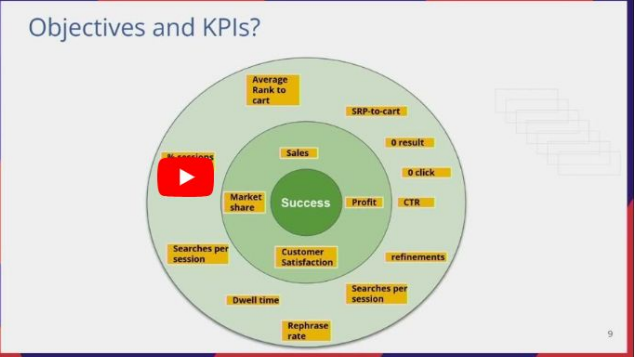


Berlin Buzzwords 2022 - The life of a search engine administrator


The life of a search engine administrator ☆

06-13, 16:00–16:40 (Europe/Berlin), Palais Atelier

Lucian Precup & Vincent Bréhin – The life of a search engine administrator



Objectives and KPIs?

Watch on  YouTube

Defining the KPIs, keeping an eye on the customer satisfaction and sales, defining the backlog, configuring the search engine, debugging relevance issues, preventing regressions ... These are a few tasks on the list of a search engine administrator. A search engine is a living thing. Seasonality, levels of stocks, lifecycle of the products, marketing events, news, etc. are a few of the many factors that force the search engine to constantly evolve. In this context, the life of a search engine manager is tough. In this talk we describe the processes and tools that we put in place and help manage a search engine. We also address the limits between what can be automated and what still needs human supervision.

Vincent Bréhin & Lucian Precup
The life of a search engine administrator
June 12-14, 2022 // Kulturbrauerei Berlin & Online



<https://pretalx.com/bbuzz22/talk/P9ZFJL/>

Agenda

A day in the life of a search engine administrator

Improving the life of a search engine administrator

NLU and LLMs to automate actions

Analytics and Machine Learning to bring insights



A day in the life of a search engine administrator

A search engine is a living entity

Seasonality

Levels of stocks

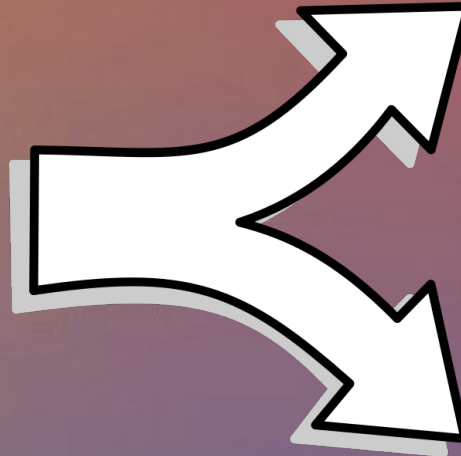
Lifecycle of documents

Marketing events

News

User's expectations
and position

...



Evolution

Regression

Different use cases, different stakes

E-commerce

- Business relevance (stocks, personalization, marketing, agenda, etc.)



Enterprise Search

- User productivity
- Information access security



Hybrid

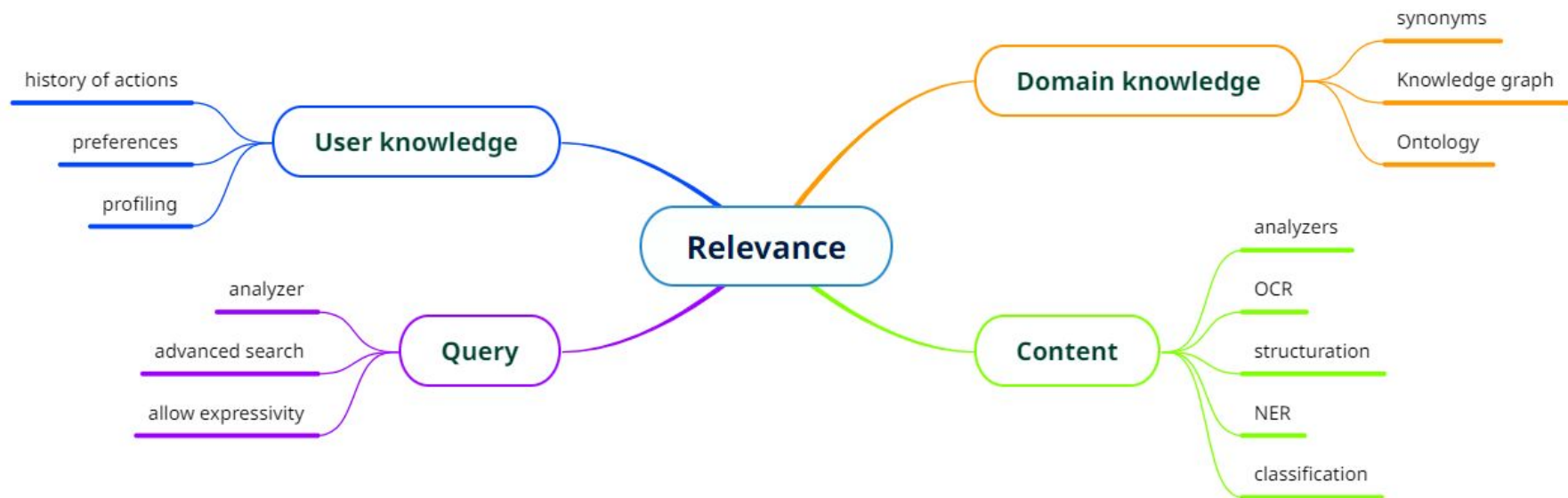
- E.g. E-commerce + Reference website



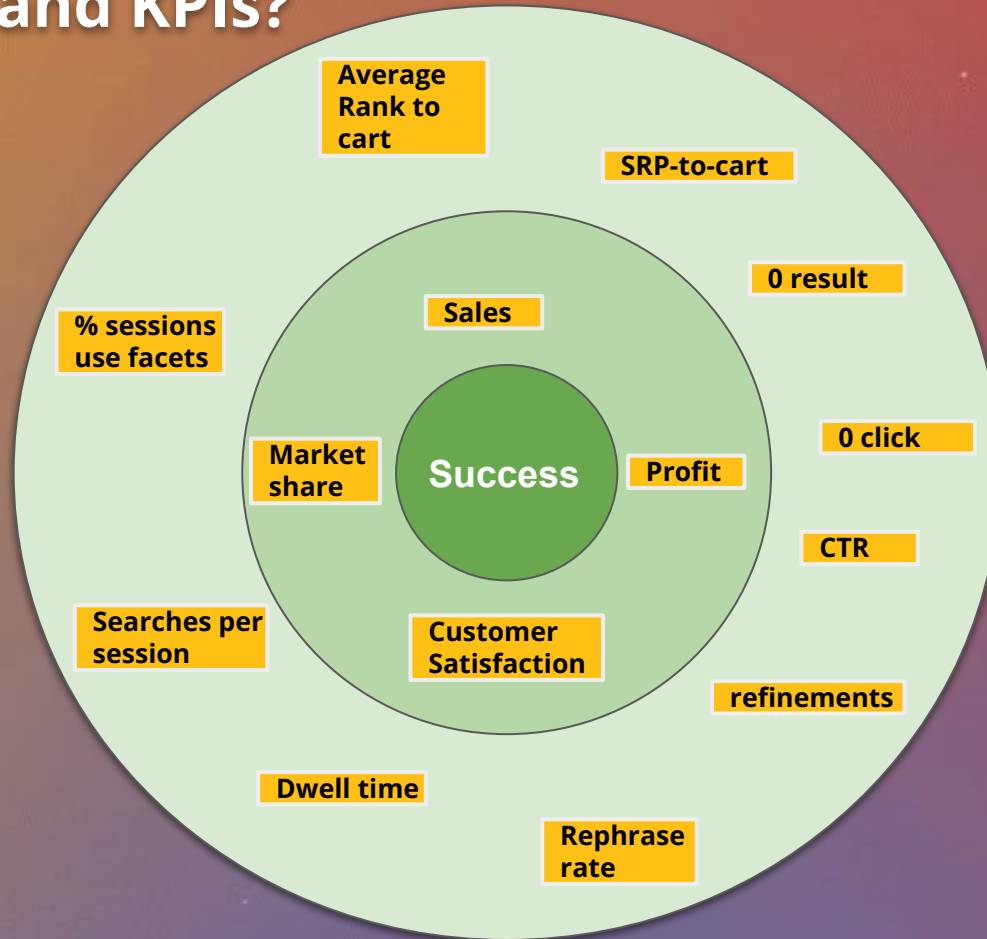
Specialized

Search is about...

Leveraging (moving) **signals** in order to fulfill a **user's need**



Objectives and KPIs?



What to monitor?

Business needs (relevance)

Data infrastructure

- Crawling errors
- Datasource availability
- Data errors

User satisfaction

...



Hypothesis: Challenges are addressed

How to define the KPIs?

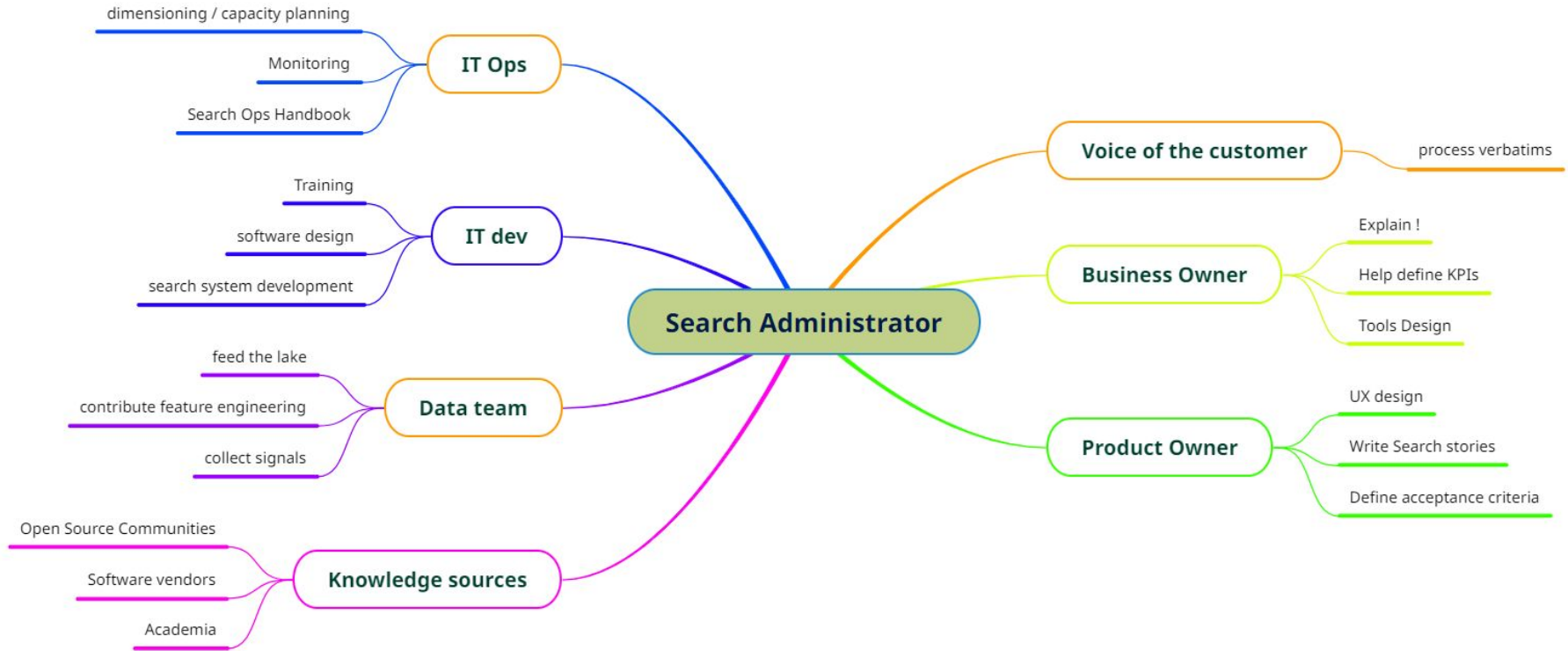
What are the domain specific issues and how to deal with them?

What to monitor?

How to define the roadmap?

...

The search administrator's roles



Day to day tools

Business people (merchandising)

- Business Console

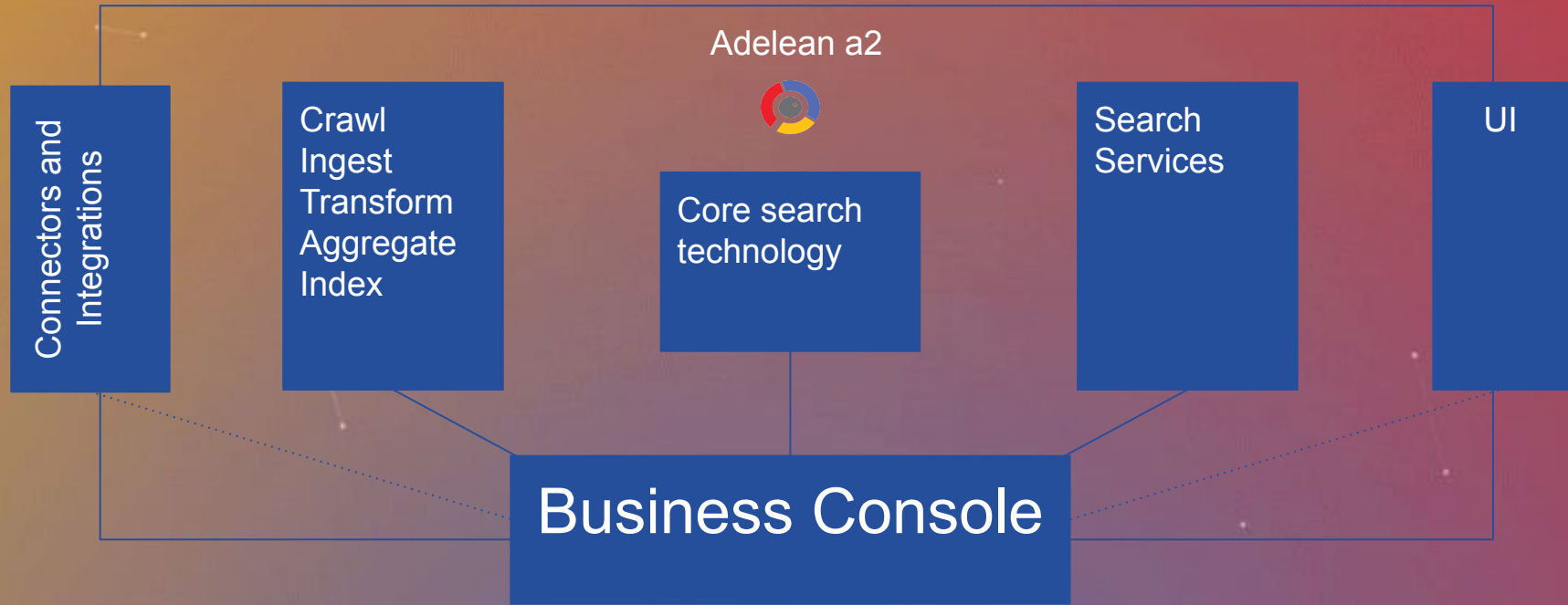
Reporting teams

- Analytics, dashboards

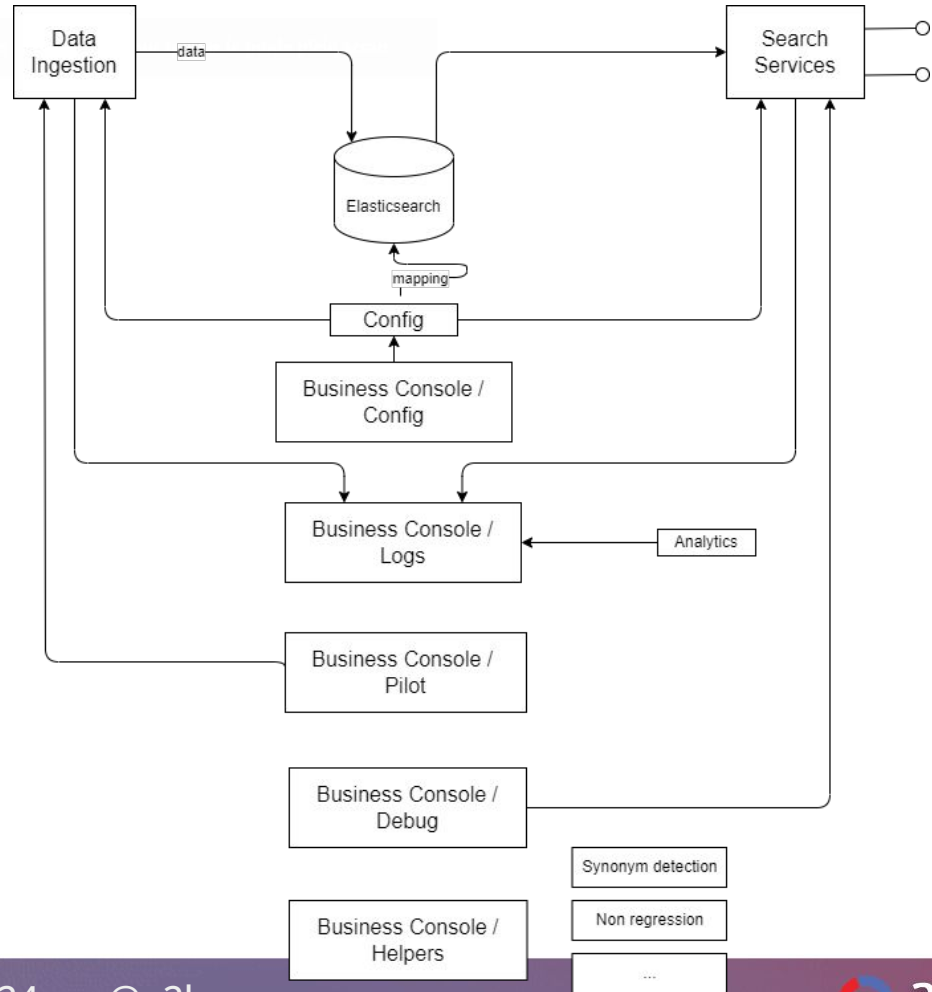
Data teams

- Automatic exports

Business Console - the ultimate tool



Business Console - the details



Business Console / Debug

Simulate a user search, visualize the data, compare the relevance scores, ensure the non-regression

Recherche

Magasin

Mode par défaut

Catalogues onecarrefour_0026-150-141

Tris

Facettes

Filtres

Requêtes

Rechercher cartable

temps de réponse: 309 ms

1		Cartable / Cartable primaire bleu 38 cm POKEMON
2		Sac à dos / Sac à dos noir EASTPAK
3		Sac à dos / Sac à dos isotherme roll up MANDINE
4		Sac à dos / Sac à dos 1 compartiment rouge lycée padded EASTPAK
5		Sac à dos / Sac à dos 1 compartiment noir lycée padded EASTPAK
6		Sac à dos / Sac à dos 1 Compartiment - L 30,5 cm EASTPAK
7		Sac à dos / Sac à dos 1 compartiment gris lycée padded EASTPAK
8		Sac à dos / Sac à dos 1 compartiment marine lycée padded EASTPAK
9		Sac à dos / Sac à dos 1 compartiment bleu denim lycée padded EASTPAK
10		Sac / Sac à dos isotherme 12 L MANDINE

EXPORTER

PRODUIT OFFRES TECHNIQUES SCORE

Filtrer

product name	Cartable primaire bleu 38 cm POKEMON	Sac à dos noir EASTPAK
product qualifier	primaire bleu 38 cm	noir
product_sub_brand_name (Marque)	POKEMON	EASTPAK
product_brand_name	POKEMON	EASTPAK
product_packaging	le cartable	le sac à dos
product_nature	Cartable	Sac à dos
product_non_food_sh or_l designation	3281640452047	3662100000723
product_gtin (EAN)	3281640452047	3662100000723
product_non_food_color (Couleur)	Bleu	
product_non_food_is_rechargeable	false	false

Lignes 20 1-10 sur 10

Recherche

Magasin

Mode par défaut

Catalogues onecarrefour_0026-150-141

Tris

Facettes

Filtres

Requêtes

Rechercher cartable

temps de réponse: 309 ms

1		Cartable / Cartable primaire bleu 38 cm POKEMON
2		Sac à dos / Sac à dos noir EASTPAK
3		Sac à dos / Sac à dos isotherme roll up MANDINE
4		Sac à dos / Sac à dos 1 compartiment rouge lycée padded EASTPAK
5		Sac à dos / Sac à dos 1 compartiment noir lycée padded EASTPAK
6		Sac à dos / Sac à dos 1 Compartiment - L 30,5 cm EASTPAK
7		Sac à dos / Sac à dos 1 compartiment gris lycée padded EASTPAK
8		Sac à dos / Sac à dos 1 compartiment marine lycée padded EASTPAK
9		Sac à dos / Sac à dos 1 compartiment bleu denim lycée padded EASTPAK
10		Sac / Sac à dos isotherme 12 L MANDINE

EXPORTER

PRODUIT OFFRES TECHNIQUES SCORE

Boost 1,5

MODIFIER

6201960 - Σ	4685939.5 - Σ
6201960 - f(x)	4685939.5 - f(x)
1181.3257 - max(a, b)	892.5599 - max(a, b)
1181.3257 - Σ	892.5599 - Σ
36.401222 - Σ	50.96398 - Σ
12.893898 - Σ	12.240915 - Σ
1132.0305 - Σ	829.356 - Σ
1109.7144 - Σ	824.5365 - Σ
5250 - min(a, b)	5250 - min(a, b)
0 - Π	0 - Π

Lignes 20 1-10 sur 10

Business Console / Config

E.g. facets

Facettes

Nature du contexte
Par défaut

+ CRÉER

ORDRE	CHEMIN	NOM D'AFFICHAGE	NATURE	VALEUR	TYPE	ORDRE DES VALEURS	ACTIVÉE	OPTIONS
1	facet_promotions	Promotions	Par défaut	--	Case à cocher	Personnalisé	✓	✎ ⚙
2	facet_marque	Marque	Par défaut	--	Case à cocher	Personnalisé	✓	✎ ⚙
3	product_categories sub_node sub_node. sub_node.name	Sous-rayon	Par défaut	--	Case à cocher		✗	✎ ⚙
4	facet_appellation	Appellation	Par défaut	--	Case à cocher		✗	✎ ⚙
5	product_categories sub_node.name	Sous-rayon	Par défaut	--	Case à cocher		✗	✎ ⚙
6	facet_nbre_lames	Nombre de lames	Par défaut	--	Case à cocher	Personnalisé	✗	✎ ⚙
7	product_categories sub_node sub_node. name	Sous-rayon	Par défaut	--	Case à cocher		✗	✎ ⚙
8	facet_nouveautes	En ce moment	Par défaut	--	Case à cocher		✓	✎ ⚙
9	facet_enseignes	Services	Par défaut	--	Case à cocher		✓	✎ ⚙
10	facet_labels1	Labels qualité	Par défaut	--	Case à cocher	Nbre de résultats	✓	✎ ⚙
11	product_categories.code	Code Rayon	Par défaut	--	Case à cocher		✗	✎ ⚙
12	product.is_para_pharmacy	Parapharmacie	Par défaut	--	Booléen		✓	✎ ⚙
13	facet_regimes_alimentaire	Préférences alimentaires	Par défaut	--	Case à cocher	Nbre de résultats	✓	✎ ⚙
14	facet_couleur	Couleur	Par défaut	--	Case à cocher		✗	✎ ⚙
15	facet_sans_substances	Sans substances	Par défaut	--	Case à cocher	Nbre de résultats	✓	✎ ⚙

Configuration avancée

GÉNÉRALE BOOST SYNONYMES DICTIONNAIRE FACETTE

Rechercher

AC_tri_ean
AC_tri_petitdej_hm_s9
AC_tri_stval_hm
EM_TRI_FQC
EM_tri_ean_casque
EM_tri_ean_deo
EM_tri_ean_ge1_douche
EM_tri_ean_jambon
EM_tri_ean_lait
EM_tri_ean_leslive
EM_tri_ean_loreal_ash
EM_tri_ean_loreal_bebe
EM_tri_ean_loreal_maquill...
EM_tri_ean_pates
EM_tri_mdc_champagnon
EM_tri_mdc_gateartiner
EM_tri_mdc_petit_pot
EM_tri_mdc_thon
EM_tri_mkp_languedoc
EM_tri_poulet_mdc
ES boost prix imbattables

NATURE

Par défaut

Nom du champ
facet_appellation

Nature du contexte
Search

Texte de recherche
vin

Paramétrage

Libellé

Afficher toujours

Activer

Afficher même si monovaluée

Mode d'affichage
Case à cocher

Sémantique
OU

ANNULER MODIFIER

ACTIVÉE ACTIONS

Lignes 10 1-6 of 6

Tools - ensure non regression

Comparaisons

Magasin ^

Mode
par défaut ^

Catalogues
onecarrefour_7879-150-353 × ^

Tris

Facettes ^

Filtres ^

Requêtes ^

+ CRÉER EXPORTER RAPPORT ANNULER VALIDER

Rechercher

temps de réponse: 578 ms (+2%)

0		CARREFOUR	↔	
7		Chou fleur	↔	▲ +2
8		Compote pomme abricot ANDROS	↔	▼ -1
9		Huitres bio creuse Chausey N3	↔	▼ -1
10		Navets ronds violet	↔	▲ +1
11		Riz long grain 10 min CARREFOUR	↔	▲ +1
12		Soda zero sucre COCA-COLA	↔	▼ -2
13		Allumettes de jambon FLEURY MICHON	↔	▲ +2
14		Gâteaux au chocolat CARREFOUR	↔	▼ -1

Lignes 20 1-20 sur 6694 < >

PRODUIT OFFRES TECHNIQUES SCORE

Filtrer

product.name	Compote pomme abricot ANDROS	Huitres bio creuse Chausey N3
product.brand_name	ANDROS	SANS
product.packaging	les 4 pots de 100g	la bourriche de 24 huitres
product.gtin (EAN)	3608580905356	3523680327801
product.warning_advice		Ces coquillages doivent être vivants au moment de l'achat.
product.non_food	[object Object]	
product.flags.is_eco_packaged	false	false
product.flags.is_salt		

Improving the life of a search engine administrator

Automatization with AI → no more Administrators?

Fine grained relevance tuning

- Boosts
- Function scores
- Synonyms
- Redirections
- Query rewriting
- User experience

Conclusion :

- Takes into account Business needs (marketing, stock, partnerships, ...)
- Easily understandable and maintainable by the Business

VS.

Learning to Rank (automated relevance tuning)

- Ranking model
- Learning algorithm
- Training data

Conclusion :

- Very expensive to implement properly
- Needs a lot of maintenance
- Does worse if not properly trained and maintained
- Black box to the Business

Beyond Artificial Intelligence for Search

It is proven that for relatively well-structured data, like in e-commerce for example, a hand tailored search configuration can easily outperform machine learning approaches for relevance. The search configuration considers the different searchable fields, a business taxonomy and ontology, some domain related synonyms, a few specific landing pages, boosts and some business numerical criteria.

In the same way, we describe an approach for relevance in the case of large-scale search engines which is not based on classical "PageRank" and machine learning approaches. We propose a model based on social interactions between communities and individuals that are using or configuring the search engine. We then compare this model with machine learning powered approaches.

Kesselhaus

15.06.2021 21:20 - 22:20

your local time:

9:20 PM - 10:20 PM ⓘ

Lightning Talk

Intermediate



<https://2021.berlinbuzzwords.de/session/beyond-artificial-intelligence-search>

Video

Lucian Precup – Beyond Artificial Intelligence for Search

The best AI :-)

A historical black and white illustration of a man in a suit sitting at a desk with a large, complex mechanical device, possibly an early computer or calculator. A play button is overlaid on the image, indicating it is a video thumbnail.

BERLIN BUZZWORDS

Search Engine Administrator's wish ...

I need my Search Engine to be like Google

The screenshot shows a Google search for "samsung galaxy s24". The search bar is at the top with the Google logo. Below the search bar, there are tabs for "All", "Products", "Product sites", "Images", "Videos", and "More". The main content area is divided into two sections: "Sponsored" and organic search results. The "Sponsored" section features a grid of product listings for various Samsung Galaxy S24 models, including prices, shipping options, and seller information. Below the sponsored listings, there are organic search results from Samsung, Amazon, and Idealo, providing detailed product information and pricing for the Galaxy S24 and S24 Ultra.

And my Business Console like ChatGPT

The screenshot displays the Microsoft Copilot interface. At the top, the Copilot logo is followed by the tagline "Your everyday AI companion". Below this, there are three AI-generated suggestions: "What are some hilarious one-liners that will brighten my day", "Create a table that helps me plan lunches for a week", and "Compose a folk song about a day in the life of a roaming cloud". A navigation bar at the bottom allows users to "Choose a conversation style" with options for "More Creative", "More Balanced" (selected), and "More Precise". A "New topic" button and a text input field "Ask me anything..." are also visible.

Examples of API calls to automate

Please add a synonym for **car** = **automobile** for field title

Can you boost by **10** the Münchner Paulaner brand from **September 21 to October 6, 2024** when my users are searching for **beer**?

Please **test** if removing the **description** field from the list of **searchable fields** increases the relevance of my search engine.

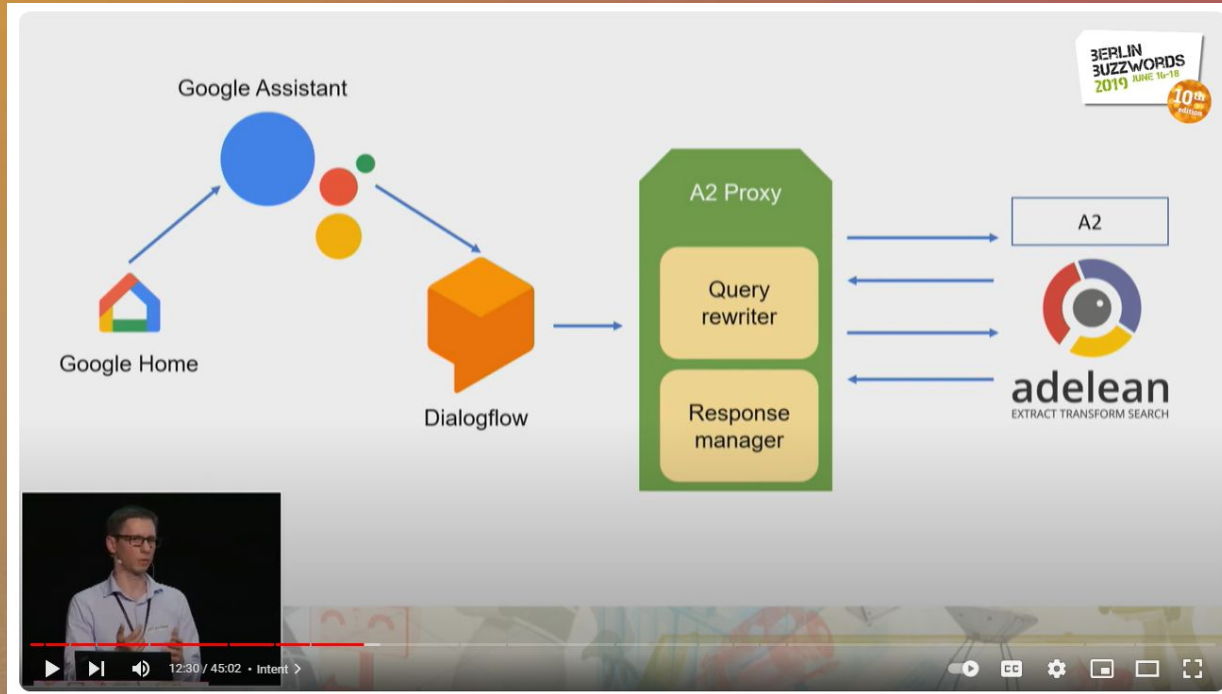
Examples of alerts and suggestions

It looks like the vast majority of your users that are searching for **sugar**, add only products from the **Baking Ingredients** category to cart. Would you like to create a filtered redirection?

Based on your product catalog and user activity, here is a synonym suggestion that would reduce zero results queries: **salad=letuce**. Would you like me to add it to your configuration?

Natural Language Understanding and Large Language Models to automate actions

Berlin Buzzwords 2019 - Integrate your Search Engine with a Voice Assistant - an AI Assistant on top of APIs



<https://2019.berlinbuzzwords.de/19/session/integrate-your-search-engine-voice-assistant.html>

What can the Large Language Models (LLMs) do?

++ Understand human language (NLU, NLP)	→ generic machine learning models
++ Perform a synthesis	→ generic machine learning models
++ Hold a conversation and understand a context	→ the very principle of LLMs, remuneration for word generation
+ Have an exhaustive knowledge base	→ if model trained on this basis
+ Adapt and learn	→ use the context (immediate), re-training of the model (long term)
(+) Allow multimodal interactivity	→ interpretation and generation of images
(+) Understand the voice	→ external module
(+) Automate tasks	→ code generation
~ Respond efficiently and relevantly	→ hallucinations

From voice to API calls



VOICE

The user speaks to his **voice** assistant



SPEECH-TO-TEXT

Voice is transformed into **text**



INTENT
UTTERANCE
SLOTS

The system recognizes an **utterance** related to an **intention** and extracts keywords from the **slots**



API

Call the **API** corresponding to the intent with the **parameters** derived from the slots

AI Assistant in OpenSearch Dashboards

for Observability and Data Analytics use cases



AI Assistant in OpenSearch Dashboards

for Observability and Data Analytics use cases

```
> source=opensearch_dashboards_sample_data_logs  
| where timestamp > DATE_SUB(NOW(), INTERVAL 1  
WEEK) | stats COUNT() AS count by response
```

AI Assistant in OpenSearch Dashboards

for Observability and Data Analytics use cases



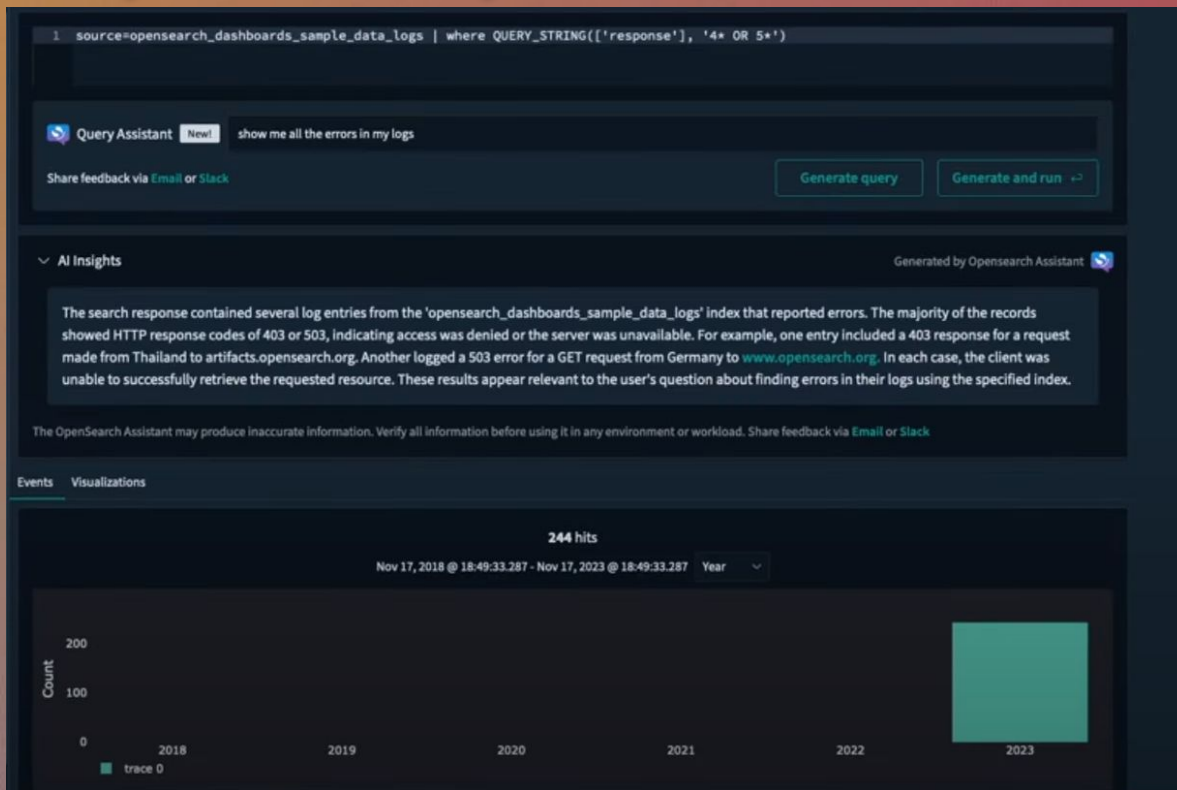
Query Assistant

New!

“show me all the errors in my logs”

AI Assistant in OpenSearch Dashboards

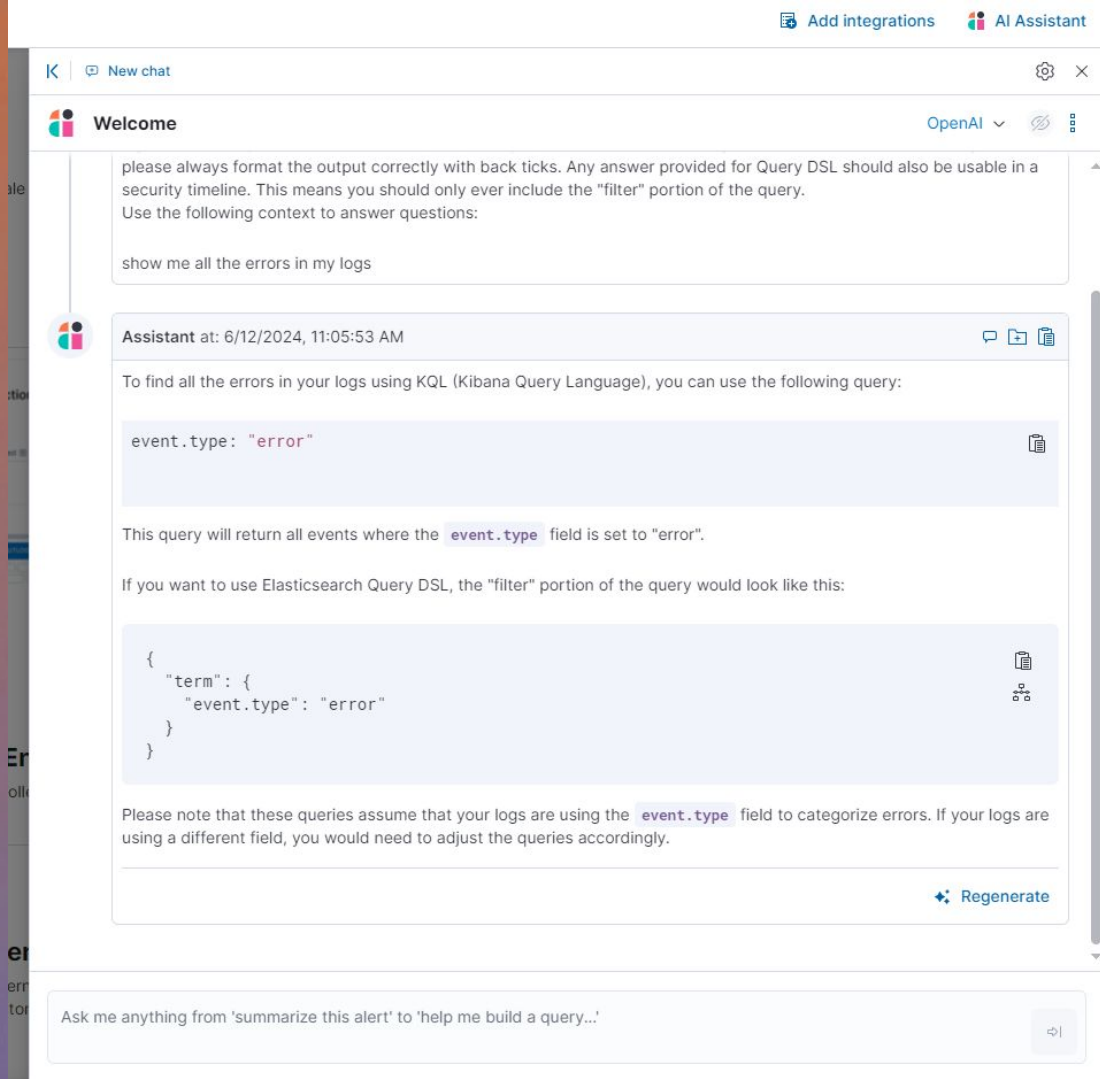
for Observability and Data Analytics use cases



https://www.youtube.com/watch?v=9r0RyB_oHkk&ab_channel=OpenSearch

AI Assistant in Kibana

for Observability and Data Analytics use cases



The screenshot shows the Kibana AI Assistant interface. At the top, there are navigation icons for 'New chat', 'Add integrations', and 'AI Assistant'. The main chat area is titled 'Welcome' and contains the following text:

please always format the output correctly with back ticks. Any answer provided for Query DSL should also be usable in a security timeline. This means you should only ever include the "filter" portion of the query.
Use the following context to answer questions:

show me all the errors in my logs

Assistant at: 6/12/2024, 11:05:53 AM

To find all the errors in your logs using KQL (Kibana Query Language), you can use the following query:

```
event.type: "error"
```

This query will return all events where the `event.type` field is set to "error".

If you want to use Elasticsearch Query DSL, the "filter" portion of the query would look like this:

```
{
  "term": {
    "event.type": "error"
  }
}
```

Please note that these queries assume that your logs are using the `event.type` field to categorize errors. If your logs are using a different field, you would need to adjust the queries accordingly.

At the bottom right of the chat area, there is a 'Regenerate' button. Below the chat area is an input field with the placeholder text: 'Ask me anything from 'summarize this alert' to 'help me build a query...'

NLU and LLMs to automate actions → implementation



Configuration avancée

GÉNÉRALE **BOOST** SYNONYMES DICTIONNAIRE FACETTE

Nombre des champs : 0
Rechercher
brand

- product
 - technical_non_food_attr...
 - it_brand_cpu
 - non_food
 - compatible_models_...
 - brand_name**
 - flags
 - is_carrefour_brand
 - sub_brand_name
 - technical_attributes
 - console_brand

Filterer
Munchner PAULANER + CRÉER

VALEUR DU CHAMP	TEXTE DE RECHERCHE	APPROXIMATIF	BOOST	MODE	DÉBUT	FIN	ACTIONS
Munch							
ner	beer		10	Précis	06/09/2024	06/10/2024	
PAULA							
NER							

```
PATCH /catalogs/config/${catalog}
{
  "op": "add",
  "path":
  "${getFieldPath({field})}.valueBoosts${ index ?
  ('.' + index) : ''}",
  "value": "index ? value : [ value ]"
}
```

NLU and LLMs to automate actions → implementation

Configuration avancée

GÉNÉRALE **BOOST** SYNONYMES DICTIONNAIRE FACETTE

Nombre des champs : 0
Rechercher
brand

- product
 - technical_non_food_attr...
 - it_brand_cpu
 - non_food
 - compatible_models_...
 - brand_name**
 - flags
 - is_carrefour_brand
 - sub_brand_name
 - technical_attributes
 - console_brand

Filterer
Munchner PAULANER + CRÉER

VALEUR DU CHAMP	TEXTE DE RECHERCHE	APPROXIMATIF	BOOST	MODE	DÉBUT	FIN	ACTIONS
Munch							
ner	beer		10	Précis	06/09/2024	06/10	
PAULA							
NER							

```
PATCH /catalogs/config/${catalog}
{
  "op": "add",
  "path": "brand_name.valueBoosts.0",
  "value": {
    "value": "Munchner PAULANER",
    "approximate": null,
    "boost": "10",
    "trigger": {
      "text": "beer",
      "textMatchingMode": "TEST_SYNONYMS",
      "beginDate": "20240920 000000",
      "endDate": "20241006 000000"
    }
  }
}
```

NLU and LLMs to automate actions → implementation

Step 1/

Prompt

- User query
- Example of API call and parameters
- Instructions on how to recognize an utterance related to an intention and extracts parameters from the slots

Result

- The code of an API call with the right parameters

Step 2/

(Optionally) Validate the syntax and validity of the API call

Launch the API call

(Optionally) Refresh the screen

NLU and LLMs to automate actions → the final result

The screenshot displays a search interface with a blue header labeled 'Recherche'. Below the header, there are radio buttons for 'Web' (selected) and 'App'. A search bar contains the text 'Rechercher'. Below the search bar, there is a section for 'Requêtes / Groupes' and a response time indicator: 'Temps de réponse : 134 ms'. The main content area shows a grid of product cards. Each card includes a title, an image, a brand name, and a price. To the right of the grid is a 'Facettes' sidebar with expandable sections for 'Rayon (8)', 'Type de produit (200)', 'Eco-score (1)', and 'Prix (10)'. At the bottom left, there is a blue circular icon with a question mark and a button labeled 'Accéder à l'assistant'.

Product Title	Brand	Price (€)
1.2.Spray - Le Balai à Plat...	vileda	20.62 €
2 Poneys Islandais et...	playmobil	10.99 €
4x4 de pompier avec lance-ea...	playmobil	39.99 €
5 Secondes MEGABLEU	megableu	26.99 €
5 secondes relais...		
Abri pour Barbecue...		
Abri de Jardin en Résine 4m2...		
Accessoire raquettes tenn...		

NLU and LLMs to automate actions → the final result

The screenshot displays a search interface with a blue header labeled "Recherche". Below the header, there are radio buttons for "Web" (selected) and "App". To the right, there are three blue buttons: "INFORMATIONS", "SCORE", and "COMPARAISON". A "Facettes" section on the right includes filters for "Rayon (8)", "Type de produit (200)", "Eco-score (1)", and "Prix (10)". The main content area shows a grid of product cards. A blue "Assistant" overlay is positioned on the left, containing a close button (X), a text prompt "Vous pouvez entrer un message en dessous pour démarrer la conversation", and a text input field with a right-pointing arrow. The product cards visible include:

Product Name	Price
2 Poneys Islandais et...	10.99 €
4x4 de pompier avec lance-ea...	39.99 €
5 Secondes MEGABLEU	26.99 €
Abri pour Barbecue...	
Abri de Jardin en Résine 4m2...	
Accessoire raquettes tenn...	

NLU and LLMs to automate actions → the final result

Recherche

Simuler Web App

INFORMATIONS

SCORE

COMPARAISON

Facettes

- Rayon (8)
- Type de produit (200)
- Eco-score (1)
- Prix (10)

Stratégies de Tri

Nom de la stratégie : defaultStrategie

Tri 1

Assistant

Can you boost by 10 the Muncher PAULANER brand from September 21 to October 6, 2024 when my users are searching for beer

34 ms

```
{ "boost_number": "10",  
  "brand_name": "Muncher PAULANER",  
  "start_date": "September 21",  
  "end_date": "October 6, 2024",  
  "search_term": "beer" }
```

EXÉCUTER

2 Poneys Islandais et...	4x4 de pompier avec lance-ea...	5 Secondes MEGABLEU
playmobil	playmobil	megableu
10.99 €	39.99 €	26.99 €

Abri pour Barbecue...

Abri de Jardin en Résine 4m2...

Accessoire raquettes tenn...

NLU and LLMs to automate actions → the final result

Recherche

Simuler Web App

INFORMATIONS

SCORE

COMPARAISON

Facettes

Rayon (8)

Type de produit (200)

Eco-score (1)

Prix (10)

Stratégies de Tri

Nom de la stratégie : defaultStrategie

Tri 1

Assistant

Can you boost by 10 the Muncher PAULANER brand from September 21 to October 6, 2024 when my users are searching for beer

Votre requête a été générée :

```
{ "boost_number": "10",  
  "brand_name": "Muncher PAULANER",  
  "start_date": "September 21",  
  "end_date": "October 6, 2024",  
  "search_term": "beer" }
```

LA REQUÊTE A ÉTÉ EXECUTÉE

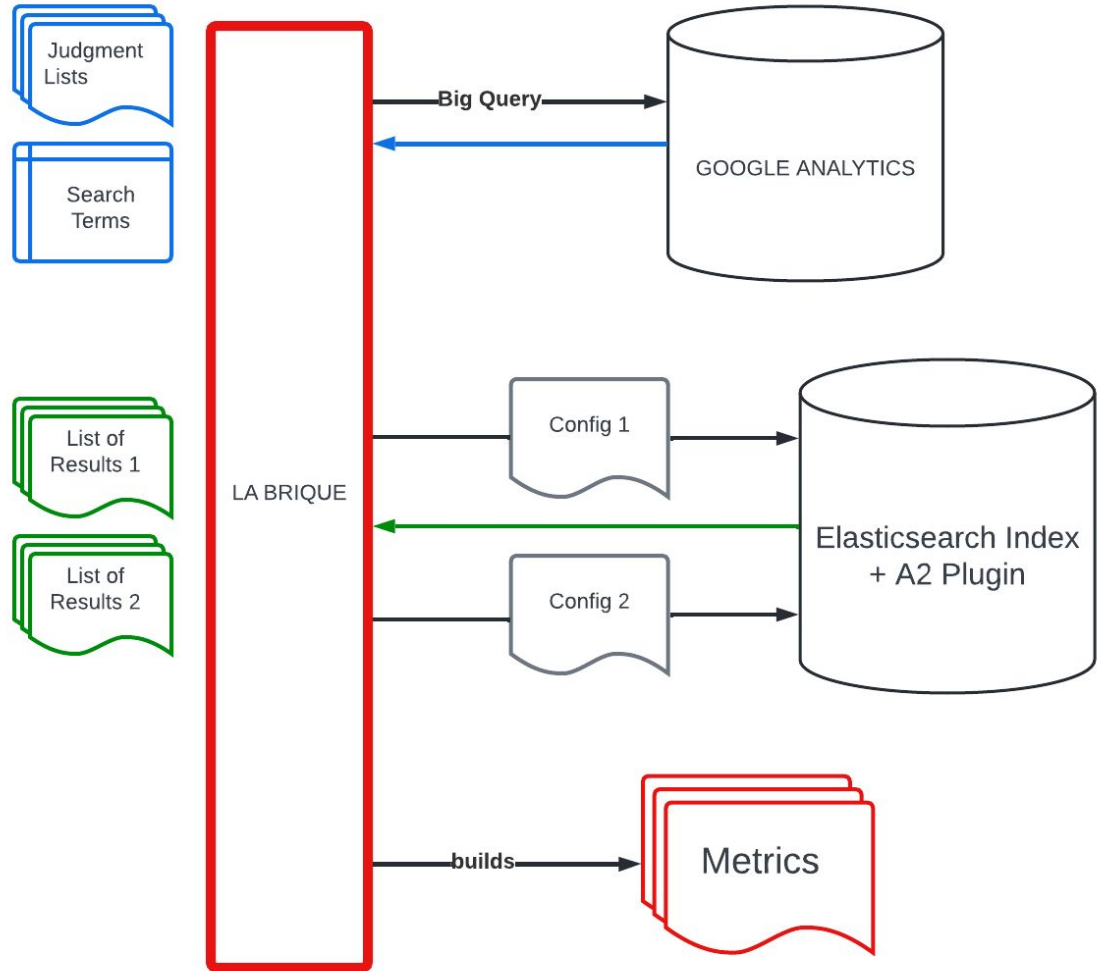
2 Poneys Islandais et... 10.99 €

4x4 de pompier avec lance-ea... 39.99 €

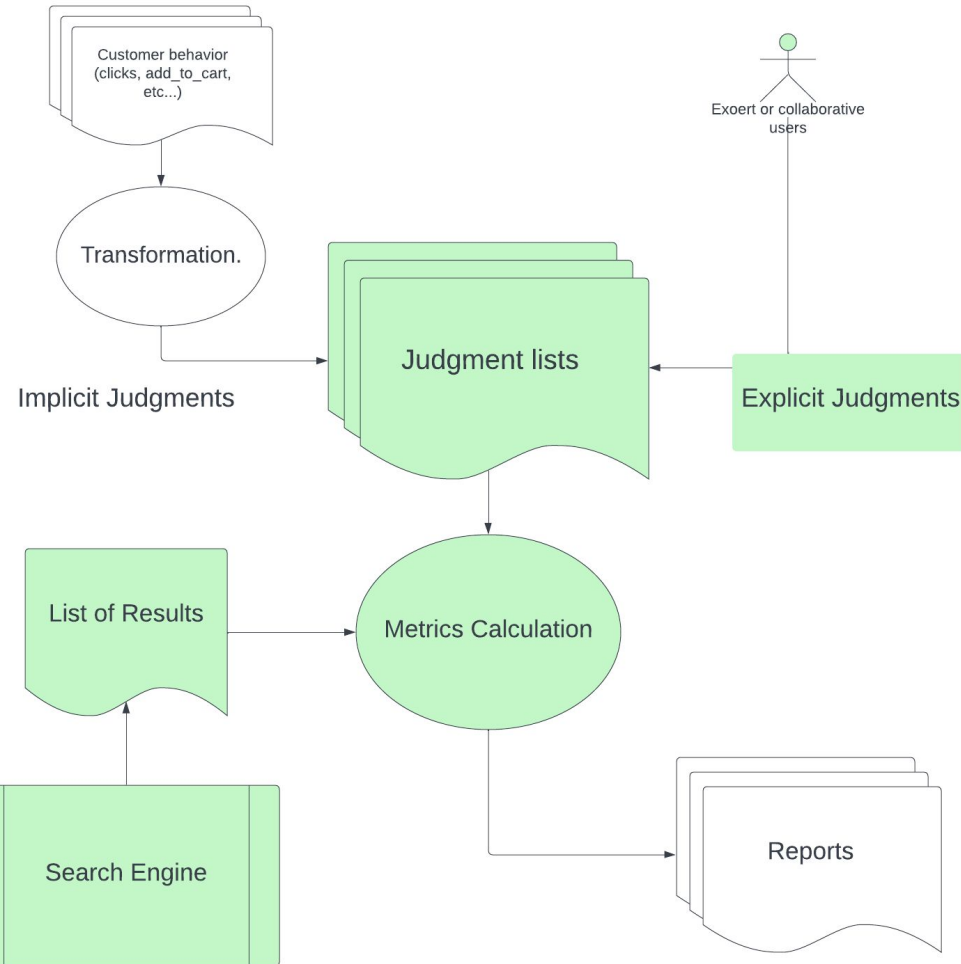
5 Secondes MEGABLEU 26.99 €

Abri pour Barbecue... Abri de Jardin en Résine 4m2... Accessoire raquettes tenn...

Testing the non-regression

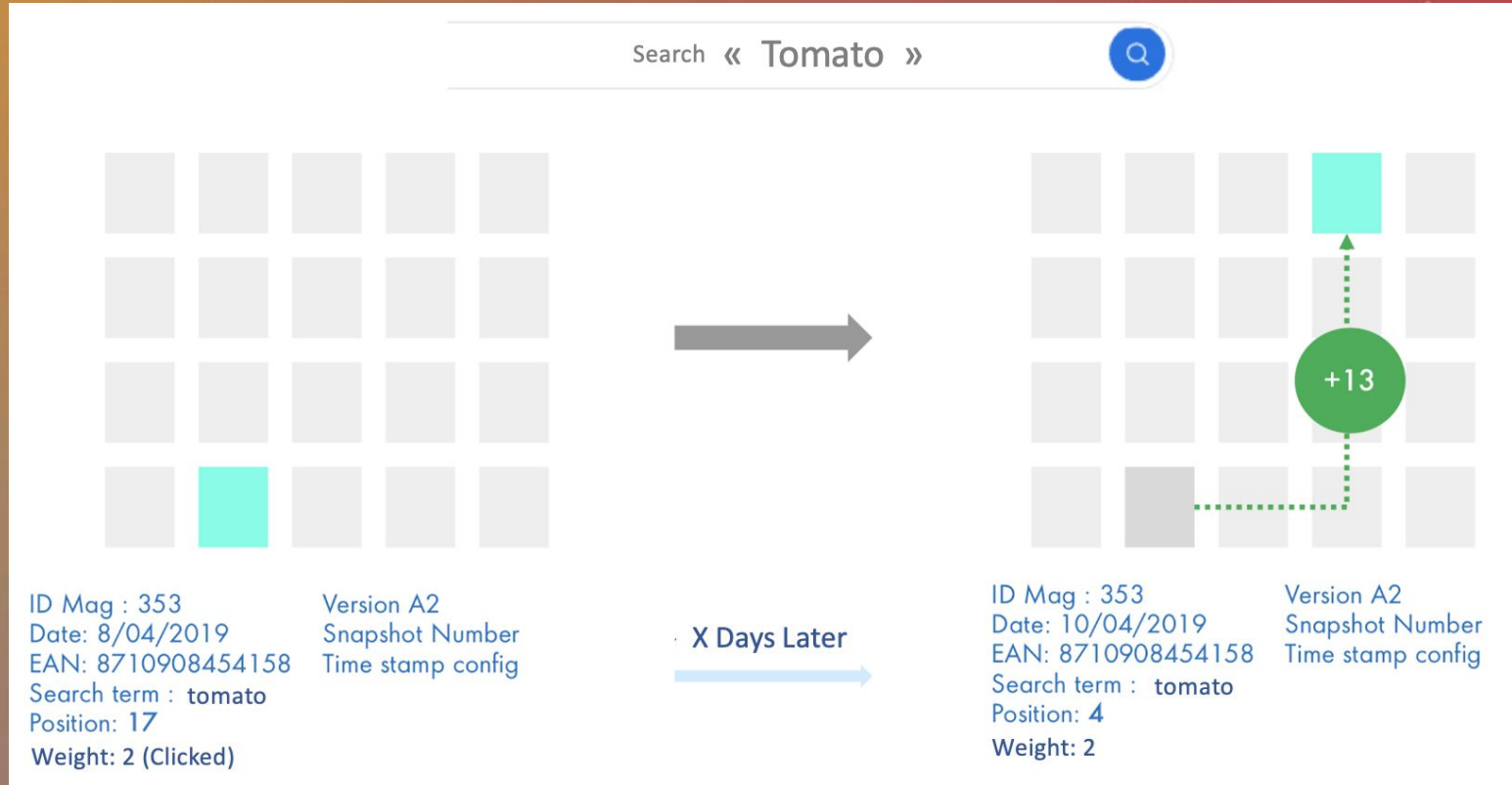


Quepid



Quepid

Testing the non-regression



Haystack EU 2022 - A practical approach to measuring the relevance and preventing regressions

HAYSTACK [Home](#) [Haystack US 2024 Schedule](#) [Tickets](#) [Travel](#) [Past Conferences](#) [Past Talks](#)

A practical approach to measuring the relevance and preventing regressions

Aline Paponaud and Roudy Khoury • Location: TUECHTIG • [Back to Haystack EU 2022](#)

In this session we present a practical approach in implementing a component that automatically measures the relevance of an e-commerce search engine, helps configuring it and prevents relevance regressions. Our component is based on the Normalized Discounted Cumulative Gain (NDCG) theory and integrated into a real-world large-scale e-commerce search engine. We give our return on experience and present all the practical details of its implementation - logs collection, integration with the search engine administration interface, deployment automation and give an idea about its return on investment (ROI).

[Download the Slides](#) [Watch the Video](#)



Aline Paponaud
Adelean

CTO of Adelean, working with search and providing consulting services and expertise around Elasticsearch, Lucene and Solr. She brings her energy to leveraging search engines, as they become more and more essential in every domain.



Roudy Khoury
Adelean

Roudy holds a Masters in Artificial Intelligence from Paris 12 University. He joined Adelean as a software engineer and has been most interested in the areas of search and natural language processing. He has hands on experience implementing Elasticsearch based search engine solutions in various sectors of activity. Roudy enjoys challenges and solving problems and has worked in a variety of industries.



<https://haystackconf.com/eu2022/talk-3/>

Analytics and Machine Learning to bring insights

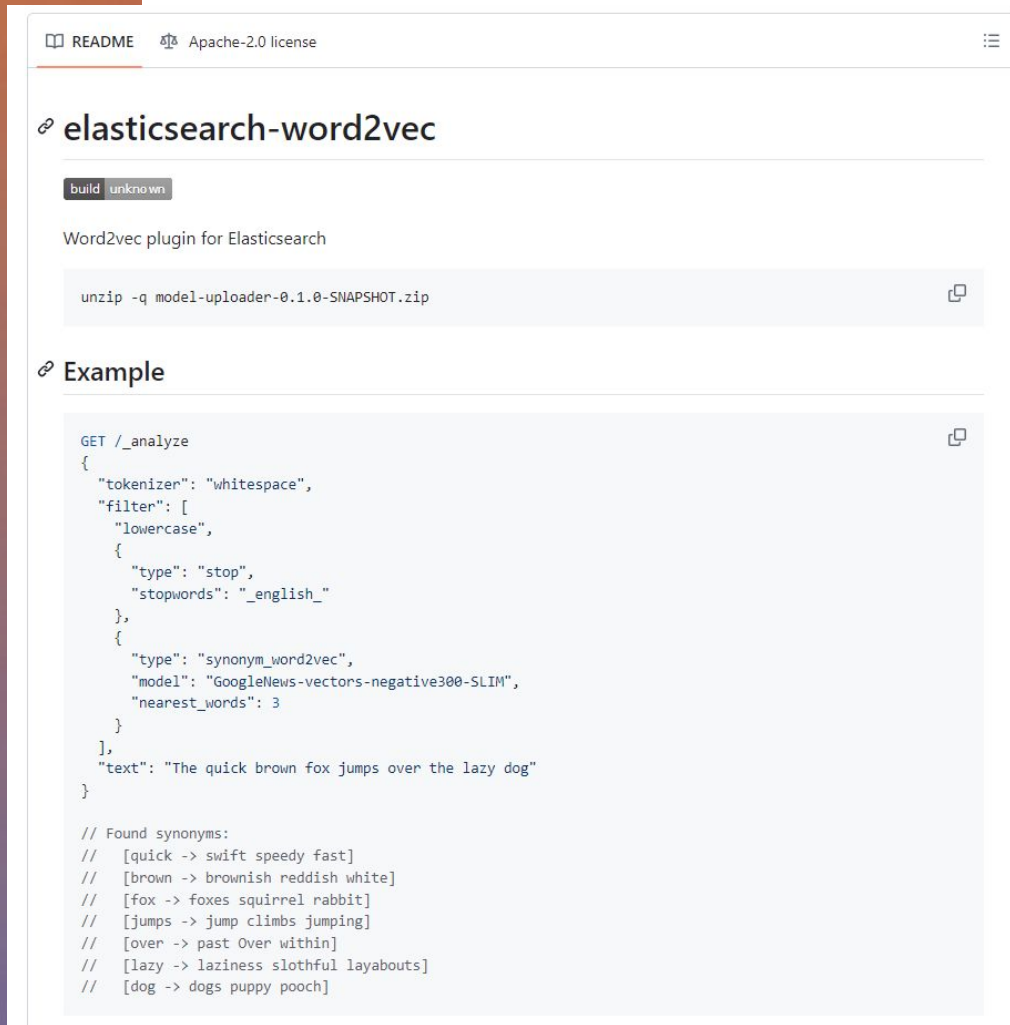
Reminder: Examples of alerts and suggestions

It looks like the vast majority of your users that are searching for sugar, add only products from the Baking Ingredients category to cart. Would you like to create a filtered redirection?

Based on your product catalog and user activity, here is a synonym suggestion that would reduce zero results queries: salad=letuce. Would you like me to add it to your configuration?

Monitor zero results queries and suggest synonyms

<https://github.com/adelean/elasticsearch-word2vec>



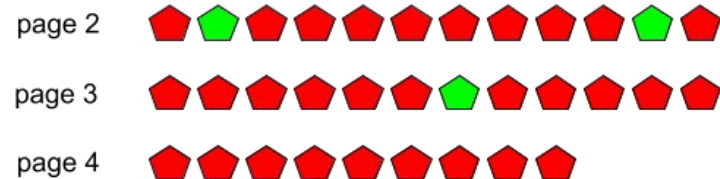
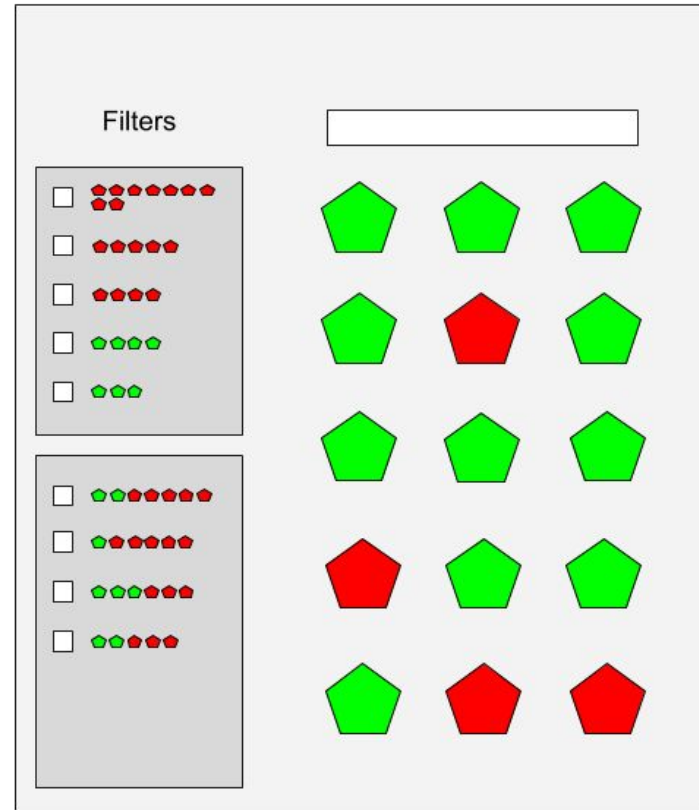
The screenshot shows the GitHub repository page for `elasticsearch-word2vec`. At the top, there are links for `README` and `Apache-2.0 license`. The repository name `elasticsearch-word2vec` is displayed prominently. Below the name, there is a status bar indicating `build unknown`. The description reads: "Word2vec plugin for Elasticsearch". A code block shows the command to unzip the model: `unzip -q model-uploader-0.1.0-SNAPSHOT.zip`. An `Example` section contains a JSON request for the `_analyze` endpoint and the corresponding response showing found synonyms for the text "The quick brown fox jumps over the lazy dog".

```
GET /_analyze
{
  "tokenizer": "whitespace",
  "filter": [
    "lowercase",
    {
      "type": "stop",
      "stopwords": "_english_"
    },
    {
      "type": "synonym_word2vec",
      "model": "GoogleNews-vectors-negative300-SLIM",
      "nearest_words": 3
    }
  ],
  "text": "The quick brown fox jumps over the lazy dog"
}

// Found synonyms:
// [quick -> swift speedy fast]
// [brown -> brownish reddish white]
// [fox -> foxes squirrel rabbit]
// [jumps -> jump climbs jumping]
// [over -> past Over within]
// [lazy -> laziness slothful layabouts]
// [dog -> dogs puppy pooch]
```

Improve precision and relevancy of facet values

Issue: how to make facet values relevant in the context of many "less relevant" results ?



- Rayons Promo, épicerie...
- Maison & High Tech Electroménager...
- Services Voyage, spectacle...

📍 Trouver un magasin 📖 Mes listes 👤 Me connecter

Marque

🔍 Marques...

- CARREFOUR 106
- TROPICANA 42
- GERBLE 36
- KARELEA 27
- JOKER 23

👇 VOIR PLUS

Préférences alimentaires

- Sans sucres ajoutés 325
- Sans sucres 100
- Végétarien 54
- De saison 5
- Halal 2



1,28€



0,67€



1,13€



1,25€



Sucre vanillé VAHINE
les 10 sachets de 7,5 g
26.40 € / Kilogramme



Sucre en poudre DADDY
la boîte de 1 kg
1.34 € / Kilogramme



Sucre en poudre DADDY
la boîte de 100 bâchettes - 500 g
4.42 € / Kilogramme



Sucre roux GRAEFFE
le paquet de 1 kg
2.14 € / Kilogramme



- Rayons Promo, épicerie...
- Maison & High Tech Electroménager...
- Services Voyage, spectacle...

Marque

Marques...

- CARREFOUR 8
- BEGHIN SAY 15
- DADDY 9
- SAINT LOUIS 9
- CANDEREL 5



VOIR PLUS

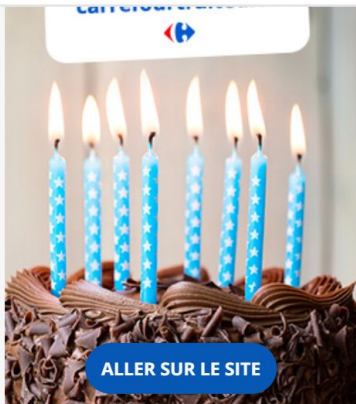
Préférences alimentaires

- Végétarien 23



Labels qualité

- Bio 14
- Commerce équitable 5



0,74€



0,69€



0,99€



Sucre en poudre DADDY
la poche de 900 g
1.34 € / Kilogramme



Sucre Cristal ERSTEIN
le paquet de 1 kg
1.28 € / Kilogramme



Sucre roux morceaux LA PERRUCHE
la boîte de 168 morceaux - 1 kg
1.78 € / Kilogramme



Sucre demi-morceaux DADDY
le paquet de 750 g
1.51 € / Kilogramme



Improve precision and relevancy of facet values: the solution

Solutions: work on your search precision



Analytics and data science have clues: for instance, when clients type “tomato”, is there a category which regroup most of the clicks ?



All you must do is prefilter some facets (or even all the results) with this category : 80% of the result set will disappear and your filters will look good !

Examples of prefiltering at Carrefour:

- 11% of results for “tomatos” are in the “Fresh vegetables” category but they represent 86% of products added to basket
- 24% of results for “rice” are in the “Pasta and Rice” category and represent 90% of purchases
- 8% of results for “sugar” are in the “Sugar and sweeteners” category and represent 90% of purchases

NLU and LLMs to automate actions → implementation

Step 1/

Monitor the clicks on search results

Identify long tails of less relevant results

Launch alert when necessary

Step 2/

Construct configuration API call

- Parameter values
- Template

Result

- The code of an API call with the right parameters

Launch the API call

Screenshot and API of filtered redirection

The screenshot displays a web application interface for managing redirections. A modal window titled "Modifier la redirection" is open, showing the following fields:

- Mots clés: tomate,tomates
- URL: /tomate
- Type: Recherche
- Mot clé: tomate
- Nom du champ: product.SVE_cdRubriqueN2

The background shows a table of redirections with columns for "DATE DE FIN", "ACTIVÉ", and "ACTIONS". The table contains 8 rows, all with "ACTIVÉ" status and "ACTIONS" icons (edit and delete).

PUT

```
.a2_data_actions_redirections/_doc/r1
{
  "keywords": "tomate,tomates",
  "url": "/tomate",
  "type": "recherche",
  "active": true,
  "searchKey": "tomate",
  "searchFilter": "product.category",
  "searchFilterValues": [
    "Sauces, condiments",
    "Légumes"
  ],
  "id": "r1"
}
```

Thank you!

@M_I_C_E_S
#MICES2024



www.adelean.com



info@adelean.com



[@a2lean](https://twitter.com/a2lean)



linkedin.com/company/adelean



adelean
search with [all.site](#)