

WE AUGMENT HUMAN INTELLIGENCE FOR BUSINESS GROWTH

Analyzing Public Data About User Queries on Search Engine to Predict Trends and Evaluate Markets

Paolo Dello Vicario, Partner @ Datrix - CEO @ ByTek



datrix
AI applications



Paolo Dello Vicario

CEO @ ByTek | Partner @ Datrix

CEO @ ByTek

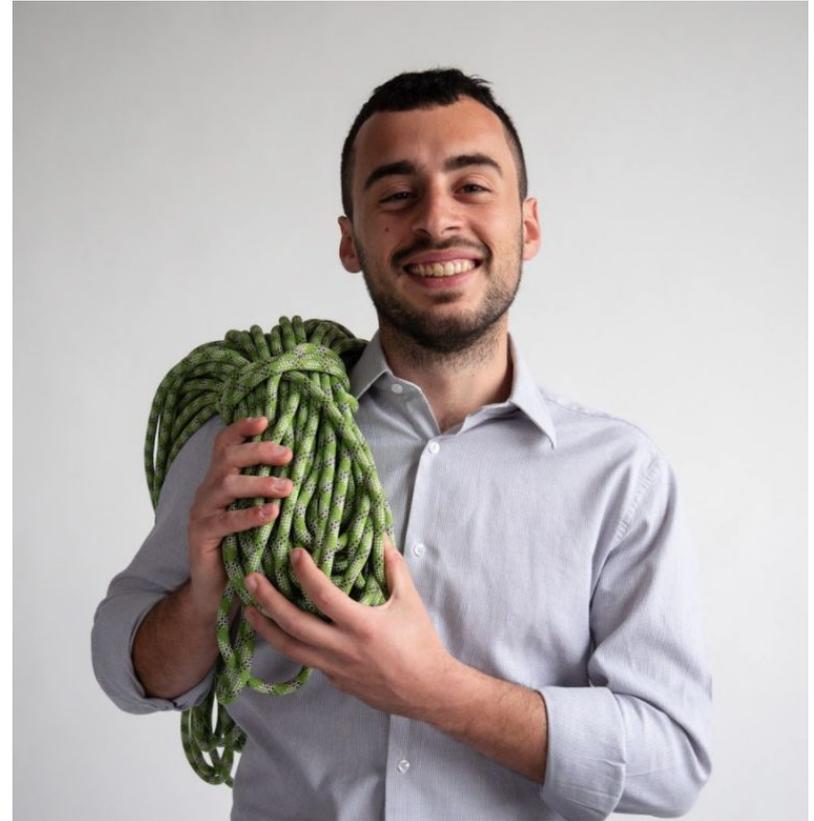
Partner @ Datrix

Coordinator @ MAIBS - University of Tuscia

Coordinator @ Growth Hacking Master - Talent Garden

www.datrixgroup.com

www.maibs.eu



SMART INTENT ANALYSIS

Analyzing the market through web research



CONTEXT

What is search data and how
“search” is changing?



DATA SOURCES

Where can we get the data?
What characteristics do they
have?

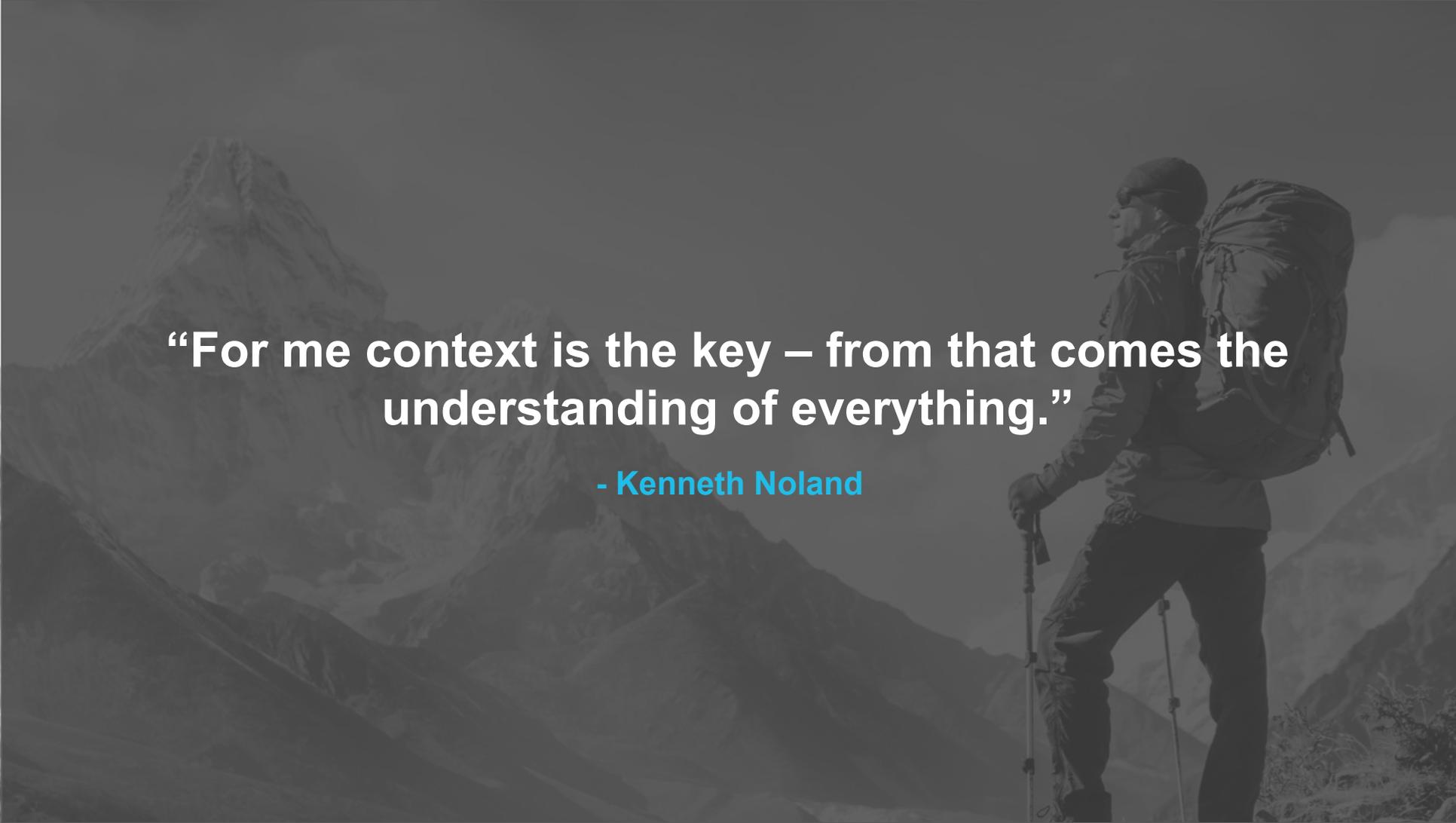


OPPORTUNITIES

What more can augmented
analytics provide?



CONTEXT

A hiker with a large backpack is standing on a mountain trail, looking out over a vast mountain range. The hiker is wearing a dark jacket, pants, and a hat, and is using trekking poles. The background shows a large, rugged mountain peak under a clear sky.

“For me context is the key – from that comes the understanding of everything.”

- Kenneth Noland



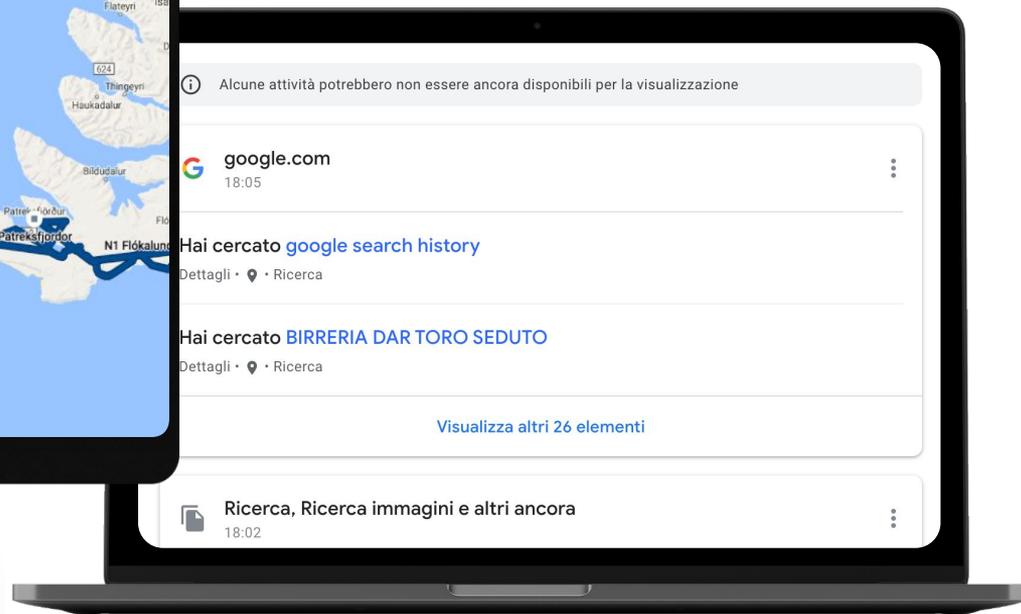
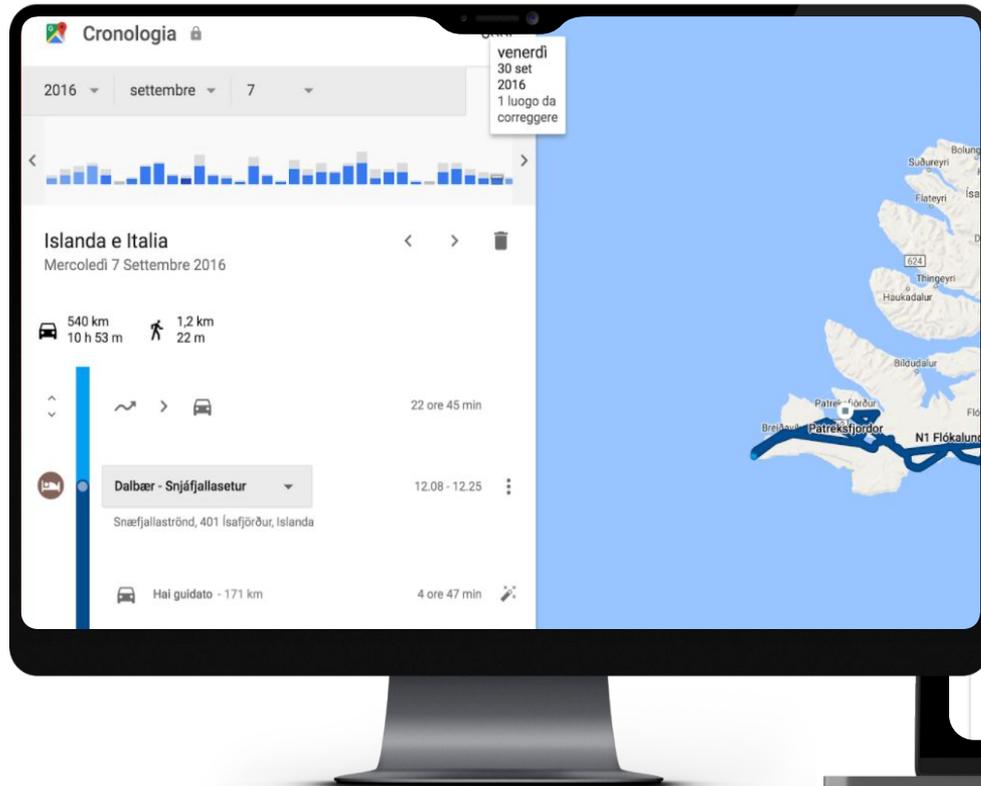












**Google is
where we
confess**

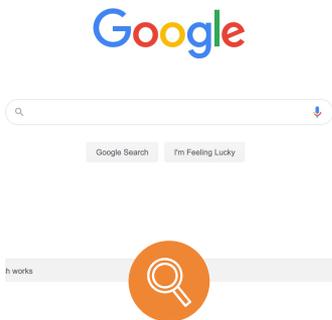




DATA SOURCES

DATA SOURCES

Where we find data



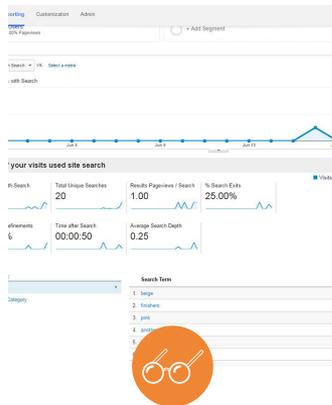
GOOGLE

Expansion of searches, GTrends, proprietary databases or GSC.



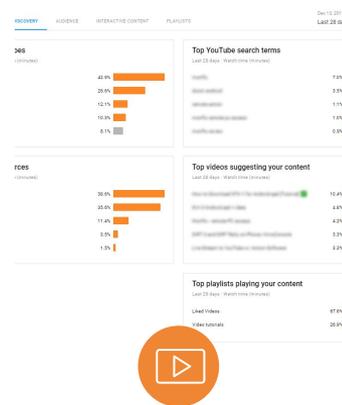
AMAZON

Through expansion with the expansion endpoint.



WEBSITE SEARCH

From Google Analytics, DataLysm and search logs.



YOUTUBE

From YouTube Analytics and the expansion endpoint.

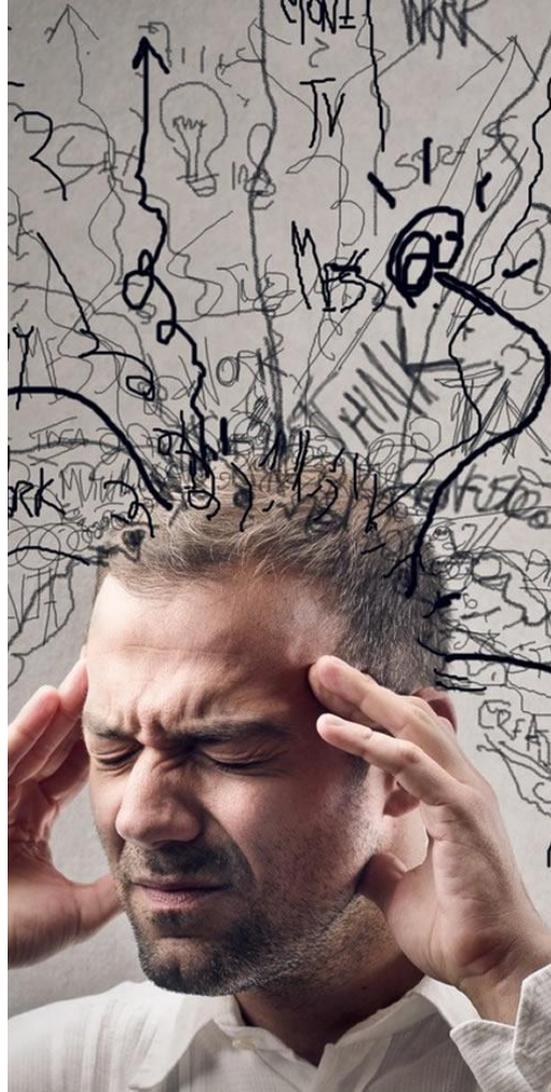
Data are difficult to analyse if you want a complete overview

1.000.000

ROWS

The number of keywords used to describe an "average" industry.

	Keyword	Volume	Vertical	Cluster	Offerta Commerciale	Search Intent	Bytek
1							
2	1st level domain	40	Academy	SEO	Academy Blog	Informational	-
3	301 moved permanently redirect	70	Academy	SEO	Academy Blog	Informational	-
4	301 redirect	22.200	Academy	SEO	Academy Blog	Informational	-
5	301 redirect code	50	Academy	SEO	Academy Blog	Informational	-
6	302 redirect	6.600	Academy	SEO	Academy Blog	Informational	-
7	404 client error	90	Academy	SEO	Academy Blog	Informational	-
8	500 Internal server error come risolvere	110	Academy	SEO	Academy Blog	Informational	-
9	a cosa servono i motori di ricerca	40	Blog	Generic	Academy Blog	Informational	-
10	a cosa sono utili le tecniche di search engine marketing	40	Blog	Generic	Academy Blog	Informational	-
11	a no follow	50	Academy	SEO	Academy Blog	Informational	1,5
12	ab testing	480	Servizi	Analytics	Landing Page Design & CRO	Transactional	-
13	accelerated mobile pages	110	Academy	SEO	Academy Blog	Informational	-
14	acronimo di seo	40	Blog	SEO	Academy Blog	Informational	-
15	ads above the fold	90	Academy	Advertising	Academy Blog	Informational	-
16	ads audience insights	90	Academy	Advertising	Academy Blog	Informational	-
17	ads che cos è	40	Blog	Advertising	Academy Blog	Informational	-
18	ads cos è	50	Blog	Advertising	Academy Blog	Informational	-
19	adsense retargeting	50	Academy	Advertising	Academy Blog	Informational	-
20	adv cos è	40	Blog	Advertising	Academy Blog	Informational	-
21	adv facebook	70	Servizi	Advertising	Facebook Ads Campaign Management	Transactional	-



MANUAL ANALYSIS

The main limits



SLOW UPDATES

Updates are extremely expensive. So they are rarely updated and you lose the speed of action.



LOSS OF STRATEGIC VISION AND OPPORTUNITIES

We limit ourselves to unrepresentative sub-sets and thus lose visibility of the opportunities. We talk about keywords and not themes.



DIFFICULT LINK BETWEEN STRATEGY AND EXECUTION

Strategy and execution thus remain on two separate tracks and the connection is made with enormous logical leaps.

1

**SLOW
UPDATES**



2

**LACK OF STRATEGIC
VISION**



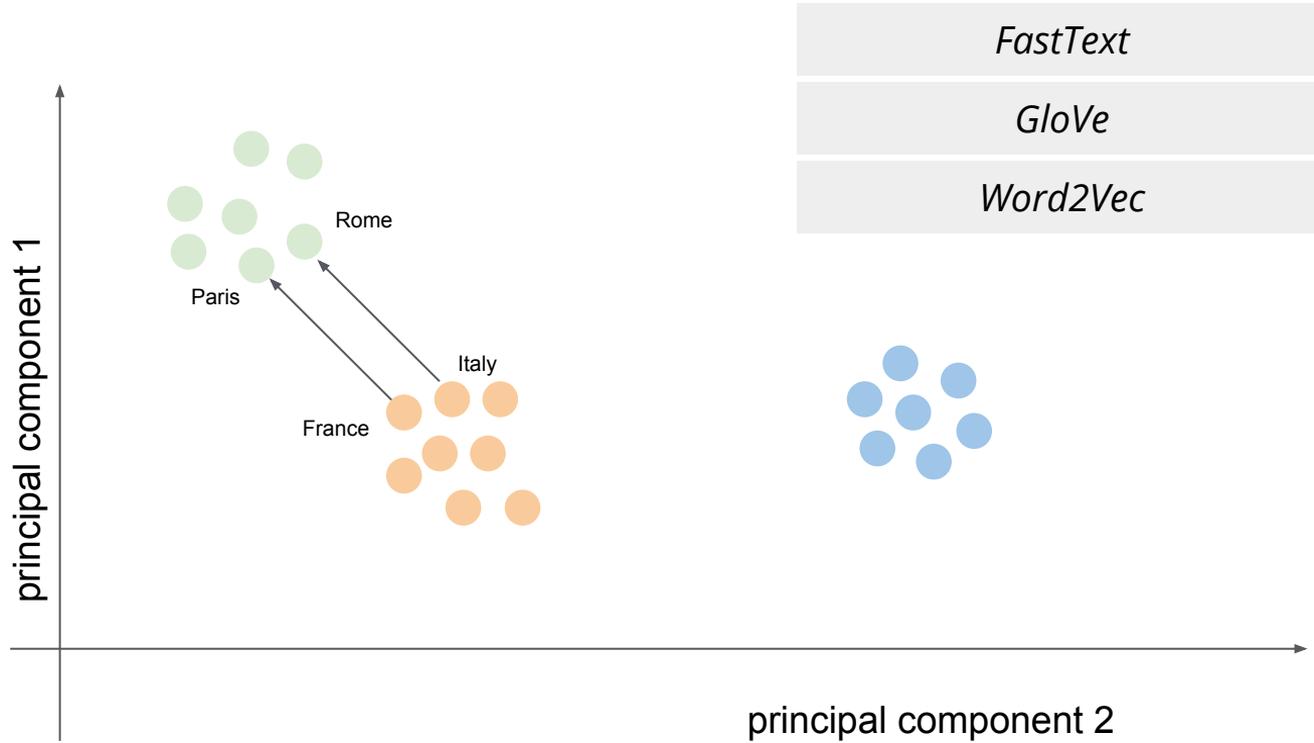
3

**NO LINK STRATEGY -
EXECUTION**





OPPORTUNITIES



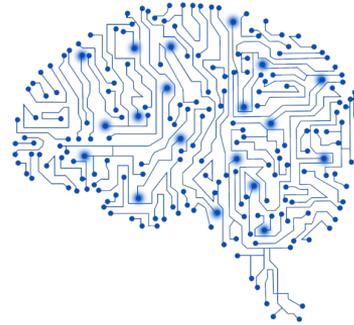


Keyword Table
<i>ID</i>
<i>Keyword</i>
<i>Volume</i>
<i>Seasonality Array</i>

DIM: 1.5 MLN

**0.5%
labeled**

**99.5%
unlabeled**



LSTM

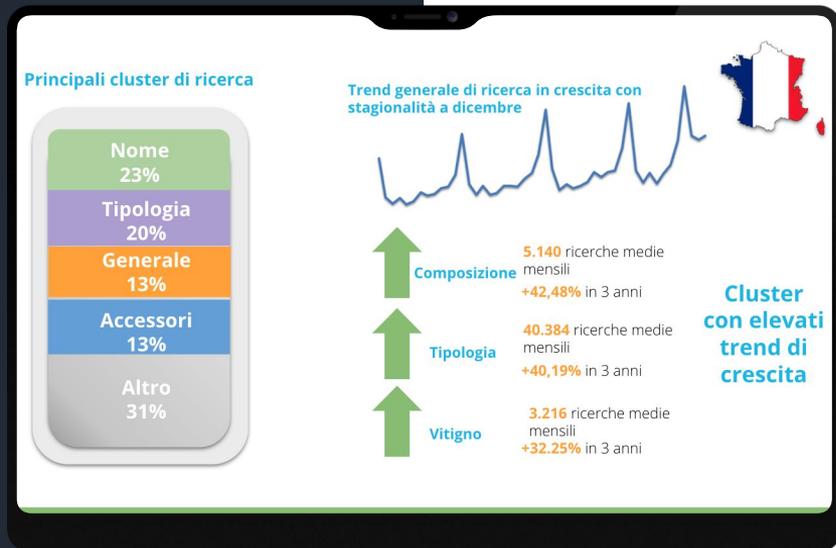
RNN starting from
Facebook FastText
Embeddings and built
with Keras



Keyword Table
<i>ID</i>
<i>Keyword</i>
<i>Cluster</i>
<i>User intent</i>
<i>Volume</i>
<i>Seasonality Array</i>

**DIM: 1.5 MLN
ACC: 0.96**

Layer (type)	Output Shape	
embedding_22 (Embedding)	(None, 9, 300)	→ FastText Embeddings
lstm_30 (LSTM)	(None, 9, 300)	
lstm_31 (LSTM)	(None, 300)	
dense_18 (Dense)	(None, 10)	→ Softmax activation



ANALYSIS

Analysis on a single market



INTEREST OVERVIEW

Static analysis of users' interests with respect to specific sub-themes



TRENDS

Seasonality analysis and trend of growing sub-topics

L'impatto del coronavirus sulle specifiche industry

Un'analisi tramite le ricerche degli utenti online

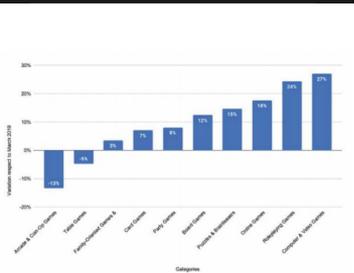
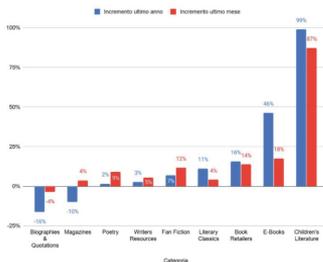
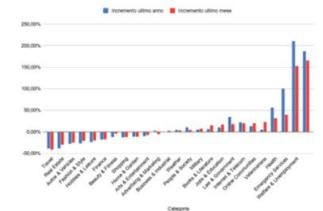


Figure 5: Variation in the online searches between March 2020 and March 2019 for "Games" categories

L'andamento per categorie

Nella Figura 1 sono evidenziate le variazioni percentuali (positive e negative) di tutti i settori sia nell'ultimo anno che nell'ultimo mese. Che il coronavirus abbia condizionato le ricerche online è evidente nella crescita marcata di settori come il welfare e la disoccupazione, i servizi per le emergenze e la salute e anche nelle diminuzioni, anche se meno marcate, di ricerche relativi a viaggi, settore immobiliare, auto e veicoli. È interessante notare la crescita delle categorie animali, online communities, Internet e telecomunicazioni, libri e letteratura.



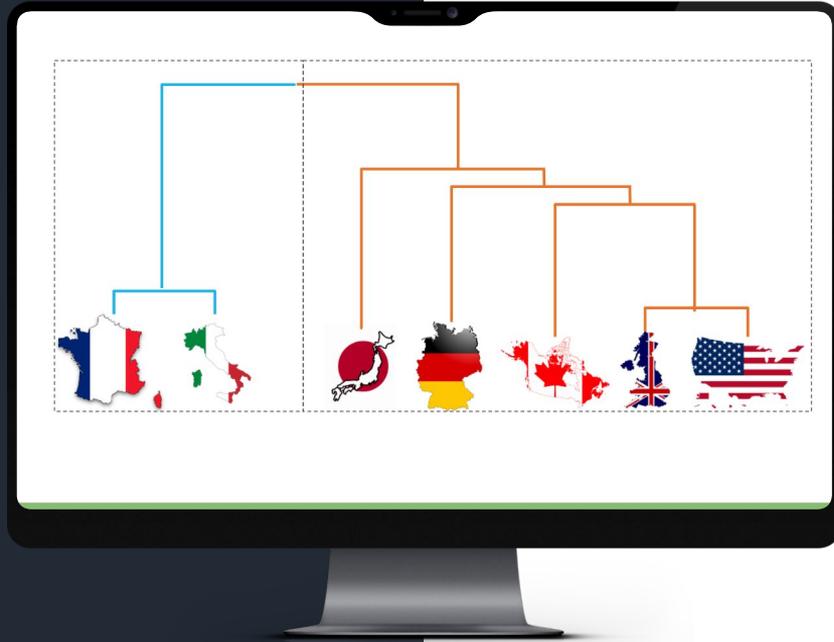
ANALYSIS

Market insights in high volatility periods

Starting from search data, if the dataset is very large, it is possible to analyse near real-time the interest for entire industries and themes.

US: <https://bit.ly/corona-search-us>

IT: <https://bit.ly/corona-search-it>



ANALYSIS

Clustering of different countries and regions



SEARCH SIMILARITY

The different countries and regions can be aggregated according to the way people are looking for.



TRENDS SIMILARITY

Or with respect to the type of emerging trends or seasonality



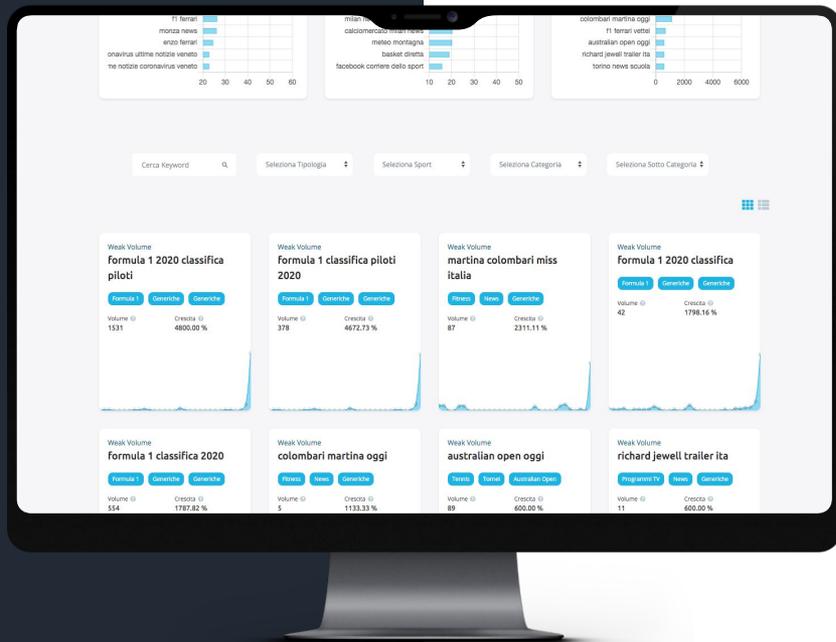
ANALYSIS

With cluster detail



GRANULARITY

Access to search details allows visibility of the individual categories in each country, immediately understanding how to act, through editorial plans for example.



ANALYSIS

Trends and anomaly detection



FROM GOOGLE DATA

Identification of Google search trends and associated news, particularly useful for editorial staff.



ON INTERNAL DATA

Identification of internal research trends in order to orient advertising and mailing choices on eCommerce.



CONCLUSIONS

TAKEAWAYS

SEARCH & AI



WE CAN DO MANY ANALYSIS

Search query data is extremely interesting, easily accessible and has a thousand different applications.



SEARCH IS EVERYWHERE

We have search systems everywhere and we can interrogate them, reading many different nuances.



THE ALGORITHMS ARE ADVANCED

We can develop query classification tasks using very advanced libraries. Let's use them.



THANK YOU!

WWW.DATRIXGROUP.COM

paolo@bytekmarketing.it