



KNOWLEDIA

Turn Data into Knowledge





ABOUT ME



Benjamin Devèze [in](#)

Expert in Artificial Intelligence, Machine Learning, Natural Language Processing, Search Engines

14 years of experience in tech start-ups as 1st employee bootstrapping & scaling technology, product and team



Founder @ Knowledia
2018 - 2020



CTO - Head of Data Science & AI @ Clustree (sold to Cornerstone)
2014 - 2018



CTO @ Yakaz (sold)
2006 - 2014

Education in Computer Science & Artificial Intelligence



TECHNOLOGY OVERVIEW

API

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS



TECHNOLOGY OVERVIEW

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

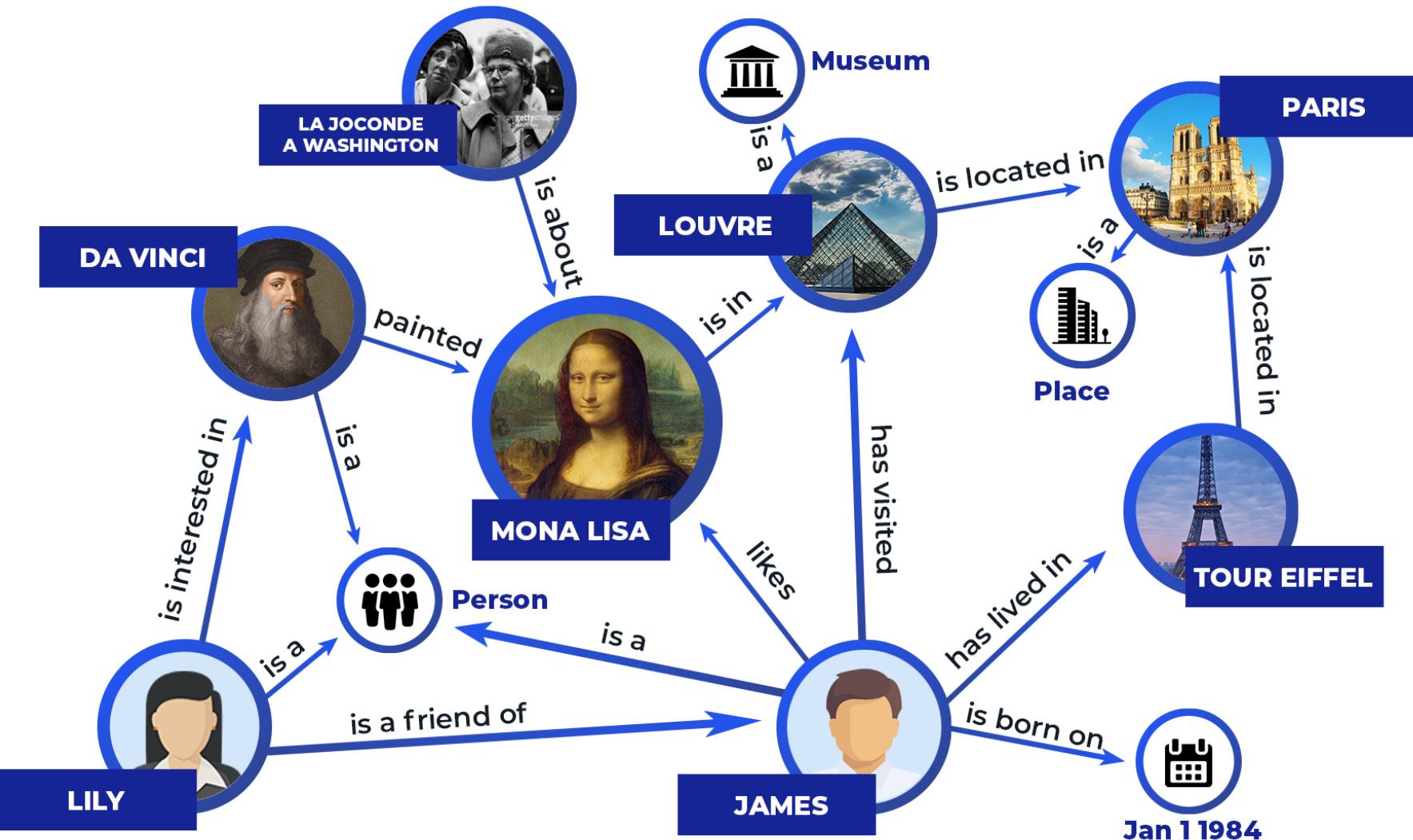
DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS

API

KNOWLEDGE GRAPH

WHAT'S A KNOWLEDGE GRAPH?





KNOWLEDGE GRAPH



NODES ARE ENTITIES



NODES ARE VALUES

MEDIA GEO COORDINATES TEXT QUANTITY TIME URL EXTERNAL IDENTIFIER

EDGES ARE STATEMENTS LINKING NODES VIA A PROPERTY

 Steve Jobs founder  Apple Inc.  Paris location $48^{\circ}51'25''\text{N}, 2^{\circ}21'5''\text{E}$

STATEMENTS ARE SOURCED

 Qasem Soleimani manner of death  Airstrike source www.nytimes.com



KNOWLEDGE GRAPH



SOURCES



Wikidata



Wikipedia



DBpedia



Freebase



LinkedIn



OpenStreetMap



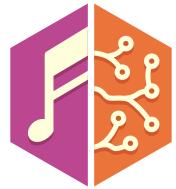
GeoNames



World Factbook



IMDb



MusicBrainz



Amazon



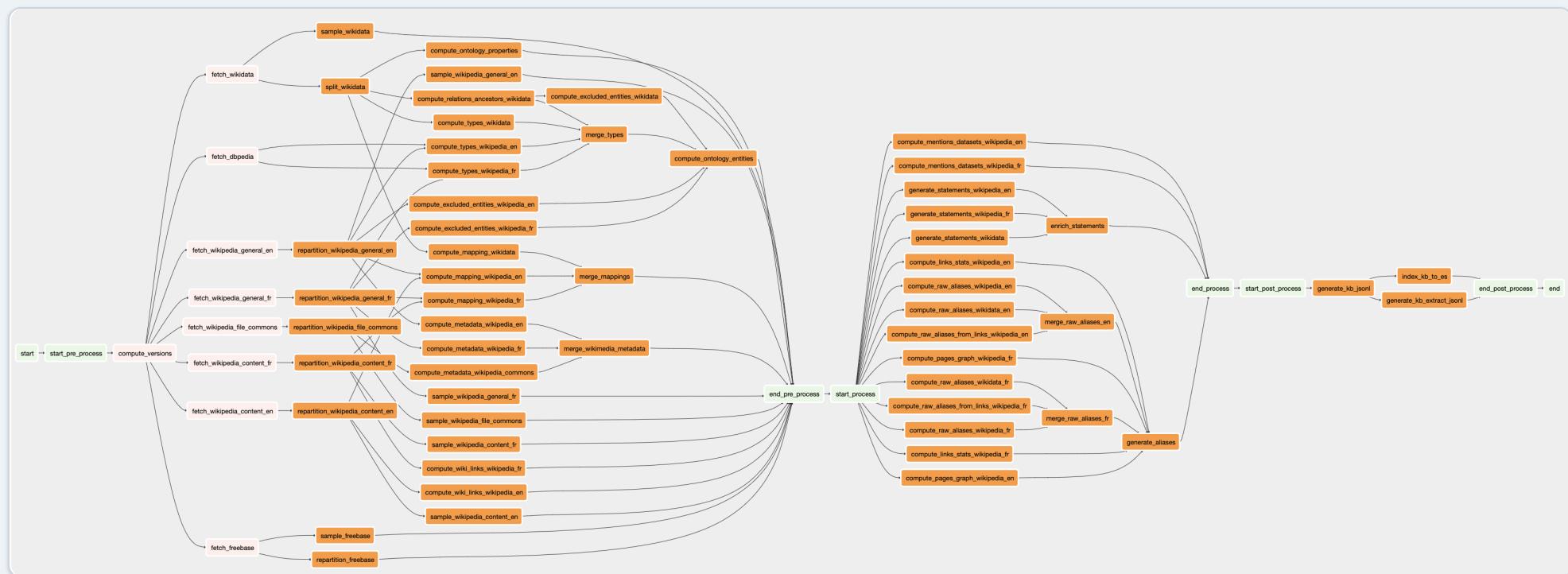
Microformats



Web

KNOWLEDGE GRAPH

COMPUTATION GRAPH





KNOWLEDGE GRAPH



CURRENT NUMBERS



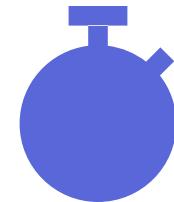
62 million entities



7200 properties



2.7 billion statements



only 12h for full recomputation



TECHNOLOGY OVERVIEW

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS

API

NATURAL LANGUAGE PROCESSING



LANGUAGE DETECTION

AVAILABLE

Steve Jobs was the CEO of Apple. →

Steve Jobs était le PDG d'Apple. →

SENTENCE SEGMENTATION

AVAILABLE

Steve Jobs was the CEO of Apple. He was born in San Francisco.

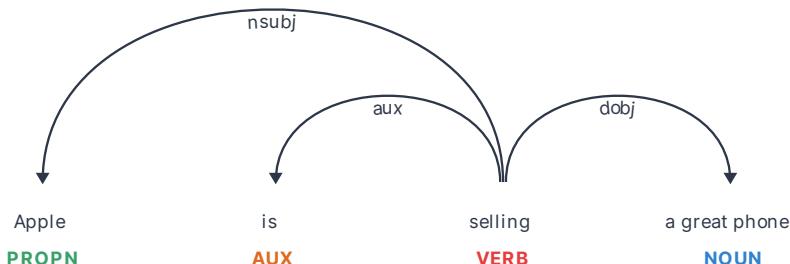
TOKENIZATION

AVAILABLE

Steve Jobs was the CEO of Apple

PART-OF-SPEECH TAGS AND DEPENDENCIES

AVAILABLE



NATURAL LANGUAGE PROCESSING



READING TIME ESTIMATION

AVAILABLE

Steve Jobs was the CEO of Apple... → 5 MINUTES READ

READABILITY ESTIMATION

AVAILABLE

Steve Jobs was the CEO of Apple. → EASY

NAMED ENTITY RECOGNITION

AVAILABLE

PERSON Steve Jobs co-founded ORGANIZATION Apple in DATE 1976

NAMED ENTITY LINKING

AVAILABLE

 Steve Jobs co-founded  Apple . He loved to eat  Apples .

 NATURAL LANGUAGE PROCESSING ••••

CO-REFERENCE RESOLUTION

AVAILABLE

Steve Jobs was the CEO of Apple. He was a perfectionist.

CROSS-LINGUAL SEMANTIC SIMILARITY

AVAILABLE

🇺🇸 Martin Fourcade wins 11th world title, equalling all-time record.

vs

🇫🇷 Fourcade un palmarès en or.



HIGHLY SIMILAR

CATEGORIZATION

AVAILABLE

Martin Fourcade wins 11th world title, equalling all-time record. → SPORTS

NATURAL LANGUAGE PROCESSING



SENTIMENT ANALYSIS

AVAILABLE SOON

We lost awfully to England yesterday due to fatigue. → NEGATIVE

The team played superbly winning all their games against some very good teams. → POSITIVE

ENTITY-BASED SENTIMENT ANALYSIS

AVAILABLE SOON

 **Satya Nadella** is a great CEO unlike  **Steve Ballmer** .

ASPECT-BASED SENTIMENT ANALYSIS

AVAILABLE SOON

The iPhone X delivers a gorgeous 5.8-inch OLED screen but I wish the battery was better. →

SCREEN

BATTERY

NATURAL LANGUAGE PROCESSING

TRANSLATION

AVAILABLE LATER

 Steve Jobs was the CEO of Apple. →  Steve Jobs était le PDG d'Apple.

INFORMATION EXTRACTION

AVAILABLE LATER

Microsoft announced their acquisition of LinkedIn. →  Microsoft subsidiary  LinkedIn

QUESTION ANSWERING

AVAILABLE LATER

South Korea reported 231 new cases of a coronavirus, taking total infections to 833.

+

How many coronavirus cases were reported in South Korea?

→ 833

SUMMARIZATION

AVAILABLE LATER

South Korea reported 231 new cases of a coronavirus, taking total infections to 833, health authorities said on Monday, a day after raising its infectious disease alert to the highest level...

→ South Korea coronavirus cases surge



TECHNOLOGY OVERVIEW

API

KNOWLEDGE GRAPH

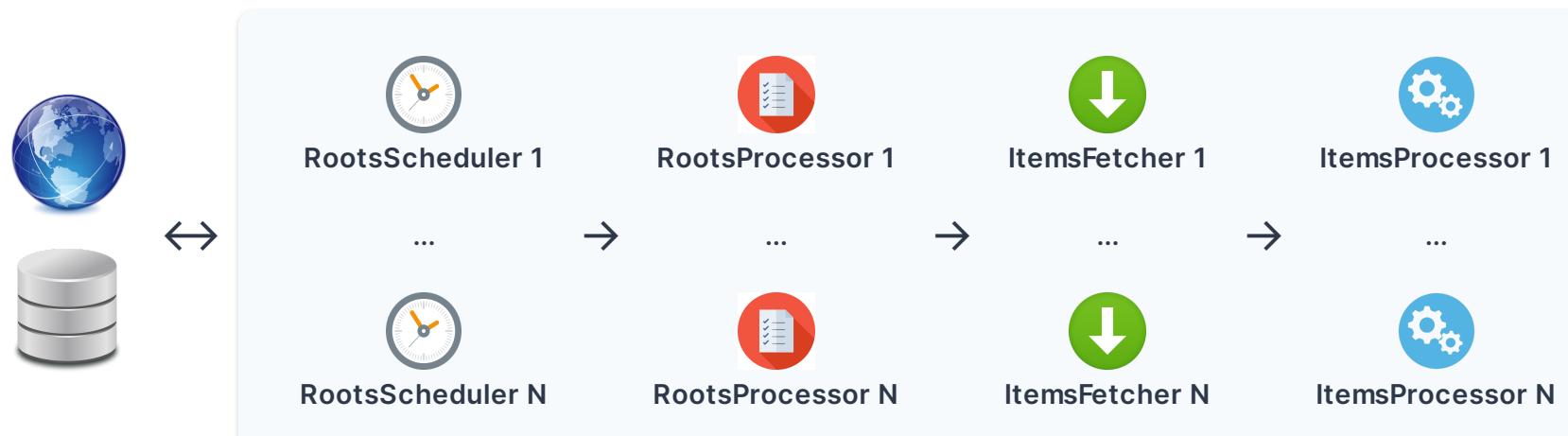
NATURAL LANGUAGE
PROCESSING

**DATA ACQUISITION &
EXTRACTION**

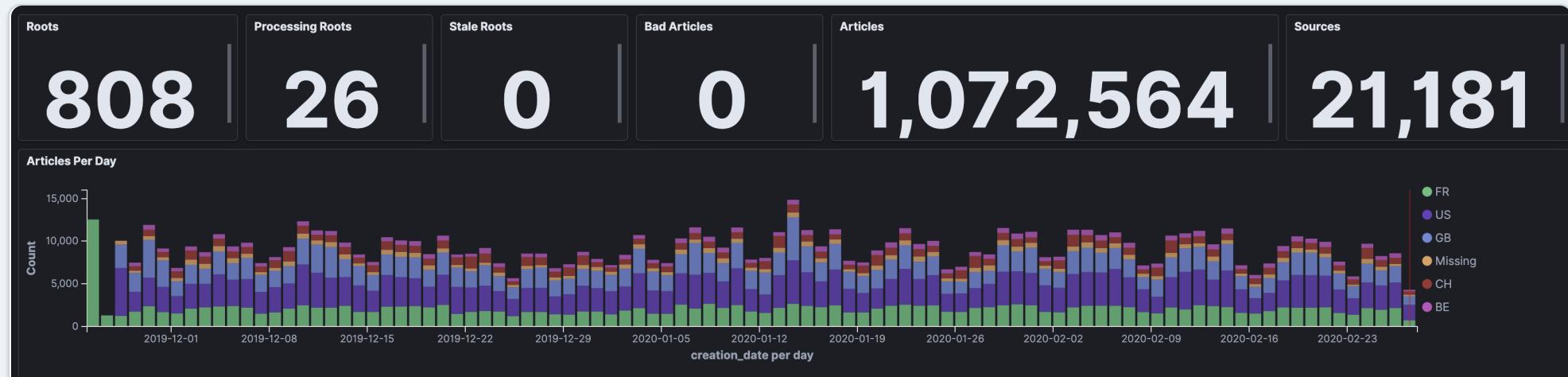
ENRICHED TEXTUAL
CORPUS

DATA ACQUISITION & EXTRACTION •

GENERIC SCALABLE DISTRIBUTED CRAWLER



CRAWLER LIVE DASHBOARD



DATA ACQUISITION & EXTRACTION •

AUTOMATIC CONTENT EXTRACTION

BGR TECH ENTERTAINMENT DEALS BUSINESS SCIENCE LIFESTYLE ABOUT Q

SCIENCE

The Toyko Olympics could actually be canceled due to the coronavirus outbreak



By Mike Wehner @MikeWehner
February 25th, 2020 at 3:12 PM

The coronavirus outbreak originated in China and has since slowly spread across many corners of the globe. Most confirmed cases of COVID-19 infection are isolated to China, but neighboring countries including Japan have seen their fair share as well.

That's bad news, specifically for Japan, as the country has a rather important date fast approaching in the form of the 2020 Summer Olympics which are to be held in Tokyo in just a few months' time. Now, with so much uncertainty surrounding the spread of the virus and the status of vaccines and other preventative measures, Olympics organizers are discussing the very real possibility that the games will be delayed or even canceled.

AMAZON'S TOP DEALS

Today's best deals: Surprise AirPods sale, MacBook Air blowout, Roku Stick+, delicious protein bars, more! >

As AP reports, longstanding International Olympic Committee member Dick Pound revealed in an interview that the group has coronavirus on the mind. At present, it's a matter of deciding how long the group could wait before ultimately deciding whether or not to delay or cancel the games as a whole.

It's a complex issue because the preparation for the games is of course already underway. The lead-up to the start of the games is, in some ways, even more work than the games themselves, and if the risk of virus spread remains high, it could pose a serious threat to public health.

"You could certainly go to two months out if you had to," Pound told AP. "A lot of things have to start happening. You've got to start ramping up your security, your food, the Olympic Village, the hotels, The media folks will be there building their studios."

Nobody wants to see what would happen if a deadly virus began spreading throughout the Olympic Village, infecting world-class athletes and completely ruining the games themselves. There's not much time left before the powers that be have to start making some very hard decisions.

As far as what would happen if the risk of virus spread began to seriously impact the planning and preparation of the games, and threatened to impact the games themselves? "You're probably looking at a cancellation," Pound told AP.

Image Source: DAYAKANT SOLANKI/EP/ET/SHUTTERSTOCK
Topic: coronavirus, covid-19, olympics, tokyo 2020

Mike Wehner has reported on technology and video games for the past decade, covering breaking news and trends in VR, wearables, smartphones, and future tech. Most recently, Mike served as Tech Editor at The Daily Dot, and has been featured in USA Today, Time.com, and countless other web and print outlets. His love of reporting is second only to his gaming addiction.

TITLE
The Toyko Olympics could actually be canceled due to the coronavirus outbreak

AUTHORS
Mike Wehner

PUBLICATION DATE
February 25th, 2020 at 3:12 PM

TEXT
The coronavirus outbreak originated in China and has since slowly spread across many corners of the globe. Most confirmed cases of COVID-19 infection are isolated to China, but neighboring countries including Japan have seen their fair share as well...

MEDIA





TECHNOLOGY OVERVIEW

API

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS

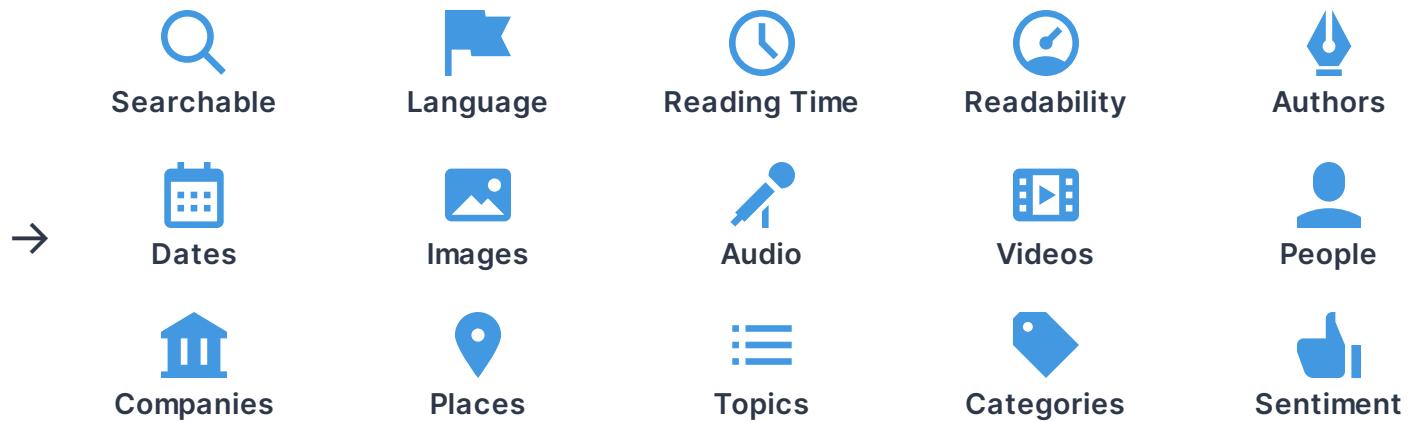


ENRICHED TEXTUAL CORPUS

WE COMBINE TECHNOLOGY MODULES



ENRICHED DOCUMENT





ENRICHED TEXTUAL CORPUS ••

SOURCES



News / Blogs



Social Media Posts



Transcripts / Books



News History

CURRENT NUMBERS



1 million documents



20 000 sources



TECHNOLOGY OVERVIEW

API

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS



API EXAMPLE

```
POST https://api.knowledia.com/v1/analyze HTTP/1.1
content-type: application/json

{
  "text": "Tim Cook is the ceo of apple.",
  "settings": {
    "features": [
      "entity_recognizer"
    ]
  }
}
```

```
10   "text": "Tim Cook is the ceo of apple.",
11   "language": "en",
12   "entities": [
13     {
14       "salience": 1.0,
15       "type": "organization",
16       "mentions": [
17         {
18           "text": "apple",
19           "start": 23,
20           "end": 28,
21           "start_index": 6,
22           "end_index": 7,
23           "score": 0.8699771167048055
24         }
25       ],
26       "metadata": {
27         "id": "LA",
28         "labels": {
29           "en": "Apple Inc.",
30           "fr": "Apple"
31         }
32       }
33     },
34   ],
35 }
```



USE CASES

API ACCESS

NEWS

FINANCE

MARKETING

API

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS



USE CASES

API ACCESS

NEWS

FINANCE

MARKETING

API

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS

AVAILABLE TODAY <https://news.knowledia.com>



psg



RELATED TOPICS



Paris Saint-Germain F.C.



Neymar



Thomas Tuchel



Kylian Mbappé



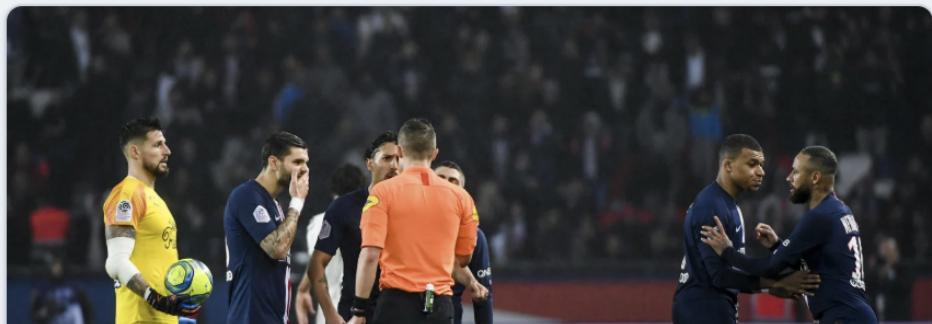
Leonardo Araújo



TRENDS



MORE THAN 10,000 ARTICLES (0.02 SECONDS)



SOCIAL NETWORKS INTEGRATION

SOCIAL NETWORKS

 Barack Obama 



A photograph showing four men sitting on stools on a stage. From left to right: a Black man in a grey suit, a white man in a dark jacket, a Black man in a plaid shirt, and Barack Obama in a dark suit. They are all smiling and appear to be engaged in a conversation or panel discussion.



A close-up portrait of Michelle Obama and Barack Obama. Michelle is on the left, smiling broadly, and Barack is on the right, looking towards her. They are indoors, with a painting visible in the background.

Tweets by @BarackObama

 Barack Obama Retweeted 

 **The Obama Foundation** 
@ObamaFoundation

The Obama Presidential Center will make Jackson Park a part of presidential history. In 1893, Frederick Douglass made it a part of Black history. For [#BlackHistoryMonth](#), we explore his visit to Chicago's World's Fair—and his call to celebrate the contributions of Black Americans.



Aerial photograph of Chicago, showing the city skyline, a river, and green spaces. A large blue play button icon is overlaid in the center of the image.

  Feb 25, 2020

 **Barack Obama** 
@BarackObama

After a lifetime of reaching for the stars, today, Katherine Johnson landed among them. She spent decades as a hidden figure, breaking barriers behind

Tweets liked by @BarackObama

 **Cam Kasky** 
@cameron_kasky

Goodnight everybody. Here's my older brother and I at an awesome [@BarackObama](#) rally we was 7 or 8.

We were so lucky to be inspired by him in pe [#NeverAgain #MarchForOurLives](#)






  F

 **Hamilton** 
@HamiltonMusical

One last time. [twitter.com/barackobama/st...](https://twitter.com/barackobama/status/6869926208)
<https://twitter.com/BarackObama/status/6869926208>

  J

KNOWLEDGE GRAPH INTEGRATION

SPOUSE



Michelle Obama

START TIME
OCTOBER 3, 1992

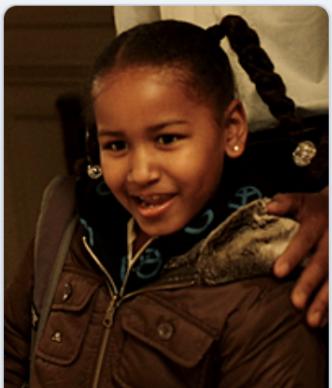
PLACE OF MARRIAGE

 TRINITY UNITED CH...

CHILD



Malia Obama



Sasha Obama

NUMBER OF CHILDREN

2

CHILD

-  SASHA OBAMA
-  MALIA OBAMA



RELATIVE

Stanley Armour
Dunham

TYPE OF KINSHIP
 GRANDFATHER



TYP
STE

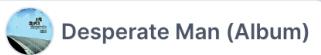
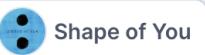
RICH EXPERIENCE

**Ed Sheeran**

English singer-songwriter, record producer, guitarist and actor



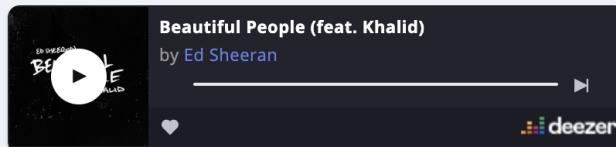
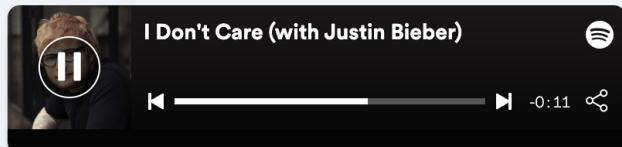
RELATED TOPICS



TRENDS



LISTEN



HOT TOPICS



Centers for Disease Control and Prevention



Ash Wednesday



Carnival Holiday



Pandemic

Ma

RELATED TOPICS



wuhan

RELATED TOPICS



Wuhan



Severe Acute Respiratory Syndrome



People's Republic of China



World Health Organization



Quarantine



Coronavirus



Centers for Disease

INSIGHTS



coronavirus



RELATED TOPICS



Epidemic



Wuhan



People's Republic of China

TRENDS



coronavirus



RELATED TOPICS



Wuhan



People's Republic of China



World Health Organization

TRENDS



wuhan



RELATED TOPICS



Wuhan



Severe Acute Respiratory Syndrome



TRENDS



TOWARDS AUGMENTED NEWS



Mentionned Entities Relations Graph

Source Political Orientation & Trustability

Deep Fake Detection (text, image, video, audio)

Events Graph

Wide Spectrum of Similar Articles (Leaving the Filter Bubble)

Related Fact Checking Articles

Related Social Media Posts

Articles Recommendation



USE CASES

API ACCESS

NEWS

FINANCE

MARKETING

API

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS

REALTIME ALERTING ON RISKS & OPPORTUNITIES ON YOUR PORTFOLIO BEYOND KEYWORDS SEARCH

- 1  Samsung Electronics is part of your portfolio
- 2 An article about potential threats in tantalum mining is posted
- 3 Exploiting the knowledge graph we find the following relationship:
 Samsung Electronics →  Mobile Phone →  Tantalum
- 4 We push you the article explaining why it is relevant to your investments



USE CASES

API ACCESS

NEWS

FINANCE

MARKETING

API

KNOWLEDGE GRAPH

NATURAL LANGUAGE
PROCESSING

DATA ACQUISITION &
EXTRACTION

ENRICHED TEXTUAL
CORPUS

ENRICHED CORPUS AS A REAL-TIME SOURCE OF BUSINESS INSIGHT



- Trends Discovery**
- Competitive Intelligence**
- Brand Monitoring**
- Products / Services / Campaigns Monitoring**



TECHNICAL STACK



GitHub



Python



Google Cloud Platform



Tensorflow



Elasticsearch



Kibana



Dgraph



Redis



Docker



Kubernetes



Apache Airflow



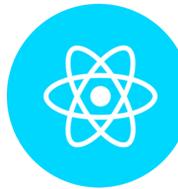
Apache Spark



TypeScript



Next.js



React