

How Artificial Intelligence and GraphQL Powers Traversals' Federated Search

Getting Closer to Make Data Speak Your Language



© 2019 by Traversals™ Analytics and Intelligence GmbH. All rights reserved. The information herein is proprietary and confidential and should not be distributed without the prior written approval of Traversals™ Analytics and Intelligence GmbH. Traversals™ is a registered trademark of Traversals Analytics and Intelligence GmbH.

Introduction **Status of Artificial Intelligence and Machine Learning**

Artificial Intelligence (AI), Machine Learning and Deep Learning are topics of great interest to any company today. The big breakthrough of these technologies came with the insight that high performance graphics cards can be used to train models and that large companies like Google are sitting on an incredible amount of data being ideal for machine learning.

For a layman, executive or CEO, however, it is becoming increasingly difficult to recognize and interpret the technical differences between various AI solutions. Executives want to know whether a technological or algorithmic approach will improve the business, provide a better customer experience and achieve operational efficiencies such as speed, cost savings and greater accuracy.

This white paper uses the Federated Search, a core feature of Traversals' [Data Fusion Platform \(DFP\)](#), to show where machine learning has proven to be useful.

Technical Overview

Need for a Federated Search

The goal of a data fusion platform is to standardize data and generate added value by processing and correlating it. It is a great advantage if data is kept in one place. This makes it easier to implement security aspects, introduce version control and integrate other services, such as processing modules, more quickly. Ideally, only one source of truth exists after deploying a data fusion system.



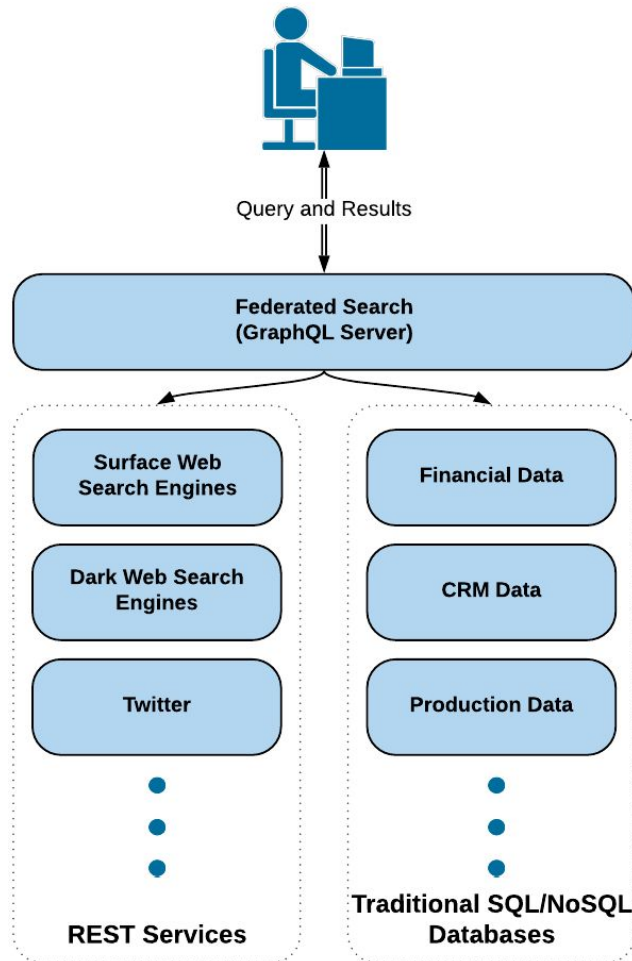
© 2019 by Traversals™ Analytics and Intelligence GmbH. All rights reserved. The information herein is proprietary and confidential and should not be distributed without the prior written approval of Traversals™ Analytics and Intelligence GmbH. Traversals™ is a registered trademark of Traversals Analytics and Intelligence GmbH.

In reality, however, data fusion systems are integrated into existing infrastructures where a landscape of many different databases exists. It has been shown that both technical and interpersonal reasons exist that prevent the consolidation of such structures.

What makes it even more difficult is that multilingual data is often available in multinational organisations, which cannot be interpreted due to a lack of language skills.

Federated Search as Smart Data Virtualization

The Federated Search was designed as a core element of DFP. It acts as a connecting element between the world of data silos and the idea of having a single source of truth. The Federated Search acts like a smart data virtualization layer. By using a fully customized GraphQL compiler, any data source such as external REST services or SQL/NoSQL databases can be addressed by DFP. The compiler offers full functionality for CRUD operations, graph traversals and extensive filtering such as fuzzy or spatial searches. GraphQL ensures that the schema and semantics are always the same for the client independent to the data source used on server side. A field always indicates where the data comes from.



Block diagram showing the Federated Search with integrated data sources.

Interesting search results can be manually or automatically imported into the DFP integrated repository, leveraging enterprise functions such as additional automatic processing or permission control. By using Federated Search, organizations can easily and quickly find and persist critical information that would otherwise remain hidden.

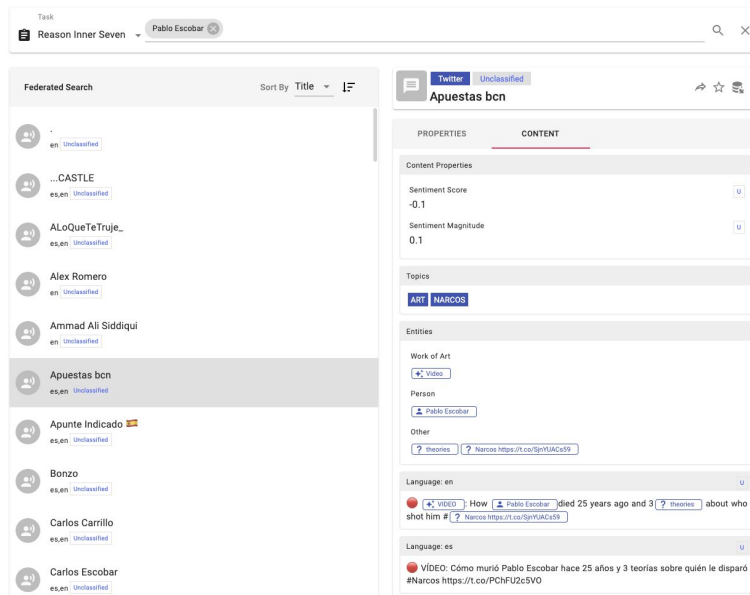


Figure showing the Federated Search with real search results. In this example the Twitter REST API was integrated as data source. The figure shows one selected Tweet being extended with NLP, such as machine translation, sentiment analysis and text categorization.

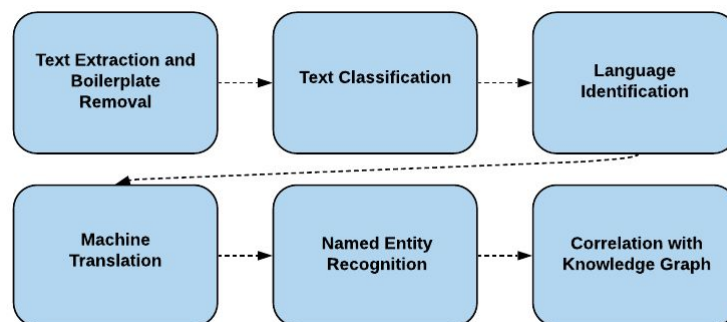
Consequences of a Federated Search

The Federated Search extends the DFP core and enables a fast and easy integration of new data sources. But it also increases the 5 Big Data properties Volume, Velocity, Variety, Veracity and Value that an analyst now has to deal with.

Natural Language Processing within Federated Search

In order to cope with the increased volume of information, the Federated Search was expanded to include elements from AI. Natural Language Processing within the Federated Search transforms unstructured text such as search results from Twitter or even PDF documents into a structured, language-independent representation.

All search results are routed through an NLP pipeline before being sent to the client.



Extension of the Federated Search result feed with an AI powered NLP pipeline.

1. **Text extraction:** Texts, which originate from web pages, are cleaned up, so that only the pure text remains.
2. **Text classification:** The text type, such as prose, short message, log files, are classified.
3. **Language identification:** The national language of the text is recognized.

4. **Machine translation:** If possible, the text will be translated into different languages.
5. **Named entity recognition:** Within the text, different entities, such as person, organization or location, are recognized.
6. **Correlation with knowledge graph:** Recognized entities are compared with the knowledge from the Knowledge Graph.

The shown pipeline is a simplified version of the pipeline used in production. The production version contains even more sub modules and various exit points to improve the performance and quality. Parallel processing is applied to decrease the latency introduced by the NLP.

The pipeline helps to put information into context, to structure results and to uncover connections. Analysts benefit from the pre-evaluation and thus achieve faster and better results.

Summary and Future

More AI and Machine Learning at Traversals

The example of Federated Search showed one use case in which AI-based methods are used at Traversals.

In practice, AI is used in many places. For example, ongoing research projects are trying to use AI to detect anomalies in data or to apply semantic processing to improve the Federated Search itself.

Traversals maintains an intensive partnership with research partners to deploy new AI technologies and improve customer experience.



© 2019 by Traversals™ Analytics and Intelligence GmbH. All rights reserved. The information herein is proprietary and confidential and should not be distributed without the prior written approval of Traversals™ Analytics and Intelligence GmbH. Traversals™ is a registered trademark of Traversals Analytics and Intelligence GmbH.

Make Data Speak Your Language

Traversals is an analysis and intelligence company based in Erlangen/Germany. It offers software products and related services for companies and authorities. Traversals plays a decisive role in making data understandable for people who are not data scientists.

For More Information

If you have questions or would like to discuss this data sheet, please contact us. As an advocate of innovative IT solutions, we are committed to keeping a dialogue open on technologies, processes and best practices that will help us to support our customers.

Contact

Web: www.traversals.com

Phone: +49 9131 92790 0

Fax: +49 9131 92790 99

Email: info@traversals.com



© 2019 by Traversals™ Analytics and Intelligence GmbH. All rights reserved. The information herein is proprietary and confidential and should not be distributed without the prior written approval of Traversals™ Analytics and Intelligence GmbH. Traversals™ is a registered trademark of Traversals Analytics and Intelligence GmbH.



© 2019 by Traversals™ Analytics and Intelligence GmbH. All rights reserved. The information herein is proprietary and confidential and should not be distributed without the prior written approval of Traversals™ Analytics and Intelligence GmbH. Traversals™ is a registered trademark of Traversals Analytics and Intelligence GmbH.