ontotext

What do David Cameron, Pedro Almodovar and Leo Messi have in common? No, the Argentinian footballer doesn't star in the Spanish director's latest movie. Neither does the UK prime minister. These three people – alongside thousands of other rich and powerful celebrities, business executives and politicians – have been linked to companies in the Panama Papers leak in recent weeks.

Listen to our webinar recording: Diving in Panama Papers and Open Data to Discover Emerging News to see how to empower your data analytics with Open Data.

'The Biggest Leak in History'

When the news of the 2.6TB of data on shell companies broke in early April, it immediately became viral and has been trending ever since. Revenue agencies and government officials around the world pledged to fight tax avoidance in tax havens, which, though not illegal, are the secret coffers the rich and powerful one-percenters have been using to reduce their tax rates.

A month later, on May 9th, the International Consortium of Investigative Journalists (ICIJ), which broke the news, released a searchable database of more than 300,000 entities from the Panama Papers and Offshore Leaks investigations.

The names of David Cameron and Lionel Messi do not appear in the Panama Papers. In the wake of the leak, though, Cameron admitted that before becoming prime minister in 2010, he had owned shares in a tax-haven fund set up by his late father.

On the other hand, Messi is believed to have avoided taxes via the company Mega Star Enterprises, which he reportedly owns together with his father Jorge Horacio Messi. And, finally, Almodovar said at the Cannes Film Festival that he was one of the least important names cited in the Panama Papers.

Panama Papers Dataset Enriched by Linked Data Portal

For two months now journalists and the general public have been wondering who's also in the Panama Papers and which shareholders are connected with which corporations in which

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

countries. A simple search of a single name or organization in a database, however, may prove tedious and enormously time-consuming.

Using the ICIJ database content and other open data sources, we, at Ontotext, created the Linked Leaks linked data Knowledge Graph database of the Panama Papers. Thus, the linked data project comes into play to enrich the data with semantics, link the dataset to other Linked Open Datasets, and provide richer findings while searching through the Panama Papers.

The Knowledge Graph portal also encourages data analytics enthusiasts, journalists and developers to dive into and dig for additional information in the Panama Papers.

Playing with Linked Leaks allows for various types of analytics queries to discover relationships between companies, shareholders, countries and chains of control. The Linked Leaks demonstration service gives an all-new perspective of the Panama Papers, linking the leaked data to open-data information about countries and geographical regions. Share on X

Linked Leaks, which contain more than 22 million RDF statements, also serve as a kind of 'Investigative Reporting Workbench', allowing for asking smart questions in SPARQL and showcasing the role of Linked Data in Investigative data journalism. Analytics enthusiasts can also freely download the Linked Leaks data in RDF for on-premise analytics and for building applications using the data.

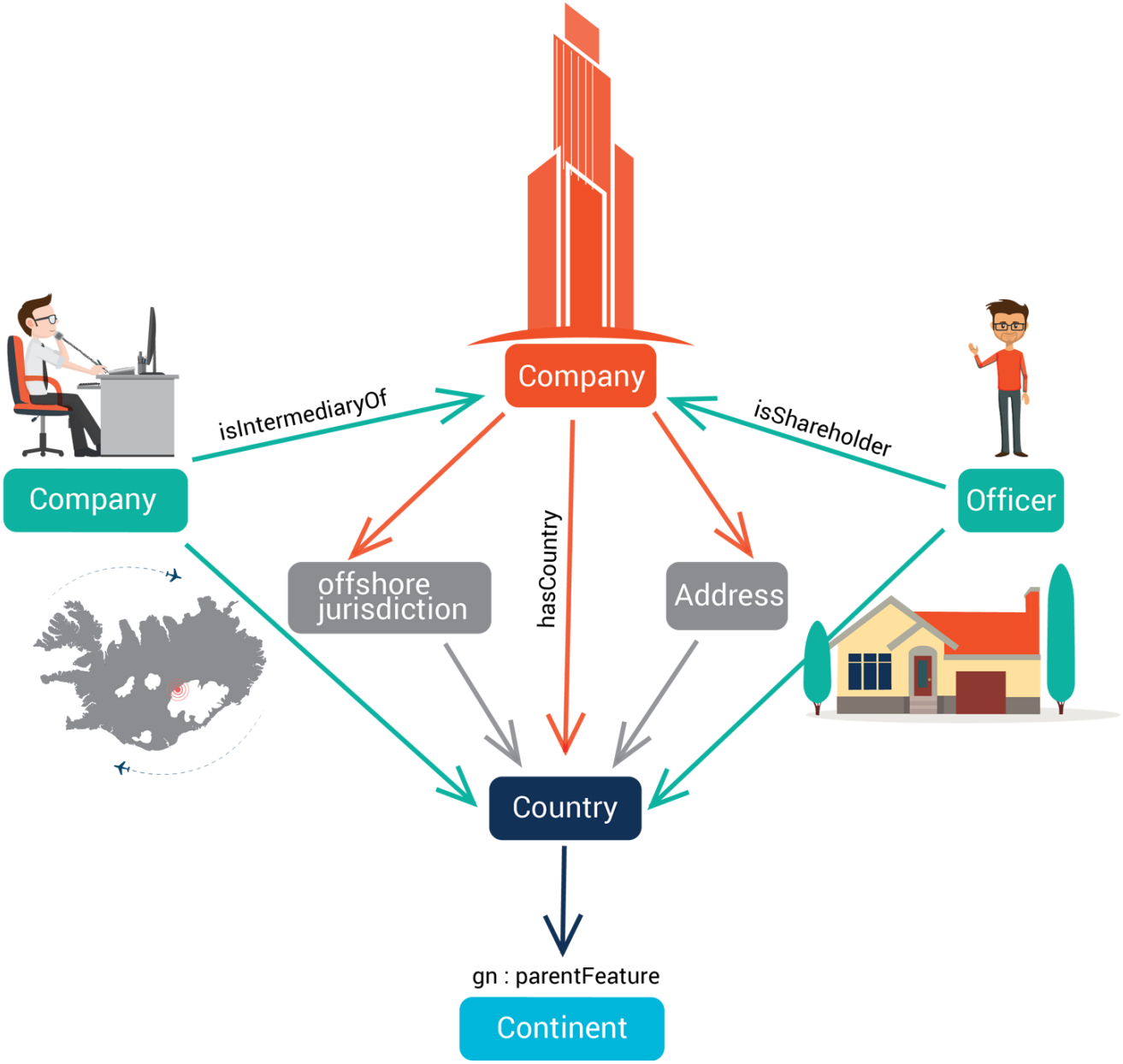
Putting the Panama Papers into Context

The Linked Leaks Knowledge Graph, published according to the Linked Open Data principles, has already been developed to link the Panama Papers to information on countries and geographical regions from the DBpedia and GeoNames resources, and links to more datasets will be added.

These datasets help all sorts of discovery and analytics queries. For example: companies related to a given shareholder (person or organization), including control relationships; companies that control other companies in the same country, through a company in an offshore zone; or most popular offshore jurisdictions.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext



‘The Game of Queries’ in Linked Leaks

By asking smart questions in SPARQL in Linked Leaks, everyone can get richer findings to their investigative search of the Panama Papers.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext

Now let's take a look at a few sample queries.

- Here's Query #1: Countries by the number of entities related to them. This query shows that Hong Kong is number one and the British Virgin Islands – number two.
- On to Q2, query #2: Country pairs by ownership statistics to search and answer the question 'Owners from which countries most often own entities in which other countries?'. The results are: the number-one pair is China-Hong Kong, followed by Hong Kong-British Virgin Islands and Taiwan-Samoa.
- Running Q3, proposed by Ivo Velitchkov, shows statistics by year of incorporation. The year 2007 was the year when the highest number of offshore companies was set up, with 26,774 numbers of incorporations.
- Next, Q4 is about officers and entities by the number of capital relations.
- While Q5 is run to show Countries in Eastern Europe by the number of owners and uses DBpedia and GeoNames resources. This query benefits from `sameAs` mapping to DBpedia and GeoNames and the basic information from those resources about countries, loaded in graph leaks. The query shows that Russia has more owners of offshore entities than all other countries in Eastern Europe combined.

As you can see, many sorts of interlinked cross-queries can be asked in the Linked Leaks graph database. Ontotext is just starting to explore the possibilities and opportunities of asking smart questions about the Panama Papers and is working to further enrich the Linked Leaks with new relations, additional mappings and new sample queries to fine-tune the raw data interpretation and analysis.

We at Ontotext also plan to map this data to the Financial Industry Business Ontology (FIBO), so that one can query and analyze the data using its semantics.

Follow #LinkedLeaks @Twitter and post your #LinkedLeaks questions and queries!

Dive in our webinar: Diving in Panama Papers and Open Data to Discover Emerging News to see a live demo of news and open data analytics!

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com