How the ICIJ unraveled the Panama Papers using Linkurious Enterprise

ICIJ

Customer

The International Consortium of Investigative Journalists (ICIJ)

Country

United States of America

Industry

Non-profit news organization

Challenge

ICIJ's 20-person team had to analyze a 2.6 terabyte data leak to unveil shady schemes in international finance with limited technical resources.

Solution

The team implemented Linkurious Enterprise to visualize and analyze connections between elements of the Panama Papers leaks. 370 journalists were able to browse through massive amounts of data to find stories of fraud, corruption, and other wrongdoing.

Benefits

- Ability to unveil complex networks of multiple entities.
- Ability to search and visualize networks of geographical data.
- Higher level of autonomy for journalists who collaborate in real time, in a secure way, from all around the world.
- Publication-ready visualizations for media outlets.

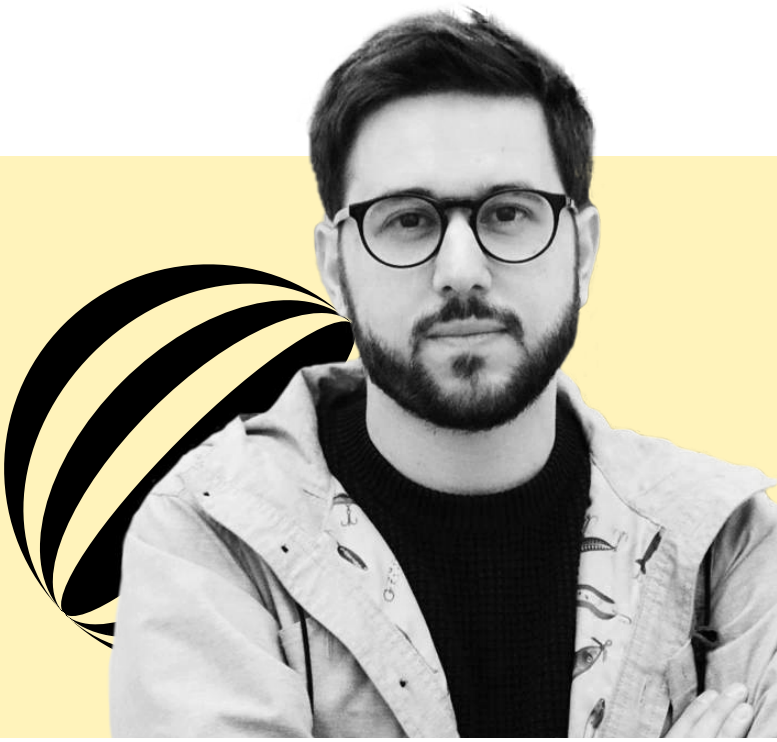
The International Consortium of Investigative Journalists (ICIJ) is a US-based non-profit news organization. It encompasses a global network of more than 200 investigative journalists and 100 media organizations in over 80 countries, working together to investigate major news stories.

Within its reporting team, the Data & Research unit is in charge of processing data gathered by journalists or coming from leaks. Over the years, the ICIJ investigation teams have exposed smuggling, tax evasion, and corruption cases. The Panama Papers investigation received the Pulitzer prize in 2017.

In 2014, an anonymous source leaked data from a shady legal firm based in Panama to the German newspaper Süddeutsche Zeitung. The leak contained more than 11.5 million documents, representing 2.6 terabytes of data and documenting 214,488 offshore structures created and administered by Mossack Fonseca over 30 years. Quickly, the newspaper approached the ICIJ which coordinated with media organizations and journalists worldwide to start investigating the leak. →

"Graph visualization technologies like Linkurious are a great asset. It's intuitive for the non-tech-savvy. Reporters just need to click on dots to expand the connections and uncover persons of interest and potential stories in a short time-frame."

- Pierre Romera
Chief Technical Officer, ICIJ



LINKURIOUS

The complexity of a data-driven journalism investigation

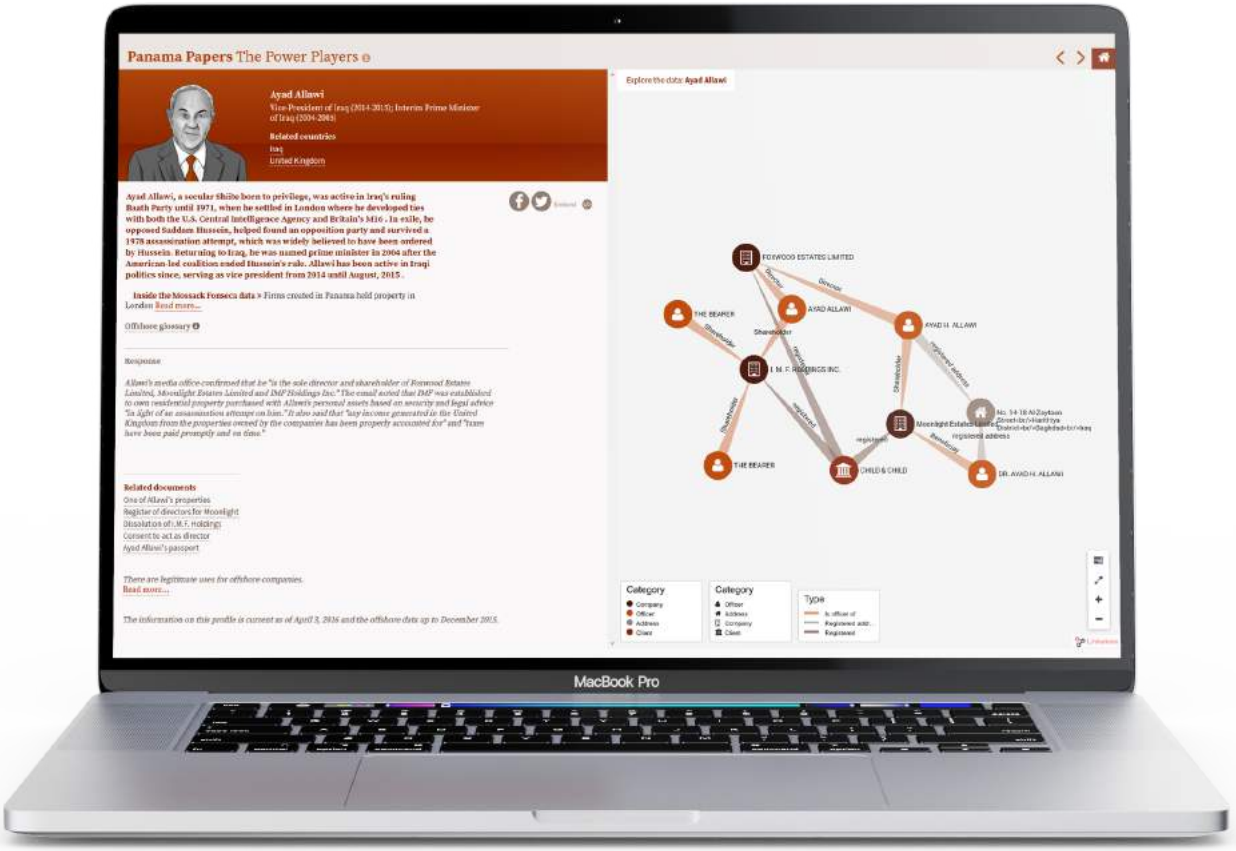
Such large data leaks are not commonplace for journalists and the nature and volume of the data create specific challenges.

The large amounts of data contained in the leak was a challenge in itself. To complicate things, the Panama Papers investigation started from raw, unstructured data. The files obtained by Süddeutsche Zeitung included millions of loan agreements, financial statements, emails, trust deeds, and other paperwork dating back nearly 50 years. To start investigating, journalists first had to turn millions of unstructured documents into computable information that could be organized, searched, and analyzed.

Another obstacle was the way the data was stored. The success of investigations is determined by the ability to find connections between entities. But in many investigation cases, data is kept in silos that make it difficult to cross-reference and highlight connections. For the Panama Papers, ICIJ's reporters conducted the investigation with data stemming not only from the leak but also from other sources. To make siloed data talk, it was essential to bring everything together.

"One of the key challenges is to make our technology user-friendly for the journalists so that everyone around the world is able to use it," explained Pierre Romera, ICIJ's Chief Technology Officer. Indeed, data-driven investigations represent an obstacle for investigators without technical knowledge. At ICIJ, making data exploration accessible to non tech-savvy reporters was both a challenge and a necessity. →

"One of the key challenges is to make our technology user-friendly for the journalists so that everyone around the world is able to use it."



LINKURIOUS

A graph approach to unveil shady connections

For some years now, the ICIJ Data Unit has understood that investigations of corruption and fraud issues were inherently graph problems and has technically treated them as such.

“While working on Offshore Leaks, I learned how important graph analysis is when investigating corruption,” said Mar Cabra, editor of the unit. “Connections are key to understanding what the real story is and who’s doing business with who. We decided early on that we needed to use a graph-based approach for the Panama Papers.”

The ICIJ decided to build the Panama Papers investigation platform using graph technologies. They proceeded in several phases to make the documents usable for the 370 journalists.

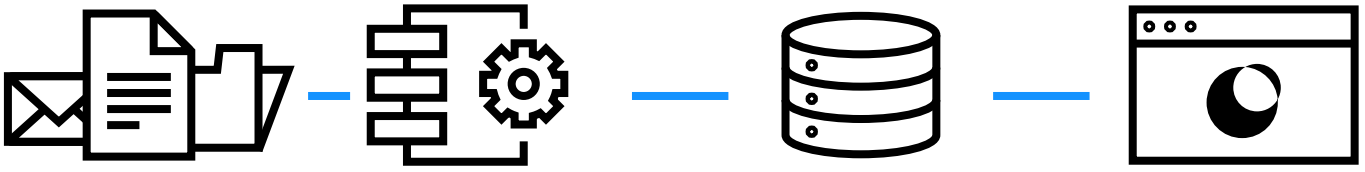
The Data & Research unit processed the documents into machine-readable formats, indexing and connecting them together through their metadata thanks to Optical character recognition, content-extraction, and document indexing technologies. The team then made use of Talend ETL (Extract, Transform, Load) tools to load the data into Neo4j, a graph database platform, creating a network of nodes and edges.

“Connections are key to understanding what the real story is and who’s doing business with who. We decided early on that we needed to use a graph-based approach for the Panama Papers.”

Once the database was created, the ICIJ made the data available to their investigative teams globally through the secure, collaborative, and intuitive exploration interface of Linkurious Enterprise. The platform gave journalists the visual investigation and analysis tool they needed to find insights in the data leak and share their stories.

“Linkurious Enterprise allowed our remote team of dozens of reporters to easily sift through complex financial data to uncover persons of interest and potential stories — all in a visual and very intuitive way,” Cabra explained. →

The technology stack set up by the ICIJ Data Unit to investigate the Panama Papers



Raw data

13,4 million files, including loan agreements, financial statements, emails, are collected.

Data processing

ICIJ Data & Research unit processes the documents into a machine readable format with OCR technologies.

Graph databse

The data is stored and indexed as a graph in the Neo4j platform.

Investigation

The data is accessible to reporters as network visualizations in Linkurious Enterprise platform.

LINKURIOUS

Shedding light on the offshore world

Using Linkurious Enterprise, journalists were able to explore the connections of the entities they were investigating and expose the complex financial structures mounted by Mossack Fonseca's clients to cover up their financial crimes.

With limited resources, the ICIJ Data Unit was able to organize an efficient and reproducible process to allow a network of 370 journalists to investigate millions of documents and uncover stories of interest. Breakthrough revelations were made possible by Linkurious Enterprise.

"Linkurious Enterprise was the best solution to make the data available to our global network of more than 370 journalists."

Linkurious Enterprise also provided security features particularly important for journalists working in countries where reporting on a political scandal could have dire consequences.

When the time came to publish the story, the ICIJ used the Linkurious Enterprise API to create and embed graph visualizations into its news platform. This way, readers from all around the world could quickly grasp the complexities of the financial networks unmasked by ICIJ.

"Linkurious Enterprise was the best solution to make the data available to our global network of more than 370 journalists. It enabled them to uncover persons of interest and potential stories in a short timeframe. Using the collaboration features and intuitive interface of Linkurious, they were able to easily expose many complex offshore structures and financial crimes," Mar Cabra, Data Editor at ICIJ explains.

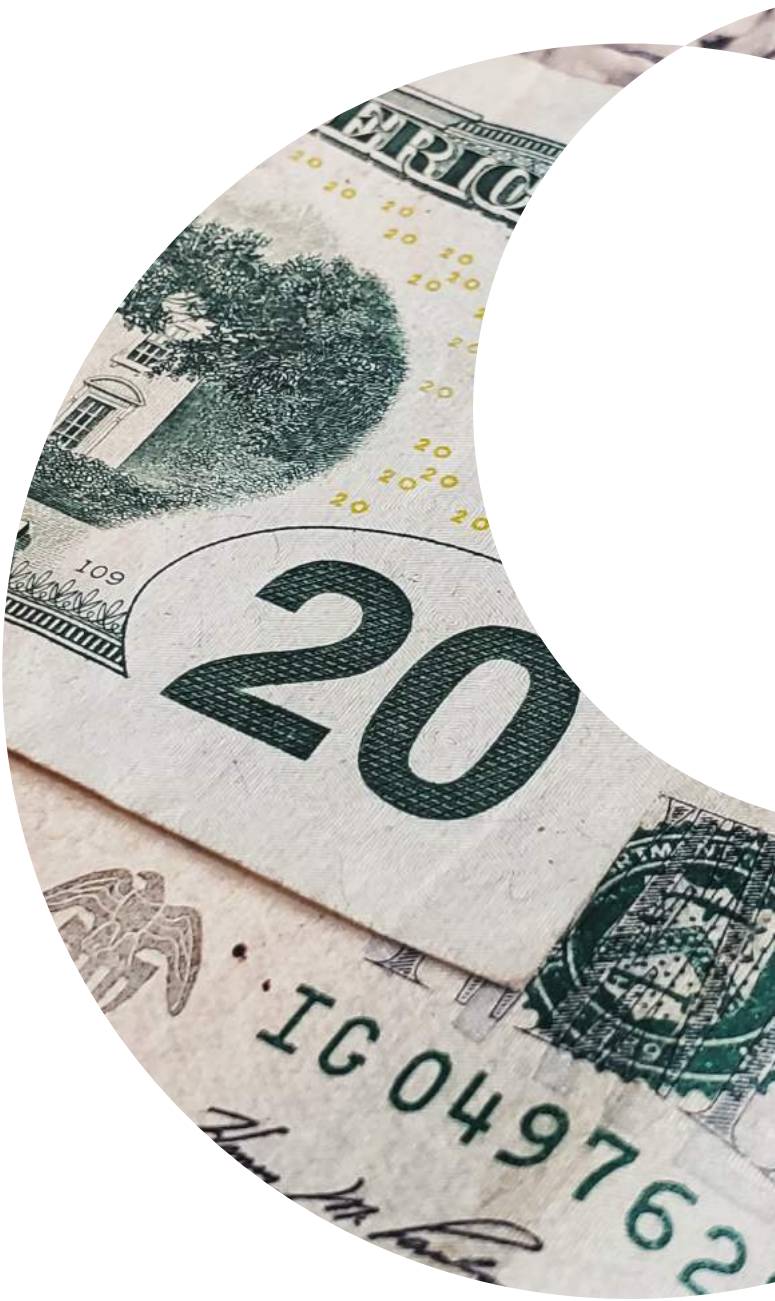


Photo by Jorge Salvador.

About us

Linkurious offers an intuitive financial crime investigation platform that enables swift and precise detection and investigation of even the most sophisticated criminal networks for optimal risk mitigation and protection. More than 3000 investigators in Global 2000 companies, public organizations, governments, and NGOs use Linkurious to harness their complex connected data and uncover fraud and money laundering schemes that would otherwise remain hidden.

Contact us

66 rue Marceau, 93100 Montreuil

www.linkurious.com

contact@linkurious.com

LINKURIOUS