

Mining Leaks and Open Data to Follow the Money

Friedrich Lindenberg
OCCRP

Abstract

How can data-driven approaches help to uncover large-scale corruption in government and business? The Organized Crime and Corruption Reporting Project (OCCRP) is a network of investigative reporters across 45 countries that uncovers cases of bribery, theft and money laundering around the world.

To support this, OCCRP has built a unique data resource covering more than a billion entities from over 400 data sources, and a suite of open-source data integration and search tools, the Aleph. This allows us to give investigative reporters visibility into large amounts of evidence, and to perform cross-referencing between databases that uncovers evidence of wrong-doing. We'll present the why, what and how of this project, and are hoping for feedback from the IR community on what our next steps could be in order to increase search quality and provide better recommendations to our investigators.

Biography

Friedrich Lindenberg leads the data team at OCCRP. He is responsible for the development of OCCRP Data and supports ongoing investigations where data analysis is needed. In 2014/2015, Friedrich was a fellow with the International Center for Journalists, working with the African Network of Centers for Investigative Reporting (ANCIR), and in 2013 he was a Knight-Mozilla OpenNews fellow at Spiegel Online in Hamburg. Prior to that, Friedrich was an open data activist, and worked to promote the release of government information about public finance, lobbying, procurement and lawmaking across the world.

Copyright © 2019 for the individual papers by the papers' authors. Copying permitted for private and academic purposes. This volume is published and copyrighted by its editors.

In: A. Aker, D. Albakour, A. Barrón-Cedeño, S. Dori-Hacohen, M. Martinez, J. Stray, S. Tippmann (eds.): Proceedings of the NewsIR'19 Workshop at SIGIR, Paris, France, 25-July-2019, published at <http://ceur-ws.org>