

# From the Panama Papers to an intelligence service for your own business

November 7 2016

---

In 2016 more than a hundred newspapers and others published revelations on tax avoidance and evasion. They were based on the Panama Papers, a collection of data that comprises 2.6 terabytes of information and 11.5 million documents. In 2015 this was leaked to the Sueddeutsche Zeitung; an international group of journalists evaluated it over the course of a year. Now, computer scientists at the Max Planck spin-off Ambiverse have analyzed the data with software in a few hours, obtaining new results.

The software is intended to help businesses analyze large amounts of text automatically. While the international group of journalists analyzed the Panama Papers in depth and focused on people like Nawaz Sharif, prime minister of Pakistan, or Petro Poroshenko, president of Ukraine, the young entrepreneurs used their software to take a broader approach. For example, they found that athletes, rather than politicians, were the largest group involved. They accounted for more than 20% of the people identified by the software, followed closely by artists. Only then came the politicians. The capability of the software to create categories automatically delivered further insights. Among the athletes, it was soccer players who were most often involved in strategies for tax avoidance and evasion. Second place was taken by tennis and basketball players (around 10%); hockey and volleyball players were in third place (about 5%). Another insight regarding politicians was found: political ideology did not influence the decision to use an offshore account. Conservative and socialist politicians are equally represented in the Panama Papers.

"It's unfortunate that we could only work with the data provided and already prepared by the International Consortium of Investigative Journalists. Our system isn't merely capable of handling raw data; it actually obtains even better results," explains Johannes Hoffart, whose PhD at the University of Saarland was based on the underlying technology. Together with four other Max Planck researchers, he founded Ambiverse just under a year ago to market the technology. Much as news services maintain armies of analysts to evaluate publicly available texts, businesses can do this with the Ambiverse software in a few minutes.

The software is so powerful because, among other things, it doesn't search by just a term. Instead, for example, a search for "Angela Merkel" also finds texts in which the chancellor is referred to only as "Angie" or "CDU chief." At the same time, the program leaves out all the documents referring to the well-known soccer coach with the same surname. Thus, businesses can find people, places, and products in large amounts of text, even when this is made more difficult by ambiguous terms or abbreviations. The search for categories even makes it possible to search for "financial companies" or "soccer players" without having to specify these more precisely. In addition, the software can be used not only for texts in German and English, but also Spanish and Chinese. This is made possible by a knowledge base developed at the Max Planck Institute for Informatics; its content was developed partly with the help of the online encyclopedia Wikipedia. With their business plan, the founders of Ambiverse won the trans-regional business plan competition "1,2,3, GO" a few weeks ago.

**More information:** Further information: [www.ambiverse.com/oh-wie-schon ... ist-offshore-panama/](http://www.ambiverse.com/oh-wie-schon...ist-offshore-panama/)

Provided by Saarland University

Citation: From the Panama Papers to an intelligence service for your own business (2016, November 7) retrieved 25 April 2024 from <https://phys.org/news/2016-11-panama-papers-intelligence-business.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.