

THE POTENTIAL OF FIGHTING CORRUPTION THROUGH DATA MINING

9 January 2015 in [Access to information](#), [Measurement and Research](#), [People power](#), [Politics and government](#), [Technology](#)

Two centuries ago, coal mining spurred the European continent's Industrial Revolution. Today, data mining is fueling the data revolution brought about by exploding streams of data. Using data mining techniques to profile customer preferences and predict purchasing patterns has become common practice in the private sector. But can data mining also be used to fight corruption? And if so, how?

Last year, [Transparency International Georgia](#) launched an open-source procurement monitoring and analytics portal, which extracts data from the government's central e-procurement website and repackages it into user-friendly formats. Users can now generate profiles of procurement transactions made by government agencies, profiles of companies bidding for public contracts, and search aggregate statistical data on government spending. If citizens suspect law violations in electronic tender processes they can submit an online report which a Dispute Resolution Board reviews within ten working days.



Data mining's potential to spot inadequacies in processes involving elected authorities and public money can be taken even further.

The European Commission, in cooperation with Transparency International, developed [ARACHNE](#) data analytics software that cross-checks data from various public and private institutions and helps to identify projects susceptible to risks of fraud, conflict of interests or irregularities.

Researchers from the [Corruption Research Center Budapest](#) have examined massive data sets of public procurement procedures from European Union countries, searching for abnormal patterns such as exceptionally short bidding periods or unusual outcomes (e.g. no competition for the winning bid, or bids repeatedly won by the same company). Using inferential statistics – analysis that can be done to draw conclusions beyond what the data actually is capturing – they identified corrupt behavior based on deviations from ordinary patterns.

Data mining can also be used to detect tax fraud and improve taxpayer compliance. In the aftermath of the [Luxleaks](#), when a whistleblower released reams of data about tax evasion schemes in Luxembourg, data mining techniques employed by New York City's former finance commissioner David Frankel may provide some inspiration: by "identifying individuals who had [businesses similar to others but who stood out](#) as outliers on taxes paid" the auditing team improved the efficiency of its investigations into companies suspected of underpaying taxes.

Similarly, data mining could be employed to fight money laundering: an algorithm reviewing banking data and comparing it with information about financial criminal data points may for example contribute to revealing illicit financial flows, an issue that ranks high on Transparency International's agenda.

The wealth of data that can nowadays be gathered through remote sensing, crowd-sourced citizen

[LATEST](#) [POPULAR](#) [COMMENTS](#) [TAGS](#)

STATE CAPTURE IN SOUTH AFRICA
14 FEBRUARY 2017



**YOUNG CHANGE MAKERS USING
TECH TO SOLVE LAND
CORRUPTION**

6 FEBRUARY 2017



WHAT'S NEXT FOR UKRAINE?
16 DECEMBER 2016



**CLEANING UP SPORT:
CONFLICTS OF INTEREST AT
THE TOP**

9 DECEMBER 2016



**OGP: FRANCE MUST DO MORE
TO TACKLE CORRUPTION**

9 DECEMBER 2016

SEARCH



CATEGORIES


Categories [Select Category](#)

reports, news media, census data, cell phone activity, and social networking sites etc., combined with traditional indicators, makes for seemingly endless opportunities. Do you want to identify issues of conflict of interest and/or revolving doors? Do you want to know what people are thinking about corruption in a specific country context? Text mining techniques analysing social media noise during a given period of time may provide you with an answer.

There are many ways non-profits and civil society organisations can benefit from data mining on a pro-bono basis. These include hackathons and advocating for the replication of tools and platforms, which not only render data public but make it relatively easy to organise and process.


The European [Citadel on the move](#) website for instance allows users to upload data sets and to create personalised applications, even for people with little experience of data management. One of the most transparent and user-friendly initiatives at the local level is the website [Checkbook NYC 2.0](#), which provides access to New York City government's US\$70 billion annual budget. It details the way money is spent, including specific information on contracts, payments, revenues, budget reports, and audits. It features an application programming interface that lets third-parties choose the data they want and then use it for their own purposes.

Data mining's nimble and purpose-oriented character can do a lot to dispel the fog in which the public sector operates. But more efforts are needed to exploit its potential to the full and make it available to the widest audience.

 You and 52 others like this. 52 people like this. [Sign Up](#) to see what your friends like.

Share and enjoy:



 [ARACHNE](#), [Checkbook NYC 2.0](#), [Citadel on the move](#), [Corruption Research Center Budapest](#), [Data mining](#), [David Frankel](#), [european commission](#), [featured](#), [Luxleaks](#), [New York](#), [open data](#), [Transparency International Georgia](#)

← [Fighting corruption above the Arctic Circle](#) [Greeting 2015: Is there any hope for those fighting corruption in Turkey?](#) →

LATEST NEWS FROM TRANSPARENCY INTERNATIONAL

- [G20, China, corruption and asset recovery](#)
- [Five key lessons from the Rio Olympics](#)
- [The week in corruption: 7 stories making the news](#)
- [Simple ways young people can fight corruption](#)
- [Brasil e os Jogos Olímpicos: três passos para revigorar a luta contra a corrupção](#)

SPACE FOR TRANSPARENCY




This blog by [Transparency International](#) provides an independent and informed viewpoint on corruption. It gives a space to start a worldwide conversation on possible solutions to overcome corruption, and on governance, transparency and accountability.

ARCHIVES

Archives [Select Month](#)

CONTACT

You can email us at press "at" transparency.org.

-  [Connect with us on Facebook](#)
-  [Follow us on Twitter](#)
-  [Add us to your circle on Google +](#)

Feel free to leave a comment on posts that interest you, too.

TI BLOGS

- [Anti-Corruption Research Network](#)
- [TI Georgia](#)
- [TI Hungary's Blog](#)
- [TI Kenya](#)
- [TI Slovakia](#)
- [TI Sri Lanka](#)
- [TI UK](#)