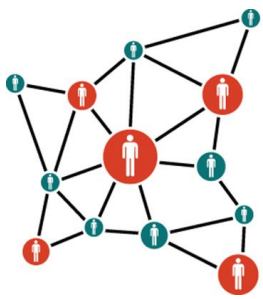
Influence Mapping

State of the Art

December 2016 (v0.3)

iilab



InfluenceMapping

Table of Contents

What is Influence Mapping? ..... p. 2

Current State of the Art ..... p. 4

Case Studies ..... p. 12

)g en^g\^ - Zi
/ i^g $ndZ
,Z &Z[kbjn^]^ ,Z ,hb
Ohebmb\Ze OZkmr 4b f ^
OZkemkZ\d
,h[[rkZ]Zk
#Zk`h`kZ ¢ZI

The Influence Mapping Toolbox ..... p. 32

Contributors: Kat Austen, Lilas Duvernois, Jun Matsushita

Published: 12/2015 (v0.1) - 01/2016 (v0.2) - 12/2016 (v0.3)

This work is published under the Creative Commons Attribution 4.0 International License

What is Influence Mapping?

Influence mapping is an emerging, diverse field. Situated somewhere between data journalism, network visualisation, political science and activism, influence mapping is still building an identity as a practice in its own right. You could be an influence mapper and not even know it yet.

Influence mapping is at a stage where there are many people doing similar work but, because the community is still fragmented, learning is not being shared as effectively or as widely as it could be. To make the most of the excellent work and expertise already in the space, we set out to create a resource where the varied expertise and in-depth knowledge of influence mappers can be shared and built upon: a toolbox that will help newcomers and experts to map influence and make the most of their work.

So, what is influence mapping? Primarily, influence mapping is about making impact. This can be through a diverse range of practices that span from journalism to activism to research. It can be about finding things out - revealing secrets, discovering links. Or it can be about getting the word out, telling a story, or even generating evidence that can be used in a legal process. Or it can be about making tools and databases that make doing all that easier.

Whether it is through identifying key people, exposing or creating transparency in power relationships, or network mapping and visualisation, influence mapping is ultimately a way to hold accountable those in power - a process that allows citizens to have oversight on governments, organisations and corporations.

Thus, we can delineate influence mapping as:

> **Projects using varied methodologies and creating tools that help to find evidence, stories and visualisations that aim to hold those in power accountable for their decisions**

We've used this loose definition to guide the design of the influence mapping toolbox. This online toolbox aims to help people learn about existing influence mapping projects, and to share learning about common practices, tools and project experiences, so that influence mapping as a practice can grow and have even more impact.

Current State of the Art

Here we provide a brief overview of the state of the art in influence mapping, with insights gleaned from our in-depth research into the projects operating within the space. Much of the project content is available in the influence mapping toolbox, along with information and advice on practices and tools.

What do influence mapping projects look like? Many projects have narrative or infovisualisation outputs, or a mix of both, such as Connected China's visualisation walkthrough of China's elite (Figure 1). On the other hand, there are projects that make data more easily accessible by collecting, collating and processing data from multiple sources to create searchable or list-based databases and APIs. These projects, such as LittleSis and Open Corporates, in addition to creating stories and visualisations from their own data, also underpin investigative work by others.

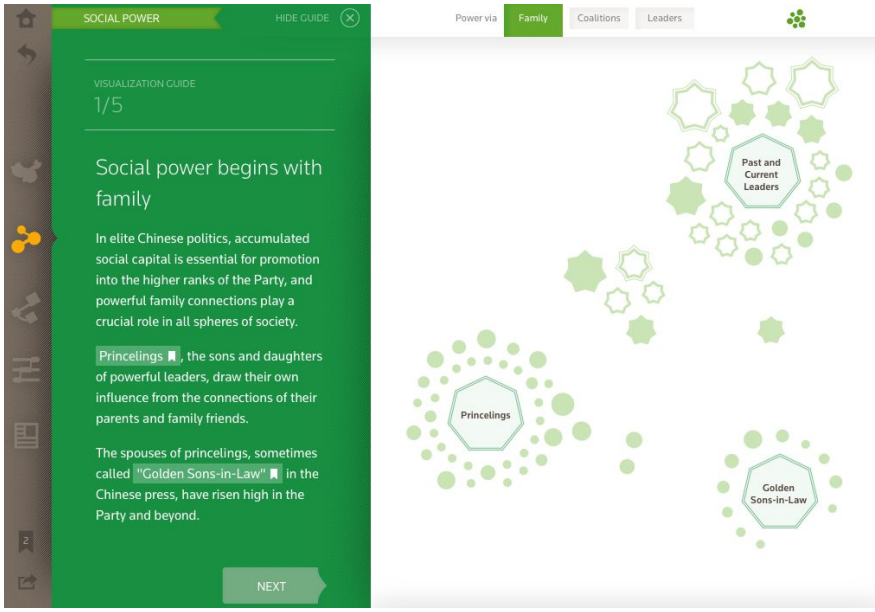


Figure 1: Connected China provides a narrative walkthrough of China's elite and powerful Chinese institutions

Projects within the influence mapping space have significant elements of data journalism and data visualisation, but as a practice as a whole it goes beyond both. While some projects focus on creating traditional journalistic reportage from data investigations, many of the projects we've come across have an element of trying to sustain a practice of shining a light on power structures, trying to create accountability mechanisms that are ongoing. They Rule, for instance, was set up in 2001 by designer Josh On (Figure 2) using static data gathered from company websites. Based on LittleSis data since 2011, the project is still active and used by a global community to explore corporate networks, 14 years later.

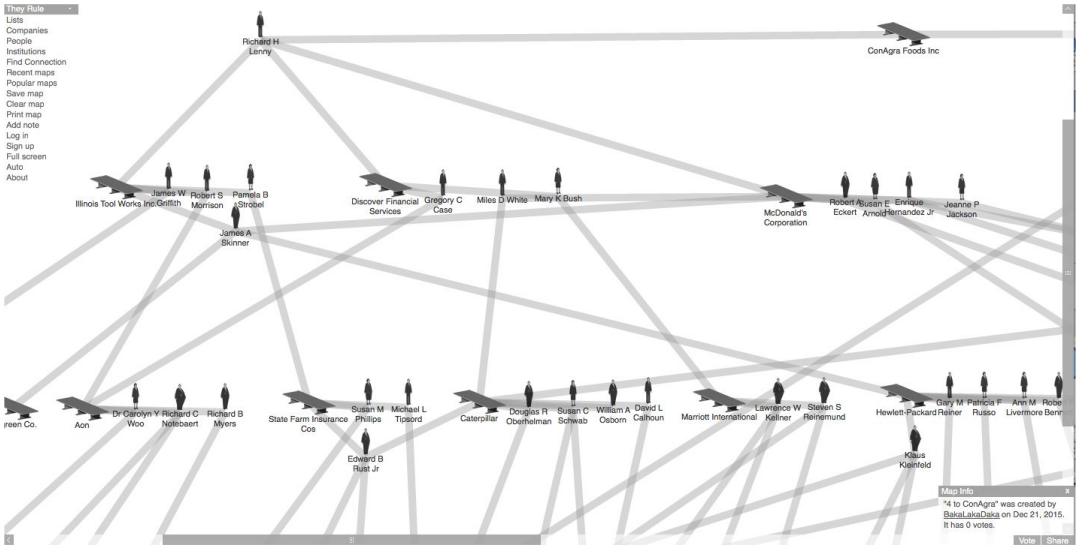


Figure 2: A recent map from the They Rule platform, 4 to ConAgra, a work in progress

Influence mapping projects usually require deep knowledge of a few different fields - in-depth knowledge of data investigative techniques, technology, visualisation, investigative and legal skills can all come into play. Often, influence mapping projects are carried out by teams of between 3 and 10 people who bring in a variety of these skills and more. The Influence Map project, for instance, which measures and lists corporate influence on climate policy, as well as bringing in a developer and a designer, also benefits from the experience of a physicist, a research director and an analyst. It is very rare to see an influence mapping project realised by only one person, something we have seen only in They Rule and Parltrack. And in both of these cases, the realisation of the project benefitted from support of an extensive and diverse network.

Influence mapping projects usually require deep knowledge of data investigative techniques, technology, visualisation, investigative and legal skills

In fact, a supportive community has benefitted many influence mapping projects, whether it is as suppliers of inspiration or information, people to bounce ideas off or to help provide guidance on a bumpy road. One fifth of the projects we've seen have been

published by an established journalistic source (eg: Al Jazeera, ProPublica, The Guardian), which has provided them with a solid infrastructure and the ability to leverage an extant communications network to increase the project's reach, something that is otherwise a challenge for projects going it alone.

Certain organisations are having a huge impact in influence mapping. ICFJ, the Sunlight Foundation and the European Commission have each contributed to funding 5 projects. Journalistic organisation ProPublica has self-commissioned three and ICIJ has self-commissioned two projects. The Knight Foundation has been involved in funding two of the projects. Around one fifth of the projects are wholly or partially self-funded.

Funding is a challenge for many of the projects, even those that, like Influence Map, are established as a business. There is a paradox here, because many of the projects are conceptualised as something having long term impact and a long term model, but they are generally reliant on short-term funding options such as grants of a few years. This means that often the projects need to be dynamic in terms of team composition and skills, frequently relying on freelancers and interns, and sometimes alongside informal contributions from the open source community.

Many projects are designed for long term impact but are reliant on short term funding

Of the projects that we studied in-depth, funding was a key issue for the project's sustainability, and a considerable amount of work in the influence mapping space is done ***ikh [hgh***, motivated by a drive to make public important findings. For example, Parltrack has been created funded by ***Z] ah*** donations, and Cargografías was created in its early stages on a voluntary basis, and then funded sporadically by micro grants. Since the funding for La Fabrique de La Loi, which tracks the evolution of the lawmaking and voting process in the French parliament, has ended, the project is

sleeping but it was updated ***ikh [hgh*** when the surveillance bill was taken through parliament.

As participants repeated during the Technology for Investigative Journalism meeting in London this year, there is not enough money in the field to create multiple versions of tools and datasets. The sector needs to become more coherent, and more interoperable. One solution for this is, for example, Who's Got Dirt, a kind of Meta API, which has the capacity to query all the APIs to the various relevant databases that are publically available.

Indeed, many of the projects rely on public or open data. These data are often scraped or downloaded from publically-available sources, sometimes in partnerships with "data providers". Much work in the influence mapping space has been done in programming scrapers to fetch the data from target websites, and in establishing workflows to deal with large tranches of difficult to parse data. In around half the cases the resulting software has been made available to the community using open licenses.

~ N There is a tension within the community between open and closed data policies, arising from different motivations and sustainability needs

Open data licensing is rarer, with only a quarter of the projects releasing their data under an open license as qualified by The Open Definition: "Open data and content can be freely used, modified, and shared by anyone for any purpose". Another quarter made their data publically available but without a valid open license. While this means the data is available for reuse, the lack of a stated open license means that there could be confusion on the part of those wanting to use it, which could hamper reuse. A handful were partially available, usually because commercial reuse was restricted, and less than half published data under a restricted license. Sometimes restrictions in reuse of data are cascade from original data licenses. However, this split between restricted and publically available data is perhaps a product of the tension within the community

between the different motivations of creating change versus the journalistic necessity to generate high-value stories from investigations - which usually means being the first or the only outlet with a story.

Another tension is the fact that some projects aim to create sustainability by generating revenue from the data, which means it needs to be exclusive. An alternative model for some projects is to sell software as a service, which is for instance the model kumu employs, but this can restrict the impact of the tool and the demographic of the user base. The choice between open and closed data and software is causing a schism within the influence mapping community, which some consider to be eroding the potential for impact from the community's work. Indeed, from our in-depth case studies we have seen that openness to reuse has reportedly increased the reach of individual projects, strengthened the community of practice and diversified and broadened the practitioner base, bringing in valuable new expertise.

Beyond reuse, assessing impact is often a challenge in projects that aim to produce broader social change. Particularly, it is challenging to reliably trace a project causing change in a broader context. Through the course of our research, our impact assessment methodology was designed to take into account both the reach of the project, and direct and indirect indicators of wider impact.

The online and offline reach of projects varies widely. The projects with the largest online reach were often carried out by established news sources, for instance The Counted from The Guardian, The Fifa Files from The Sunday Times, China's African Spending Spree from Al Jazeera and Dollars for Docs from ProPublica. However, there were some exceptions, notably both Open Corporates and their Financial Visualisation, which have a high Alexa page rank of 34,328 and high OSE domain authority ranks.

Other projects have been enormously important in facilitating other influence mappers. LittleSis data, for instance, is used by paradigm shifting project TheyRule, which

changed how visualisations were used in activism, and has also received acclaim as an artistic work. LittleSis data has also been cited in over 500 times in news media.

Alongside online statistics and page rankings, we have looked for documented instances of projects being targetted in some way, implying impact. Some projects have been targetted by legal action, as in the case of the OCCRP's Proxy Platform, which was the target of an unsuccessful libel case. Similarly, anti-corruption NGO K-Monitor have been targetted for investigation by the Hungarian government.

Outcomes for the projects vary from academic papers through to bankers' resignations after ICIJ's Secrecy for Sale: Inside the Global Offshore Money Maze revelations. Non-profit Explorer has been cited as a tool to direct public donations, and perhaps the project with the most traceable impact is The National Institute on Money in State Politics's Follow the Money, which has been cited in legal proceedings over 1,000 times in the past decade.

As discussed at the Technology for Investigative Journalism conference - in this realm, you know you're having an impact if someone wants to stop you from doing it. This is an important lesson from investigative journalism that applies directly to influence mapping. If your project is successful in making an impact the requirements for rigorousness in the data analysis and investigation are extremely high. The outcome must be watertight to either be used as evidence or to generate a story that can be published without putting you at risk of being subject to legal action.

Successful influence mapping projects could be targets for legal action so itôs important to ensure practices, background data and findings are watertight to safeguard the project and its intended impact

When embarking on a project, therefore, from the beginning it is important to have trust in the data you are using - to trust the sources, the quality and the cleanliness of the data, to ensure that the conclusions that are drawn from it are also trustworthy. Furthermore, it is paramount to follow rigorous procedures and best practice in data modelling and analysis, and to know the law and the region you are working and publishing in and how it applies to your work.

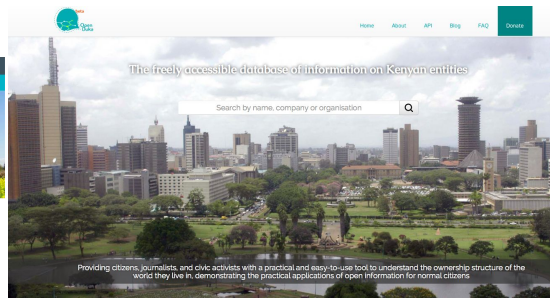
The trustworthiness of the tools used, and communications methods, are also important here from a digital security standpoint, and highlights the importance of planning and threat modelling when embarking on an influence mapping project.

This being said, experienced practitioners, when asked to give advice to newcomers considering tackling an influence mapping project, advised to go for it, even if it would be in a small way - just to start and get something out there which can then be built upon. Influence Mapping projects are still part of an experimental field. There are myriad opportunities for learning: projects have to invent their own models, their own objectives and be as modular as possible to deliver an output. While there are many challenges in terms of funding and sustainability, the community is diverse, expert and growing. Key now to the development of influence mapping as a community of practice is to approach the community development strategically, focus on building a coherent and collaborative community, avoiding unnecessary repetition of projects and tools, and furthering understanding on how to create the most impact from the work.

Case Studies

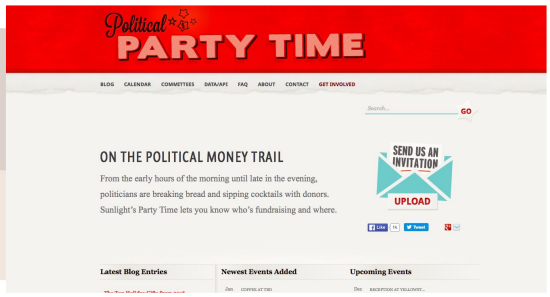
We picked a handful of projects for deeper investigation, and interviewed project representatives to explore further the influence mapping space. These projects represent diverse practices, geographical and thematic focus, and technological approaches.

Influence Map

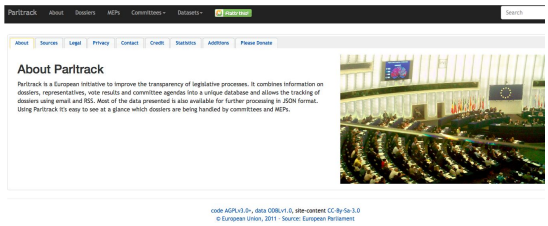


Open Duka

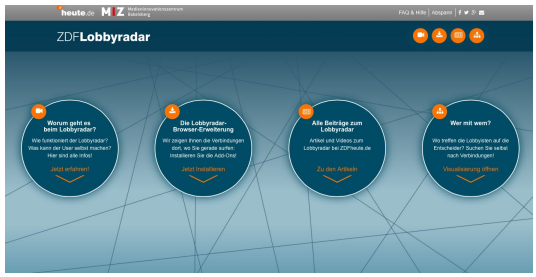
La Fabrique de la Loi



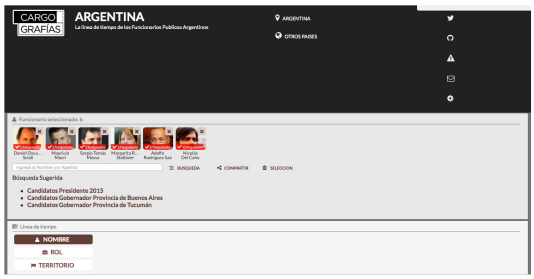
Political Party Time



Parltrack



Lobbyradar



Cargografías

Influence Map

/k`ZgblZmbhgZe 3mkn\mnk^: Non profit Community of Interest Company (CIC)
3mZmnl: Currently updated
3mZkm]Zm^h_ma^ ikhc^\m: 2015
4^Z f : 10 people
4ri^h__ng]bg`: Funded until now through mixed funding (self-funds and UK-based Foundations)
"n]`^mZ f hngm: Not disclosed
,h\Zmbhg: London (United Kingdom)
52,: http://influencemap.org
4ri^h_hnmi nm: A scoring system ranks and categorises corporations according to how they influence climate policy.

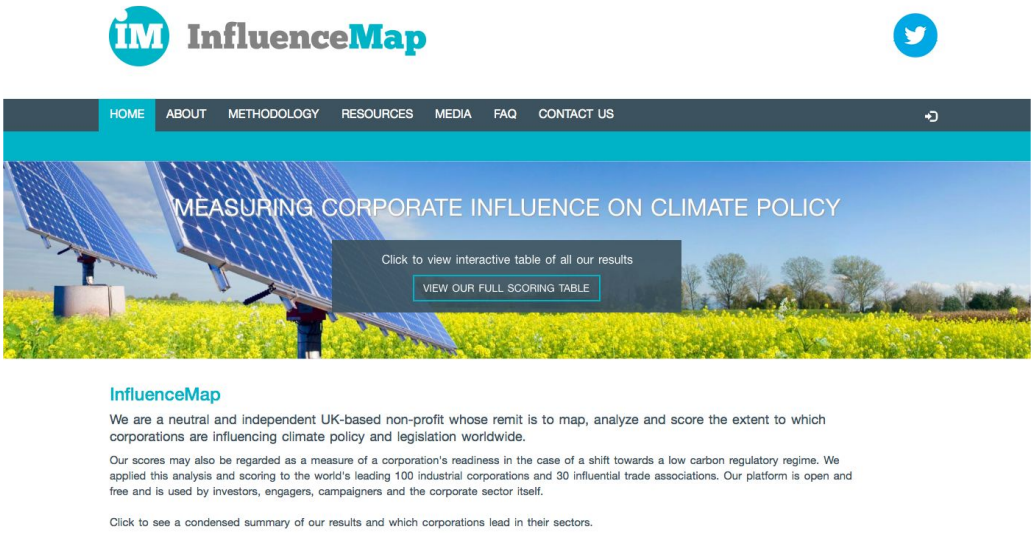
Influence Map measures and lists corporate influence on climate policy worldwide. The project was created to provide a ranking system that is, in their own words, "as objective as possible", as a tool to address gap in how corporations influence climate legislation.

The project outputs to date have been a list of rankings of corporations released in 2015. This report created a good deal of traction in the media, with almost 30 citations of the report in the press. The project also reports engagement from stakeholders such as investor groups and policymakers:

"We had opinions from some very, very influential, some senior climate policy legislators who said that our, for example, our report on oil and gas industry lobbying was very timely and very helpful in balancing similar propaganda coming from the oil and gas industry." Dylan Tanner, Founder and VP.

To generate the score for each corporation, Influence Map aggregates information from 8 different publicly available data sources, querying these sources with 12 questions relating to climate policy. To constitute the score, the Influence Map's team uses a traditional method encountered in other Influence Mapping projects, relationship mapping, to weight the data, which was then processed by the team's algorithm to generate a numerical score to feed into the Influence Map rankings. (More detailed information on the project's methodology).

The project's aim to be "as objective as possible", and their quantified-approach, provides more leverage to stakeholders, who might have had to deal with atomised sources of data about climate policy and corporate influence in the past.



The impact of the project has been amplified by their project management strategy, which involved from its early stages engagement with a core user group of specialists in the sector who have used the tool and organically disseminated their findings. This is combined with a strong and strategic media engagement strategy, timing important reports to coincide with important events or reports in the same sector and formatting

them in an easily accessible manner to bring the most newsworthy information to the top line.

The team of 10 combines expertise from a variety of different disciplines, including design, software development, physics and analysis, aiming now to bring in visualisation expertise to increase legibility of the outputs. In terms of sustainability of the project, the aim is to increase the project's links to specialist media. However, plans for the future of Influence Map will be affected by the financial resources available to the project.

Open Duka

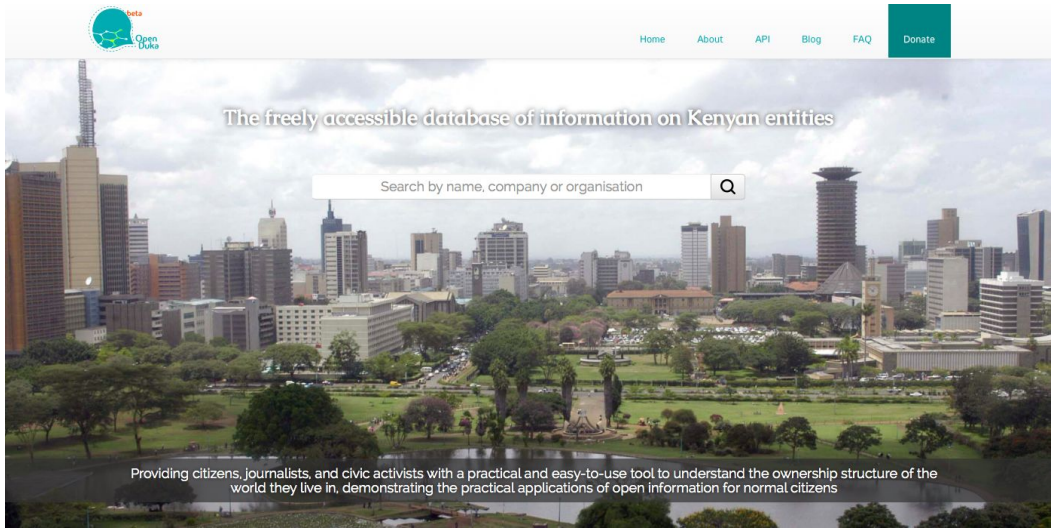
/k`ZgblZmbhgZe 3mkn\mnk^: Project incubated at the Open Institute
3mZmnl: Currently updated
)gm^kob^p^^: Benjamin Charagu, Social-Innovation Lead of Open Institute
3mZkm]Zm^h_ma^ ikhc^\m: 2012
4^Z f : Size variable (from Small to Medium)
4ri^h_ng]bg`: Grants Budget amount: Not disclosed
,h\Zmbhg: Nairobi (Kenya)
52,: http://openduka.org
4ri^h_hnminm: A free database about Kenyan entities to browse between networks of relationships linking people, organisations, tenders, contracts, cases and land parcels together.

Open Duka was started to gather information in one repository to create a searchable database of institutions and actors in power and commerce in Kenya, in order to foster more accountability and to facilitate oversight by citizens, journalists and civic activists.

Incubated by the Open Institute, the project is open source and the database is freely available and can be queried via the Open Duka API. Data has been scraped from online sources that provide shareholder information, procurement information, legal cases and company information, from data providers and partners The Kenyan Gazette and Kenya Law. Open Duka's site is coded in PHP and the database is managed with MyTML library. Visualisation is now based on D3 javascript library.

The project has faced challenges since its inception, most notably access to data and tools. For example, the lack of a proper API to the Kenya Law data has slowed down access to their content. The problem with access to data has been exacerbated by lack of funds for tools to convert between different data formats. On occasion, Open Duka has

found it necessary to translate data that was not machine readable with the tools available and to enter data manually into the database.



Alongside technological challenges in managing the database and visualisation tools, the team came across specific challenges when they attempted to use Document Cloud to extract information from documents: “Document Cloud couldn't pick up the entities, for example, the African names, can't pick, or it doesn't know how to identify a person and an institution based on our locality,” says Charagu.

At this stage, the impact of the project has been not been formally assessed, and the online outreach is fairly small.

Benjamin Charagu and his team at the Open Institute are looking to change the model of Open Duka in order to potentially improve its impact by offering the database to organisations which might need entities mapping. However, these plans hang in the balance as the project’s funding situation is critical, making the future of Open Duka uncertain.

La Fabrique de la Loi

/k`ZgbIZmbhgZe 3mkn\mnk^ç Project led in collaboration between 2 research labs at SciencesPo and the NGO Regards Citoyens
)gm^kob^p^^ç Benjamin Ooghe-Tabanou
3mZmnIç Maintained but not regularly updated
3mZkm]Zm^h_ma^ ikhc^\mç 2011
4^Z fç Large
4ri^h__ng]bg`ç Grant
"n]`^mZfhngm: ¸150,000 grant from the Ile de France Region fund over 3 years
,h\ZmbhgçParis (France)
52,ç http://www.lafabriquedelaloi.fr
4ri^h_hnminmç 290 law texts analysed and visualized (from proposal bill to its published version) showing how parliamentary process in France works.

La Fabrique de la Loi publishes and tracks the evolution of bills through French Parliament from their inception to the final publication, showing through visualizations the chronological changes to the text and the related debates history between the different MPs. Aside from the web interface, the data is also made accessible through an API.

The goal of La Fabrique de la Loi is “to easily understand and visualize the French parliamentary process”. The site and API are the products of a three year long research project between two research labs from Sciences Po in Paris: the médialab and the Center for European Studies (CEE) and Regards Citoyens, an NGO urging for transparency in the public sphere.

The project started in 2011, funded by a grant of ¸150 000 from The Région Île-de-France to sponsor and support innovation and research for citizens. The grant was used to support the team during three years, including funding the salary of a PhD candidate and the designers who worked on the visualization part of the website, and to support conferences and workshops.

To retrieve the texts and the different MPs votes, the team programmed scrapers to extract data from different websites and reformat it. However, as Benjamin Ooghe-Tabanou reported “this process did not succeed in 30% of the cases” because of the wide diversity of the formats used by the National Assembly and the Senate to publish their documents.

There has not been an impact assessment regarding the project so far and Ooghe-Tabanou even considers it ““mostly unknown by the broad public””. However, according to Ooghe-Tabanou, the project was well known by some MPs assistants and had been used by activists during the campaign against the surveillance bill in France to track the parliamentary processes. Looking at it with hindsight Ooghe-Tabanou even

hopes that the project contributed to raise awareness about political open-data in France.

The tools used by the project team are all free software, which is a choice and a statement according to Ooghe-Tabanou: “transparency has to be applied even to the code”. The scrapers were programmed with bash, perl and Python; D3.js was used to compute the visualisations. Finally, La Fabrique de la Loi tried to document as much as possible their technological choices on their website.

Today, La Fabrique de La Loi is no longer actively updated and the project is, as Ooghe-Tabanou qualified it, “sleeping”, but the team voluntarily chose to upload information from to time to time regarding “symbolic and key” texts such as the surveillance bill which was voted by the Parliament in July 2015. If time and opportunities allow it, they would like in the future to complete the automated extraction of the data to include all bills live during their parliamentary discussion.

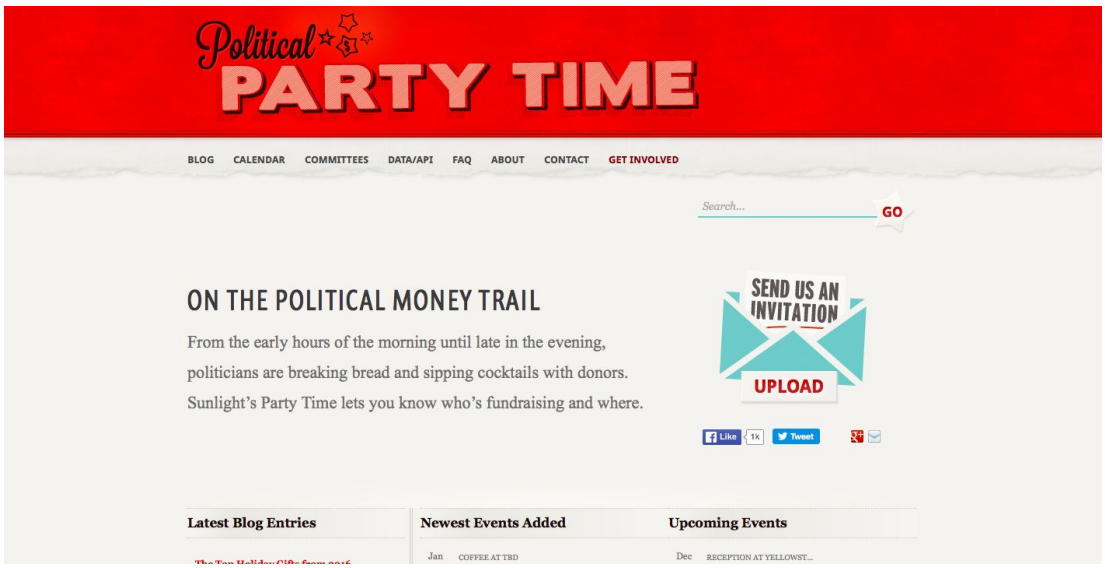
Political Party Time

/k`ZgblZmbhgZe 3mkn\mnk^: Project incubated at the Sunlight Foundation
)gm^kob^p^^l: Jen Topper, Communications Director at Sunlight Foundation and Joshua Stewart, Deputy Communications Director at Sunlight Foundation
3mZkm]Zm^h_ma^ ikhc^\m: 2007
4^Z f : Size variable (from Small to Medium)
4ri^h_ng]bg`: Donations
"n]`^mZ f hngm: Not disclosed
,h\Zmbhg: Washington D.C. (USA)
52,: http://politicalpartytime.org
4ri^h_hnmi nm: A tracking system to shed light on cocktail parties and informal meetings organized to raise money for political campaigns or influence candidates through lobbying

Political Party Time helps to track political fundraising events and the movement of political actors from event invitations. This information can be used to map the relationships between candidates and private interests from different sides. Launched in 2007 by the Sunlight Foundation, the project is still maintained by the organisation but no longer funded.

Invitations to fundraising events are contributed by users who can submit an invitation by emailing or by sending it anonymously through the website interface. A small reporting team actively contributes to Political Party Time by tracking these invitations. As Topper and Stewart put it, Political Party Time aims for “content generation”. Its objective is to lead storytelling based on the fundraising event information collated on the website.

The site is fairly simple and it uses open source tool programmed by Sunlight Foundation. Stewart explains that "they created a standardized form that allows admin, with basic data entry training, to enter fundraising invite data. The key is to tag the post with as much information as possible from the invite. This allows for our database to connect to previous hosts, lawmakers, and even venues." The metadata is made available as bulk downloads "so any individual or organization can download our entire invite database". The data collection is entirely done manually. Sunlight Foundation is currently working on scripts to automate the data collection (scraping from PDFs) but so far this has not been released.



No precise metrics for impact assessment have been reported, however the two interviewees mentioned how one of the recent achievement of the project was to uncover Jeb Bush's presidential campaign launch. By tracking the body of evidence (the invitations), the project managed to infer his political candidacy.

Parltrack

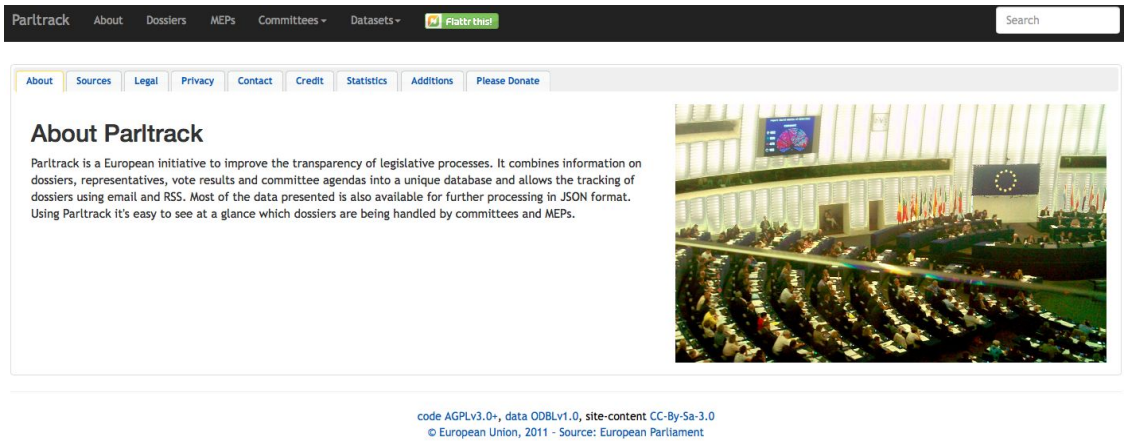
/ k`ZgbIZmbhgZe 3mkn\mnk^: Self-initiative/independent project
)gm^kob^p^^: Stef, founder and project manager
3mZkm]Zm^ h_ ma^ ikhc^\m: 2011
4^Z f : One person
4r i ^ h_ _ng]bg`: Donations
"n]`^m Z f hngm: Not disclosed
,h\Zmbhg: Not disclosed
52 , : http://parltrack.euwiki.org
/ nm i nm: A database to track dossiers, representatives, vote results and committees agendas in the European Parliament.

Parltrack is a free database of European Union dossiers, committee meetings and the people who contributed to them. The project started in 2011, during the campaign against ACTA (the Anti-Counterfeiting Trade Agreement).

Since its launch, Parltrack has been lead by one person, who maintains the project and manages the website. Parltrack is built using open source tools: Python for the scrapers and Mongo DB for the data management. The project has benefitted from open source development through coding contributions from the wider community. Parltrack also benefitted from crowdsourcing campaigns and donations, the most recent reached ¸ 11,000 to pay the maintenance of the website.

Parltrack uses web scrapers to gather data automatically, predominantly from the website of the European Parliament. There are frequent changes to this site, however, and as a result frequent changes to the web scrapers are necessary to maintain consistent data collection. This is often a time consuming exercise: “I had to debug the site for like half a day to find out why one of my scrapers didn’t work, and then it turned

out that the European parliament changed the URL of one of their search interfaces." Stef, Parltrack. This dynamic environment informed the decision to use Mongo DB, which was considered a good choice of database to store data from the European Parliament "which are constantly changing".



In terms of impact, the Parltrack dataset has been used by other projects such as Political Memory, LobbyPlag and Score EP. "I really like to think of it [Parltrack] as a data provider, and I have all these cool NGOs and projects who use my data and do really really cool stuff with it that I'm really proud of, so, to support these people and projects." Stef, Parltrack. The project is lauded by hacktivists as one of the cornerstones of the transparency movement, according to European political news site Politico, and has underpinned work by Transparency International on lobbying.

Paltrack will remain an independent project and will host freely the database, an ideological choice to keep the data available and open so that others can build on Parltrack and challenge the existing power structures - to "make the powerful people sweat".

Lobbyradar

/k`ZgbIZmbhgZe 3mkn\mnk^: Project incubated at ZDF (Zweites Deutsches Fernsehen)
)gm^kob^p^^: Dominik Wurnig and Jan Schneider, project managers and journalists
3mZkm]Zm^ h_ ma^ ikhc^\m:2015
4^Z f : Size variable (from Small to Medium)
4ri^h__ng]bg`: Funded by ZDF
"n]`^mZ f hngm: ¸150,000
,h\Zmbhg: Germany
52,: http://lobbyradar.de (currently offline), archived at http://lobbyradar.org
4ri^h_ hnmi nm: A visualization and a plugin which shows the different ties between lobbyists and politicians in Germany.

Lobbyradar uncovers the lobbying groups and private interests in German politics, by enabling users to explore network visualisations that show links between institutions and people of influence within the german political sphere. There is also a browser plugin, which, once installed, supplies information about politicians and lobbyists via a tooltip when their name appears on a website.

The project started as a cooperation between ZDF (Zweites Deutsches Fernsehen) and the Medieninnovationszentrum Babelsberg (MIZ) with the aim to tackle lobbying and its intersection within the political sphere in Germany. The project took shape during a scholarship of the three journalists Michael Hartlep, Dominik Wurnig and Jan Schneider at MIZ and included a series of brainstorming workshops organized between journalists, coders and designers.

Lobbyradar's team was spread across multiple cities in Germany and mainly composed of the three journalists, working on the project part-time over six months alongside programmers from Open Data City. These disparate locations were challenging for the