

JOTA

About JOTA

JOTA IS AN INFORMATION PROVIDER THAT SEEKS TO GIVE TRANSPARENCY TO PUBLIC INFORMATION, PROVIDE GREATER LEGAL SECURITY TO THE MARKET, AND TRANSFORM SOCIETY THROUGH SPECIALIZED, TECHNICAL, IMPARTIAL, AND NIMBLE JOURNALISM.

LAVA JOTA

The Case

*Since 2014, Brazilian politics have been enmeshed in its largest corruption scandal ever. **Carwash Operation** (Operação Lava Jato) revealed a large corruption scheme involving private contractors and public entities that promoted illegal campaign financing and facilitated the embezzlement of billions.*

The scale of the operation was unprecedented in Brazil. Following the scandal, there were more than 1100 Carwash Operation cases under review at federal, regional, and Supreme courts; meanwhile, the federal police had been forced to build new software to process data, as more than 1.2 million gigabytes had been collected during the investigations.

The Case

*How many countries are involved in the **Carwash** investigations?*

A) 13

B) 27

C) 49

The Project

*To address the challenge of analyzing such a body of information, Jota launched “**Lava Jota**”, a public database of judicial information related to Carwash. We put together all of Carwash’s legal actions on a single website and built an integrated search engine that allows you to search for keywords within each of the documents attached to the processes.*



*Everything you need to know about
Carwash Operation, but did not
know who to ask*

#LAVAJOTA

Tudo que você sempre quis saber sobre a **Operação Lava Jato**, mas nunca soube a quem perguntar.



PROCESSOS



ENVOLVIDOS



VIDEOS

*Cases; People Involved;
Videos*

Buscar em Processos



+OPÇÕES



BUSCAR POR TERMO(S) EXATO(S)

Exact term search

Brazil's Judicial System, more transparent than ever

Behind the Scenes

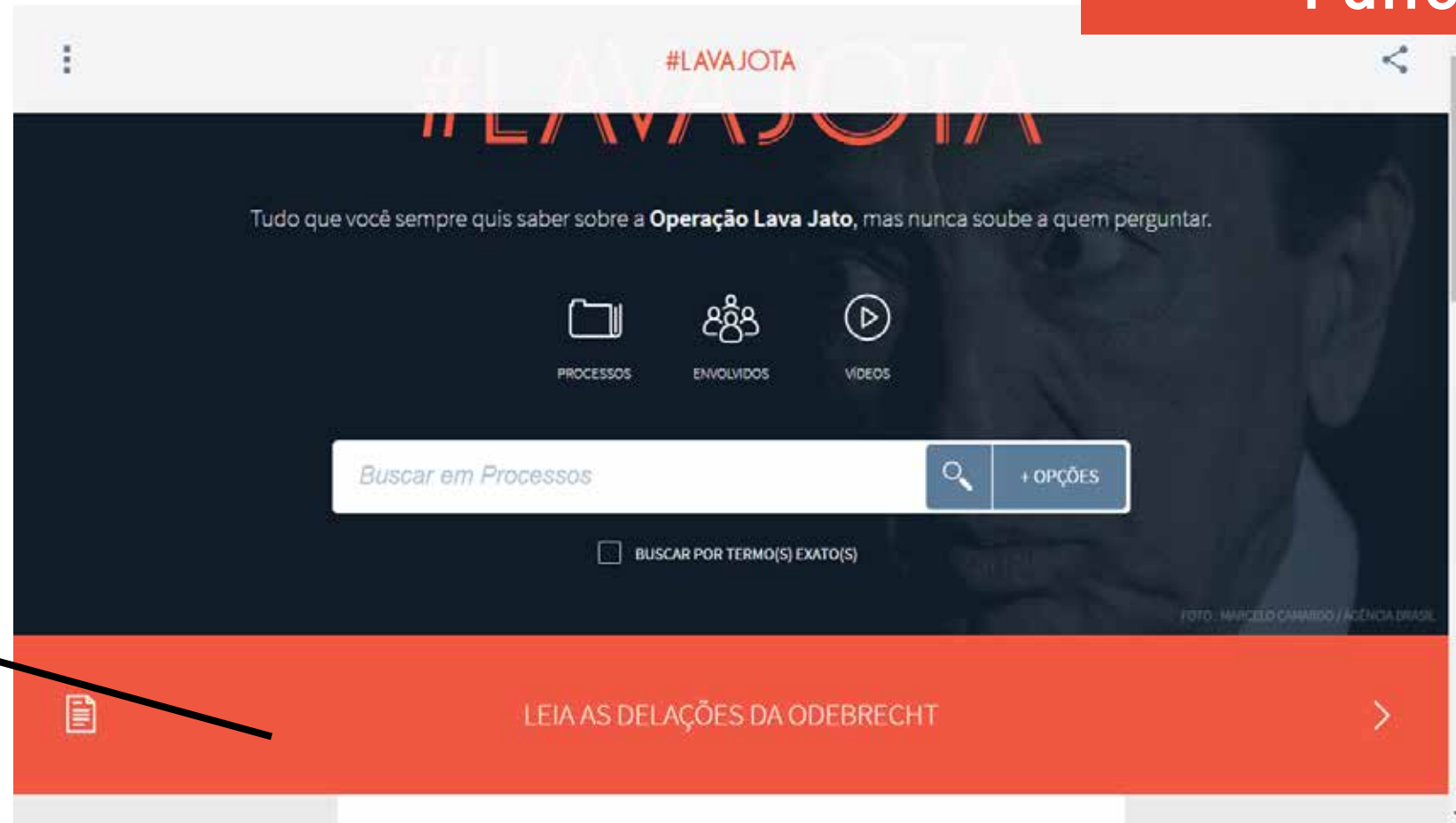
Techniques:

- *Web scraping script for downloading data.*
- *We gathered the documents in a **database management system** and made them available with searching functionalities on our webpage.*
- *We transformed all documents into searchable **text** and put them on our database.*
- *We utilized **technology** to transcribe and classify **videos of depositions**, and then had our reporters check the result.*
- *We structured part of the data and **generated a list of the involved people**.*

Tools:

- ***Python** for coding the crawlers.*
- ***R language** to analyze structured data.*
- ***Elasticsearch** system for handling the documents.*
- ***Microsoft Video Indexer** for indexing and transcribing videos.*
- ***d3.js** for the generation of data visualizations to display some of our structured data.*

Example: Read here testimonies of all of the Odebrecht executives that entered into cooperation agreements with Brazilian authorities



Every case that came from the Carwash Operation, with all their respective files and documents, searchable in one place

Free access, available for all society

** Lançamento previsto para o dia da divulgação das delações da Odebrecht.*



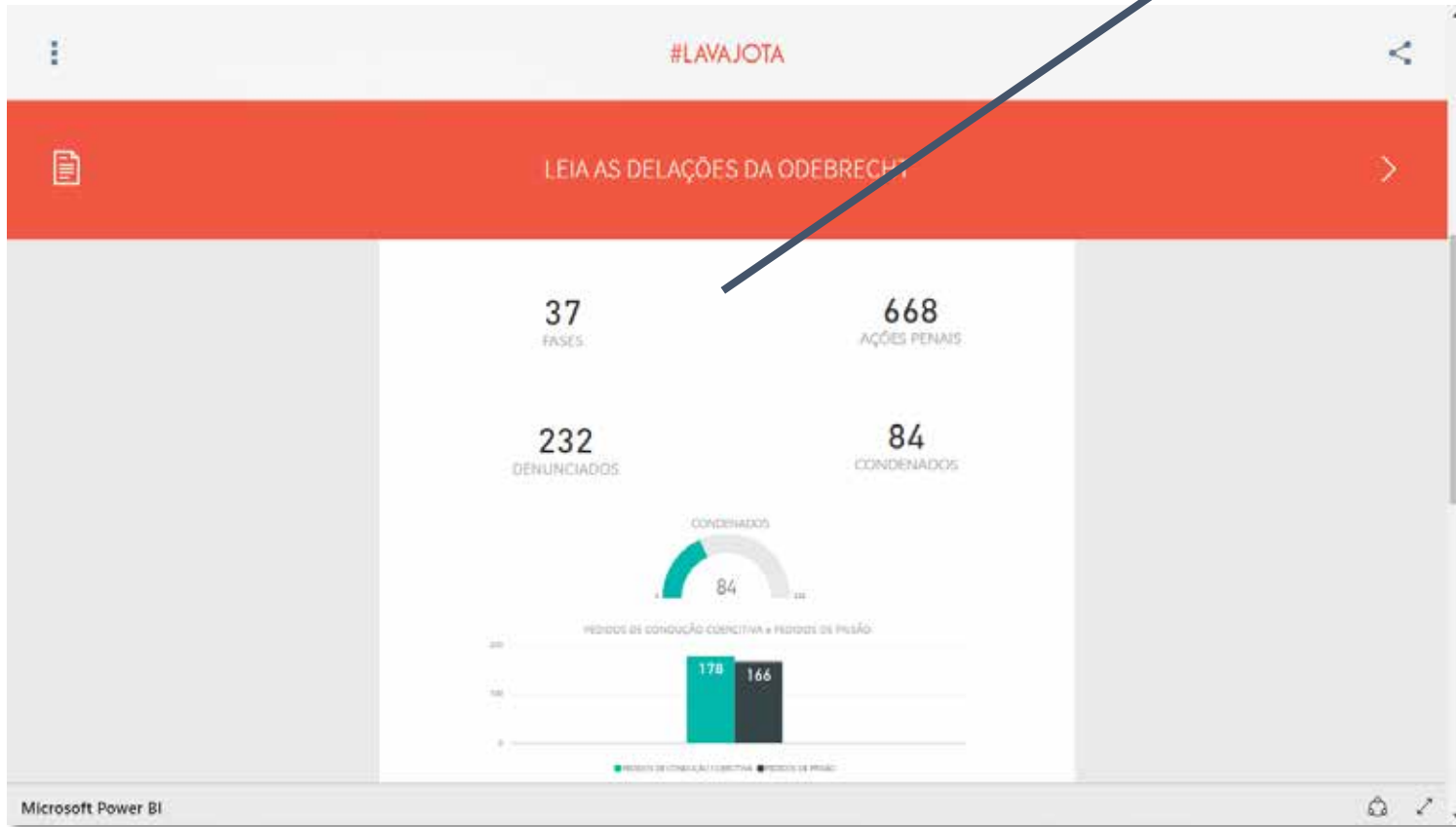
Over 1100 cases on file – all the legal processes filed in Brazilian courts, including regional courts, the Superior Court of Justice, and the Supreme Court -
UPDATED DAILY



UPDATED database with names of **all parties** involved and the **status of each case**

Functions

Example: Analysis of results of cases stemming from the Odebrecht cooperation agreement, with number of criminal suits, convictions, charges, and more



First phase of *business intelligence* review of data brought in by the cases, allowing a deeper analysis of trends in the investigations

The Case

Do you have an equivalent tool in your country?

A) YES

B) NO

C) YES BUT NOT WITH THE SAME REACH

Measuring impact: pageviews around the world

| Página | Visualizações de página | Visualizações de páginas únicas | Tempo médio na página | Entradas | Taxa de rejeição |
|---|---|---|---|---|---|
| | 310.404 Porcentagem do total: 2,65% (11.712.537) | 251.005 Porcentagem do total: 2,39% (10.518.410) | 00:01:24 Média de visualizações: 00:03:02 (-53,68%) | 83.778 Porcentagem do total: 0,96% (8.754.825) | 58,70% Média de visualizações: 84,99% (-30,94%) |
| 1. /lavajota/jota.uol.com.br | 69.291 (22,32%) | 52.240 (20,81%) | 00:01:04 | 35.501 (42,38%) | 44,19% |
| 2. /lavajota/videos.html | 39.618 (12,76%) | 33.434 (13,32%) | 00:01:54 | 28.765 (34,33%) | 85,35% |
| 3. /lavajota/index.html | 18.480 (5,95%) | 12.036 (4,80%) | 00:01:06 | 8.116 (9,69%) | 43,03% |
| 4. /lavajota/reus.html | 17.087 (5,50%) | 11.953 (4,76%) | 00:01:09 | 993 (1,19%) | 56,90% |
| 5. /lavajota/processos.html?pg=1 | 10.198 (3,29%) | 7.743 (3,08%) | 00:00:40 | 578 (0,69%) | 47,92% |
| 6. /lavajota/graficos.html | 5.780 (1,86%) | 5.056 (2,01%) | 00:01:18 | 358 (0,43%) | 71,23% |
| 7. /lavajota/odebrecht.html | 1.286 (0,41%) | 936 (0,37%) | 00:00:54 | 552 (0,66%) | 42,57% |
| 8. /lavajota/odebrecht.html?odb=1 | 1.250 (0,40%) | 954 (0,38%) | 00:00:37 | 70 (0,08%) | 57,14% |
| 9. /lavajota/busca.html?buscar=Lula&pg=1 | 1.182 (0,38%) | 1.032 (0,41%) | 00:00:53 | 51 (0,06%) | 78,43% |
| 10. /lavajota/busca.html?buscar=lula&pg=1 | 1.141 (0,37%) | 1.000 (0,40%) | 00:01:00 | 29 (0,03%) | 58,62% |

Lava Jota has received more than **310 thousand** pageviews since its launching, with **251 thousand** unique pageviews.

Measuring impact: pageviews around the world

| | | | |
|-----|---------------------------|--|-----------------|
| 1. | /lavajota/jota.uol.com.br |  Brazil | 55.842 (82,84%) |
| 2. | /lavajota/jota.uol.com.br |  United States | 3.237 (4,80%) |
| 3. | /lavajota/jota.uol.com.br |  Dominican Republic | 2.936 (4,36%) |
| 4. | /lavajota/jota.uol.com.br |  Peru | 884 (1,31%) |
| 5. | /lavajota/jota.uol.com.br |  Panama | 566 (0,84%) |
| 6. | /lavajota/jota.uol.com.br |  Argentina | 389 (0,58%) |
| 7. | /lavajota/jota.uol.com.br |  Mexico | 336 (0,50%) |
| 8. | /lavajota/jota.uol.com.br |  Germany | 318 (0,47%) |
| 9. | /lavajota/jota.uol.com.br | (not set) | 265 (0,39%) |
| 10. | /lavajota/jota.uol.com.br |  Switzerland | 212 (0,31%) |
| 11. | /lavajota/jota.uol.com.br |  France | 195 (0,29%) |
| 12. | /lavajota/jota.uol.com.br |  Portugal | 177 (0,26%) |
| 13. | /lavajota/jota.uol.com.br |  Italy | 159 (0,24%) |
| 14. | /lavajota/jota.uol.com.br |  United Kingdom | 142 (0,21%) |
| 15. | /lavajota/jota.uol.com.br |  Chile | 124 (0,18%) |
| 16. | /lavajota/jota.uol.com.br |  Spain | 124 (0,18%) |
| 17. | /lavajota/jota.uol.com.br |  Philippines | 124 (0,18%) |
| 18. | /lavajota/jota.uol.com.br |  Turkey | 124 (0,18%) |
| 19. | /lavajota/jota.uol.com.br |  Uruguay | 124 (0,18%) |
| 20. | /lavajota/jota.uol.com.br |  Canada | 88 (0,13%) |

Lava Jota has been utilized in more than **30** countries;

the **top 5** are Brazil, the United States, the Dominica Republic, Peru, and Panama.

Challenges

- **Financial:** *JOTA built Lava Jota as a pro bono project. The financial resources received from sponsors and partners, like law and technology firms and Transparency International, only paid for the production costs. JOTA invested a lot of personnel time in order to make this information available to society.*
- **Operational:** *The magnitude of the Carwash Operation required a tremendous operational effort in a team of 25 editorial staff total. For example, uploading videos of court sessions and testimonies to YouTube can take hours, and there are dozens of such videos here. Also, automatic video transcription can take hours and is not perfect. A reporter must be assigned to read the transcript, identify key issues, and then choose the most important tags in each submission. We haven't yet been able to do this for even half the material available.*

There's still plenty to do.



**And we are just getting started!
Lava Jota will continue applying
intelligence to the data and
offering a never-before seen
window into the Brazilian
Justice System**

The Case

Would an equivalent tool be useful to you?

A) YES

A) NO

JOTA