

Fighting data deserts with civic engagement

Natalia Mazotte, JSK fellow Stanford 2020

@nataliamazotte

10th NIC.br Annual Workshop on Survey Methodology





SCHOOL OF **DATA**



OPEN KNOWLEDGE
BRASIL

OPERAÇÃO

SERENATA DE AMOR

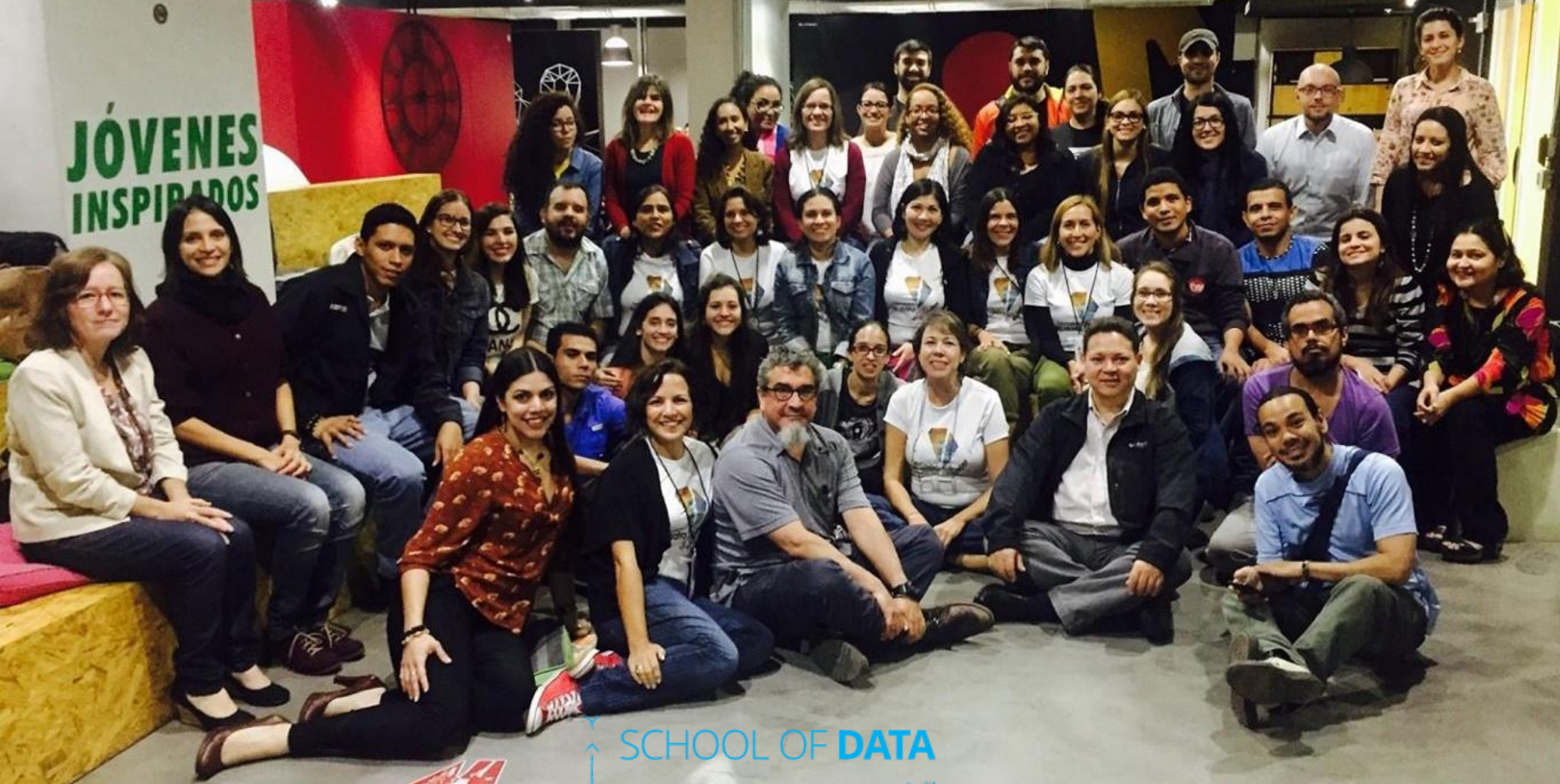
ORDEM E PROGRESSO



OPEN
KNOWLEDGE
BRASIL

Associação de Desenvolvimento UFRJ

JÓVENES
INSPIRADOS



SCHOOL OF DATA

*"How might we help local
journalists to better use technology
and apply digital methods in order
to overcome **data deserts?**"*





COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)



Global Cases

35.675.704

Cases by Country/Region/Sovereignty

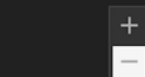
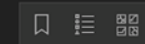
- 7.495.674 US
- 6.685.082 India
- 4.927.235 Brazil
- 1.231.277 Russia
- 869.808 Colombia
- 829.999 Peru
- 825.410 Spain
- 824.468 Argentina
- 789.780 Mexico
- 683.242 South Africa

Last Updated at (M/D/YYYY)
10/6/2020 9:23 PM

- Cumulative Cases**
- Active Cases
- Incidence Rate
- Case-Fatality Ratio
- Testing Rate

188
countries/regions

Lancet Inf Dis Article: [Here](#). Mobile Version: [Here](#). Data sources: [Full list](#). Downloadable database: [GitHub](#), [Feature Layer](#).
Lead by [JHU CSSE](#). Technical Support: [Esri Living Atlas team](#) and [JHU APL](#). Financial Support:



Esri, FAO, NOAA

Global Deaths

1.047.220

- 210.784 deaths US
- 146.675 deaths Brazil
- 103.569 deaths India
- 81.877 deaths Mexico
- 42.535 deaths United Kingdom

US State Level

Deaths, Recovered

- 33.219 deaths, **77.192** recovered New York US
- 16.510 deaths, **687.277** recovered Texas US
- 16.208 deaths, **recovered** California US
- 16.147 deaths, **35.102** recovered

Daily Cases

How did we get here?

75,523

DEATHS

1,970,909

CASES

*If they cannot fight
the problem, they
will fight the data*

Covid-19 data blackout

- Official statistics website inconsistency
- Incomplete datasets
- Lack of details and padronization (Centralized Health System, but no centralized data infrastructure)
- Unjustifiable changes in the methodology
- Delays in the release of statistics
- Lack of open format datasets



Covid-10 Data Desert

*Covid-19 data
collection*



Álvaro Justen

@turicas



DADOS: CASOS DE [#COVID19](#) POR MUNICÍPIO POR DIA!

A população precisa saber o quanto antes a quantidade de casos em seu município, bem como sua evolução *diária*, só assim teremos agilidade e efetividade nas ações de combate à propagação do vírus. Quer os dados? Segue o fio:

[Translate Tweet](#)

3:24 PM · Mar 20, 2020 · [Twitter Web App](#)

 View Tweet activity

378 Retweets **1.1K** Likes



42



378



1.1K



Brasil 15/7/2020

Casos









































1.971.591

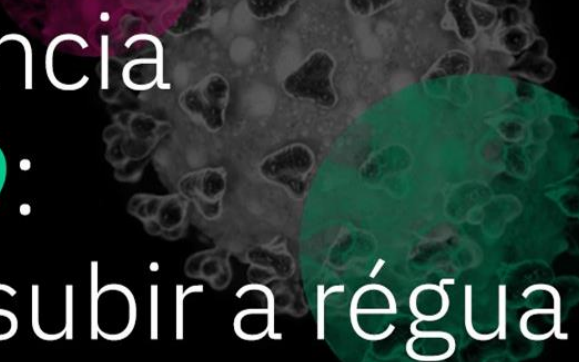
40.474 novos

Mortes

75.548

1.321 novas

UF	data do boletim	data dos dados	casos ▾	mortes
SP	15 de jul. de 2020	15 de jul. de 2020	393.176 	18.640 
CE	15 de jul. de 2020	15 de jul. de 2020	141.832 	7.081 
RJ	15 de jul. de 2020	15 de jul. de 2020	134.449 	11.757 
PA	15 de jul. de 2020	15 de jul. de 2020	133.039 	5.385 
BA	15 de jul. de 2020	15 de jul. de 2020	112.993 	2.638 
MA	15 de jul. de 2020	15 de jul. de 2020	102.469 	2.572 
AM	15 de jul. de 2020	15 de jul. de 2020	86.944 	3.080 
MG	15 de jul. de 2020	15 de jul. de 2020	82.010 	1.752 
DF	15 de jul. de 2020	15 de jul. de 2020	75.381 	1.001 
PE	15 de jul. de 2020	15 de jul. de 2020	74.960 	5.772 
ES	15 de jul. de 2020	15 de jul. de 2020	66.352 	2.096 
PB	15 de jul. de 2020	15 de jul. de 2020	63.939 	1.383 
SC	15 de jul. de 2020	15 de jul. de 2020	47.976 	569 
AL	15 de jul. de 2020	15 de jul. de 2020	47.864 	1.331 
PR	15 de jul. de 2020	15 de jul. de 2020	47.124 	1.200 
RS	15 de jul. de 2020	15 de jul. de 2020	42.239 	1.101 
RN	15 de jul. de 2020	14 de jul. de 2020	40.767 	1.479 
SE	15 de jul. de 2020	15 de jul. de 2020	40.139 	1.054 
GO	15 de jul. de 2020	15 de jul. de 2020	39.520 	959 
PI	15 de jul. de 2020	15 de jul. de 2020	36.542 	1.043 



Transparência
COVID-19:
é hora de subir a régua

Covid-19
Transparency:

It's time to set
the bar higher

30
ABR

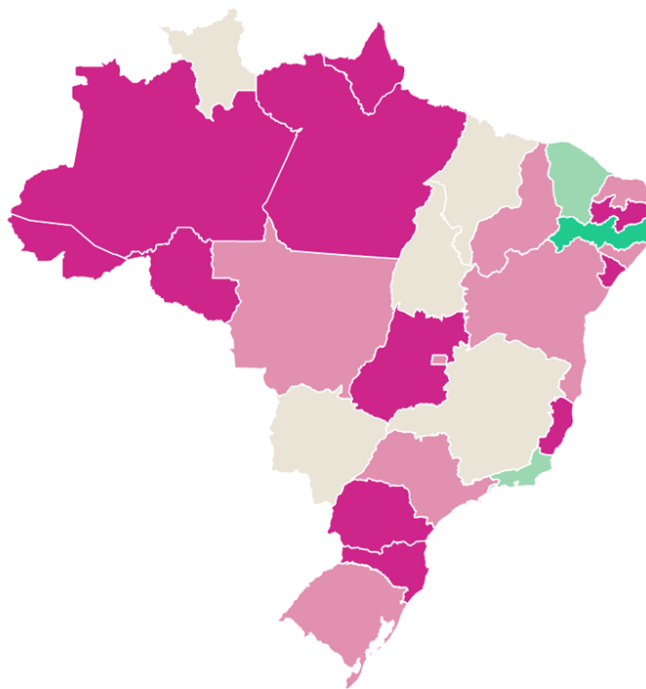
23
ABR

16
ABR

9
ABR

3 ABRIL 2020
Período de 02abr - 08abr

Avaliação por estado da federação



NÍVEIS DE TRANSPARÊNCIA



Ranking

TABELA GRÁFICO

ESTADO	PONTUAÇÃO	VARIAÇÃO	NÍVEL
1º Pernambuco (PE)	81	↻ 0	ALTO
2º Ceará (CE)	69	↻ 0	BOM
3º Rio de Janeiro (RJ)	64	↻ 0	BOM
4º Tocantins (TO)	50	↻ 0	MÉDIO
5º Minas Gerais (MG)	48	↻ 0	MÉDIO
6º Maranhão (MA)	45	↻ 0	MÉDIO
6º Mato Grosso do Sul (MS)	45	↻ 0	MÉDIO
7º Roraima (RR)	40	↻ 0	MÉDIO
8º Governo Federal*	36	↻ 0	BAIXO
8º Rio Grande do Sul (RS)	36	↻ 0	BAIXO
9º Alagoas (AL)	33	↻ 0	BAIXO
9º Bahia (BA)	33	↻ 0	BAIXO
10º Mato Grosso (MT)	31	↻ 0	BAIXO
10º São Paulo (SP)	31	↻ 0	BAIXO
11º Rio Grande do Norte (RN)	29	↻ 0	BAIXO
12º Distrito Federal (DF)	21	↻ 0	BAIXO
13º Distrito Federal (DF)	21	↻ 0	BAIXO

11 JUNHO 2020
Período de 11Jun - 17Jun

4
JUN

28
MAI

21
MAI

14
MAI

7
MAI

30
ABR

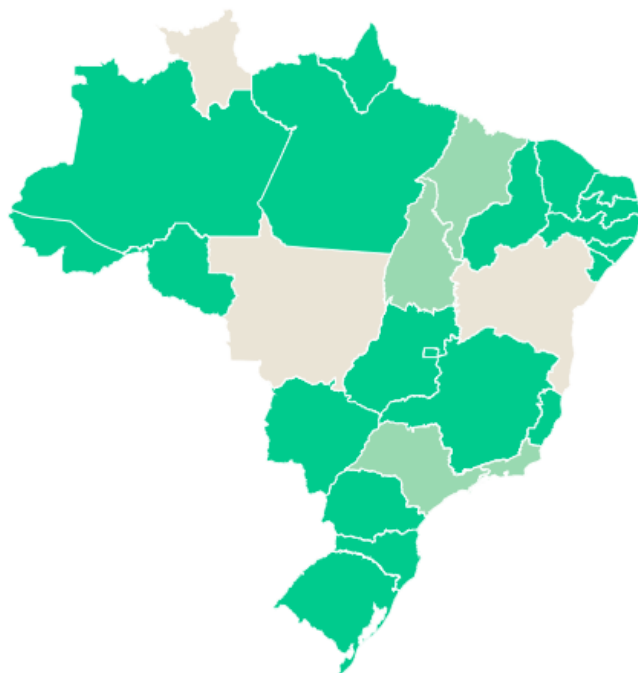
23
ABR

16
ABR

9
ABR

3
ABR

Avaliação por estado da federação



NÍVEIS DE TRANSPARÊNCIA

OPACO

0-19 PTS

BAIXO

20-39 PTS

MÉDIO

40-59 PTS

BOM

60-79 PTS

ALTO

80-100 PTS

Ranking

TABELA

GRÁFICO

ESTADO	PONTUAÇÃO	VARIAÇÃO	NÍVEL
1º Goiás (GO)	100	○ 0	ALTO
1º Minas Gerais (MG)	100	○ 0	ALTO
1º Rondônia (RO)	100	○ 0	ALTO
1º Rio Grande do Sul (RS)	100	↑ +5	ALTO
2º Acre (AC)	98	↑ +5	ALTO
2º Alagoas (AL)	98	○ 0	ALTO
2º Amazonas (AM)	98	↑ +53	ALTO
2º Ceará (CE)	98	↓ -2	ALTO
2º Pernambuco (PE)	98	○ 0	ALTO
2º Paraná (PR)	98	↑ +8	ALTO
3º <i>Governo Federal*</i>	95	↑ +5	ALTO
3º Distrito Federal (DF)	95	○ 0	ALTO
3º Santa Catarina (SC)	95	↑ +5	ALTO
4º Amapá (AP)	93	○ 0	ALTO
4º Espírito Santo (ES)	93	↓ -5	ALTO
4º Pará (PA)	93	○ 0	ALTO
5º Mato Grosso do Sul (MS)	88	○ 0	ALTO

***"Acabou matéria
no JN" (Jair Bolsonaro)***

"No more stories in JN"

News organizations team up to provide transparency to Covid-19 data

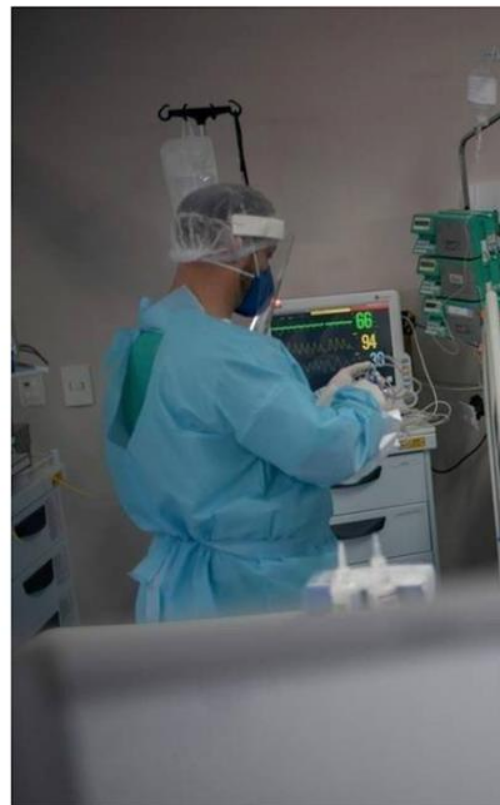
Due to limitations imposed by the federal government, journalists from O GLOBO, Extra, G1, Estadão, Folha and UOL will jointly collect and publish figures from the state health departments on deaths and those who are infected by the new coronavirus

O Globo

08/06/2020 - 13:42 / Atualizado em 08/06/2020 - 14:41



| Newsletters 



Nurses take care of a Covid-19 patient at a

WHAT IS SERENATA?

TECHNOLOGY + ENGAGEMENT



WHAT IS SERENATA?

Artificial intelligence for civic engagement and social control of public administration

Data science using open data

Multidisciplinary team: Scientists, programmers, marketers, data journalists

Open source: more than 700 volunteers in our telegram group



SO MUCH DATA WE COULD TEACH A ROBOT

THIS IS ROSIE!

ARTIFICIAL INTELLIGENCE → SOCIAL CONTROL AND ACCOUNTABILITY

DATA PROVIDED BY HUMANS POWERS AI ALGORITHMS



SO MUCH DATA WE COULD TEACH A ROBOT

THIS IS JARBAS!

Jarbas Dashboard

Início > Cota para Exercício da Atividade Parlamentar > Reembolsos

Selecione reembolso para visualizar



Buscar

7945 resultados (1619213 total)

REEMBOLSO		NOME DO PARLAMENTAR	ANO	SUBQUOTA TRANSLATED	FORNECEDOR	VALOR	SUSPEITO
6235571		LIDERANÇA DO PT	2017	Fornecimento de alimentação do parlamentar	RESTAURANTE BORGES E MAIA EIRELI - ME 21.659.052/0001-44	R\$ 569,00	
6235567		LIDERANÇA DO PT	2017	Fornecimento de alimentação do parlamentar	RESTAURANTE BORGES E MAIA EIRELI - ME 21.659.052/0001-44	R\$ 490,00	

Rosie @RosieDaSerenata Following

 Gasto suspeito de Dep. @renatoandrademg (MG). Você pode me ajudar a verificar? jarbas.serenata.ai/layers/#!/docum... #SerenataDeAmor na @CamaraDeputados

Translate from Portuguese

1:00 PM - 6 Mar 2018

26 Retweets 87 Likes

9 26 87

Tweet your reply

King ReD @redmaster · Mar 6
Replying to @RosieDaSerenata @camaradeputados @renatoandrademg
14 mil reais na gráfica seu deputado? @renatoandrademg

Translate from Portuguese

1 3

Allan. @allanunesjr · Mar 19
Replying to @RosieDaSerenata @renatoandrademg @camaradeputados
Quanto gasto desnecessário!!! Vagabundos comendo nosso dinheiro !

Translate from Portuguese

1 1

James Bispo @JamesBispo · Mar 6
Replying to @RosieDaSerenata @renatoandrademg @camaradeputados
Vamos esclarecer ?

25/04/2017 11:30:00
481117436 - INCO - BRASIL - BRFS


COMPROVANTE DE PAGAMENTO COM CDB, BARRA

CONVENIO GRU-CAIXA RECIBO EM UNIAO
Codigo de Barras 6951000001-0 21260001010-0 60000213433-3
95523162888-3


Data do pagamento 25/04/2017
Nº de Referência 914
Competência M/R/AAA 07/2016
Data de Vencimento 24/04/2017
CPF 404.804.706-00
Valor Principal 121,26
Valor em Dinheiro 121,26
Valor em Cheque 8,00
Valor Total 121,26

NR.AUTENTICACAO 7.844.000.640.648.1FC

Gerado a partir de http://consulta.tesouro.fazenda.gov.br/mg_inco/legislimpa_jarbas.asp SR. CONTRIBUINTE: ESTA DIXA NÃO PODERÁ SER LIQUIDADA COM CHEQUE

	MINISTÉRIO DA FAZENDA	Código de Recolhimento	28886-1
	SECRETARIA DO TESOUREO NACIONAL	Numero de Referência	914
	Guia de Recolhimento da União - GRU	Competência	07/2016
		Vencimento	24/04/2017
Nome do Contribuinte / Prestador	CPF ou CPF do Contribuinte	404.804.706-00	
Nome da Unidade Fiscalizadora	UO / Gestão	010090 / 00001	
FUNDO ROTATIVO DA CAMARA DOS DEPUTADOS			
Instruções: As informações inseridas nessa guia são de exclusiva responsabilidade do contribuinte, que deverá, em caso de dúvidas, consultar a Unidade Favorecida dos recibos.		(*) Valor do Principal	121,26
SR. CAIXA: NÃO RECEBER EM CHEQUE		(1) Descontos/Abatimentos	
		(2) Outras deduções	
		(*) Muta / Multa	
		(*) Juros / Encargos	
		(*) Outras Atribuições	
		(*) Valor Total	121,26

89910000001-0 21260001010-0 95523162888-3 60000213433-3

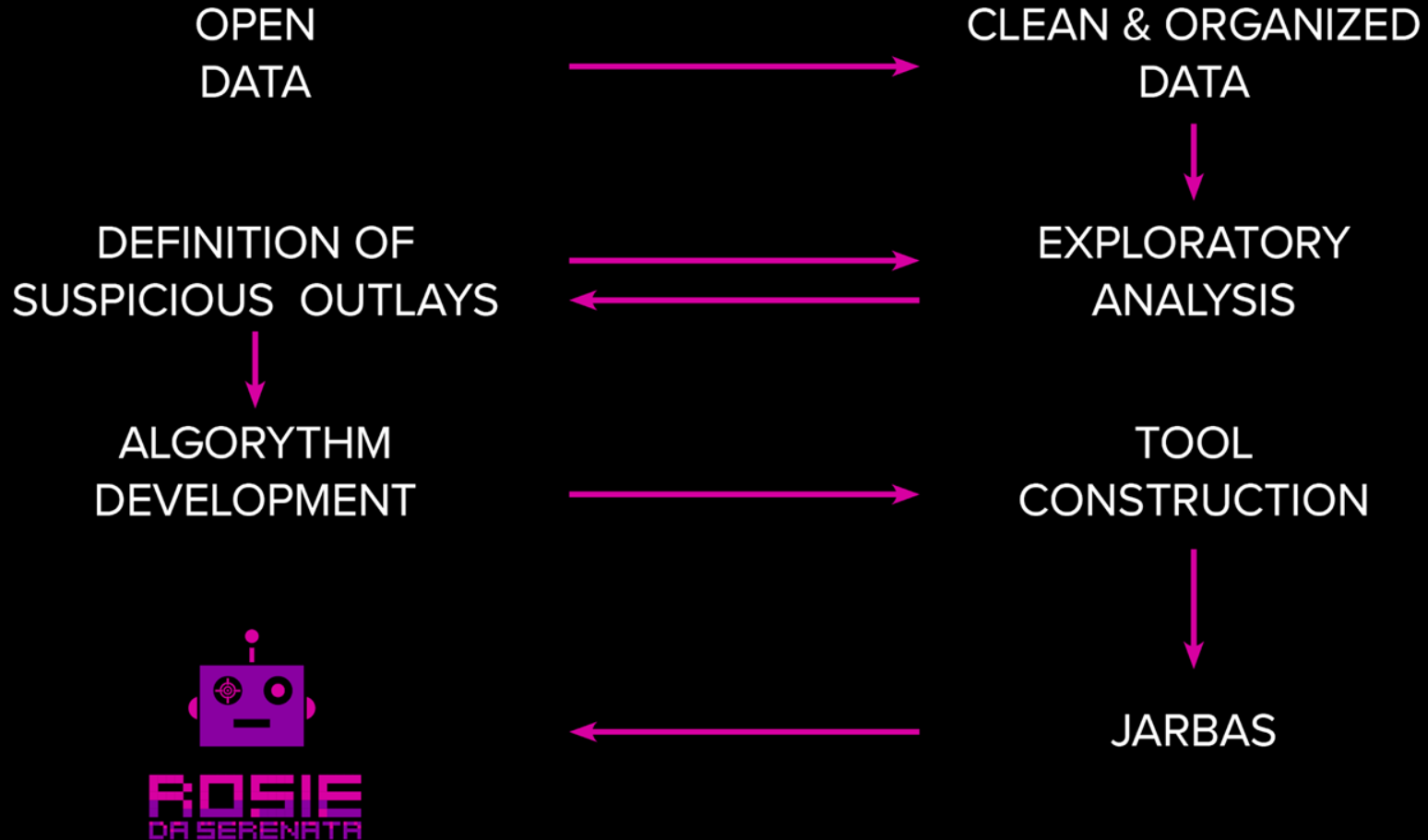


Laudívio Carvalho @laudivio

Replying to @RosieDaSerenata @camaradeputados
Constatado erro na emissão da nota e despesas foram ressarcidas conforme GRU anexa. Ofício 143/2017, foi enviado à sra. Rosineide C. Lopes

4:46 PM - Aug 11, 2017

14 See Laudívio Carvalho's other Tweets



OPERAÇÃO

**SERENATA
DE AMOR**

WHERE HAVE YOU SEEN SERENATA?



- + **DE 50** LECTURES, WORKSHOPS, SPRINTS
- + **DE 60** NEWS, MAGAZINES & ONLINE ARTICLES
- + **DE 15** PODCASTS
- + **DE 10** TV APPEARANCES

WHERE HAVE YOU SEEN US?

VICE

G1



EL PAÍS

FOLHA DE S.PAULO
★ ★ ★ UM JORNAL A SERVIÇO DO BRASIL

CBN

JN

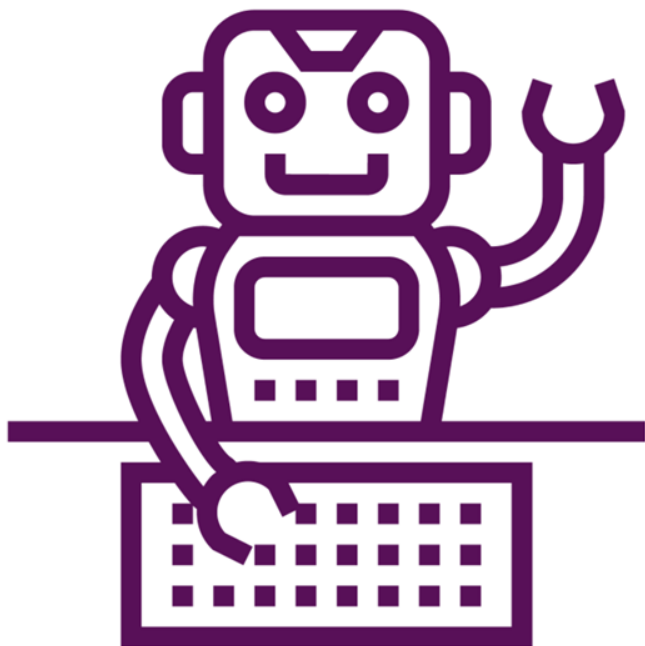
NEXT STEP: MUNICIPALITIES





QUERIDO DIÁRIO

INSCREVA-SE



UMA INTELIGÊNCIA ARTIFICIAL PARA ACOMPANHAR AS COMPRAS PÚBLICAS MUNICIPAIS

Com a **Operação Serenata de Amor**, já conseguimos mostrar que a inteligência artificial pode ser uma aliada no combate à corrupção ao analisar os gastos da cota parlamentar dos deputados federais. Agora queremos uma fatia maior do bolo.

A tecnologia usada para acessar os dados dos diários fará com que qualquer pessoa possa encontrar informações de seu interesse com facilidade, além de saber em quais outras cidades os dados também aparecem e verificar se alguma empresa citada possui irregularidades.

*Transparency as the only way
to keep track of reality and of
the policies that guide our lives*

Thank you!

Twitter: @nataliamazotte

nmazotte@stanford.edu