Using Big Data In Evaluations
What exists and how useful is it?

Jerusha Govender

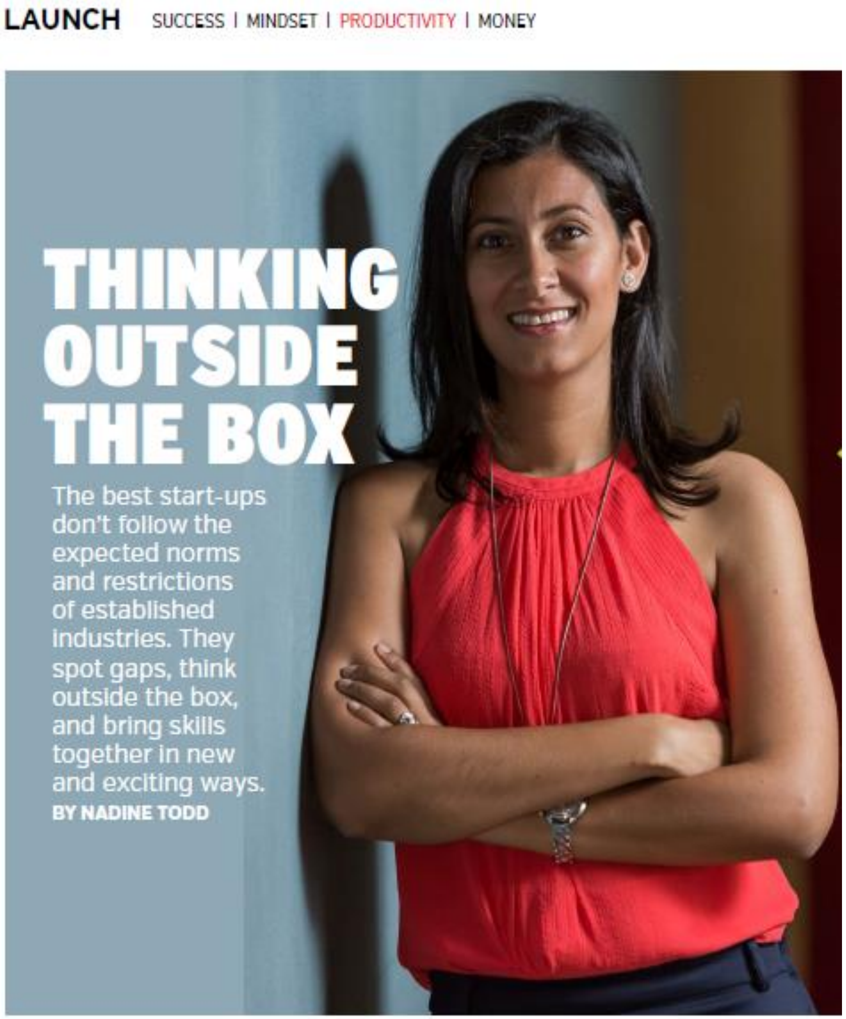
Managing Director – Data Innovator

27 June 2018

DPME – National Evaluation Seminar

planning, monitoring & evaluation
Department:
Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

DATA INNOVATOR



LAUNCH SUCCESS | MINDSET | PRODUCTIVITY | MONEY

THINKING OUTSIDE THE BOX

The best start-ups don't follow the expected norms and restrictions of established industries. They spot gaps, think outside the box, and bring skills together in new and exciting ways.

BY NADINE TODD

DATA INNOVATOR
strategic information management,
analysis & communication

SAMEA
South African Monitoring and Evaluation Association

MERL TECH

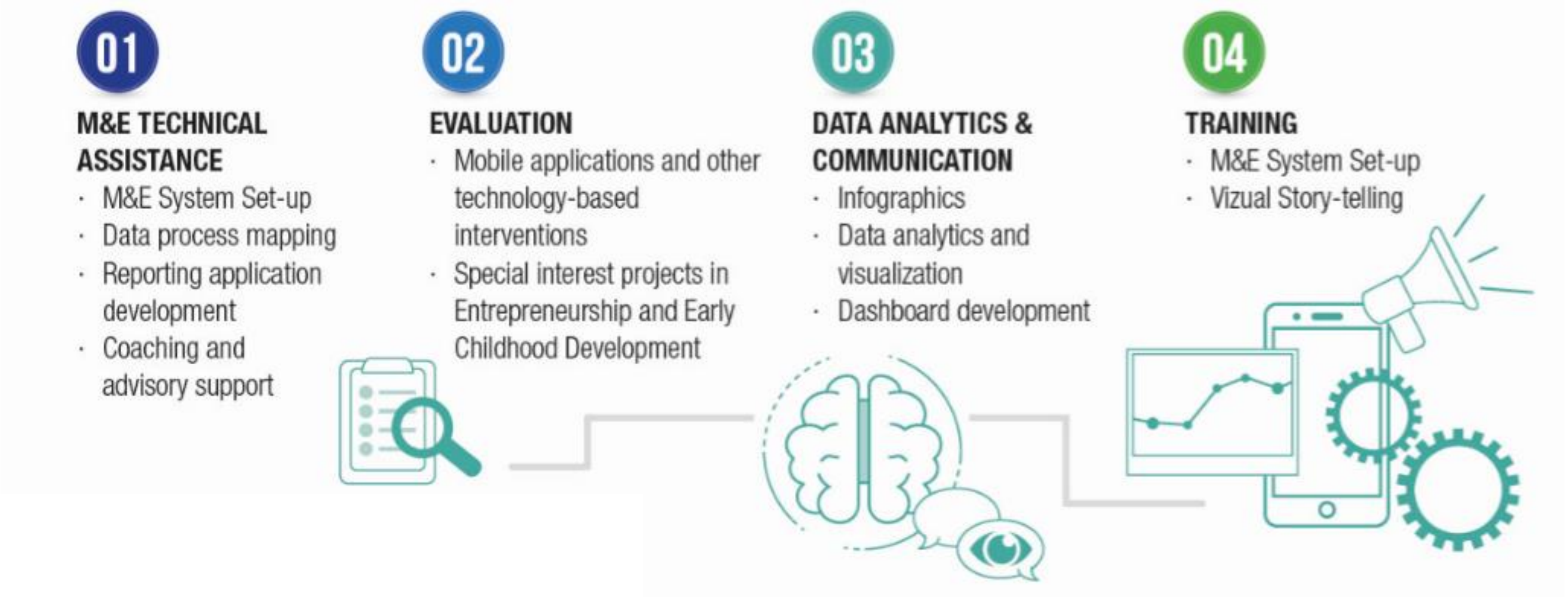
We help organizations tell their story through innovative analysis, data visualization & strategic communication

Founded in 2015

BEE Level 1

Led by team of young, African women

Socially-conscious model



01

M&E TECHNICAL ASSISTANCE

- M&E System Set-up
- Data process mapping
- Reporting application development
- Coaching and advisory support

02

EVALUATION

- Mobile applications and other technology-based interventions
- Special interest projects in Entrepreneurship and Early Childhood Development

03

DATA ANALYTICS & COMMUNICATION

- Infographics
- Data analytics and visualization
- Dashboard development

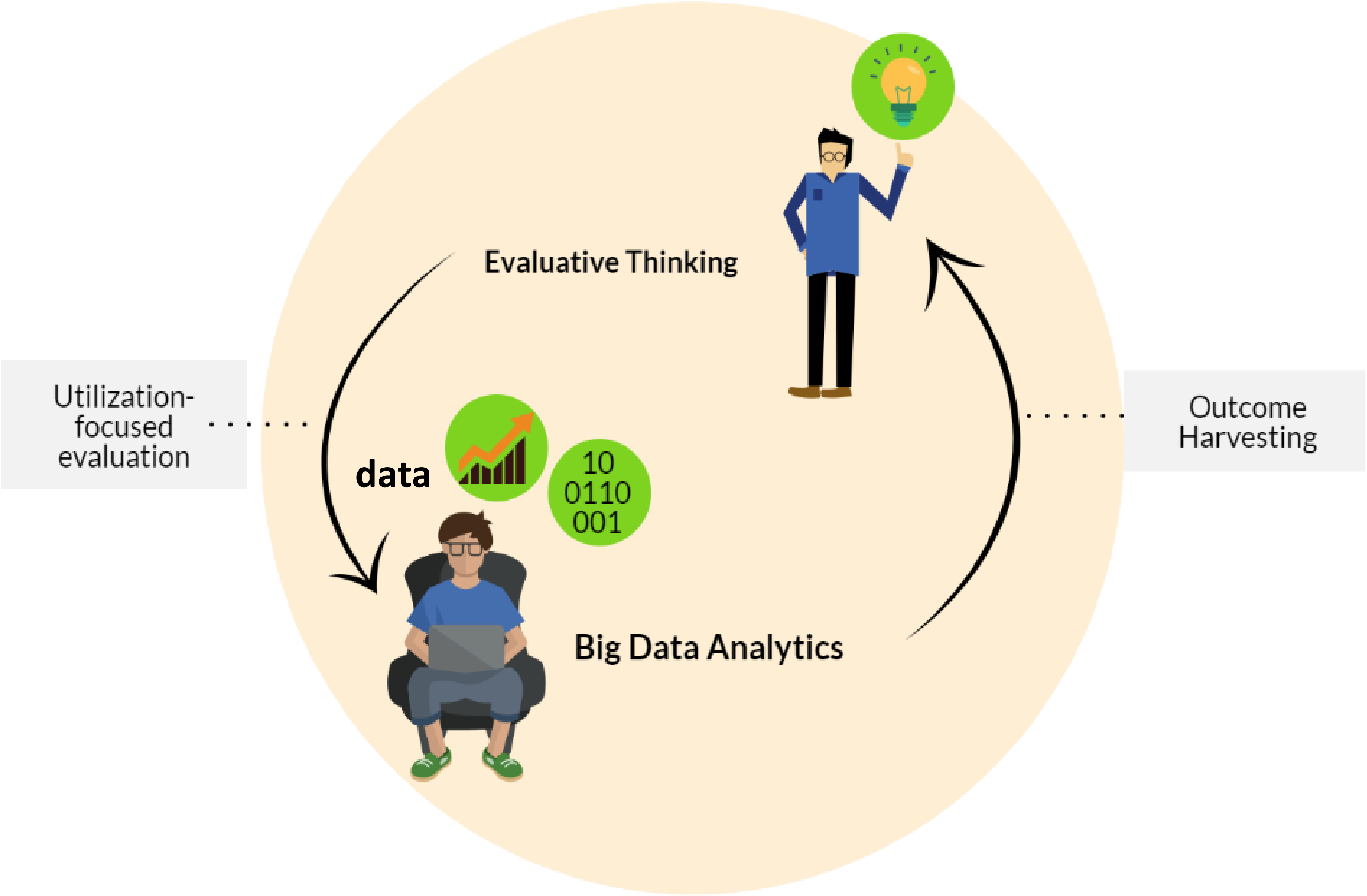
04

TRAINING

- M&E System Set-up
- Vizual Story-telling

What exists and how useful is it?

What exists and how useful is it?



Evaluative Thinking

Utilization-focused evaluation

data

Big Data Analytics

Outcome Harvesting

Big Data Ecosystem

BANK
Shop
NO PARKING

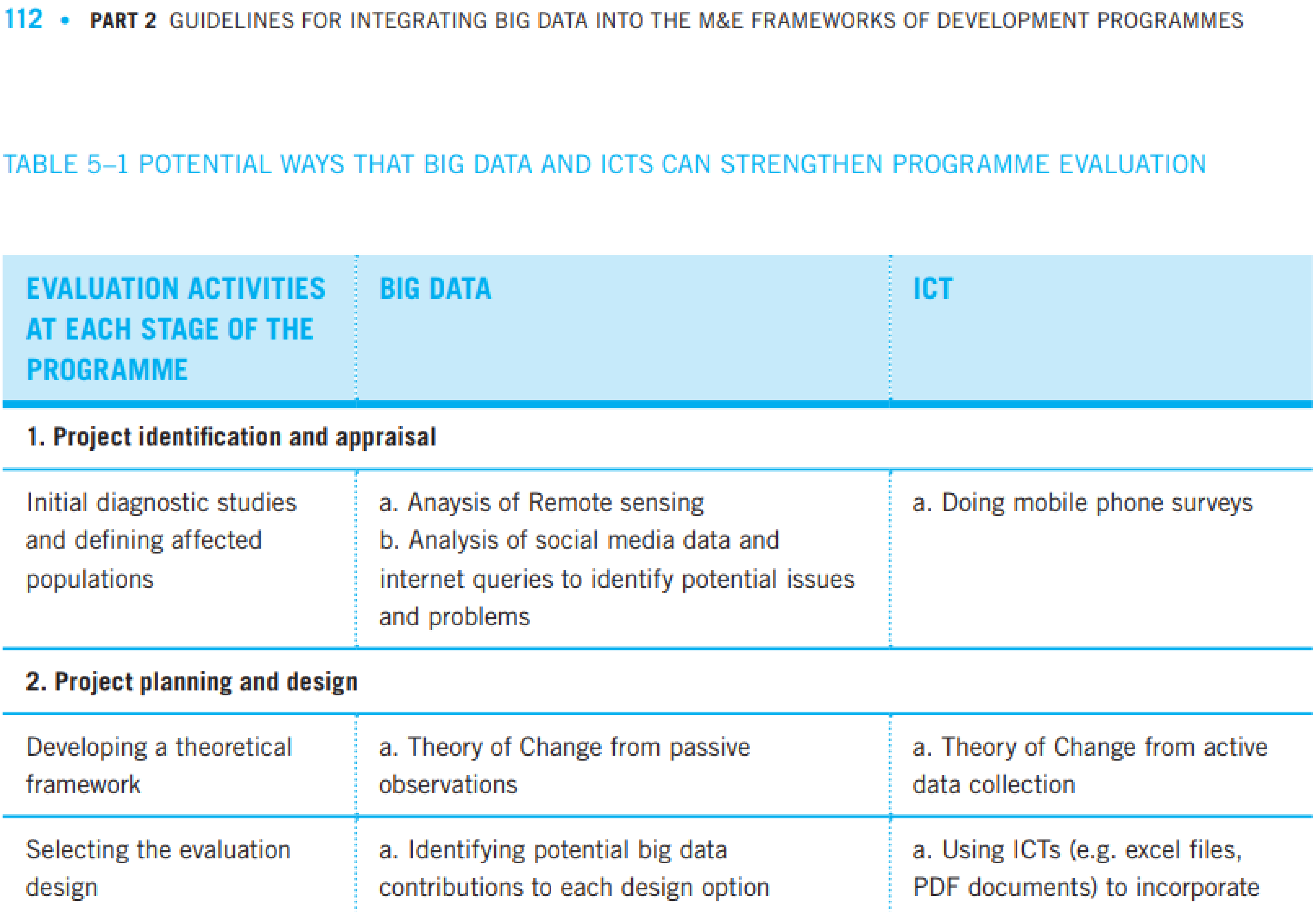
Individuals

Private Sector

Public Sector

http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf

Open Data ≠ Big Data



112 • **PART 2** GUIDELINES FOR INTEGRATING BIG DATA INTO THE M&E FRAMEWORKS OF DEVELOPMENT PROGRAMMES

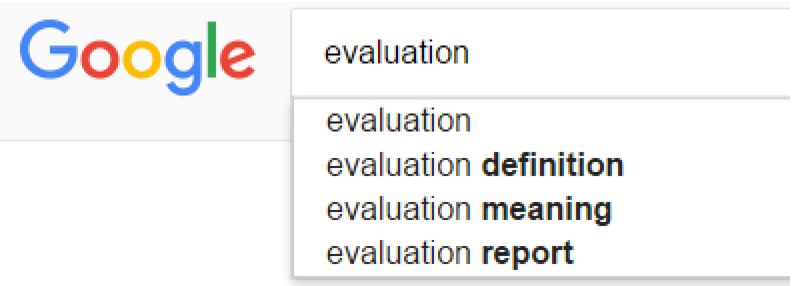
TABLE 5–1 POTENTIAL WAYS THAT BIG DATA AND ICTS CAN STRENGTHEN PROGRAMME EVALUATION

EVALUATION ACTIVITIES AT EACH STAGE OF THE PROGRAMME	BIG DATA	ICT
1. Project identification and appraisal		
Initial diagnostic studies and defining affected populations	a. Anaysis of Remote sensing b. Analysis of social media data and internet queries to identify potential issues and problems	a. Doing mobile phone surveys
2. Project planning and design		
Developing a theoretical framework	a. Theory of Change from passive observations	a. Theory of Change from active data collection
Selecting the evaluation design	a. Identifying potential big data contributions to each design option	a. Using ICTs (e.g. excel files, PDF documents) to incorporate

UN Global Pulse, ‘Integrating Big Data into the Monitoring and Evaluation of Development Programmes,’ 2016

Categories

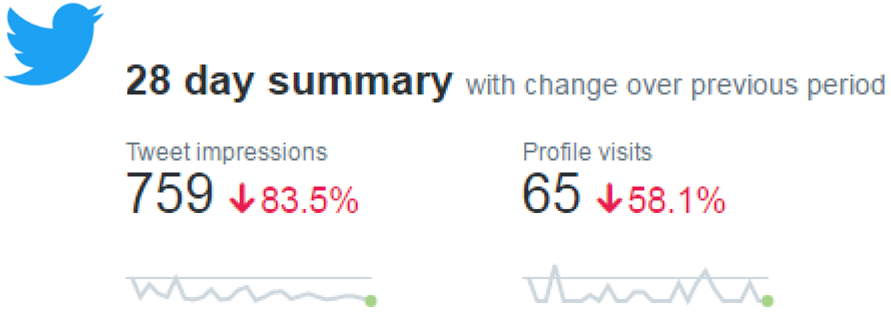
1. Data Exhaust



evaluation

- evaluation
- evaluation **definition**
- evaluation **meaning**
- evaluation **report**

2. Online Information

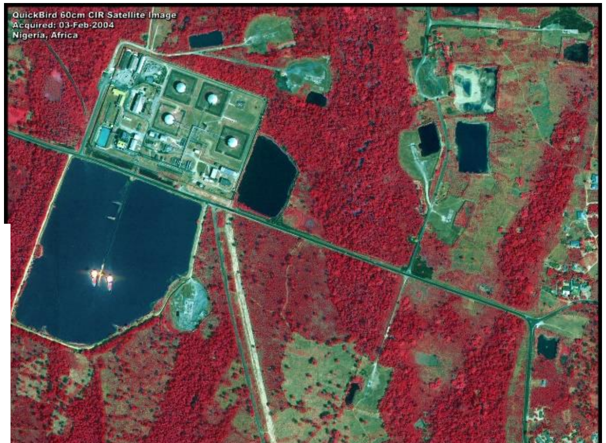


28 day summary with change over previous period

Tweet impressions
759 ↓83.5%

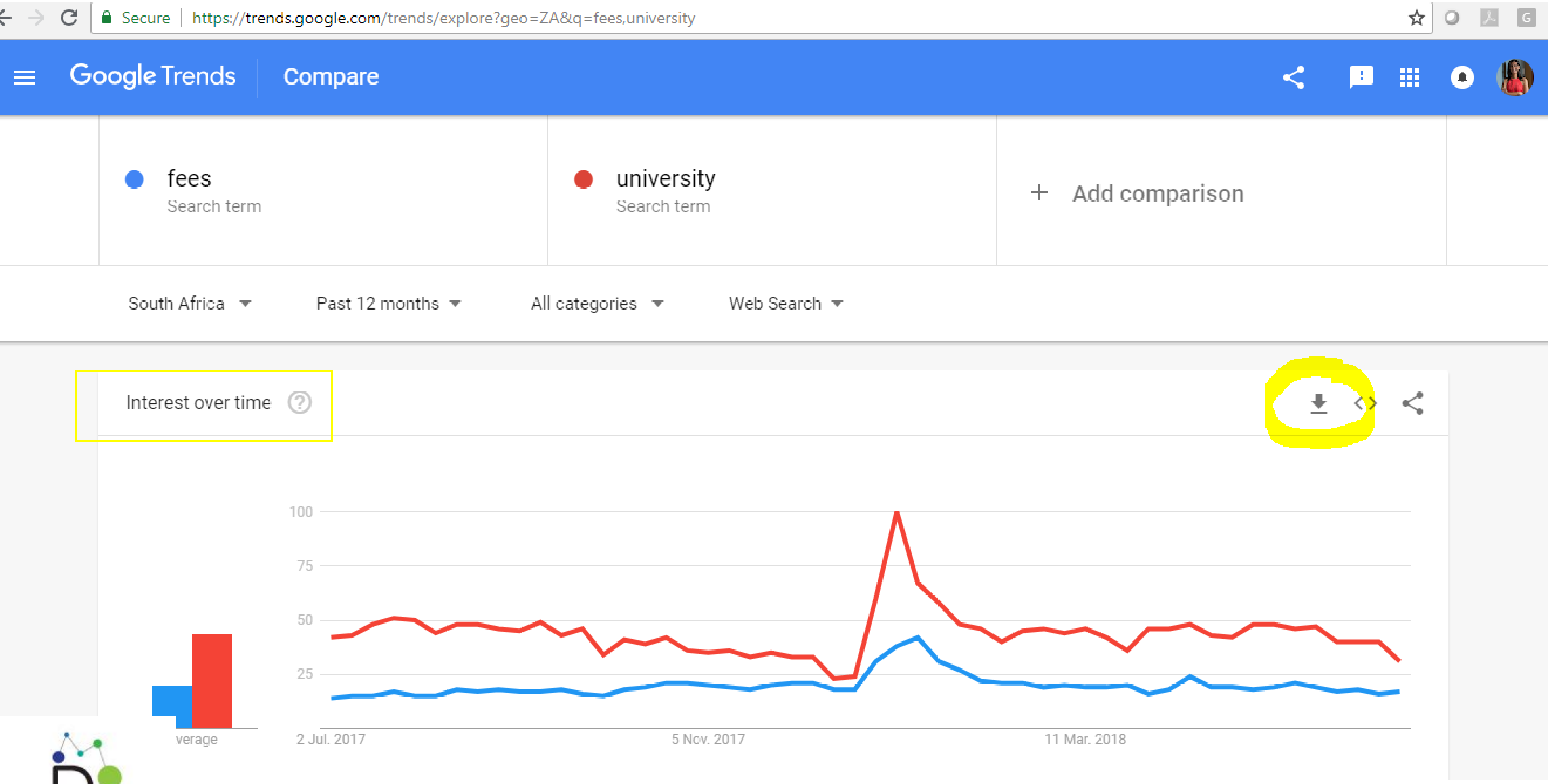
Profile visits
65 ↓58.1%

3. Physical Sensors


QuickBird 60cm CIR Satellite Image
Acquired: 03-Feb-2004
Nigeria, Africa

4. Citizen Reporting or Crowd-sourced Data

1. Data Exhaust



https://trends.google.com/trends/explore?geo=ZA&q=fees,university

● fees
Search term

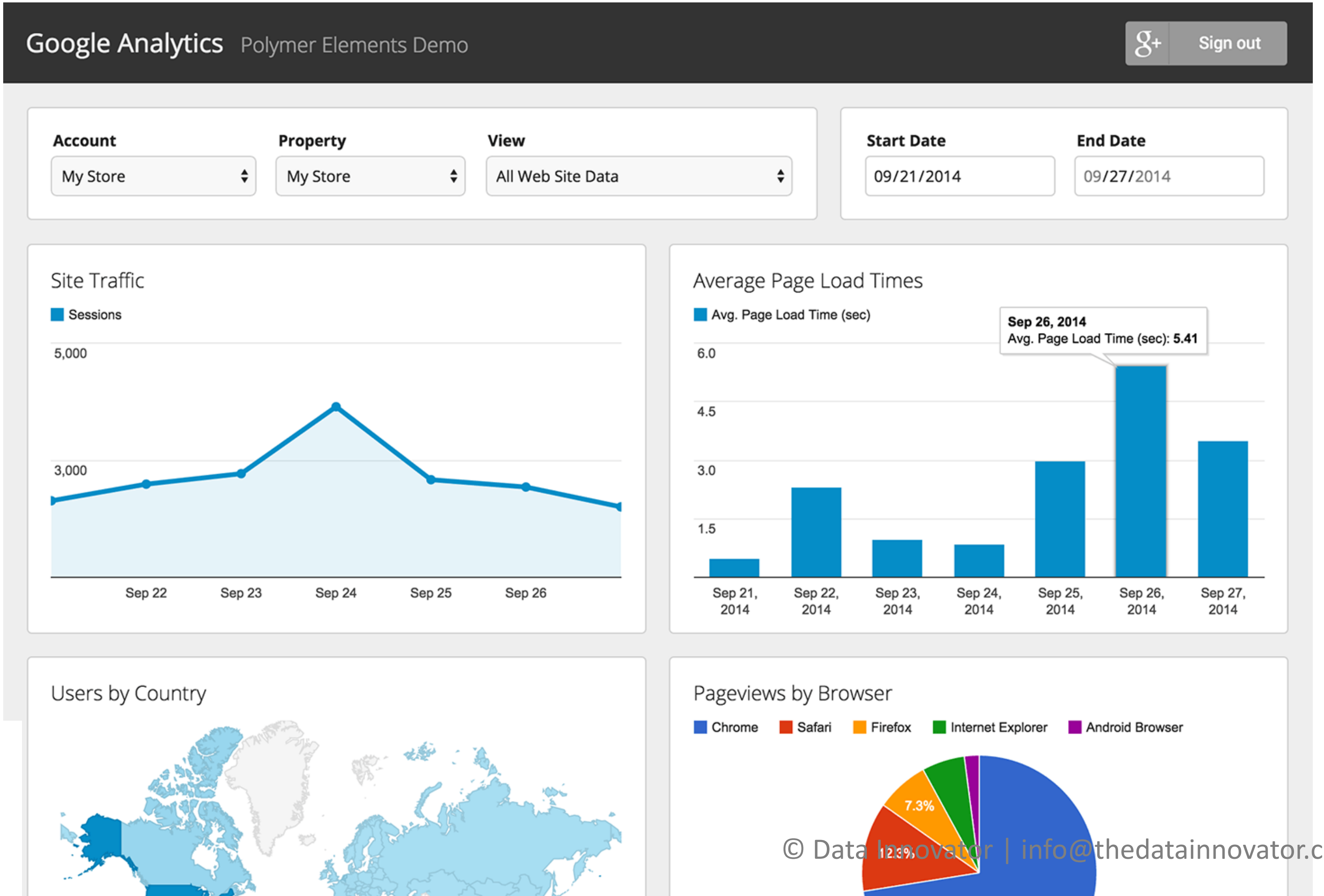
● university
Search term

+ Add comparison

South Africa
Past 12 months
All categories
Web Search

Secure
Google Trends
Compare
Interest over time
100
75
50
25
verage
2 Jul. 2017
5 Nov. 2017
11 Mar. 2018

1. Data Exhaust


Google Analytics
Polymer Elements Demo
Sign out

Account My Store

Property My Store

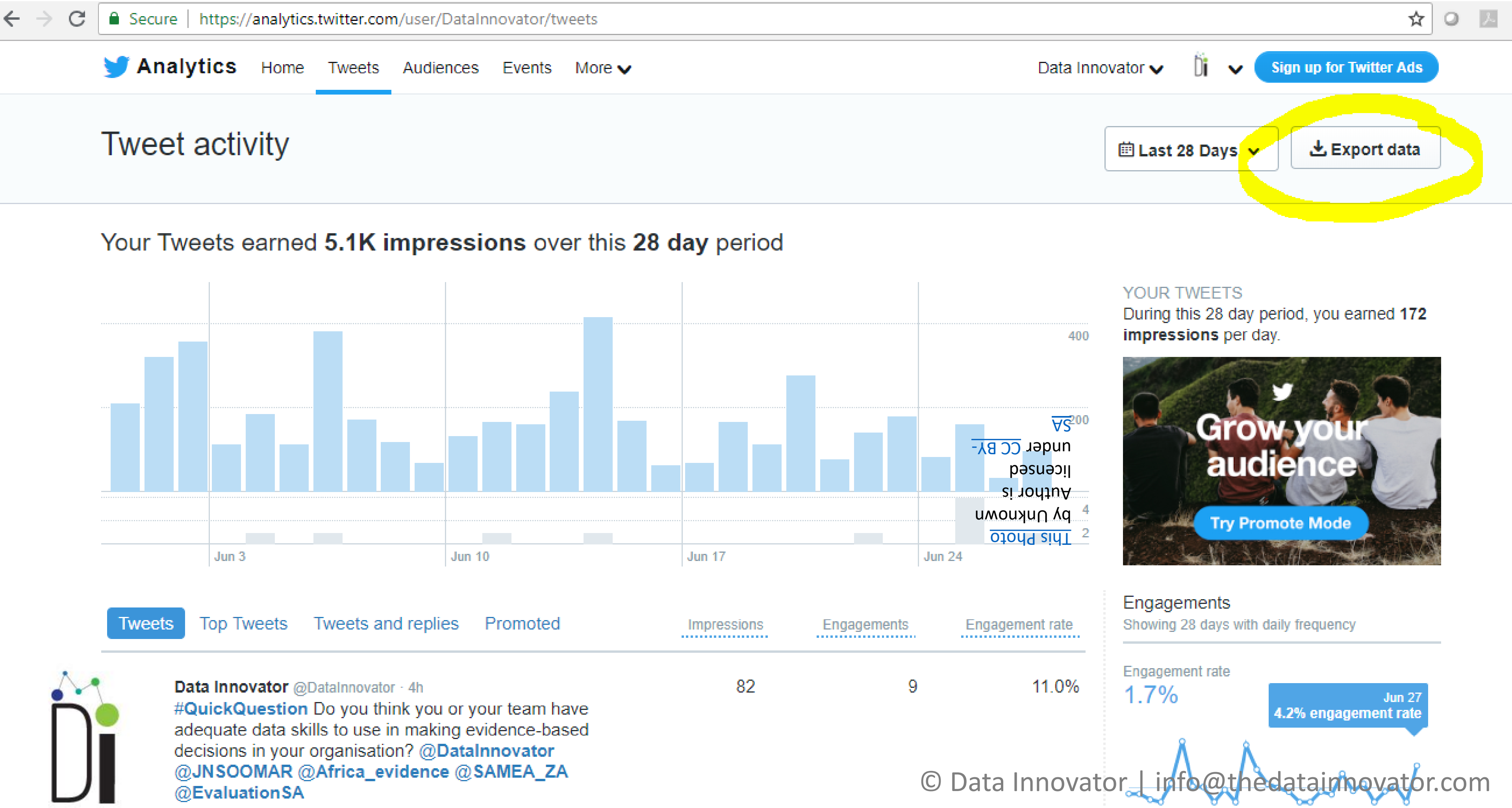
View All Web Site Data

Start Date 09/21/2014

End Date 09/27/2014

Site Traffic
Sessions
5,000
3,000
Sep 22
Sep 23
Sep 24
Sep 25
Sep 26
Average Page Load Times
Avg. Page Load Time (sec)
Sep 26, 2014
Avg. Page Load Time (sec): 5.41
6.0
4.5
3.0
1.5
Sep 21, 2014
Sep 22, 2014
Sep 23, 2014
Sep 24, 2014
Sep 25, 2014
Sep 26, 2014
Sep 27, 2014
Users by Country
Pageviews by Browser
Chrome
Safari
Firefox
Internet Explorer
Android Browser
7.3%
12.9%

2. Online Information


Secure
https://analytics.twitter.com/user/DataInnovator/tweets
Analytics
Home
Tweets
Audiences
Events
More
Data Innovator
Sign up for Twitter Ads
Last 28 Days
Export data
400
200
4
2
Jun 3
Jun 10
Jun 17
Jun 24
Grow your audience
Try Promote Mode
Tweets
Top Tweets
Tweets and replies
Promoted
Impressions
Engagements
Engagement rate

Tweet activity

Your Tweets earned **5.1K impressions** over this **28 day** period

YOUR TWEETS

During this 28 day period, you earned **172 impressions** per day.

Data Innovator @DataInnovator · 4h

#QuickQuestion Do you think you or your team have adequate data skills to use in making evidence-based decisions in your organisation? @DataInnovator @JNSOOMAR @Africa_evidence @SAMEA_ZA @EvaluationSA

Impressions	Engagements	Engagement rate
82	9	11.0%

Engagements

Showing 28 days with daily frequency

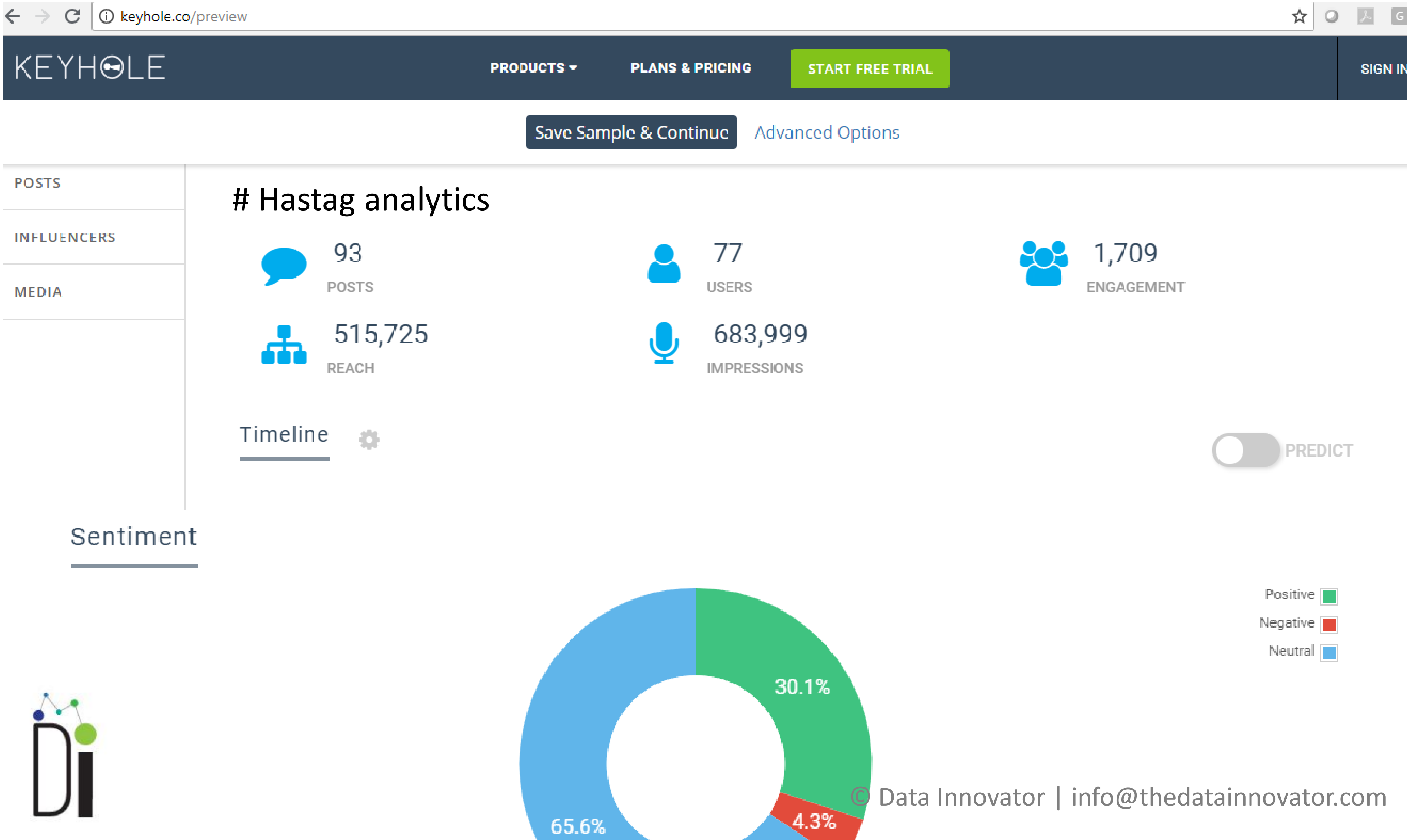
Engagement rate

1.7%

Jun 27
4.2% engagement rate

This Photo by Unknown Author is licensed under CC BY-SA

2. Online Information



keyhole.co/preview
KEYHOLE
PRODUCTS ▾
PLANS & PRICING
START FREE TRIAL
SIGN IN
Save Sample & Continue
Advanced Options
POSTS
INFLUENCERS
MEDIA
Timeline
PREDICT
Sentiment
Positive 🟩
Negative 🟥
Neutral 🟦
30.1%
4.3%
65.6%

# Hastag analytics

93 POSTS

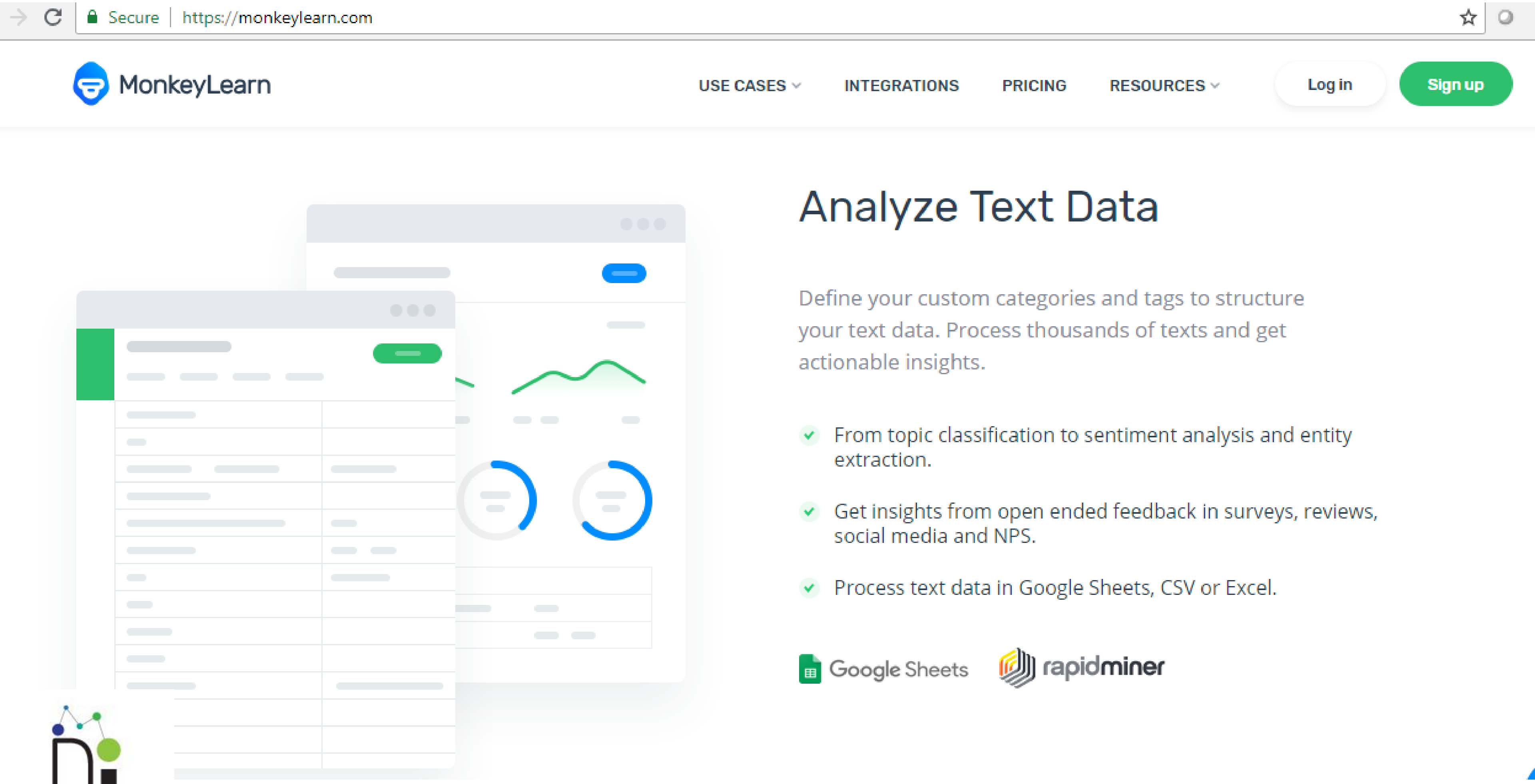
77 USERS

1,709 ENGAGEMENT

515,725 REACH

683,999 IMPRESSIONS

2. Online Information


Secure
https://monkeylearn.com
MonkeyLearn
USE CASES
INTEGRATIONS
PRICING
RESOURCES
Log in
Sign up
Google Sheets
rapidminer

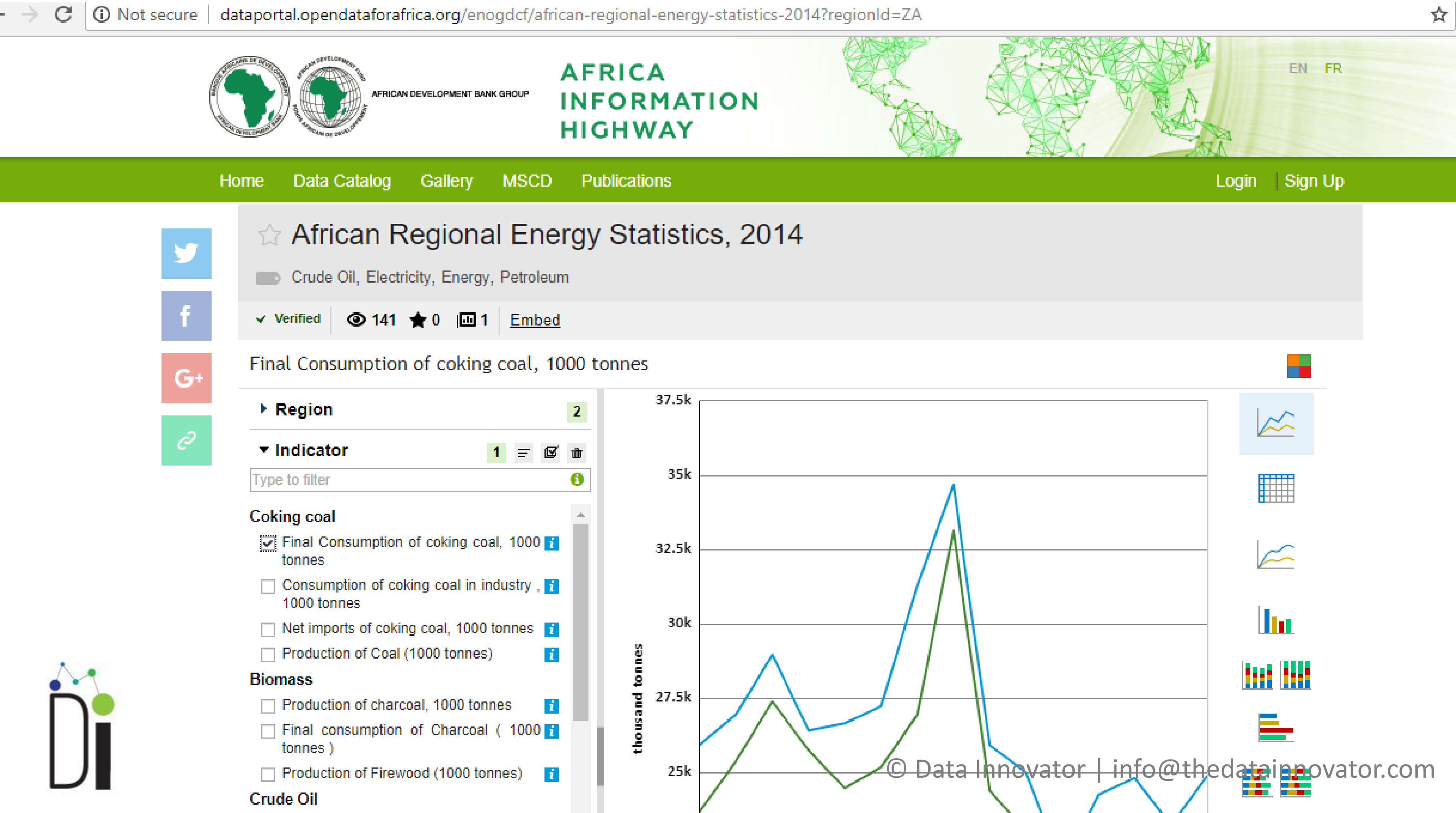
Analyze Text Data

Define your custom categories and tags to structure your text data. Process thousands of texts and get actionable insights.

- ✓ From topic classification to sentiment analysis and entity extraction.
- ✓ Get insights from open ended feedback in surveys, reviews, social media and NPS.
- ✓ Process text data in Google Sheets, CSV or Excel.

2. Online Information

Open Data



Not secure
dataportal.opendataforafrica.org/enogdcf/african-regional-energy-statistics-2014?regionId=ZA
AFRICAN DEVELOPMENT BANK GROUP
AFRICA INFORMATION HIGHWAY
EN FR
Home
Data Catalog
Gallery
MSCD
Publications
Login
Sign Up
Verified
141
0
1
Embed
Region 2
Indicator 1
Type to filter
37.5k
35k
32.5k
30k
27.5k
25k
thousand tonnes

African Regional Energy Statistics, 2014

Crude Oil, Electricity, Energy, Petroleum

Final Consumption of coking coal, 1000 tonnes

Coking coal

- ☑ Final Consumption of coking coal, 1000 tonnes
- ☐ Consumption of coking coal in industry , 1000 tonnes
- ☐ Net imports of coking coal, 1000 tonnes
- ☐ Production of Coal (1000 tonnes)

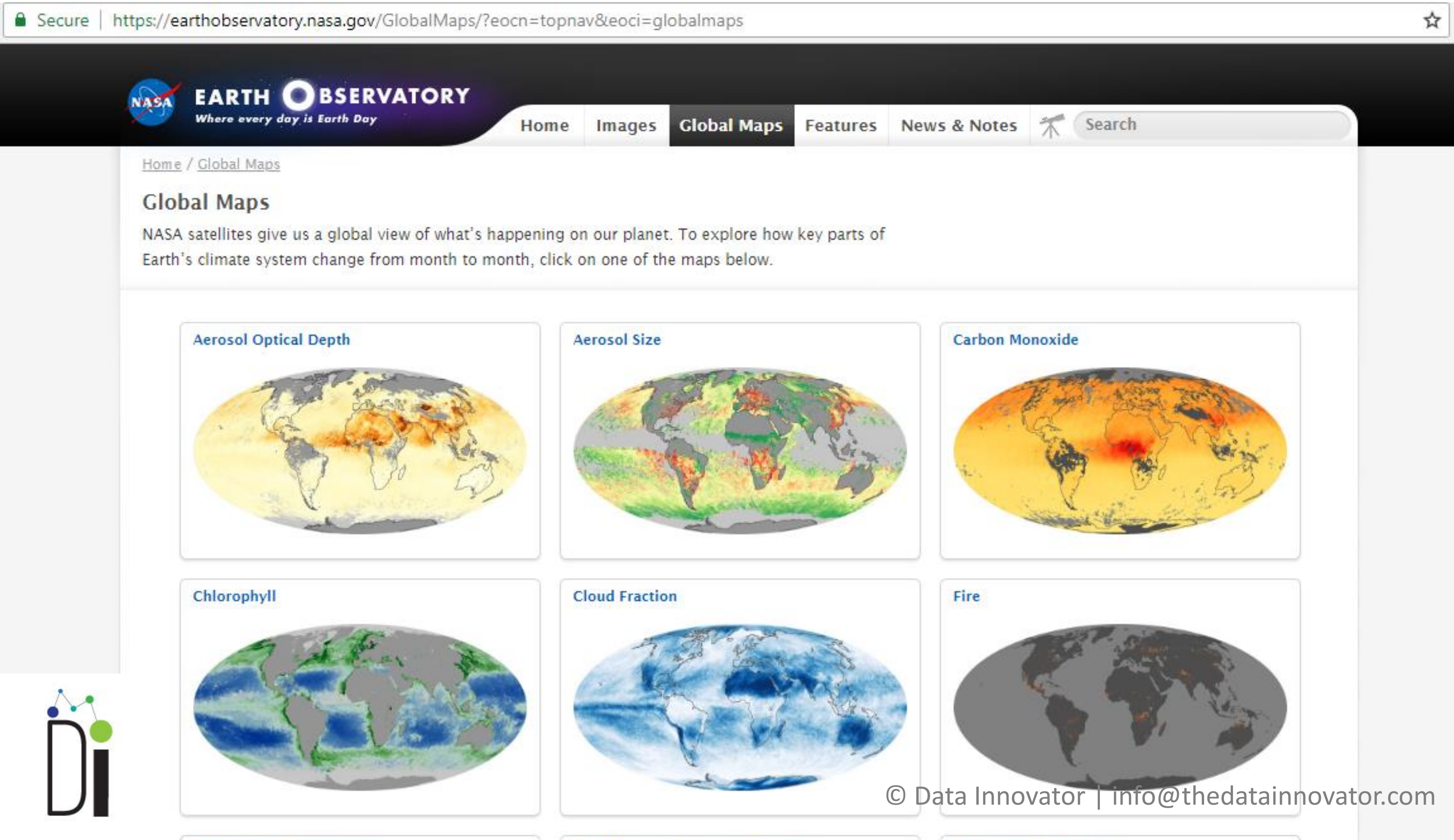
Biomass

- ☐ Production of charcoal, 1000 tonnes
- ☐ Final consumption of Charcoal (1000 tonnes)
- ☐ Production of Firewood (1000 tonnes)

Crude Oil

3. Physical Sensors

Proxies/Direct



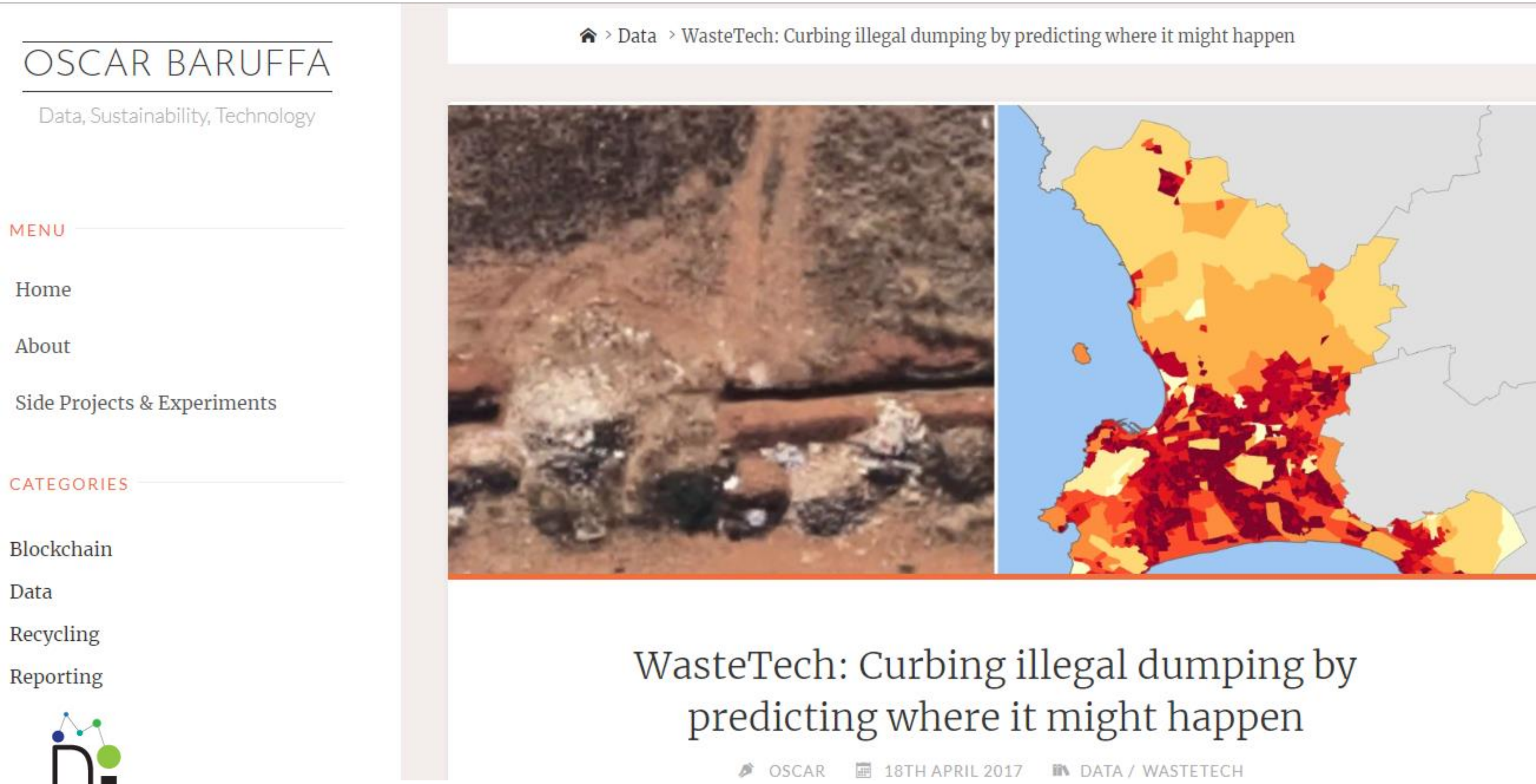
Secure
https://earthobservatory.nasa.gov/GlobalMaps/?eocn=topnav&eoci=globalmaps
NASA
EARTH OBSERVATORY
Where every day is Earth Day
Home
Images
Global Maps
Features
News & Notes
Search
Home / Global Maps

Global Maps

NASA satellites give us a global view of what's happening on our planet. To explore how key parts of Earth's climate system change from month to month, click on one of the maps below.

Aerosol Optical Depth
Aerosol Size
Carbon Monoxide
Chlorophyll
Cloud Fraction
Fire

3. Physical Sensors



OSCAR BARUFFA

Data, Sustainability, Technology

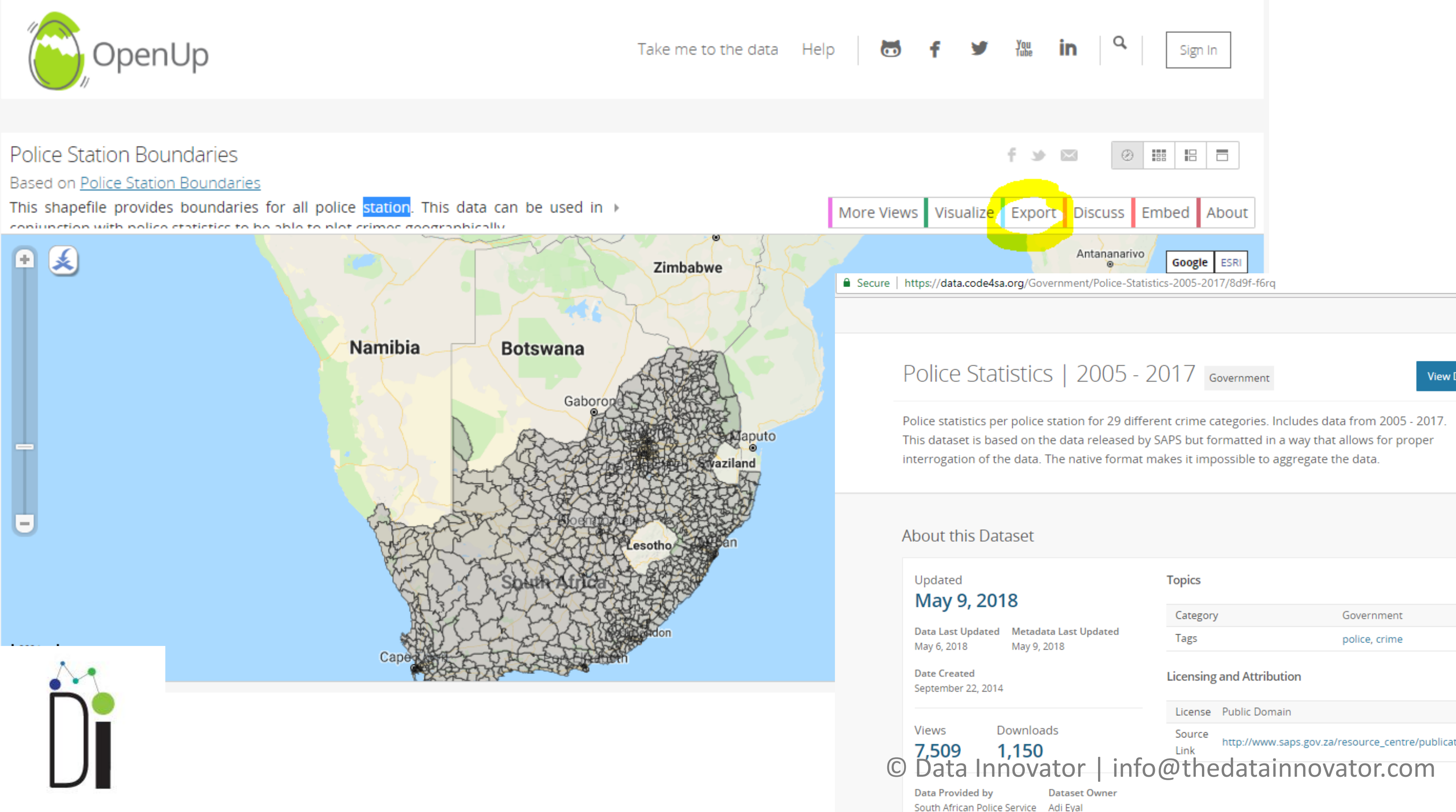
MENU
Home
About
Side Projects & Experiments
CATEGORIES
Blockchain
Data
Recycling
Reporting

> Data > WasteTech: Curbing illegal dumping by predicting where it might happen

WasteTech: Curbing illegal dumping by predicting where it might happen

OSCAR 18TH APRIL 2017 DATA / WASTETECH

3. Physical Sensors


OpenUp
Take me to the data
Help
Sign In
More Views
Visualize
Export
Discuss
Embed
About
Google
ESRI
Zimbabwe
Namibia
Botswana
Gaborone
Maputo
Swaziland
Lesotho
South Africa
Antananarivo
Cape Town
Secure

Police Station Boundaries

Based on Police Station Boundaries

This shapefile provides boundaries for all police station. This data can be used in

https://data.code4sa.org/Government/Police-Statistics-2005-2017/8d9f-f6rq

Police Statistics | 2005 - 2017 Government

View D

Police statistics per police station for 29 different crime categories. Includes data from 2005 - 2017. This dataset is based on the data released by SAPS but formatted in a way that allows for proper interrogation of the data. The native format makes it impossible to aggregate the data.

About this Dataset

Updated
May 9, 2018

Data Last Updated
May 6, 2018

Metadata Last Updated
May 9, 2018

Date Created
September 22, 2014

Views
7,509

Downloads
1,150

Data Provided by
South African Police Service

Dataset Owner
Adi Eyal

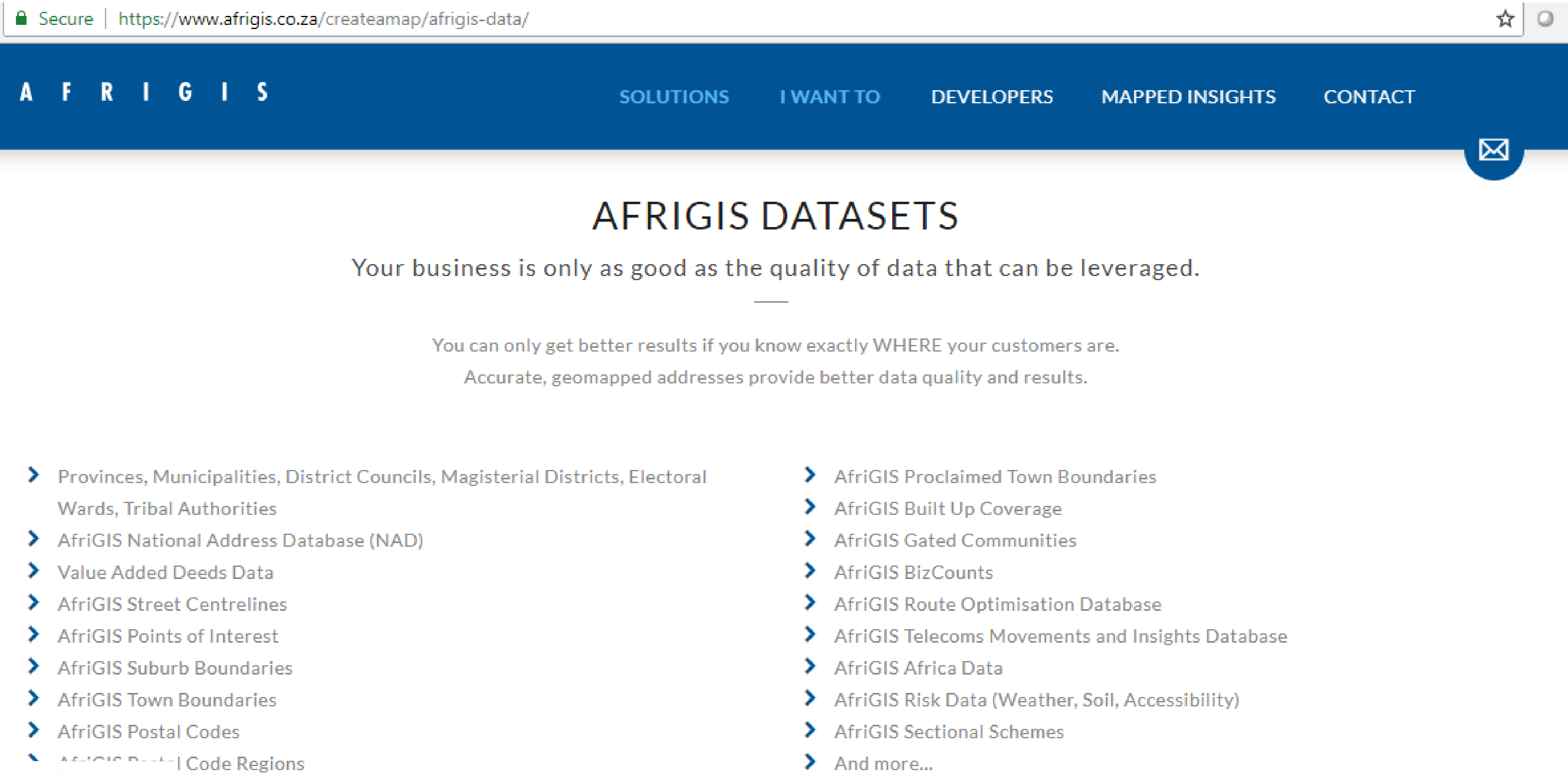
Topics

Category	Government
Tags	police, crime

Licensing and Attribution

License	Public Domain
Source Link	http://www.saps.gov.za/resource_centre/publicat

3. Physical Sensors


Secure
https://www.afrigis.co.za/createamap/afrigis-data/
AFRIGIS
SOLUTIONS
I WANT TO
DEVELOPERS
MAPPED INSIGHTS
CONTACT

AFRIGIS DATASETS

Your business is only as good as the quality of data that can be leveraged.

You can only get better results if you know exactly WHERE your customers are.
Accurate, geomapped addresses provide better data quality and results.

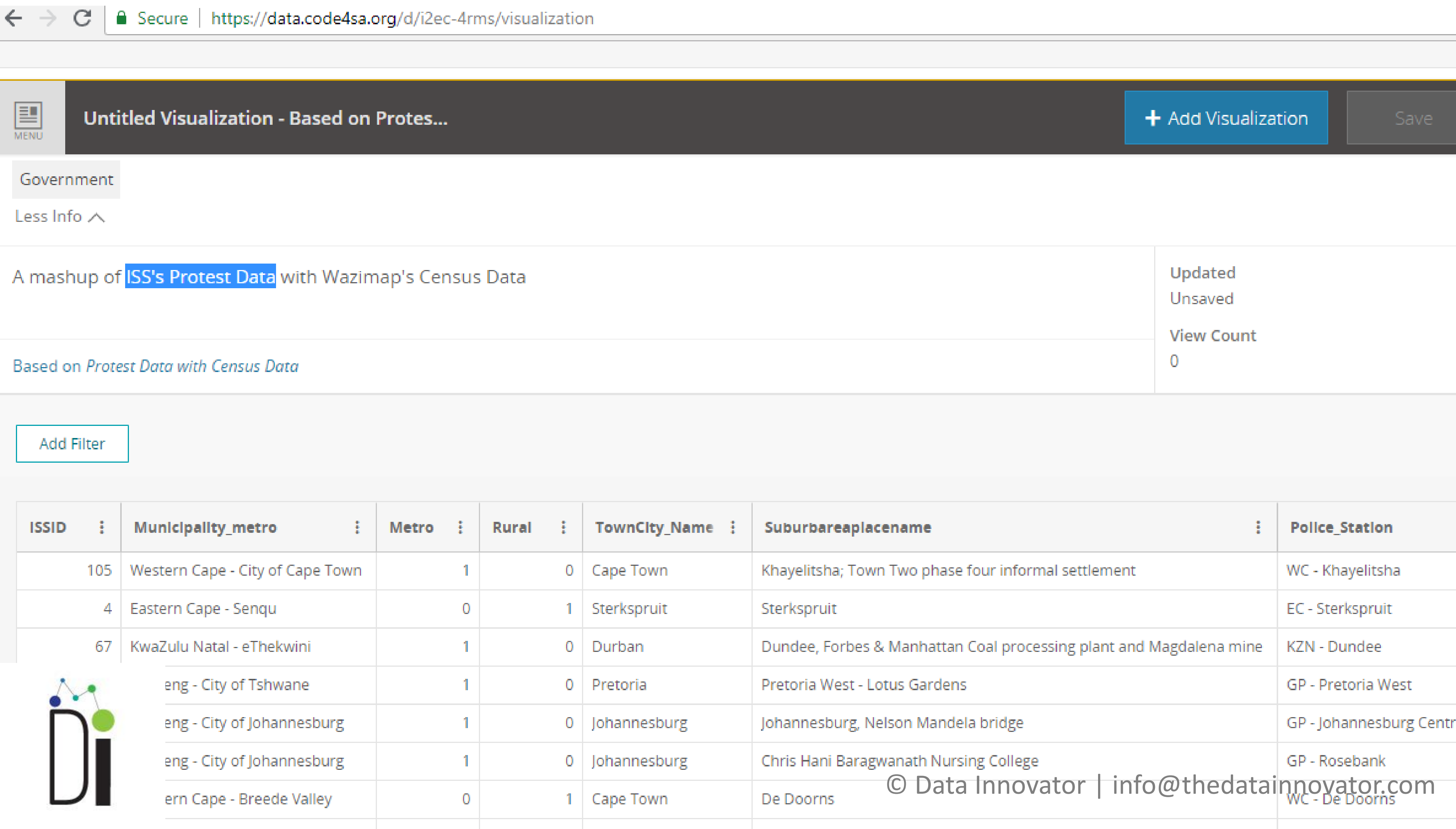
- Provinces, Municipalities, District Councils, Magisterial Districts, Electoral Wards, Tribal Authorities
- AfriGIS National Address Database (NAD)
- Value Added Deeds Data
- AfriGIS Street Centrelines
- AfriGIS Points of Interest
- AfriGIS Suburb Boundaries
- AfriGIS Town Boundaries
- AfriGIS Postal Codes
- AfriGIS Postal Code Regions
- AfriGIS Proclaimed Town Boundaries
- AfriGIS Built Up Coverage
- AfriGIS Gated Communities
- AfriGIS BizCounts
- AfriGIS Route Optimisation Database
- AfriGIS Telecoms Movements and Insights Database
- AfriGIS Africa Data
- AfriGIS Risk Data (Weather, Soil, Accessibility)
- AfriGIS Sectional Schemes
- And more...

4. Citizen Reporting/Crowd-sourcing

Online Polls

Hotlines – e.g. **National Commuter Hotline**

4. Citizen Reporting/Crowd-sourcing



Untitled Visualization - Based on Protes...

Government

Less Info

A mashup of ISS's Protest Data with Wazimap's Census Data

Based on *Protest Data with Census Data*

Updated
Unsaved

View Count
0

Add Filter

ISSID	Municipality_metro	Metro	Rural	TownCity_Name	Suburbareaplacename	Police_Station
105	Western Cape - City of Cape Town	1	0	Cape Town	Khayelitsha; Town Two phase four informal settlement	WC - Khayelitsha
4	Eastern Cape - Senqu	0	1	Sterkspruit	Sterkspruit	EC - Sterkspruit
67	KwaZulu Natal - eThekwini	1	0	Durban	Dundee, Forbes & Manhattan Coal processing plant and Magdalena mine	KZN - Dundee
	eng - City of Tshwane	1	0	Pretoria	Pretoria West - Lotus Gardens	GP - Pretoria West
	eng - City of Johannesburg	1	0	Johannesburg	Johannesburg, Nelson Mandela bridge	GP - Johannesburg Centr
	eng - City of Johannesburg	1	0	Johannesburg	Chris Hani Baragwanath Nursing College	GP - Rosebank
	ern Cape - Breede Valley	0	1	Cape Town	De Doorns	WC - De Doorns

What exists and
how useful is it?

> “…Big data also means new challenges involving **complexity, security, and risks** to privacy, as well as a need for new technology and human skills.”

Kim, G. Trimi S. and Chung J. Big-data applications in the government sector. 2014. Communications of the ACM. Vol 57. No. 3

Is the data **reliable**?

Does this data support my **evaluation question**?

Is the **format** of the data appropriate or useable?

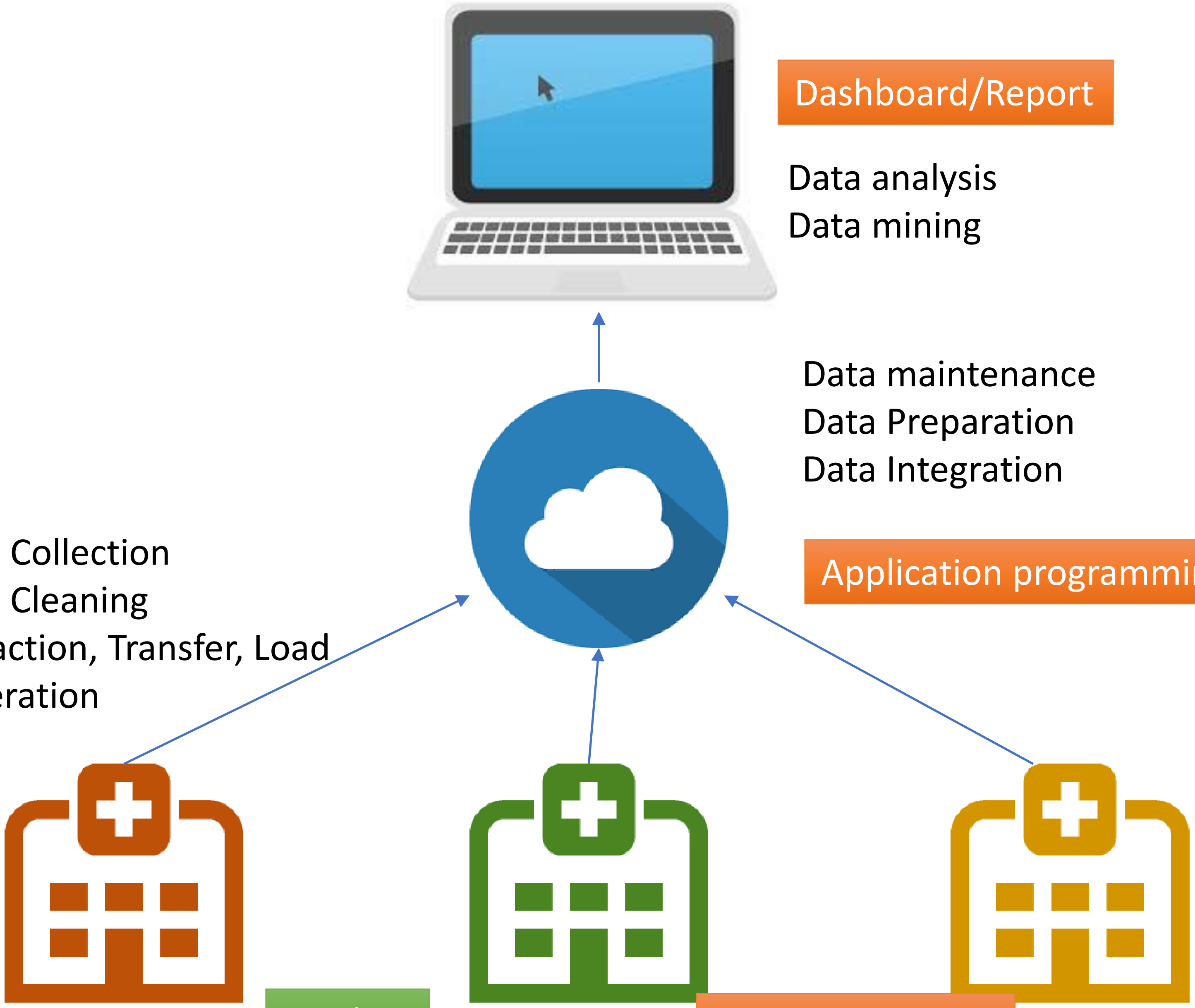
Can we **access** the data?

Is it **ethical** to use this data?

://exceednutrition.com/the-shopping-trolley-conundrum/

“Big data: Not a product, but a collection of **processes**”

Source: Irina Shklovski, Technologies in Practice (TiP), IT University of Copenhagen, presentation at ILO Big Data Symposium



Dashboard/Report

Data analysis
Data mining

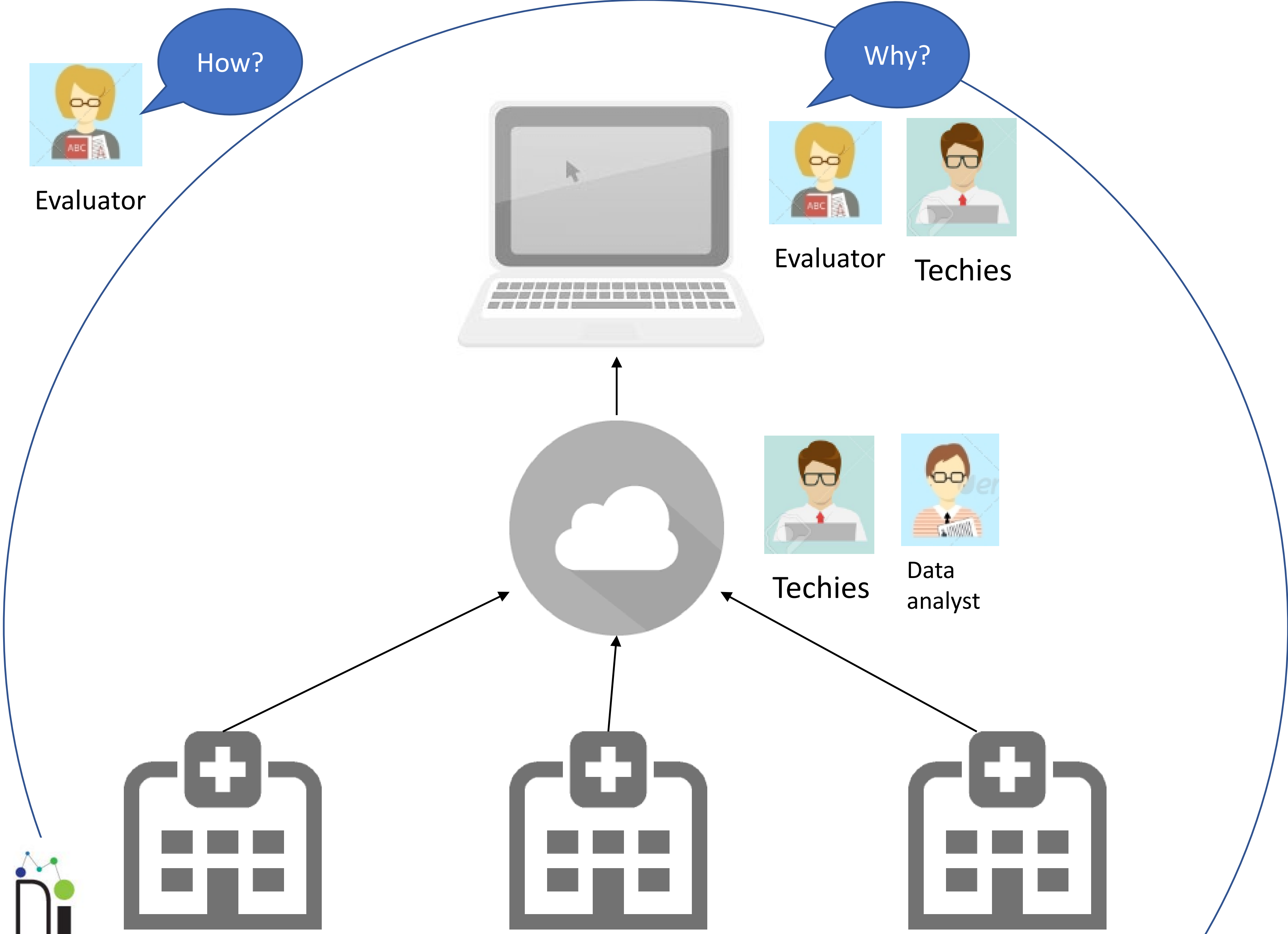
Data maintenance
Data Preparation
Data Integration

Application programming interface

Data Collection
Data Cleaning
Extraction, Transfer, Load
Federation

Big data

Various platforms



How?

Evaluator

Why?

Evaluator

Techies

Techies

Data analyst

Big Data Analytics is useful…
the data is sometimes useful

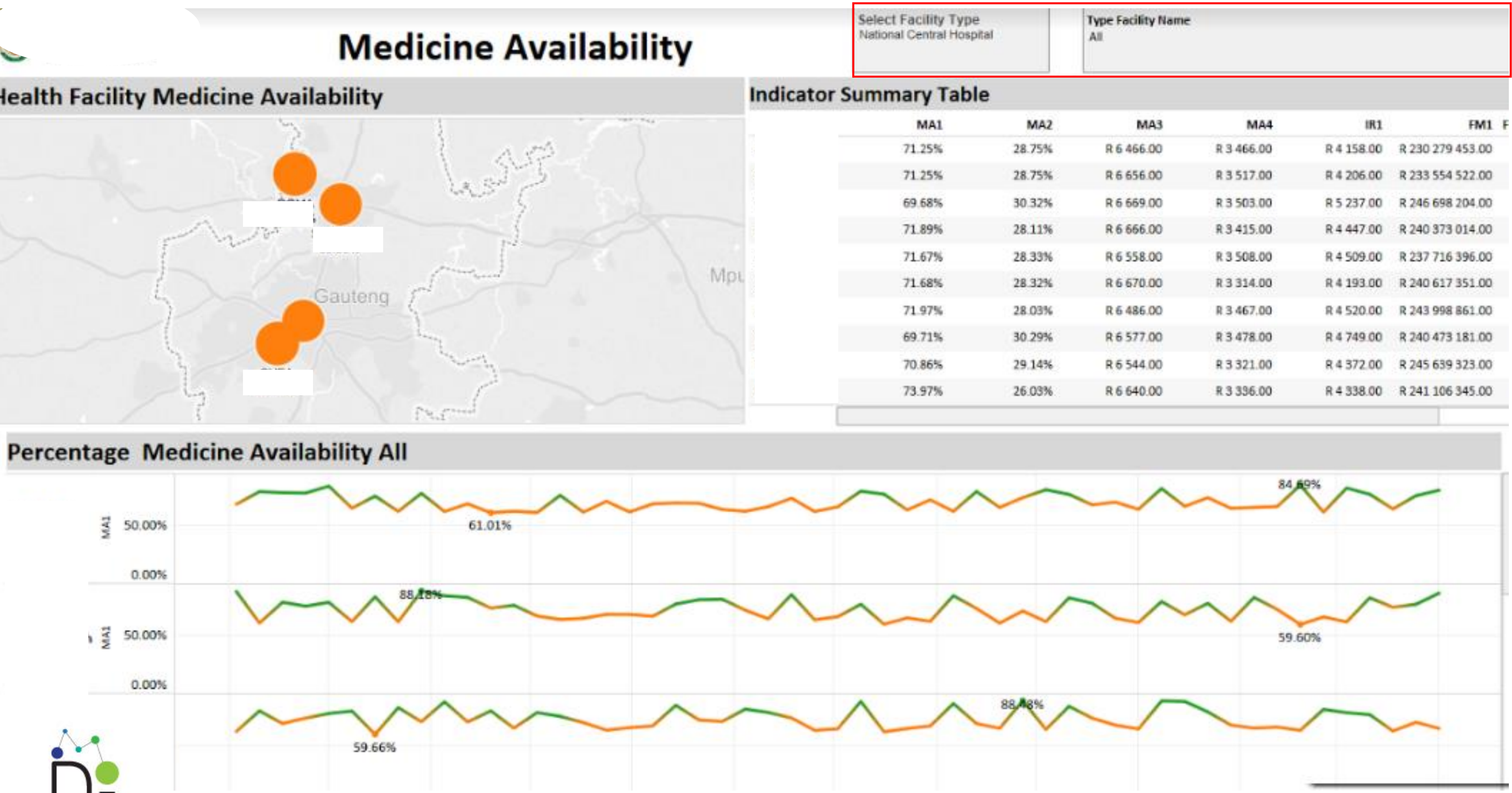
Novel forms and tools of analysis

Real-time data analysis - improved decision making

Supplementary data sources – triangulating

Interactive data visualizations – evaluative thinking

....implementing in Public Sector



Medicine Availability

Select Facility Type: National Central Hospital

Type Facility Name: All

ealth Facility Medicine Availability

Indicator Summary Table

MA1	MA2	MA3	MA4	IR1	FM1
71.25%	28.75%	R 6 466.00	R 3 466.00	R 4 158.00	R 230 279 453.00
71.25%	28.75%	R 6 656.00	R 3 517.00	R 4 206.00	R 233 554 522.00
69.68%	30.32%	R 6 669.00	R 3 503.00	R 5 237.00	R 246 698 204.00
71.89%	28.11%	R 6 666.00	R 3 415.00	R 4 447.00	R 240 373 014.00
71.67%	28.33%	R 6 558.00	R 3 508.00	R 4 509.00	R 237 716 396.00
71.68%	28.32%	R 6 670.00	R 3 314.00	R 4 193.00	R 240 617 351.00
71.97%	28.03%	R 6 486.00	R 3 467.00	R 4 520.00	R 243 998 861.00
69.71%	30.29%	R 6 577.00	R 3 478.00	R 4 749.00	R 240 473 181.00
70.86%	29.14%	R 6 544.00	R 3 321.00	R 4 372.00	R 245 639 323.00
73.97%	26.03%	R 6 640.00	R 3 336.00	R 4 338.00	R 241 106 345.00

Percentage Medicine Availability All

….learning from Private Sector

http://www.smallbusinessconnect.co.za/news/wi-fi-taxi-project-launches-first-phase-gauteng.html#.WzOCXaczbIU

www.wi-taxi.co.za
With Wi-Fi Onboard
WI-TAXI

http://ewn.co.za/2014/06/12/1500-taxis-to-be-fitted-with-free-WiFi

Big Data **exists** in multiple sources, and can be **generated**

Big Data has the **potential** to be useful in evaluation

Big Data **Analytics** is very useful in monitoring

Jerusha Govender | jerusha@thedatainnovator.com | @datainnovator