

**Hi CCLab!**



# Elena Poughia

MD at Dataconomy Media  
Founder at Data Natives  
Evangelist at polypoly  
Occasional Art curator

# About Data Natives

---

- We are Europe's largest network for the data-driven generation with 106,000 + data experts worldwide.
- We have been relevant in the data science and tech industry sector for over 5 years.
- We are fast, dynamic and interdisciplinary and bring the know-how for online hackathons and other educational event formats.
- We know how to build and cultivate online and offline communities and understand all aspects of community building and digital infrastructure.
- We give higher visibility & connections to established brands and provide access to a pool of talented techies & a global tech ecosystem.



# Our year

---



## Data Natives Conference

7 day event, 3.500 attendees, 182+ speakers, 29 partners, 8.9 Million Social Media impressions



## Transparency by Design Summit with bothAnd

3 day event, 200 + active participants in discussion groups, 7.420+ people reached with Livestreams

**#HACK  
CORONA**

## #HackCorona Event Series

4.166 enthusiasts on slack  
864 active participants  
105 teams  
128 experts & mentors  
1.956,8 Million Social Media impressions  
60.000 + live stream impressions



## #Eu vs Virus

21.000+ enthusiasts on Slack  
2.150+ project ideas  
120 winning teams  
Supported by 2.400 mentors, 800 partner, 380 volunteers

# #OpenData

# Future Histories





# The Open Data Opportunity

THE WALL STREET JOURNAL  
Open Data a Boon  
for **Entrepreneurs**

McKinsey & Company  
MCKINSEY GLOBAL INSTITUTE REPORT  
**\$3 trillion a Year** in Value Can Be  
Unlocked Through Smarter Use  
of Open Data.

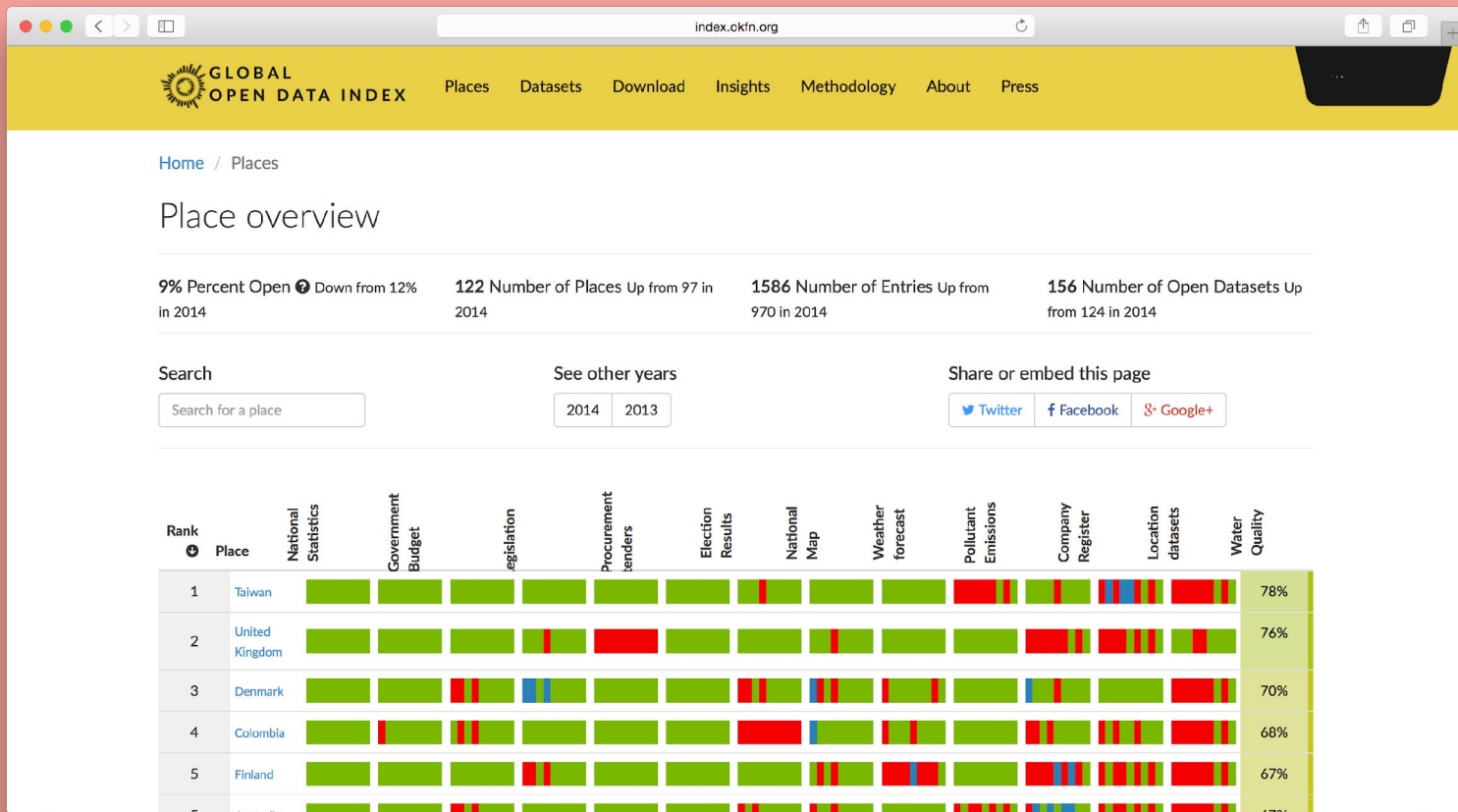
FORRESTER Forrester Research  
Open Data Plays a  
Key Role in  
Advancing **Better  
Governance.**

Gartner®  
Big Data Makes Organizations  
Smarter, but **Open Data  
Makes Them Richer.**



# Open Data is a choice





# Tunisia: Open Data Index 66 Rank

# The 4 Pillars of Value

## Open Data Value Framework



“

**Data science is the  
discipline of making  
data useful.**

Cassie Korzykov



DATA NATIVES

**Large, well  
organized,  
data sets**



DATA NATIVES

# What kind of data exists?



## Business critical data

Data that is indispensable for the economic continuity and growth of a company. It has been structured, which is why we also speak of "clean data".

14%

## ROT data

The acronym stands for "redundant, obsolete and trivial", i.e. for data that exists several times, is not needed for other reasons and is therefore worthless.

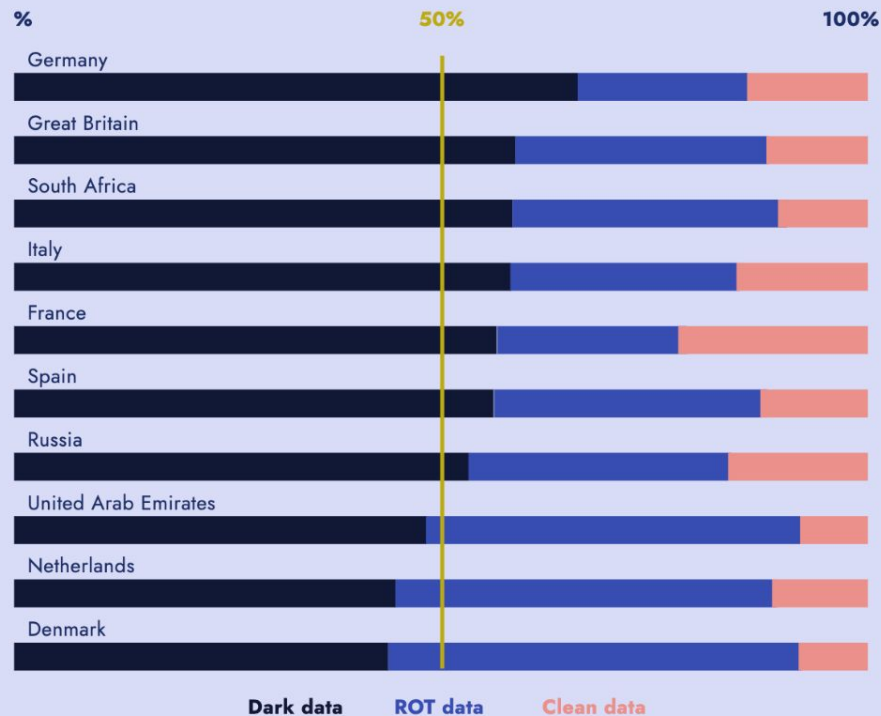
32%

## Dark data

Dark data is unclassified, i.e. unused data whose content and use is unknown. This group is growing faster and faster.


54%

# We are masters of collecting dark data





# A truly expensive “hobby”



In a smaller company that has accumulated about 250 terabytes of data, the estimated total cost of annual data storage is **1.25 Million US-Dollars.**

**175,000 \$**

for clean data

**400,000 \$**

for ROT data

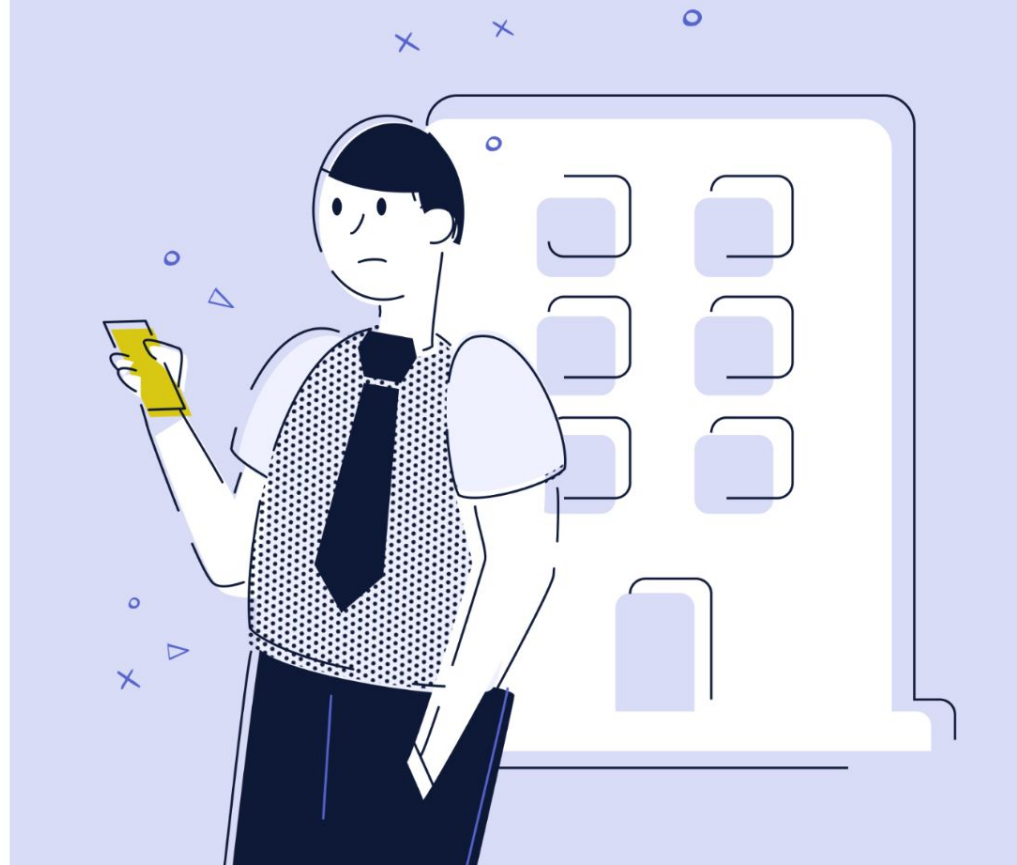
**675,000 \$**

for dark data

# With unpredictable legal risks

The assumed 250 terabytes corresponds to roughly 580 million files. The content of more than half, i.e. just over 300 million files, is completely unknown. Many of these files contain personal information, i.e. information relevant to the GDPR.

**Welcome to privacy hell!**



**The belief that you only need to collect as much data as possible to find a hidden fortune in it.**

**A costly mistake**

## BI after BI in the organisation

### Source systems



### Reports



### Decision makers



# Referendum: Urban planning #Berlin







Image © 2010 AeroWest

© 2009 Go

Bildaufnahmedatum: 5. Mai 2006

52°28'27.39" N 13°24'21.58" E Höhe 0 m

Sichthöhe





Picture: Nic  
Simanek2013\_CC-BY-NC-ND

*NIC SIMANEK*









# Post-Pandemic world: using geographic information system



@elena\_poughia



DATA NATIVES



# The India example



[Tracking Home](#)[Data Visualizations](#) ▾[Global Map](#)[U.S. Map](#)[Data in Motion](#)[Tracking FAQ](#)

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



11/10/2021 4:21 PM

**Cases | Deaths by**  
Country/Region/Sovereignty

**US**  
28-Day: **2,104,503** | 39,312  
Totals: **46,726,060** | 758,007

**United Kingdom**  
28-Day: **1,142,003** | 4,205  
Totals: **9,451,717** | 142,772

**Russia**  
28-Day: **1,012,844** | 30,112  
Totals: **8,764,713** | 245,794

**Turkey**  
28-Day: **781,190** | 5,905  
Totals: **8,290,135** | 72,510

Total Cases  
**251,205,221**

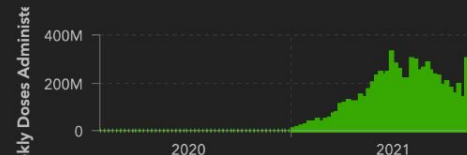
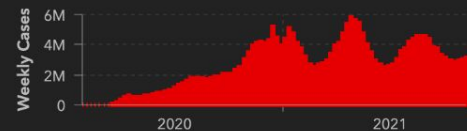
98 Percent  
**12,115,248**

Total Deaths  
**5,069,673**

98 Percent  
**197,256**

Total Vaccine Doses Administered  
**7,344,526,031**

98 Percent  
**743,877,276**



Weekly 28-Day

Admin0

[28-Day](#)[Totals](#)[Incidence](#)[Case-Fatality Ratio](#)[Global Vaccinations](#)[US Vaccinations](#)[Terms of Use](#)

**Epidemic/disease outbreak management  
has been an important application of GIS**



@elena\_poughia



DATA NATIVES



In three days time when Covid hit India for the first time, we helped several states establish a control room/dashboard for monitoring the pandemic spread and containment,” says Agendra Kumar, MD, Esri India.



@elena\_poughia



DATA NATIVES

**DATA** is the new **INTERNET**



@elena\_poughia



DATA NATIVES

We need more **TECHNOLOGISTS**  
in the **GOVERNMENT**



@elena\_poughia



DATA NATIVES

# Citizen-data:

We aren't the product, we're the  
**PRODUCERS**



@elena\_poughia



DATA NATIVES

# A Regulatory Superstate: **DATA SHARING & DATA ACCESSIBILITY**



@elena\_poughia



DATA NATIVES

Both/And:  
**PRIVATE** but still **TRANSPARENT**



@elena\_poughia



DATA NATIVES

# IMPACT is the NEW MONEY



@elena\_poughia



DATA NATIVES



# Access vs Abuse



@elena\_poughia



DATA NATIVES

# Fake vs Real



@elena\_poughia



DATA NATIVES

# Truth vs Illusion



@elena\_poughia



DATA NATIVES

# Us vs Them



@elena\_poughia



DATA NATIVES

# We need Data Managers in municipalities



DATA NATIVES

# #Private Data for Public Data Research: A Promising Model?



# #HackCoronaGreece

LIVE Special Award Ceremony

18:30 CEST / 19:30 GMT+3



Sponsored by



# #HackCoronaGreece Winning teams success story

**3rd place: Assistant Volunteer** - a volunteer resource management system designed to support citizens in need and allow for planning volunteer missions and operations, registry management and volunteer communication.

Powered by



# Why should you care?



@elena\_poughia



DATA NATIVES

# Making data science accessible

# Removing unconscious bias

Making it open, inclusive and  
transparent



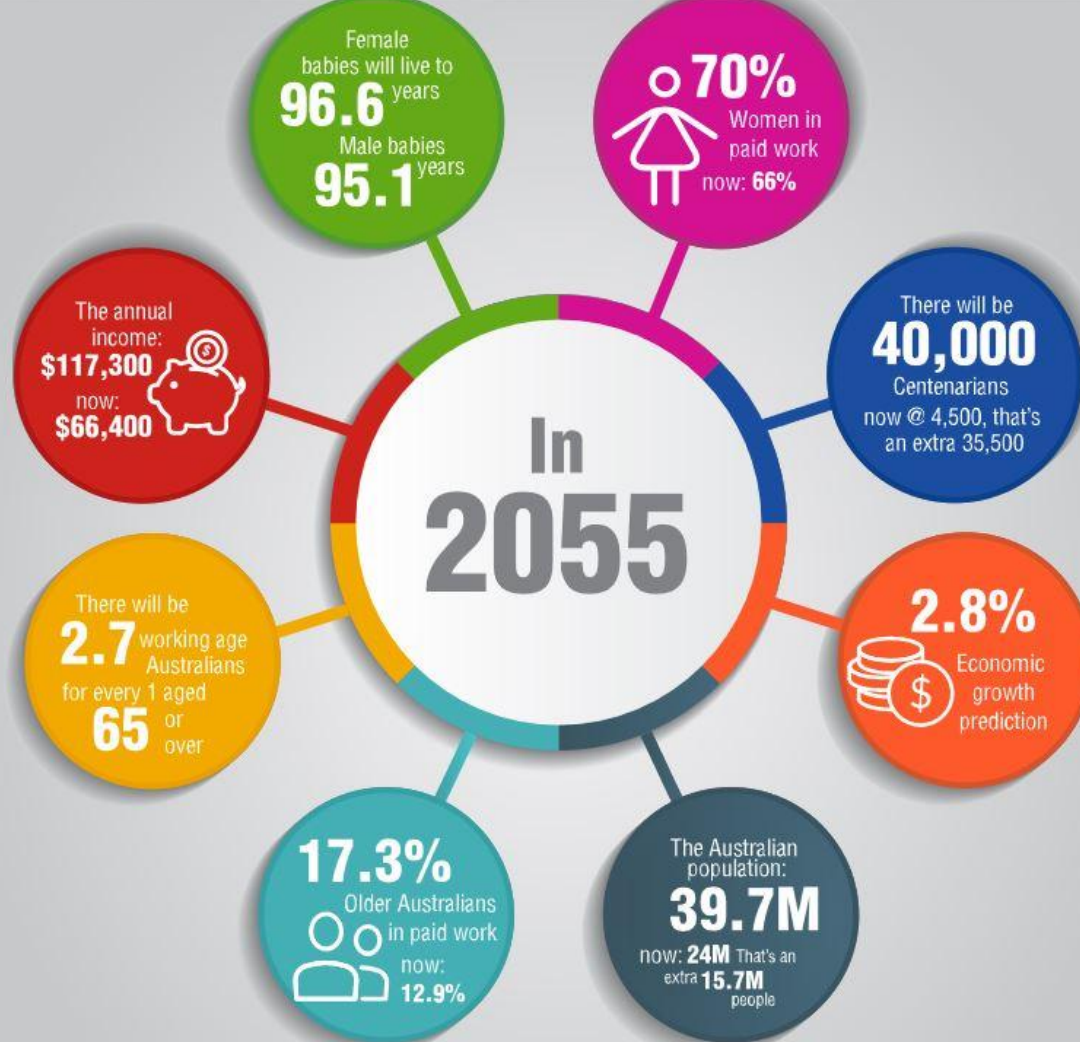
DATA NATIVES



# #FutureofWork



DATA NATIVES



# Data is all about culture



DATA NATIVES

**How can you  
build a data  
culture in your  
company?**



DATA NATIVES

# 1. Employees must be data literate



## 2. Easy-to-use data infrastructure



DATA NATIVES



**3. Simple metrics lexicon;  
data dictionary should be  
developed**



**4. Data should be a part  
of any important  
discussion and decision**



# 5. BI/DS should not become a bottleneck



**Thank you!**