DATA DRIVEN GOVERNANCE

1st CES Public Leaders Summit
May 29, 2019

Emmanuel C. Lallana, PhD
Chief Executive, ideacorp

GUESSTIMATE

GUESSTIMATE

An estimate based on a mixture of guesswork and calculation

DATA-DRIVEN GOVERNANCE

intensive & extensive use of data by societies in defining & achieving their common future.

DATA-DRIVEN GOVERNMENT

where, for all critical decisions, actionable information is available when and where needed.

Using data to enhance governance and improve government is not new.

The history of statistics is closely tied with the emergence of the modernizing 19th century state.

What makes the present different is that the data revolution is helping realize the ideal.

In 2002, researchers estimated that the world produced 5 exabytes of new information

In 2013, 4.4 zettabytes of data — *as many bits as stars in the physical universe* — were created worldwide

By 2025, the world will create 163 zettabytes of data

A Zettabyte in Numbers

1 KB	1,000
1 MB	1,000,000
1 GB	1,000,000,000
1 TB	1,000,000,000,000
1 PB	1,000,000,000.000,000
1 EB	1,000,000,000,000,000,000
1 ZB	1,000,000,000,000,000,000,000

DATA REVOLUTION

an explosion in the volume of data, the speed with which data are produced, the number of producers of data, the dissemination of data, and the range of things on which there is data.

2 data-driven governance use cases

.Data-driven decision-making

.Evidence-based policymaking

2 data-driven governance use cases

data-driven decision-making

the process, which involves collecting data, extracting patterns and facts from that data, and utilizing those facts to make inferences to make decisions

Data Driven Decisions

1) *Nascent*
2) *Basic*
3) *Intermediate*
4) *Advance*
5) *Datavore*

Data Driven Decisions

1) *Nascent*

Rich in data, poor in intelligence.

Data is not a key part of decision-making processes.

Data Driven Decisions

2) *Basic*

Data is used in reports but in a cursory way and with little reference to decisions which have to be made.

3) *Intermediate*

Data analysis is usually requested for decision making, but can be inadequate because the right data is not available or analysis is not of high quality

Data Driven Decisions

4) *Advanced*

Some decisions are informed by data on both the frontline and at senior levels, but is not consistent across the organization

Data Driven Decisions

5) *Datavore*

Data is available in a timely fashion.

Data is analyzed specifically for the purposes of key decisions

Rich in data insight.

Data Driven Decision Making
Where are you? Where is your Agency?

- ***Nascent***
- ***Basic***
- ***Intermediate***
- ***Advanced***
- ***Datavore***

2 data-driven governance use cases

evidence-based policymaking

the use of the best available research and information to guide decisions at all stages of the policy process

4 principles of EBP

1. Build and compile rigorous evidence about what works, including costs and benefits;
2. Monitor program delivery and use impact evaluation to measure program effectiveness;

4 principles of EBP

3. Use rigorous evidence to improve program, scale what works, and redirect funds away from consistently ineffective programs; and

4. Encourage innovation and test new approaches

REALIZING DATA DRIVEN GOVERNANCE
THREE CHALLENGES

I. Data Sources

II. Data Governance and Data Management

III. Data Analytics

RDDG CHALLENGES

I. Data – Old & New

Official Statistics

numerical data-sets, produced by government agencies mainly for administrative purposes

RDDG CHALLENGES

I. Data Sources

Official Statistics

Census data

Survey

Administrative Data

RDDG CHALLENGES

I. Data Sources

Despite attempts at being comprehensive there is a lack of official data on important concerns like gender in official statistics.

RDDG CHALLENGES

I. Data Sources

BIG DATA

datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.

RDDG CHALLENGES

I. Data Sources

BIG DATA - information asset characterized by 3Vs

VOLUME - **amount**

VARIETY – **different types**

VELOCITY- **processing speed**

RDDG CHALLENGES

I. Data Sources

BIG DATA typology

Exhaust data

Passively collected data from people's use of digital services like mobile phones or web searches

Call Detail Record (CDR) contains various attributes of the call, such as time, duration, completion status, source number, and destination number.

RDDG CHALLENGES

I. Data Sources

BIG DATA typology

Sensing data

Actively collected data from sensors, e.g. in smart cities or from wearables and also through remote sensing and satellite images.

basic smartphone includes:

1) Photo sensor;

2) Camera lenses;

3) One or more microphones;

4) Touch sensor;

5) GPS

RDDG CHALLENGES

I. Data Sources

BIG DATA typology

Digital content

content actively produced by users of social media and in internet-mediated transactions

FB posts, Tweets, Blogs, Dating sites, YouTube videos

digital content is intentionally edited, thus subjective or even deceptive, depending on the intentions of the author

RDDG CHALLENGES

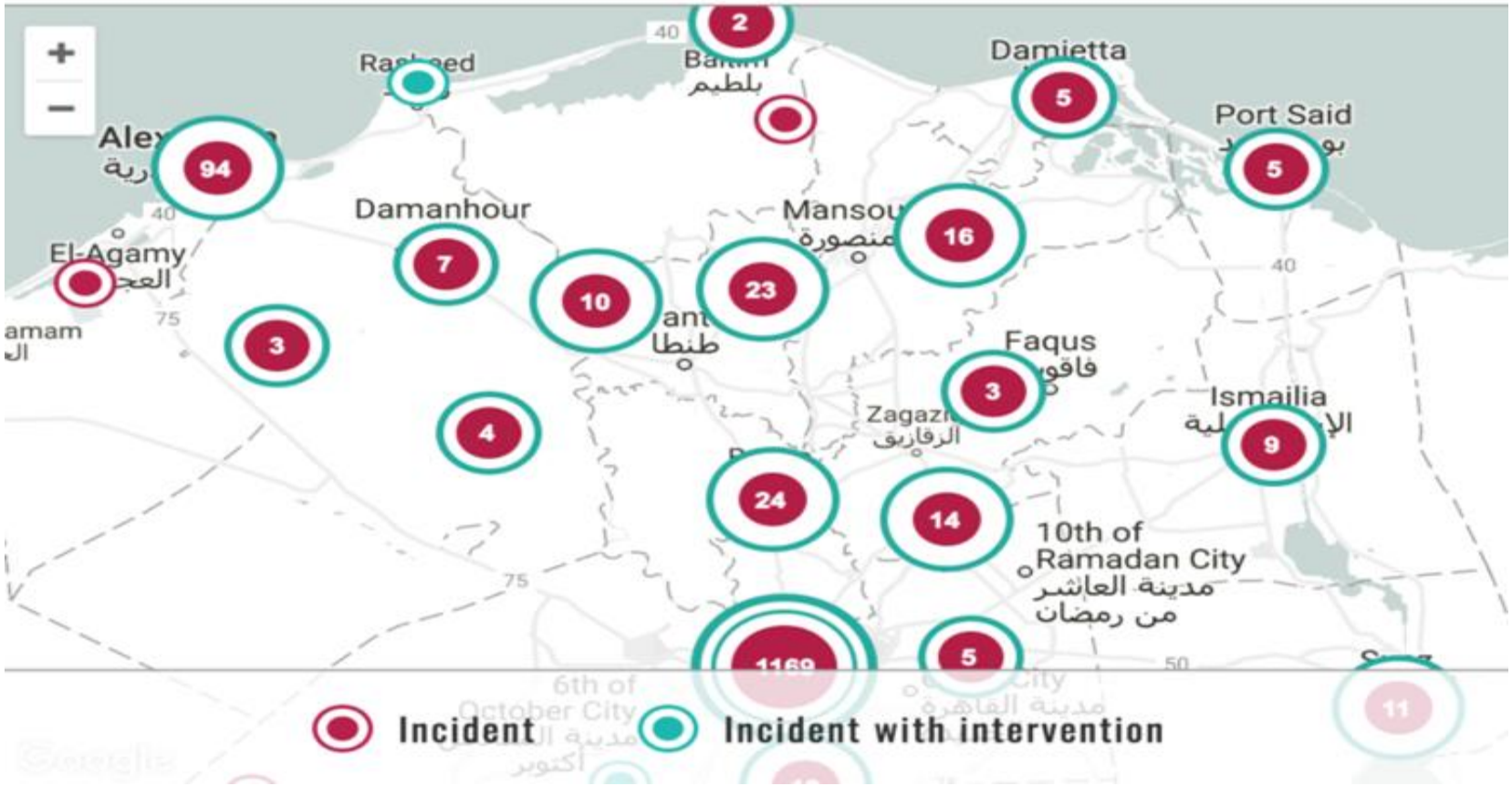
I. Data Sources

CITIZEN GENERATED DATA

data that people or their organizations produce to directly monitor, demand or drive change on issues that affect them

HARASSMAP

Stop sexual harassment, together



SEXUAL HARASSMENT MAP of CAIRO

RDDG CHALLENGES

I. Data Sources

CITIZEN GENERATED DATA

. citizen created data on air quality (Beijing)

. updated water point statuses (Tanzania).

RDDG CHALLENGES

I. Data Sources

CITIZEN GENERATED DATA

particularly useful in trying to understand communities where data is lacking (like IPs), and the vulnerable and/or marginalized sectors of the population (urban poor youth)

RDDG CHALLENGES

.Data – Old & New

.Data Governance and Data Management

.Data Analytics

RDDG CHALLENGES

Data Governance & Data Management ensure that the data generated or collected are easily accessible, effortlessly shared and smoothly reusable.

The distributed nature of data assets in the public sector is even more complex than that in the private sector.

DATA GOVERNANCE

a system of decision rights and accountabilities for information-related processes

DATA GOVERNANCE

comprehensive process for controlling the integrity, use, availability, usability, and security of all data 'owned' or controlled by an organization.

Data Governance Elements

.A clear vision of the purpose of the data and how it serves the organization's mission.

.Organizational policies, processes, and principles on how data will be managed to achieve the vision.

Data Governance Elements

- **Identification of & advocacy for sufficient resources (money, time, skills, etc.) needed to manage the data to meet the vision.**

Data Governance Elements

.Monitoring & Evaluation indicators for data management and compliance with legal requirements

Data Governance Elements

- **Setting expectations & policies on what happens if an incident is suspected**
  - **a clear chain of command & authorization for action so that if there is harm, it can be stopped & responded to quickly.**

- Siobhan Green "What Does Responsible Data Look Like?" ICTWorks May 14, 2019 https://www.ictworks.org/responsible-data-journey/#.XOYuytMzai4

RDDG CHALLENGES

Data Management

the execution of processes, architectures, policies, practices and procedures in order to manage the data owned and controlled an organization

RDDG CHALLENGES

Data Management includes developing policies, strategies, standards and programs for the following:

- **Data Architecture;**
- **Data Modeling & Design;**
- **Data Storage & Operations;**
- **Data Security;**

RDDG CHALLENGES - 2

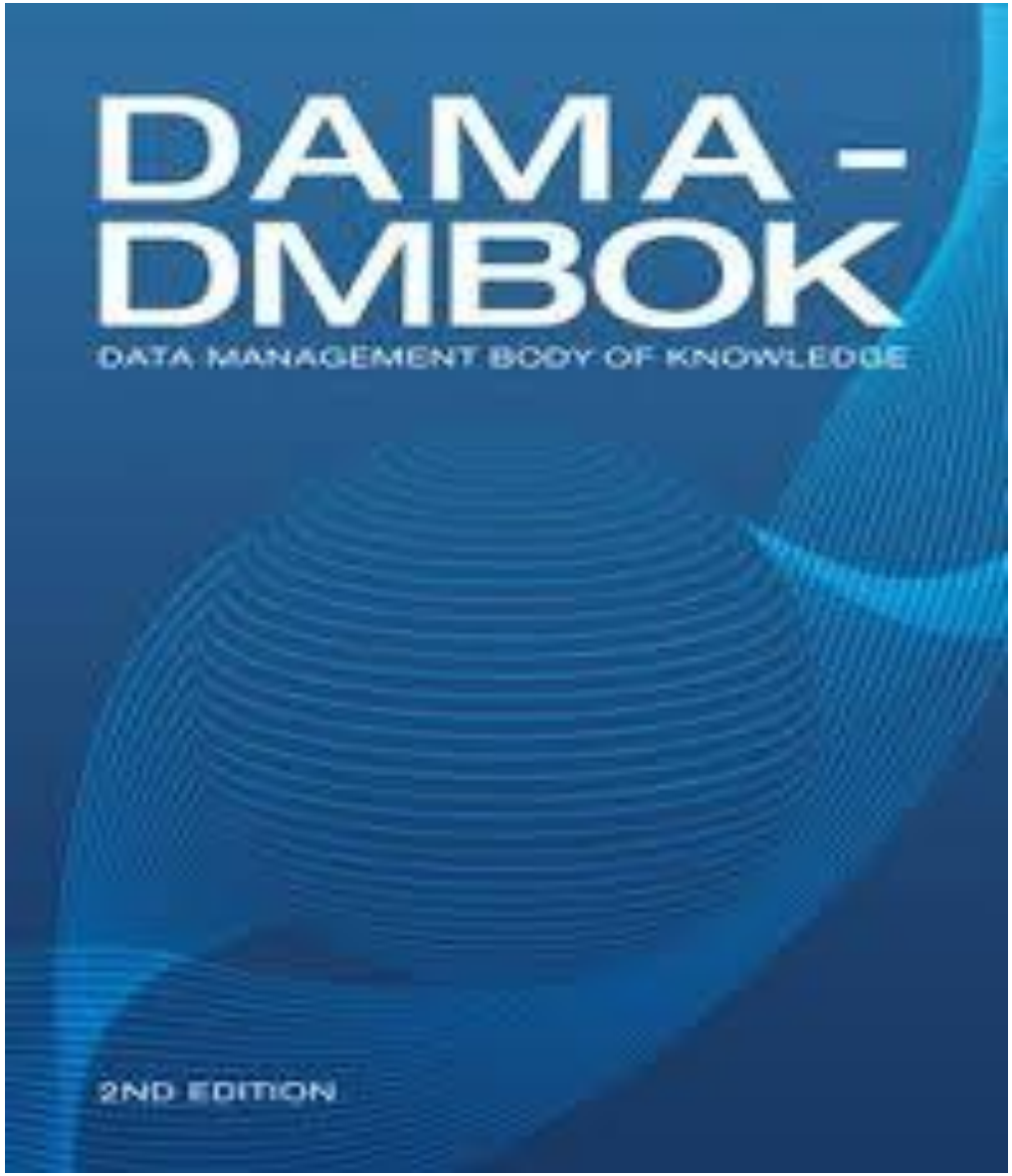
Data Management includes developing policies, strategies, standards and programs for the following:

- **Data Integration & Interoperability;**
- **Reference & Master Data;**
- **Data Warehousing & Business Intelligence;**

RDDG CHALLENGES - 2

Data Management includes developing policies, strategies, standards and programs for the following:

- **Metadata;**
- **Data Quality**


DAMA-
DMBOK
DATA MANAGEMENT BODY OF KNOWLEDGE
2ND EDITION

Data Governance is about how decisions are made about data and how people and processes are expected to behave in relation to the data

Data Management implements goals and values defined by governing bodies

RDDG CHALLENGES

.Data – Old & New

.Data Governance and Data Management

.Data Analytics

RDDG CHALLENGES

Data Analytics

extracting meaning from raw data using specialized computer systems… that transform, organize, & model the data to draw conclusions & identify patterns

RDDG CHALLENGES

III. Data Analytics

Four types of analytics:

- Descriptive
- Diagnostic
- Predictive
- Prescriptive

RDDG CHALLENGES

III. Data Analytics

Four types of analytics:

<u>Descriptive analytics</u> provide insight to answer: *What has happened?*

Data Analytics

Descriptive analytics

Books from Project Gutenberg were fed into the "The Hedonometer" a program developed by computer scientists at the University of Vermont and the University of Adelaide.

Data Analytics

Descriptive analytics

After the Hedonometer finished "reading" the books it plotted "the emotional trajectory" of all of the stories using a "sentiment analysis to generate an emotional arc for each work."

Data Analytics

Descriptive analytics

It found were 6 broad categories:

1) **Rags to Riches** *(rise)*;
2) **Riches to Rags** *(fall)*;
3) **Man in a Hole** *(fall then rise)*;
4) **Icarus** *(rise then fall)*;
5) **Cinderella** *(rise, fall, rise)*;
6) **Oedipus** *(fall, rise, fall)*

RDDG CHALLENGES

III. Data Analytics

Four types of analytics:

<u>Diagnostic analytics</u>

measure historical data against other data to answer *why something happened;*

Data Analytics

Diagnostic analytics

health insurance providers are interested in risk factors for abusing opiods

Blue Cross / Blue Shield, using years of insurance & pharmacy data, identified 742 risk factors abusing opioids.

https://www.datapine.com/blog/big-data-examples-in-healthcare/

RDDG CHALLENGES

III. Data Analytics

Four types of analytics:

<u>Predictive analytics</u> use statistical models and forecasting techniques to answer: *What could happen*?

Data Analytics

Predictive analytics

A customer at a Target store outside Minneapolis complained to the manager

"My daughter is still in high school, and you are sending her coupons for baby clothes and cribs? Are you encouraging her to get pregnant?"

Data Analytics

Predictive analytics

Target is interested in identifying unique moments in consumers' lives when their shopping habits become particularly flexible and the right advertisement or coupon would cause them to begin spending in a new way.

Data Analytics

Predictive analytics

Using their data, Target was able to identify about 25 products that, when analyzed together, allowed them to assign each shopper a "pregnancy prediction" score.

Data Analytics

Predictive analytics

More important, Target could also estimate her due date to within a small window, so Target could send coupons timed to very specific stages of her pregnancy.

Data Analytics

Predictive analytics

A few days after the costumer complained, he called the Target manager to say:

“I had a talk with my daughter. It turns out there's been some activities in my house I haven't been completely aware of.”

RDDG CHALLENGES

III. Data Analytics

Four types of analytics:

<u>Prescriptive analytics</u> suggests *what action to take*.

Data Analytics

Prescriptive analytics

Otto, a German e-commerce merchant, "is deploying technology to make decisions at a scale, speed and accuracy that surpass the capabilities of its human employees".

Data Analytics

Prescriptive analytics

OTTO's IT system analyses around 3 billion past transactions and 200 variables (including past sales, searches on Otto's website and weather information) to determine with 90% accuracy what customers will buy in the coming week.

Data Analytics

Prescriptive analytics

OTTO's IT system purchases around 200,000 items a month from third-party brands with no human intervention.

THREE CHALLENGES

- **Data – Old & New**
- **Data Governance and Data Management**
- **Data Analytics**

MEETING THE CHALLENGES

Policies and programs needed to realize data-driven governance.

MEETING THE CHALLENGES

- strengthen data collection;
- institutionalize data governance;
- gear up for Big Data

MEETING THE CHALLENGES

The key condition for data-driven governance is a public sector culture that embraces the enormous value of its data.

in the public sector, data culture means

a deep, government-wide comfort level with using metrics to maximize social impact

Thank You

eclallana@ideacorpphil.org