“Thick” data for development
PULSE LAB JAKARTA

In late 2012 Pulse Lab Jakarta was established through a partnership between the United Nations and the Ministry of National Development and Planning (Bappenas). The first innovation lab of its kind in Asia, Pulse Lab Jakarta brings together experts from United Nations agencies, the Indonesian government, non-governmental organisations and the private sectors to research and, most importantly, facilitate the adoption of new approaches for applying new, digital data sources and real-time analysis techniques to social development. The Indonesian cross-government partners include the Ministry of...
• The era of the mutants
• The promise
• Reality check
• Where to start?
“If the rate of change on the outside exceeds the rate of change on the inside, the end is near”

- Jack Welch
**INSTRUCTIONS**
Please answer each question clearly and completely. **Type or print in ink.** Read carefully and follow all directions.

**UNITED NATIONS**
PERSONAL HISTORY

<table>
<thead>
<tr>
<th>No.</th>
<th>Field Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Family name</td>
</tr>
<tr>
<td>2</td>
<td>Date of Birth</td>
</tr>
<tr>
<td>3</td>
<td>Place of birth</td>
</tr>
<tr>
<td>4</td>
<td>Nationality (ies) at birth</td>
</tr>
<tr>
<td>5</td>
<td>Present nationality (ies)</td>
</tr>
<tr>
<td>6</td>
<td>Sex</td>
</tr>
<tr>
<td>7</td>
<td>Height</td>
</tr>
<tr>
<td>8</td>
<td>Weight</td>
</tr>
<tr>
<td>9</td>
<td>Marital status: Single □ Married □ Separated □ Widow(er) □ Divorced □</td>
</tr>
<tr>
<td>10</td>
<td>Entry into United Nations service might require assignment and travel to any area of the world in which the United Nations might have responsibilities. Have you any disabilities which might limit your prospective field of work or your ability to engage in air travel? YES □ NO □ If &quot;yes&quot;, please describe.</td>
</tr>
<tr>
<td>11</td>
<td>Permanent address</td>
</tr>
<tr>
<td>12</td>
<td>Present address (if different)</td>
</tr>
<tr>
<td>13</td>
<td>Office Telephone No. Office Fax. No E-mail:</td>
</tr>
<tr>
<td>14</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Have you any dependents? YES □ NO □ If the answer is “yes”, give the following information:</td>
</tr>
<tr>
<td>16</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>
That’s why Valve is flat. It’s our shorthand way of saying that we don’t have any management, and nobody “reports to” anybody else. We do have a founder/president, but even he isn’t your manager. This company is yours to steer—toward opportunities and away from risks. You have the power to green-light projects. You have the power to ship products.
Project Sunroof

Mapping the planet’s solar potential, one roof at a time.

Enter your address to see how much you can save

Project Sunroof is currently helping people save with solar in the following cities:
San Francisco Bay Area, Fresno CA and Greater Boston MA.
We aim to design litterless cities using artificial intelligence.

Takanome is an AI-based program that maps and measures litter using visual recognition technology. We aim to contribute to designing litterless cities by evaluating anti-litter efforts through our cutting-edge research and providing other consulting services related to urban development.
Data Mining for flood and disaster management

output Jakarta

Fonte: Floodtags
FIND THE DATA YOU NEED
Discover and Analyze the World's Open Data

REQUEST A DEMO

BETTER DATA. SMATER DECISIONS
We process millions of unstructured datasets every day & use our advanced processing capabilities to clean & enrich data.
You + Your Data
We’re creating the apps and infrastructure to put your small data to good use for you.
Landscape of Innovation Approaches
A tentative overview of innovation approaches for the public good. Where do you play?

EMPOWERING PEOPLE
TALENT SPACE
Developing skills and competencies, mobilising and engaging people to make change happen.

Change Management
Talent Management
GovJams/Service Jams
Innovation Labs/Teams
Challenge Prizes

GENERATING UNDERSTANDING
INTELLIGENCE SPACES
Gathering insights, generating understanding about the challenges, evaluating solutions, identifying opportunities.

Sensemaking
Ethnographic Research
Evidence-based Policy
Lead User Innovation
Organisational Culture
Ambidextrous Leadership

Action Research
Behavioural Insights
Systems Thinking
Design Thinking
Randomised Control Trials
Speculative Design
Foresight

CREATING SOLUTIONS
CREATION SPACE
Translate ideas into solutions, bringing resources and people together, making it happen, generate outcome.

Prototyping
Innovation/ Creative Spaces
Hackathons
FabLabs
Gamification
Social Innovation Camps
Impact Investment

ENABLING ACTION
TECHNOLOGY SPACE
Tools, infrastructure, systems and platforms supporting people and their activities to make innovation happen.

Open Policy Making
Open Policy Spaces
Open Data
Open Data
Big Data
Sensors
"Digital"
Internet of Things

The promise
Would you cross the street based on information that is five minutes old?
The promise

• High-frequency, realtime data collection
• Real time monitoring of projects
• Cutting data collection costs
• Predicting trends
• Going subnational
• New understanding of complex social phenomena
• Making the invisible visible
BPP

The Billion Prices Project is an academic initiative that uses prices collected from hundreds of online retailers around the world on a daily basis to conduct economic research.

This page shows our most recent research leveraging high-frequency price data, as well as the US daily inflation index (updated monthly on this page).

**BPP Geographic Coverage**
Dengue trends - Indonesia

We've found that certain search terms are good indicators of dengue activity. Google Dengue

Attractions forecast

Provide national 5A scenic predict the future on the 2nd passenger and comfort for your short-term tourist travel information

Click to enter
This AI Can Accurately Predict the Outcome of Human Rights Trials

October 23, 2016 // 08:00 PM EST

In 2015, the European Court of Human Rights, which rules on alleged human rights
a) Abidjan
b) Liberian border
c) Roads to Mali and Burkina Faso
d) Road to Ghana

Real time measurement of poverty?

Ref: arxiv.org/abs/1309.4496: Evaluating Socio-Economic State Of A Country Analyzing Airtime Credit And Mobile Phone Datasets
MAPPING SOCIOECONOMIC LEVELS OF POPULATIONS

Ground-Truth Data (Official)

Map of 11 regions showing Multidimensional Poverty Index

Using Mobile Phone Data

Finer granularity map estimated from communication patterns

Figure 1: Niamey 2003–2004 brightness and measles cases.

(A) Left: map of Africa, Niger shaded grey; right: map of Niger showing three largest cities, each is a health district. (B) Reported measles cases by commune for Niamey’s 2003–2004 outbreak. Vertical dashed line indicates start of reactive immunization campaign. Inset: Pixels of Niamey: communes outlined in black, faded pixels outside communes show peripheral areas. Colors indicate communes for (B,C). (C) Solid lines: brightness within each commune’s boundaries (bright pixels in (B)). Dashed lines: brightness for each commune including associated peripheries (bright and faded pixels in (B)). Vertical line: start of reactive immunization campaign. Maps are GADM shapefiles^33 drawn and edited in R version 1.7.1 (https://cran.r-project.org)^47, finished in Adobe Illustrator CS3 (http://www.adobe.com/products/illustrator.html).
The computer model’s predictions were surprisingly accurate when compared with survey data.
Fig. 2.
(a) Weekly number of unique subscribers in Chittagong City. There is a clear increase in the number of unique subscribers (SIM cards) in Chittagong City after the cyclone (18th May). The increase amounts to approximately 50,000 unique subscribers (see Section 2 and S1). (b) Visualisation of the netflow (inflow minus outflow) between Chittagong City and all Bangladeshi districts two months after the cyclone (May–July 2013). Positive netflow: blue, negative netflow: red. (c) Chittagong City’s monthly relative change in subscriber numbers during 2012 (grey) and 2013 (blue). The changes from May (cyclone landfall 18 May, 2013) to July are virtually identical during the two years, rendering a relationship to Cyclone Mahasen highly unlikely. Correlation between data points: $r = 0.977, p < 0.001$. (d) Seasonality of migration patterns for all districts. Each circle represents a particular district in a particular type of month (“February”, “March”, “April” etc). Y-axis shows the netflow in each month (inflow minus outflow) for 2013 and x-axis for...
Our findings closely match results from expensive & infrequent transportation surveys; are cheaper & can be produced as needed.
Reality check
• Proof of concepts
• Religious wars (big vs. open vs small, “old” vs new)
• Still relies on official statistics for validation
• “I want answers, not data”
• Lack of data skills
• Lack of governance structures (e.g. data collaboratories)
• Access to data
• Privacy
• Discrimination
Testing assumptions

Data/evidence → Action
Passive listening → No reaction/adaptation
Real time data → Real time orgs
Machine → No bias
Big data → No need for local context
Organizations engage in many rituals to look compliant and modern. They buy new technological tools, create new departments, and swear that they will change how they do things. But transforming the actual practices of the people who work in a given organization is a different ballgame. Workers often have good reasons to continue doing what they have always done. No algorithm will change that overnight, unless it is forced upon the employees.

Angele Christin, *Models in Practice*
CASE STUDY

Developmental Evaluation for Public Sector Innovation
How can one ambitious program teach practical lessons to a global community: innovators?

Absurdity and the Sensible Decision
implementation of Danish labour market policy

Nina Holm Vohnsen
SAVE MONEY
SAVE THE PLANET
BE A GOOD CITIZEN
YOUR NEIGHBORS ARE DOING BETTER
INNOVATION

Turn On, Tune In, Transcribe: U.N. Develops Radio-Listening Tool

October 31, 2016 · 5:17 PM ET

CAMILA DOMONOSKE
One such tactic for seamless data is obfuscation. AdNauseam is an obfuscation tool I produced in collaboration with Daniel Howe and Helen Nissenbaum. It

https://vimeo.com/111943439
Translation: "Dmitry Peskov's watch is a 'Richard Mille' worth 37,000,000 Roubles ($620,000)"
Just because you have real time data it does not mean that you have real time organisations
'NOWCASTING' COMMODITY PRICES IN INDONESIA

Can estimates of food prices be extracted from social media?

How do these estimates compare to official statistics?

Pulse Lab Jakarta conducted a research project to investigate the feasibility of 'nowcasting' commodity prices in Indonesia, a country famously addicted to social media. This page presents near to real-time price information on three staple food commodities: beef, chicken and onion, using tweets about food prices in Indonesian language Bahasa Indonesia.

This project has been done in collaboration with the Indonesian Ministry of National Planning Development (Bappenas) and the World Food Programme Indonesia.

For more information on this project, please visit: http://www.unglobalpulse.org/nowcasting-food-prices
I find it helpful to think of algorithms as a dim-witted but extremely industrious graduate student, whom you don't fully trust.

...You want them to go through ten million photos and find every picture of a horse? Perfect.

You want them to draw conclusions on gender based on word use patterns? Or infer social relationships from census data? Now you need some adult supervision in the room.

- Maciej Ceglowski
When Bushmen Race Aerial Robots to Protect Wildlife

WeRobotics  @WeRobotics · Oct 9
UAV pro pilot Khadija is quick learner. She wants to use flying robots to map Zanzibar, which is where she's from. We'll be there next week

WeRobotics  @WeRobotics · FOLLOWS YOU
We scale the impact of humanitarian, development and environmental project by making appropriate robotics solution available to outstanding local partners.

Global
Nesta
Why Big Data Needs Thick Data

Originally published on May 13, 2013 for Ethnography Matters, I’m re-publishing the post for the launch of the new Ethnography Matters Medium channel. I’ve added an updated article with a case study from my time at Nokia where I worked on their over-dependence on quantitative data. I’m continuously blown away by how much my original post in 2013 sparked a discussion about the integration of Thick Data and Big Data. Since then, I talked about the need to integrate Data with Thick Data at EPIC conference, and Strata. In 2015 I gave a keynote at Republica where I expanded on this article by tracing the roots of origins of valuing numbers over stories, How to Avoid Curses in the Ethnographic Data. Just last year, Word Spy created an entry of Thick Data, referencing my...

Where to start?
“Humans at the beginning, humans at the end”

- D. Snowden
“There is no such thing as a smart sensor – a sensor is dumb, it does one thing. It reads a temperature, water height or whatever. A human being can look around, smell, hear and use a whole variety of their sensory and mediated relationships to relay information in an emergency.”
Mapping social & biophysical flood vulnerability anywhere around the world in a browser within seconds

- **Computational Hydrology**: Mapping the floodplain
- **Social Vulnerability**: Identifying the most vulnerable & least likely to recover
- **Google Earth Engine**: Leveraging Big Data GIS

= Smarter Planning
Greater Resilience
Data Empowered Communities
Explore some of the ways Premise data is used to better understand our world.

**China: Staples**

<table>
<thead>
<tr>
<th>Index</th>
<th>30 Day</th>
<th>7 Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>101.6</td>
<td>0.13%</td>
<td>0.09%</td>
</tr>
</tbody>
</table>

View a real-time monitor of consumer staples prices in China. Click components to access thousands of granular daily observations. Explore

**Argentina: Bread Onions**

<table>
<thead>
<tr>
<th>Index</th>
<th>30 Day</th>
<th>7 Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>117.5</td>
<td>3.32%</td>
<td></td>
</tr>
</tbody>
</table>

View a real-time map of bread prices in Argentina. Explore

**Components**

<table>
<thead>
<tr>
<th>Components</th>
<th>INDEX</th>
<th>30 DAY</th>
<th>7 DAY</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barrackpore, West Bengal (1765)</td>
<td>35.6</td>
<td>-2.6%</td>
<td>-13%</td>
<td></td>
</tr>
<tr>
<td>Chennai (2627)</td>
<td>102.9</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Chennai, Tamil Nadu (2235)</td>
<td>40.3</td>
<td>-2.9%</td>
<td>-14.5%</td>
<td></td>
</tr>
<tr>
<td>Kanchrapara, West Bengal (1960)</td>
<td>73.1</td>
<td>-19.7%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Kolkata (1790)</td>
<td>105.6</td>
<td>0.8%</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Onions, a core staple for 1.23 billion people, have seen unprecedented price

http://www.premise.com/solutions.html
Incubatee Insight - WeatherSafe Ltd

Forecasting the future of agricultural success

Coffee is the second most traded commodity in the world. Rwanda is a country who has trees which produce high quality coffee and has remained one of the most important crops in the country's agricultural landscapes. However, changes in climate such as heavy rainfall can cause a disease called Hemileia Vastatrix, or coffee rust, which has devastating effects on the crops, and consequentially its yield. WeatherSafe Ltd, a new company in the ESA Business Incubation Centre at Harwell, have created the WeatherSafe ‘Coffee Farmer Edition’ mobile application. They have recently developed new links with Rwandan government to implement their innovative idea to help save the Rwandan economy, which is still largely based on agriculture.

Using data architecture which has been researched by ESA, WeatherSafe Ltd use hyper-local weather forecasting methods to alert farmers of potential breaks in rain and any diseases that might occur. Being only 26,338 km² in size, Rwanda is a manageable size for using this satellite data. The sensitivity of this prediction enables the forecast to be tailored to each farmer to enable farmers to make the best decisions for their crops to speed up their move from being subsidence farmers to professional.
Where to start

1. What’s the distance between the collection, interpretation and action on data?
2. Who is *really* using the data?
3. How can we empower those closer to the problem to use data to take better decisions over time?
4. What do those who produce data get in exchange?
5. If we do not always act rationally, how can data help?
Data empowered villages

Data Correlation for Human Resource Rationalization

Ojek Call Tree

Pregnant mother calls Ojek driver 1

Ojek driver 1 responds to the pregnant mother’s phone call and then calls up two other Ojek drivers

Ojek driver 1 calls ojek drivers 2 and 3

Ojek driver 2 and Ojek driver 3 respond to Ojek driver 1’s phone call

Ojek driver 3 calls pregnant woman and takes her address

Ojek driver 3 going to pick up pregnant woman

Nesta
“Traditionally, it has been a government social worker who administers a survey and then takes that information with them. Poverty Stoplight is the opposite. Here, a family assesses their level of poverty in 50 specific indicators, and the results are visualised in a dashboard for the family to use. So instead of being an index for policymakers, the Poverty Stoplight is a tool for a very different kind of decision maker: the head of the household.”

- Martin Burt, CEO, Fundación Paraguaya
Can staff act on the data?

Figure 17: Decentralisation and Data-Driven Decision Making (percentage who agree or disagree with the statement)

Source: Nesta, The Rise of the Datavores
Using remote sensing to empower local communities in forestry management

A new monitoring system in Ethiopia makes satellite data available at the local level to promote participatory forest management
Sens-Us
The two day ‘data studio’ took an explicitly design approach to creating some ideas for research projects focused on the strategic and operational challenges of increasing food bank use in the UK. The event incorporated techniques from the design school it was situated in: exploring food poverty from different human perspectives; making physical mock-ups of aspects of food poverty; wire-framing potential research designs. The data studio
1. What’s the distance between the collection, interpretation and action on data?
2. Who is really using the data?
3. How can we empower those closer to the problem to use data to take better decisions over time?
4. What do those who produce data get in exchange?
5. If we do not always act rationally, how can data help?
DATA HUMANISM

small / big data
imperfect / infallible data
subjective / impartial data
inspiring / descriptive data
serendipitous / predictive data
data conventions / possibilities
data to simplify complexity / depict
data processing / drawing
data driven design
data / is numbers / people
data will make us more efficient / human.

Source: Giorgia Lupi
Thank you!
It was a hypothesis – but data from this early work showed a correlation between people’s mobility and the spread of Ebola.

Mobility data pointing at the south ferry line as the most common route of travel from an infected area in Kaffu (in blue) to Freetown area.

The Zika epidemic allows us to take this system one step further. Influenced by recent academic research on the spread of Dengue ([5]) and new methods for using social media to assess emergencies ([6]), Magic Box v.0.5 was built as a collaboration between Google engineers and UNICEF’s Office of Innovation. The work around Zika, which is still ongoing, and in very early stages, has allowed us to bring new partners into the effort to combat global risks. Amadeus, which provides more than 40% of global travel booking joined as an initial data provider, while also generously offering support from their core business and engineering teams. IBM recently entered the collaboration with support from its weather data teams. In addition to these three incredibly connected and technologically savvy partners, UNICEF continues to work with UN Global Pulse and through the UN Innovations Network, which it co-chairs with UNHCR and WFP to drive a research agenda that will benefit other parts of the United Nations.

Connecting aggregate data sources from private sector partners is not easy. We have to understand issues of privacy, different corporate languages, and a variety of technology stacks. We also need to link data to action. But this early stage prototype points to a new set of possibilities: if we can bring more collaborators into the rapidly-developing epidemic response data ecosystem, we may be able to use this kind of data to help people in need.
Satellite Manufacturing Index and Official PMI

UnionPay and Official Real Estate Sales

Fonte: Bloomberg
300,000 answered a typical OkCupid question

GALLUP

3,050 answered nationwide Gallup Poll in 2008 election
Which place looks livelier?

For this question: 129,134 clicks collected

<table>
<thead>
<tr>
<th>RANK</th>
<th>CITY</th>
<th>CLICKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>London</td>
<td>6334</td>
</tr>
<tr>
<td>2</td>
<td>New York</td>
<td>7783</td>
</tr>
<tr>
<td>3</td>
<td>Paris</td>
<td>5824</td>
</tr>
</tbody>
</table>

About StreetScore

StreetScore is an algorithm that assigns a score to a street view based on how safe it looks to a human — but using a computer (see FAQ). This website is a collection of map visualizations of perceived safety of street views from cities in the US as predicted by StreetScore. We will be releasing a map of perceived safety for a new city each week. The StreetScore algorithm was created by Nikhil Naik as part of a collaboration between the Macro Connections group and the Camera Culture group at MIT Media Lab. Jade Philipoop created the visualizations presented in the StreetScore website.

Please send your questions/comments to streetscore@media.mit.edu
Quantifying social contacts in a household setting of rural Kenya using wearable proximity sensors, Tizzoni, Catutto et al, 2016
Data

Project participants have collected data on garbage disposal through ghost wipes as well as using smartphones to map garbage collection routes. You can also access the data on the Digital Matatus platform.

Collaborative Research and Mapping for Nairobi
Citizenscience.gov is an official government website designed to accelerate the use of crowdsourcing and citizen science across the U.S. government. The site provides a portal to three key assets for federal practitioners: a searchable catalog of federally supported citizen science projects, a toolkit to assist with designing and maintaining projects, and a gateway to a federal community of practice to share best practices.
Mobile Phone Data Reveals Literacy Rates in Developing Countries

A machine learning algorithm has learned to identify individuals who can read or write by analyzing their mobile phone records.

by Emerging Technology from the arXiv July 1

One of the millennium development goals of the United Nations is to eradicate extreme poverty by 2030. That’s a complex task that can be addressed by leveraging data from mobile phones.
Digital footprints of unemployment