World Data Lab

International Workshop on Data Disaggregation for the SDGs

What we do

We specialize in developing granular **economic and demographic forecasts** for every country and any point in time.

HOMI KHARAS

WDL CO-FOUNDER & LEAD ECONOMIST

“we need to harness the transformative power of the data revolution.”

WDLco-founder, Homi Kharas is a Senior Fellow and Deputy Director in the Global Economy and Development program at the Brookings Institution in Washington D.C. He has served as the **lead author and executive secretary** supporting the High Level Panel for the U.N. Secretary General regarding the post-2015 development agenda.

BROOKINGS

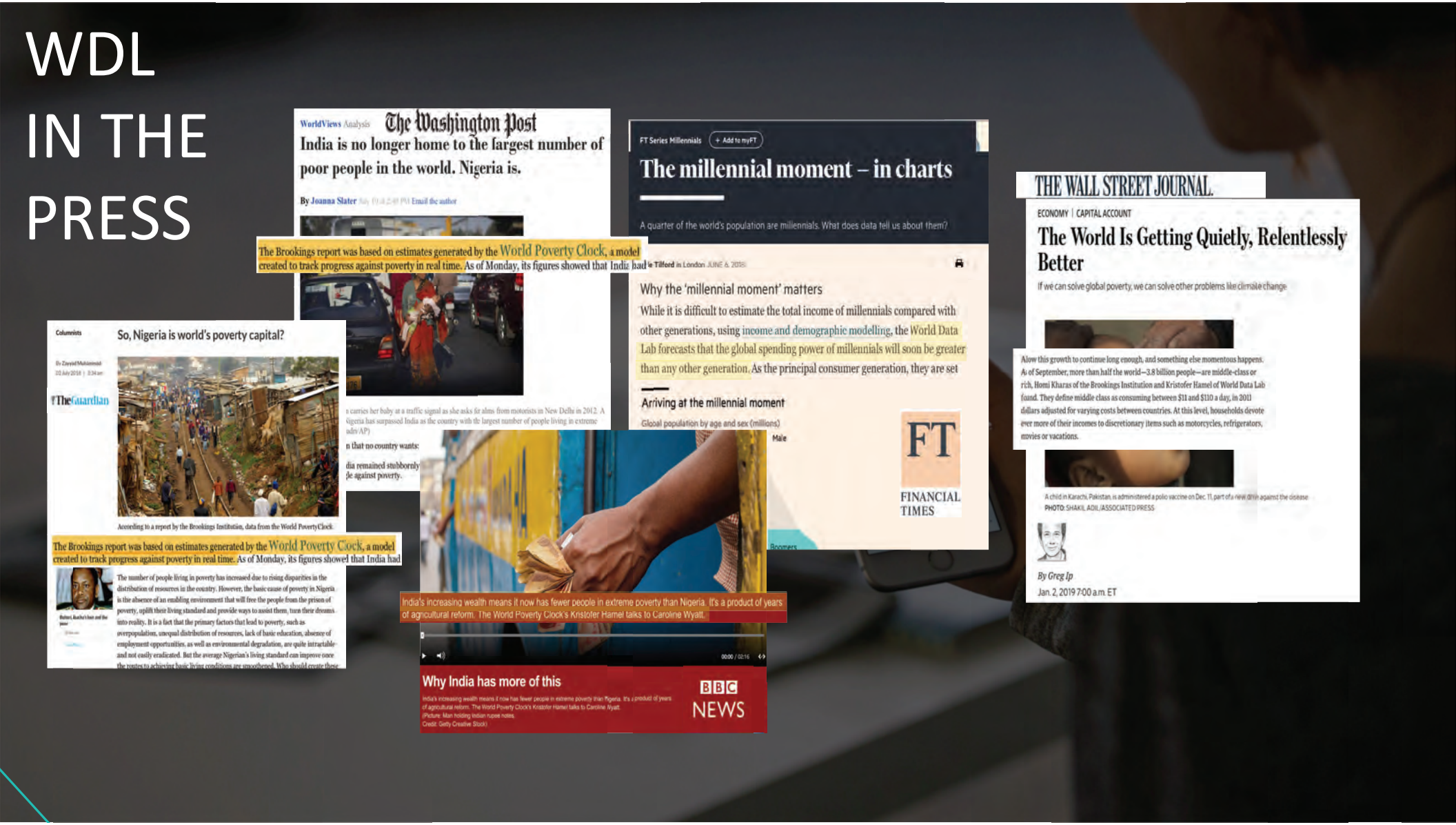
Expanding Global Network

New Approach to data modeling

	Traditional approach	World Data Lab approach
Timeliness	4 years old (narrative of rapid poverty reduction)	Real-time (poverty in 2016 may have increased)
Direction of analysis	Past (Stating the obvious)	Present and Future (focus on progress and benchmarking against needs)
Unit of analysis	Percentages	Actuals
Modelling	Economic, mainly linear	Integrated economic, demographic, climate in consistent shared socio-economic pathways (SSPs)
Communication	General messages for a small number of experts	Focus on raising awareness at global level and sub-national data for policy makers in countries
Illustration	*[In Mali] "between 2001 and 2010, GDP growth averaged 5.7% per year. During the period, the GINI index fell 7 points. The income of the bottom 40 grew, while the mean contracted."* *(World Bank, GMR 2016)*	*Today, Kenya has 11.1M poor people. If Kenya keeps reducing poverty at the expected rate, there will be 5.0M Kenyans in extreme poverty in 2030. It needs to double the rate of poverty reduction.*

worlddata.io

WDL IN THE PRESS



Funders

Partners

World Poverty Clock

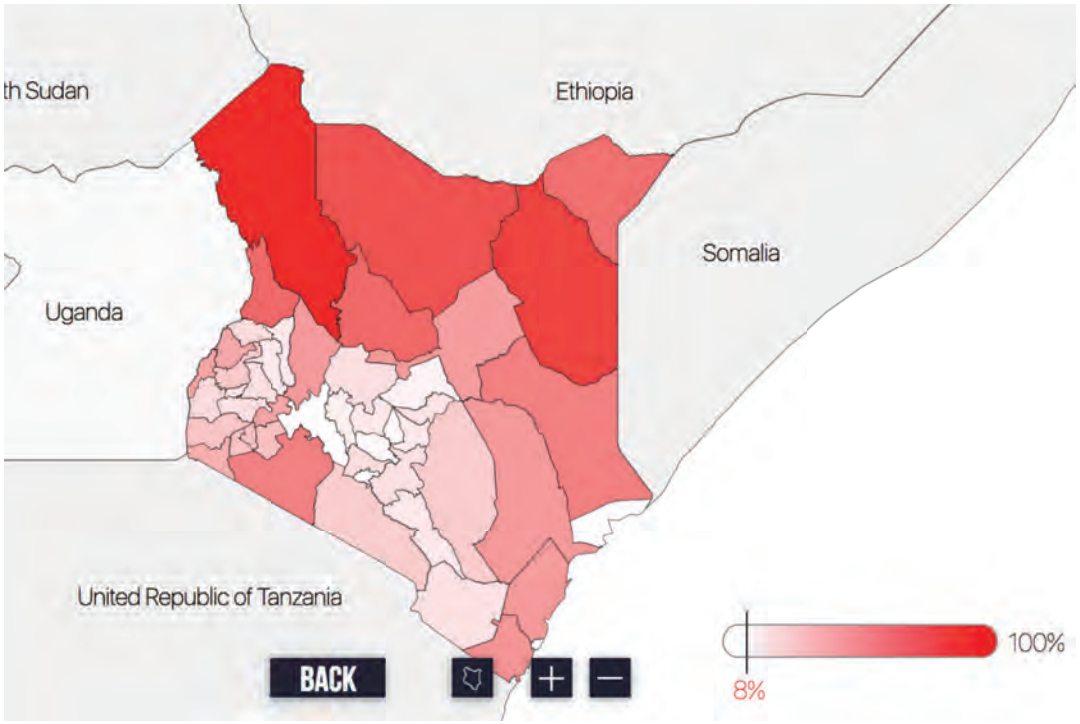
Survey-based poverty modeling

- The World Poverty Clock (WPC) is a global model that tracks poverty in real time.
- This tool measures how well countries are progressing toward SDG 1
- The WPC tool covers 99.7 percent of the world population.
- Peer-reviewed and published methodology on *Nature*'s website

worlddata.io

Subnational Modeling

KENYA



- WDL has developed the first ever, real-time subnational income model that monitors poverty dynamics for all 47 counties of Kenya.
- The model includes numbers for every year until 2030.
- With this tool you can track each region's performance towards SDG 1.

World Poverty Clock

Upgrades in 2019

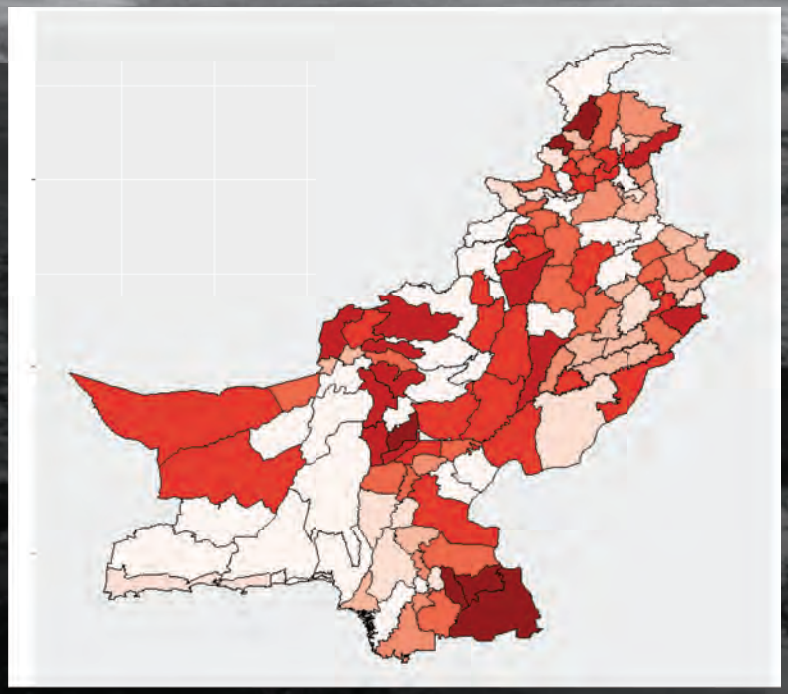
Disaggregation of poverty headcounts by several additional dimensions:

- Gender
- Age
- Rural / Urban

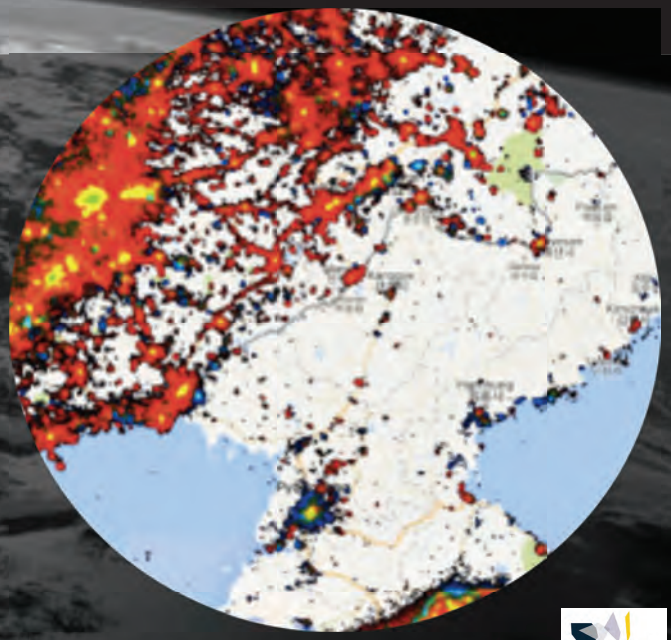
“How many girls live in extreme poverty in rural Niger?”

worlddata.io

Hybrid Projects:
"Nightlights" can improve subnational econometric models



Pakistan



North Korea


Spatial Demography

Spatial Demography

Spatial demography is based on

- new types of geospatial data, and
- advanced methods and technologies.

Spatial demography creates

- new insights into demographic processes, and
- demographic information at an unprecedented level of granularity.

Spatial Demography:
Population Mapping

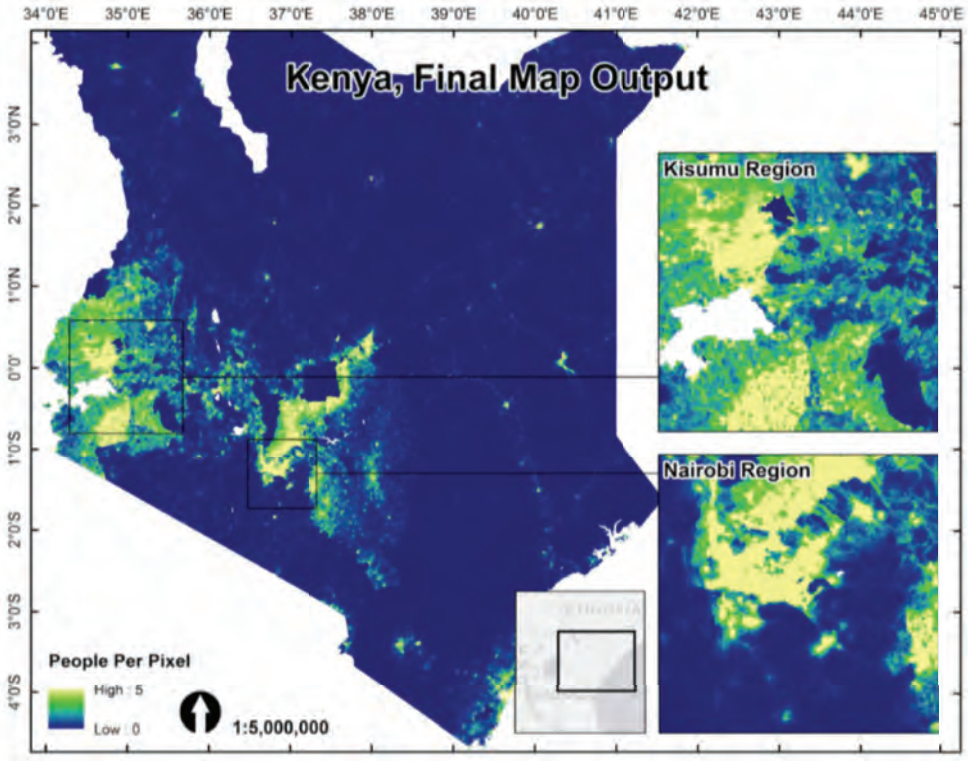
- Create population estimate and forecasts at the 100 x 100 sqm level for Thailand and Philippines

- Method:
  - Create a Random Forest of decision trees (each of which only considers random subsets of covariates and input data)
  - Conduct a covariate selection process
  - Use refined Random Forest to predict population densities
  - Average over estimates of individual trees
  - Obtain robust grid-level results

- Large set of input data: census, landcover, night lights, climatic spatial variation, human presence on landscape

Example: Results

National-scale population data on 100 by 100 meters grid-level



Applications

Survey support
- Validate/benchmark census and other results
- Gap filling for difficult coverage areas

Planning and policy development:
- Forecast key elements of a population (growth)
- Model humans / environment interactions
- Estimate other characteristics (health)

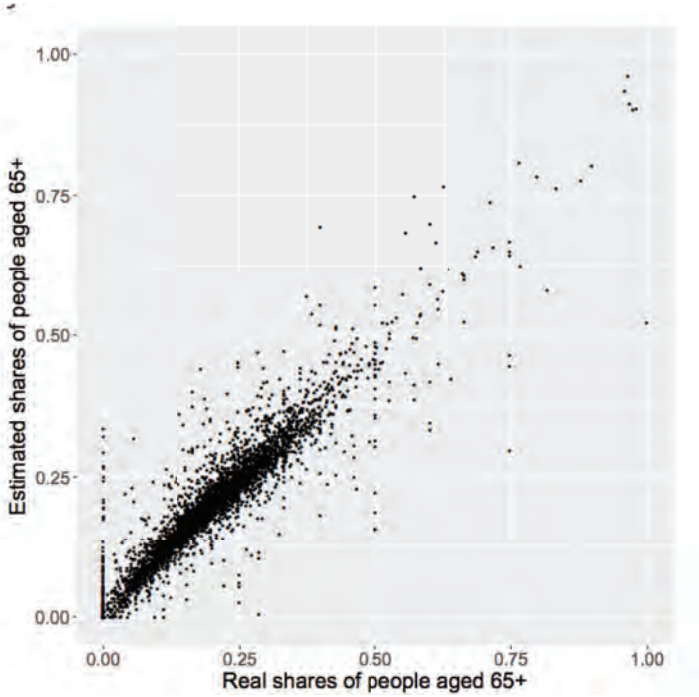
Spatial Demography:
AgeSpot

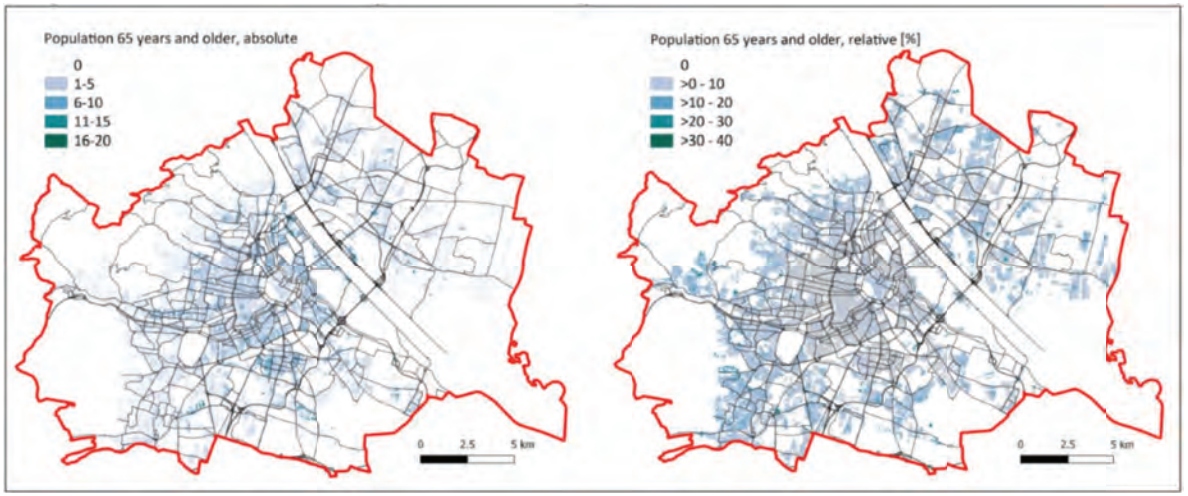
Estimate **number of people by age group** on a granular level (up to 50 by 50 meter grids)

Methodology:

- Bayesian Model Averaging to estimate shares of age groups
- Urban Growth Model to forecast population density

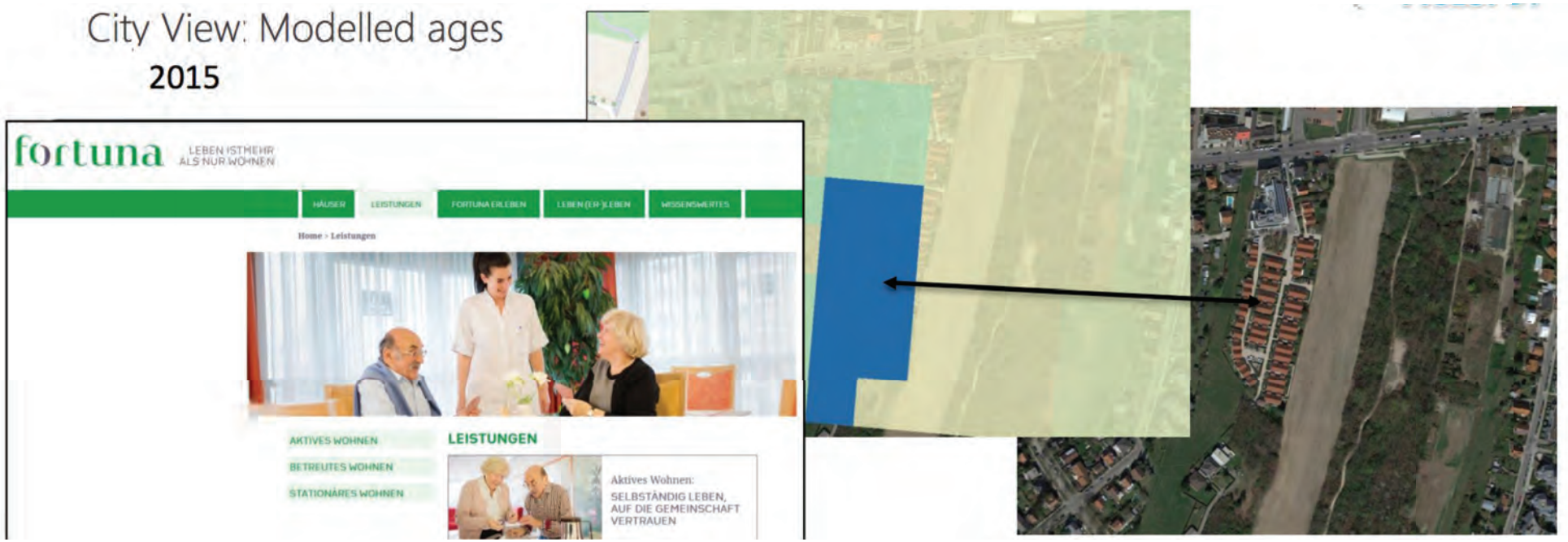
AgeSpot


1.00
0.75
0.50
0.25
0.00
Estimated shares of people aged 65+
0.00
0.25
0.50
0.75
1.00
Real shares of people aged 65+


Population 65 years and older, absolute
0
1-5
6-10
11-15
16-20
0
2.5
5 km
Population 65 years and older, relative [%]
0
>0 - 10
>10 - 20
>20 - 30
>30 - 40
0
2.5
5 km

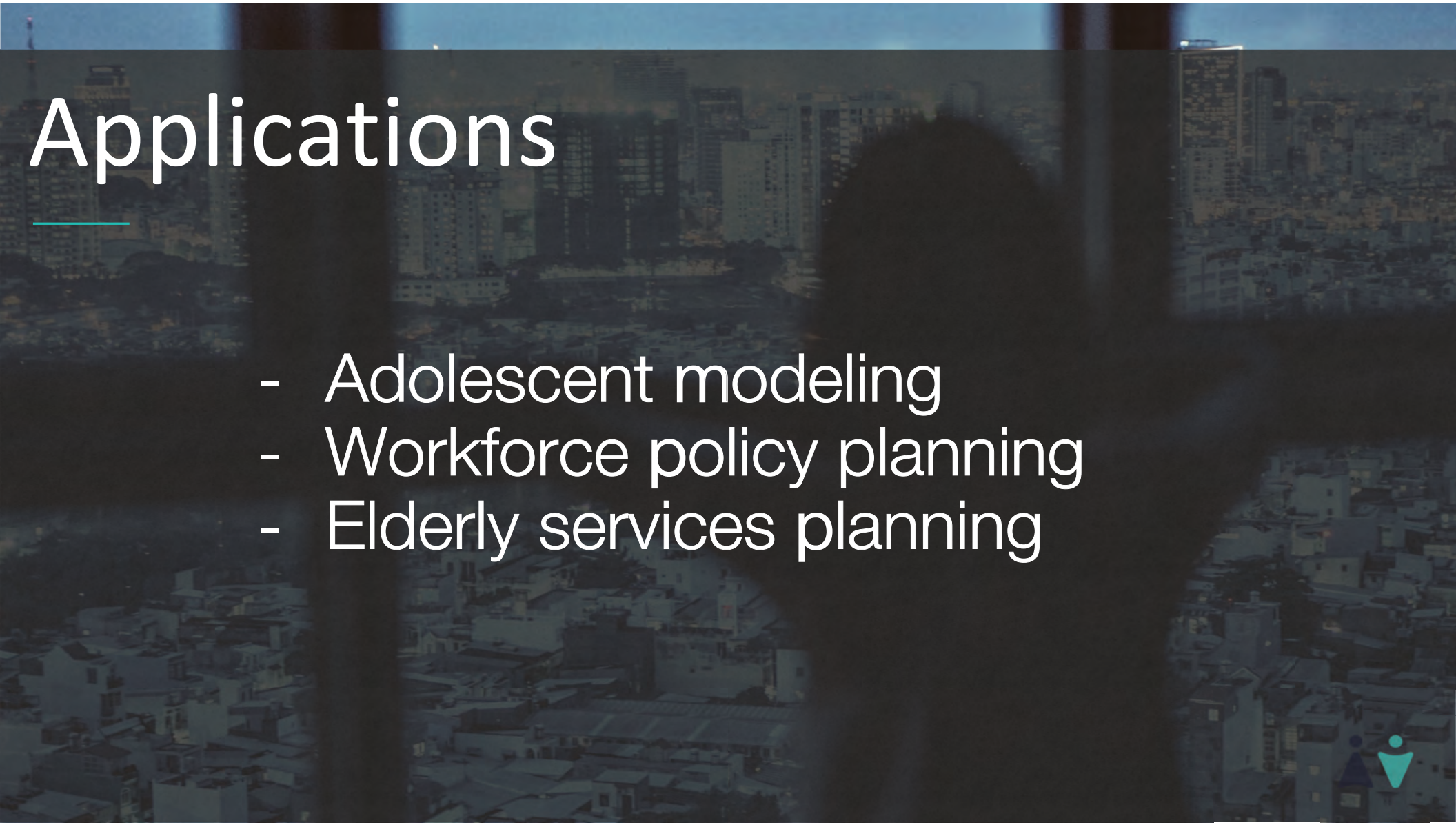
AgeSpot

European Space Agency



Applications

- Adolescent modeling
- Workforce policy planning
- Elderly services planning



Applications

Health Center Modeling

Concept: Develop a model to estimate the expected/optimal location for health centers at a highly granular level.

Approach: Use a variety of inputs (survey input, OpenStreetMap, land-use data and satellite imagery) to model where health centers are most often situated in a given context and to then predict where they could be expected to exist but currently do not

Neural Networks:

Poverty mapping:

Develop ”village-level” poverty estimates and forecasts for Philippines and Thailand.

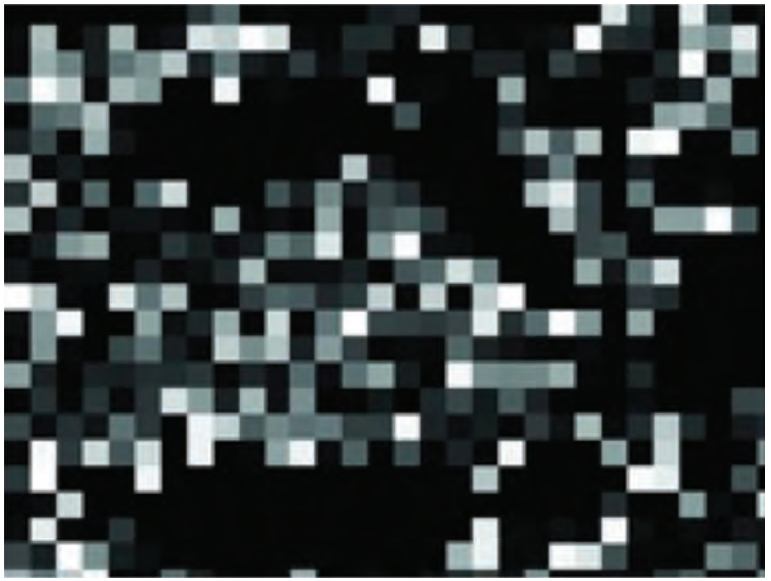
In addition to poverty/income, experiment with modeling and mapping other related variables as well (such as unemployment).

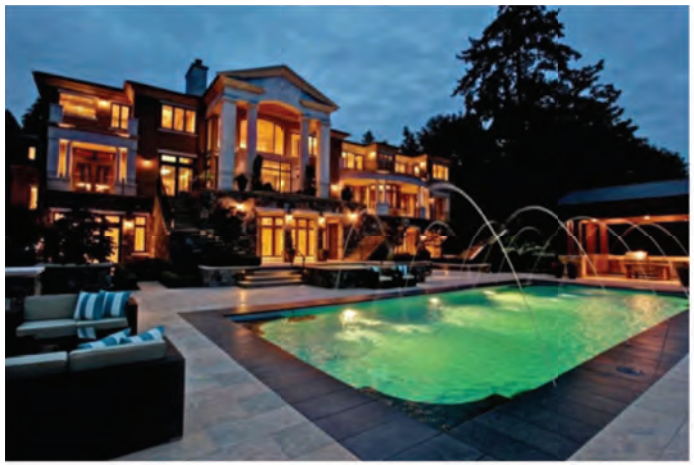
Neural Networks:

Poverty mapping:

- There is a clear relationship between nighttime luminosity data and wealth
- Difficult to distinguish between poor, densely populated and wealthy areas
- Difficult to distinguish differences in economic activity in areas with populations earning less than $1.90 a day

Which place is poorer?



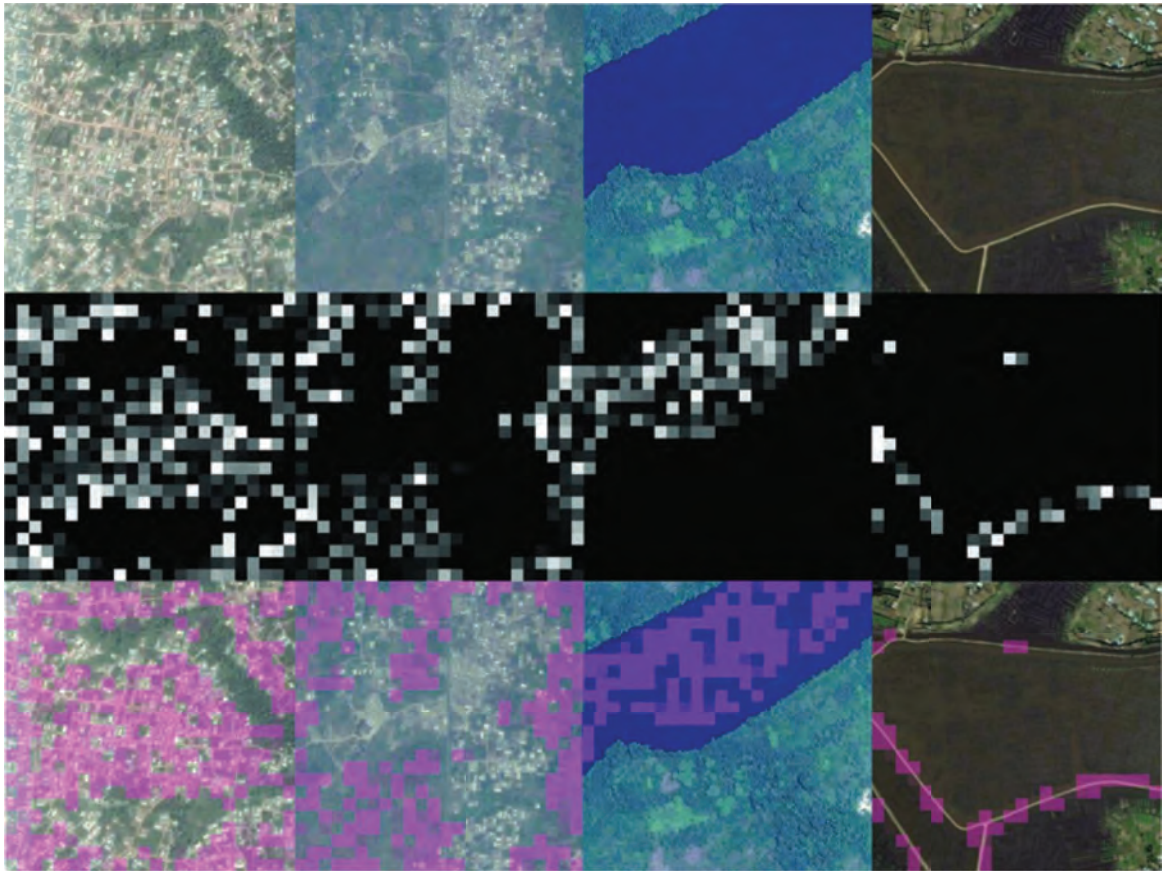


Which place is poorer?

Daylight images

Nightlights

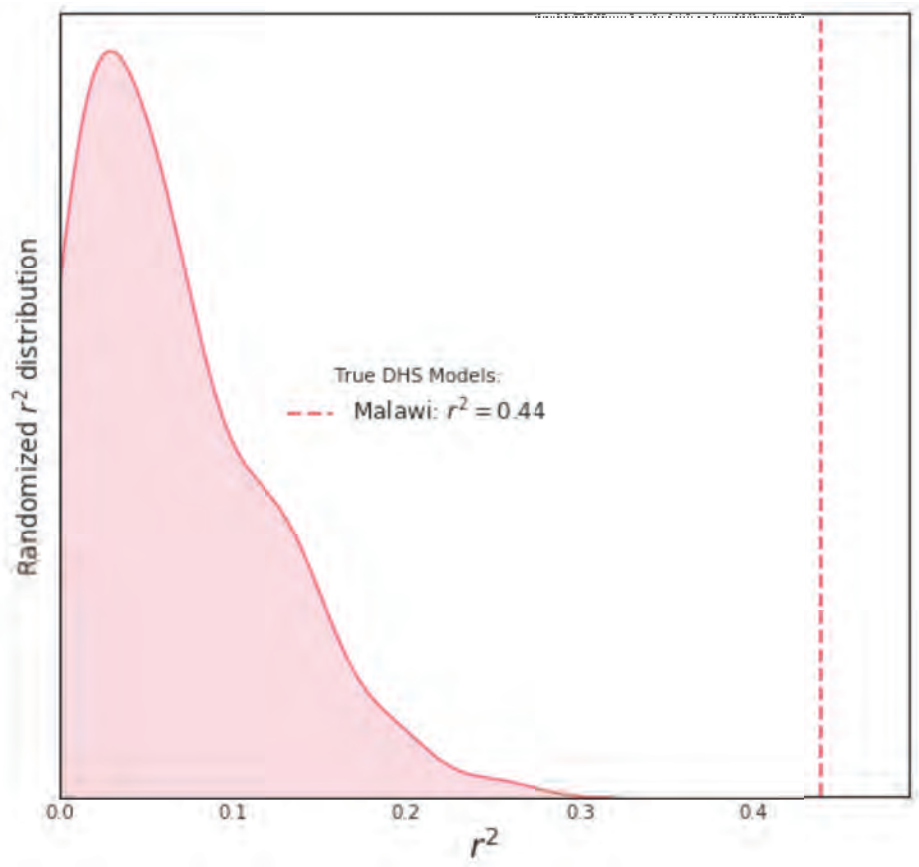
Pattern recognition



Science, 2016

Pilot project: **Malawi**

Method will improve with 10x more images



Applications

- Development policy planning
- Development policy evaluation

Applications

Inequity Mapping

Concept: Estimate inequality at the village level.

Approach: We would use Beta-Lorenz curves to separate the shape and scale of the income distributions as well as Convergence Modeling techniques to forecast mean income/consumption.

World Data Lab

International Workshop on Data Disaggregation for the SDGs