Using 'Insight Tiles' to Map Modern Slavery Risk

Myriad untapped, non-traditional data sources exist in some low-income countries, from communication data to mobile money patterns. Such countries, while being infrastructurally poor, are data rich, having experienced recent digital revolutions often surpassing those in industrialised economies. These data streams can serve as 'proxies' for traditional data sources, such as survey data, which is often unavailable, out of date and costly to collect, to enable social vulnerability and different forms of risk to be quantitatively assessed.

Insight Tile analysis has been developed at University of Nottingham's Rights Lab in order to use these novel data streams (such as Call Detail Records (CDR), Mobile Financial Services (MFS) and Earth Observation (satellite) imagery). Combining them with primary survey data on social vulnerability, and applying Artificial Intelligence (AI) techniques, this approach **generates fine-grained geospatial maps that contain novel insights into key drivers of risk and vulnerability - and can be used specifically to estimate vulnerability to modern slavery at a hyperlocal level.**

Anti-slavery practitioners can use Insight Tiles to:

- Predict risk of modern slavery with a high degree of accuracy using 'non-sensitive' datasets, e.g. features from demographic, social, behavioural and environmental factors that do not directly measure incidence of modern slavery and related forms of exploitation - but which have been recognised as indicators of risk in a particular context;
- Model vulnerability to modern slavery at fine-grained, local levels in order to uncover new, previously unconsidered predictors of risk at the level of local neighbourhoods ($1km^2$);
- Support local actors to develop intervention activities that target specific geographical locations of high risk;
- Measure the effectiveness of interventions – Insight Tiles can be refreshed at regular intervals, not only to examine changing circumstances, but also to help enhance the impact of subsequent interventions.

Using Insight Tiles to identify modern slavery vulnerability in Tanzania: A case study

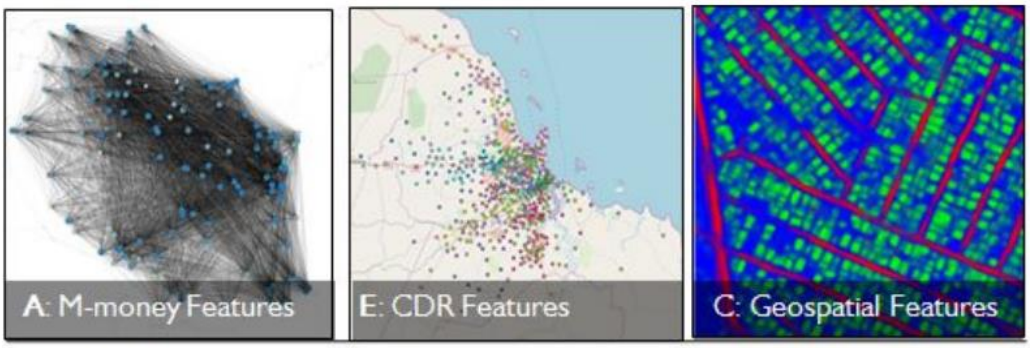
Identifying vulnerable areas in cities can lead to more accurate identification of citizens who are at high risk of modern slavery, and guide pre-emptive interventions. Dar es Salaam, Tanzania, is the largest city in Tanzania, serving as a significant economic hub and gateway for freight to neighbouring landlocked countries. However, owing to rapid urbanisation, the city has doubled in size in just a decade, with 70% of its approximately 6.7 million inhabitants living in vast, informal slums, with poor infrastructure, services and security. A complete dearth of accurate data, even about local population levels, renders traditional risk analysis approaches ineffective; surveys are prohibitively expensive and logistically challenging; there's a lack of political backing; and even when data is collected, it rapidly becomes out of date.

However, using cutting-edge AI techniques, key socio-demographic, infrastructural and descriptive features can be extracted from both satellite imagery and digital data streams (mobile money; call data records; transport patterns; drone surveying). The goal of Insight Tiles analysis is to determine how novel Machine Learning techniques can be best combined with these mass proxy data streams, obtained from commercial data partners, in order to derive otherwise unobtainable fine-grained population data across Tanzania, in order to identify those at high risk of modern slavery. An overview of the research method and its application in Dar es Salaam is set out below.

Methodology	Application in Dar es Salaam, Tanzania
STEP 1: Data collection	To build Insight Tiles models, “ground-truth” data is essential. To this end, 190 local facilitators conducted an extensive street survey across all 452 sub-wards of Dar es Salaam, producing measurements for risk to forced labour, along with a range of “non-sensitive” variables. This extensive survey was co-created by N/LAB and the Rights Lab in collaboration with the in-country team - translators, local volunteers and data collection experts active locally. Surveyed topics covered a host of socio-demographic, infrastructural and environmental factors relating to each subward and its residents. Questions covered a wide range of Sustainable Development Goal-related indicators, from poverty and social vulnerability, to local medical facilities and road conditions.
STEP 2: Primary Survey Analysis	Survey responses were analysed and used to identify a ground-truth risk score for forced labour for each subward. Of the 452 subwards surveyed, 8 indicated the highest possible level of risk. Responses were also analysed to identify co-varying socio-demographic factors that are relevant to risk of forced labour, covering local infrastructure (e.g. building quality, road conditions, healthcare facilities, street lighting); socio-demographics (e.g. age levels, employment, pay levels, school attendance, family living arrangements, the origin of residents); local safety (e.g. daytime safety; night safety; crime; poverty); and behaviour (e.g. mobility; vehicle usage; mobile phone usage; weekend activity; etc.). Analysing these factors allowed for an assessment of risk without the need to have to ask intrusive questions.
STEP 3: AI Feature engineering from Call Detail Records and Mobile Financial Services Data	Informed by the surveys, a host of variables were then extracted from streams such as anonymized Call Data Records (CDR), Mobile Financial Services (MFS) data and M-money (XDR). A CDR is logged every time a network event such as sending an SMS or making a phone call takes place. The anonymised data used as part of this study covered a total of 450.2m calls and SMS events for 330k mobile phone subscribers in the Dar es Salaam region over 122 days. The rich variables that were extracted include the SMS activity in a subward (reflecting population), uptake of mobile money (reflecting financial inclusion), the mobility of residents (reflecting physical isolation), the mean call distance (social isolation), as well as measures of activity at different points in the day, estimates of income, and spend from mobile money transactions. With estimates of as much as 25% of Gross Domestic Product (GDP) reported to be passing through mobile money mechanisms in some East African countries, such data provides a powerful indication of economic activity in the region.
STEP 4: Geospatial Features	Further variables were then generated from data collected via drone and satellite imagery. The features extracted from this dataset include subward area (in km^2), AI-driven land-use classes (the residential area, slum, urban, industrial and unused, connectivity), road and building conditions, and the location of subwards relative to geospatial features such the coast, the central business district, the port and predominant slum and industrial zones.

Step 5: Experimental Setup	Together with the forced labour risk score derived from Step 2, the novel data features (variables) derived in Steps 3 and 4 can then be input into AI models. This allowed the research team to not only predict modern slavery risk from these remote data streams, but helped them to assess the relationships between those variables and modern slavery risk.
Step 6: Results	Results of applying these methods have provided strong evidence that a range of previously unconsidered features effectively predict forced labour risk. The following factors in particular were positively associated with risk of vulnerability to forced labour in an area: - Dense populations (high CDR activity), indicating potential surplus of workers, adjacent to industrial subwards - Low levels of official ID ownership for, specifically, teenage girls and older men and women, as well as female mobile phone ownership - High level of illness and disease, and limited access to medical facilities - A majority of households being extended rather than immediate families - Local pollution and litter, indicating lack of services - Low levels of schooling for children and teenagers

Examples of the data layers than inform the Insight Tiles approach are illustrated below.



M-money makes up ~25% of GDP - it is ubiquitous, and can unearth financial exclusion, dependence on remittances, flows & poverty.

Everyone in Tanzania owns a cell phone. Anonymized **CDR data** can reveal mobility issues, isolation activity and demographics drivers.

Analysis of Drone/Satellite Imagery can reveal land use, infrastructural isolation and environmental problems key to vulnerability.

Insight Tiles can be used in a multitude of ways and contexts to support anti-slavery activities. If you think this technology may be able to assist you in your work, or you would like further information about Rights Lab, please contact rightslab@nottingham.ac.uk