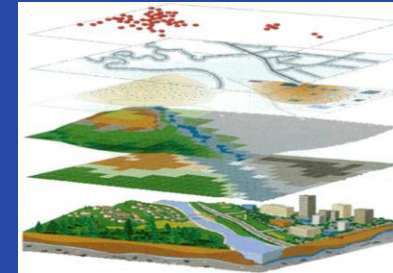
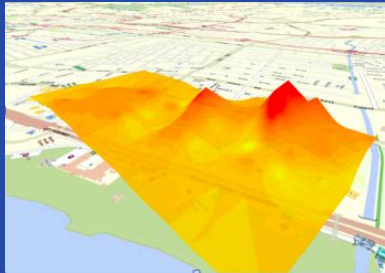


# Integrating Official Statistics and Geospatial Information: Issues and Challenges



**Professor Paul Cheung**  
**Professor, National University of Singapore**

# Location Information Framework

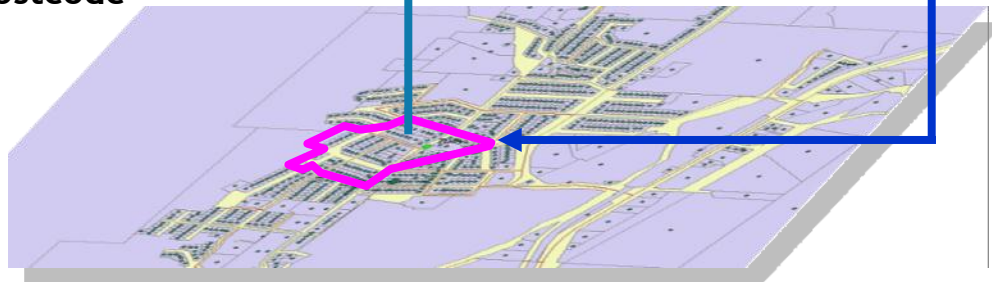
Analysis and aggregation across geographies



Aggregated to Local Government area or higher



Aggregated to suburb or postcode



Location information at address level

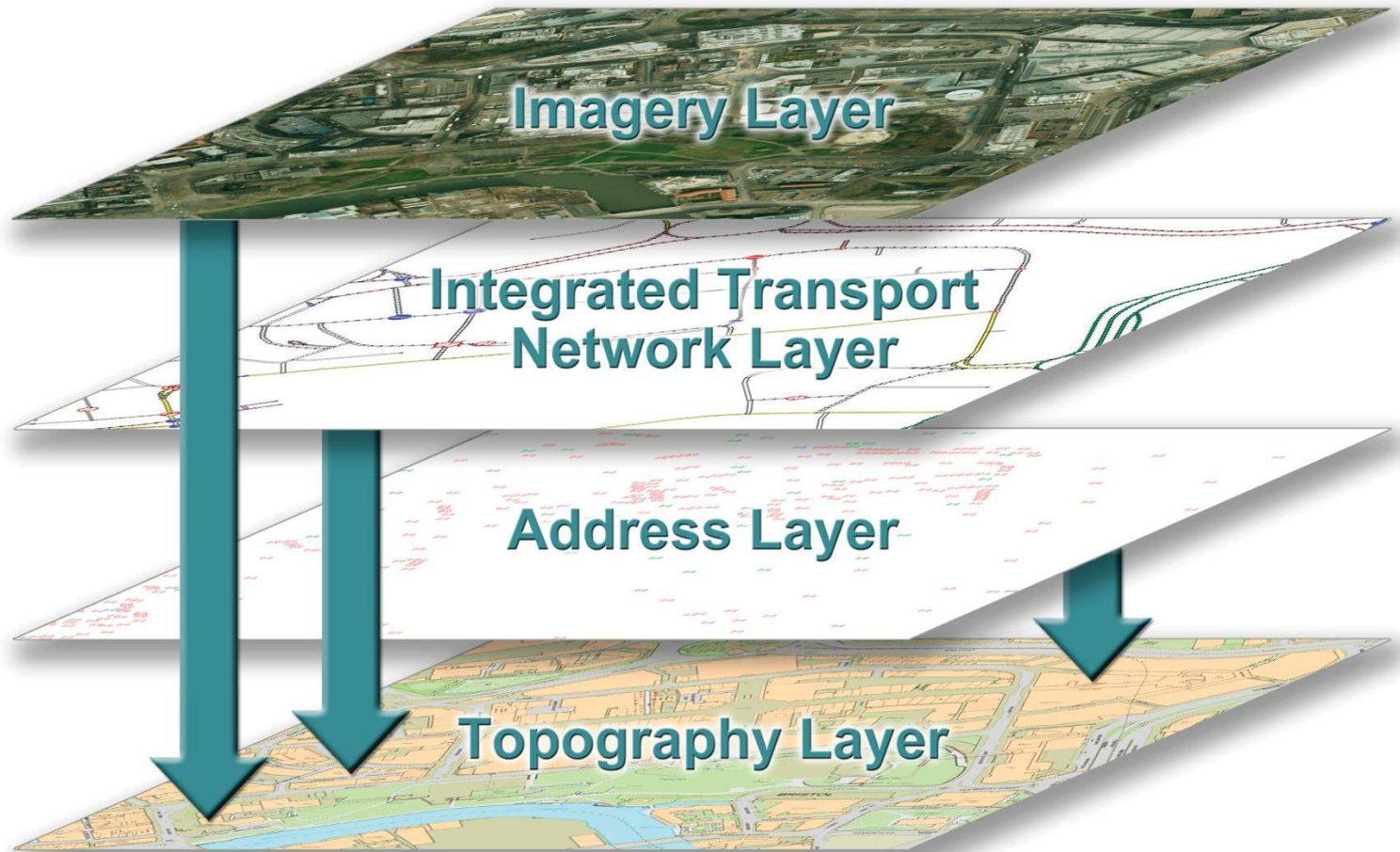


Geocoded unit level data

25 Smith St = x,y: 35.5676, 135.6587



# Mapping layers; Connecting Information



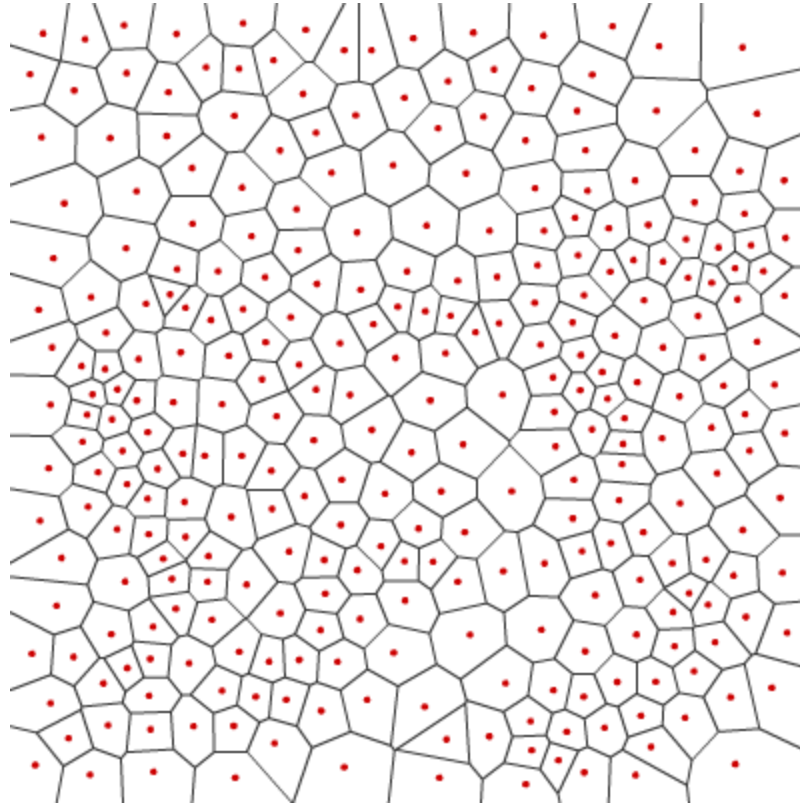
# Two Sources of Information

- Two communities (Official Statistics and Geospatial Information) operating on different analytical schemes and data structures, with minimal overlap;
- Distinct culture, languages and practices;
- Comfortable as distinct professional communities;
- But now compelled by emerging trends to look for the common ground.

What is the Common Ground? How to get there??

# Polygons as Nuclei in Mapping Data Structure

## But they are not the Basic Unit



# Hierarchical Data Structure : Location as Basic Unit of Observation

**25 Smith St, Town Z**  
**x,y: 35.5676, 135.6587**

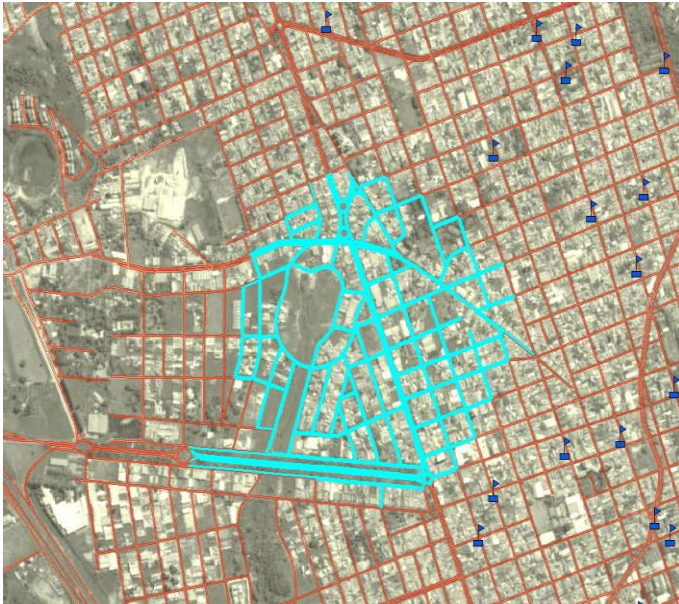


Address / Geocode

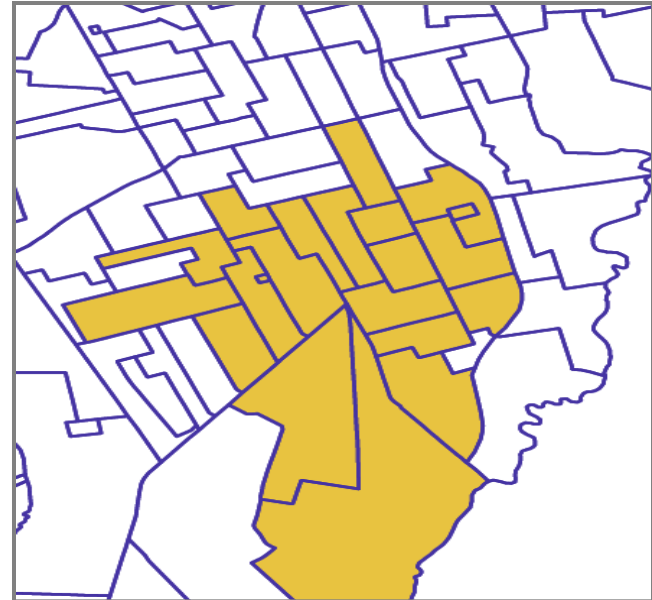


Cadastral property parcels

# Polygons Representing a Unit or Groupings of Units



Block Face

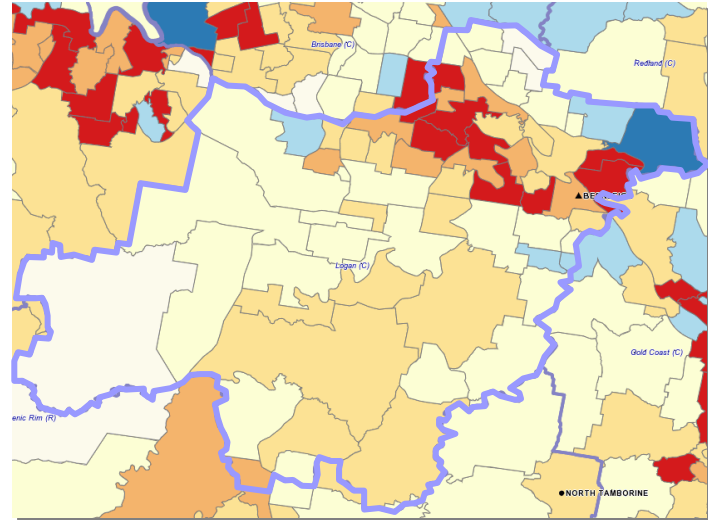


Mesh Blocks

# Higher Level Aggregations



Census Districts/Post codes

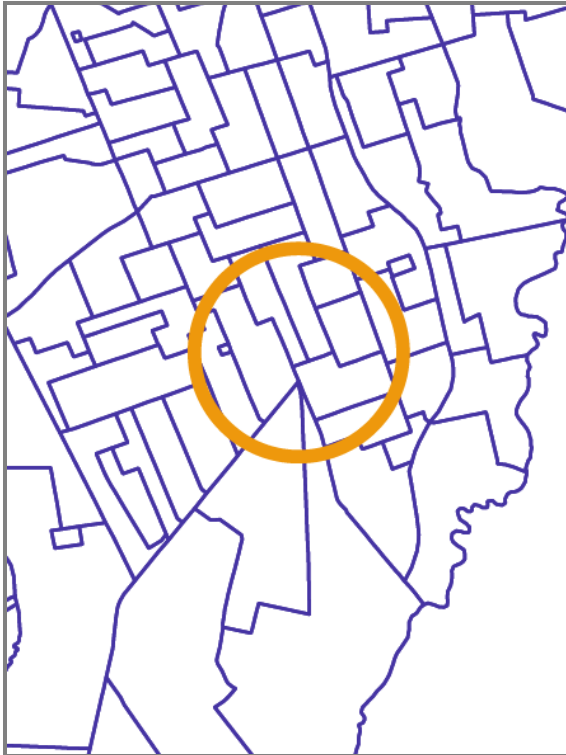


Local Government Areas

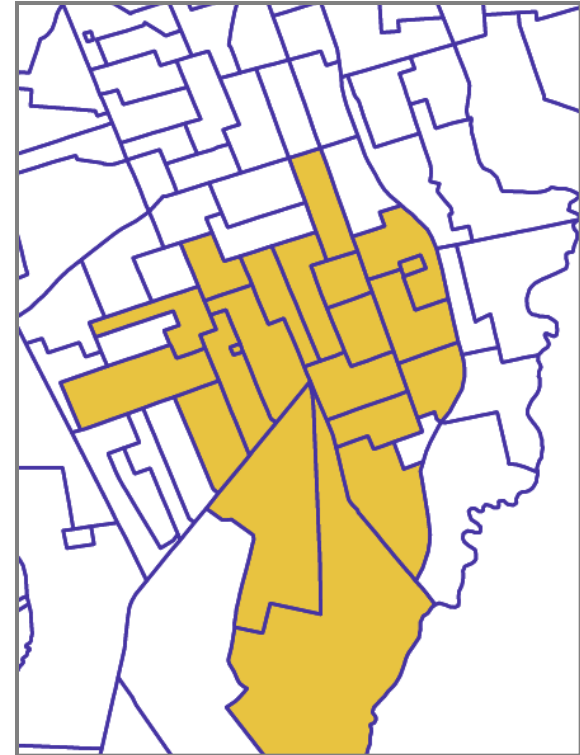


# From Polygons to Points of Relevance (POR)

Users demand increasing precision.  
What is the smallest spatial unit possible??



**area of interest**



**intersection result**

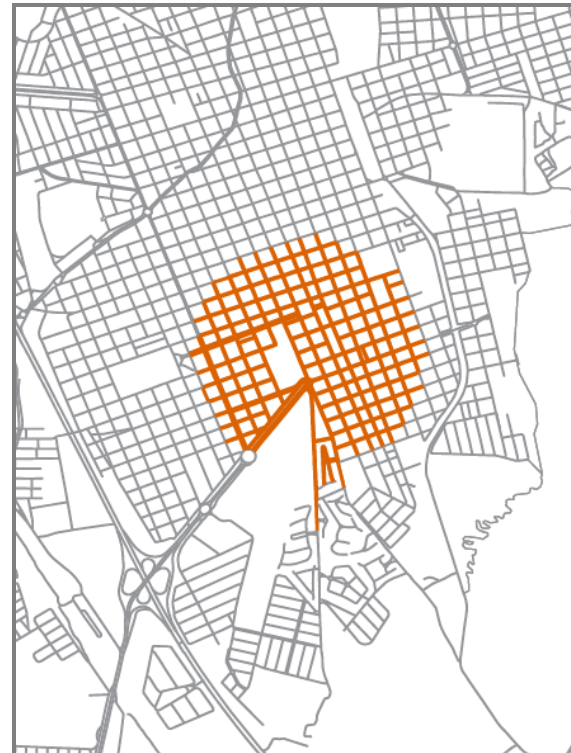
# Smaller Polygons, More Precise Data

Confidentiality the key constraint

But users demand (and will supply) POR data



**area of interest**

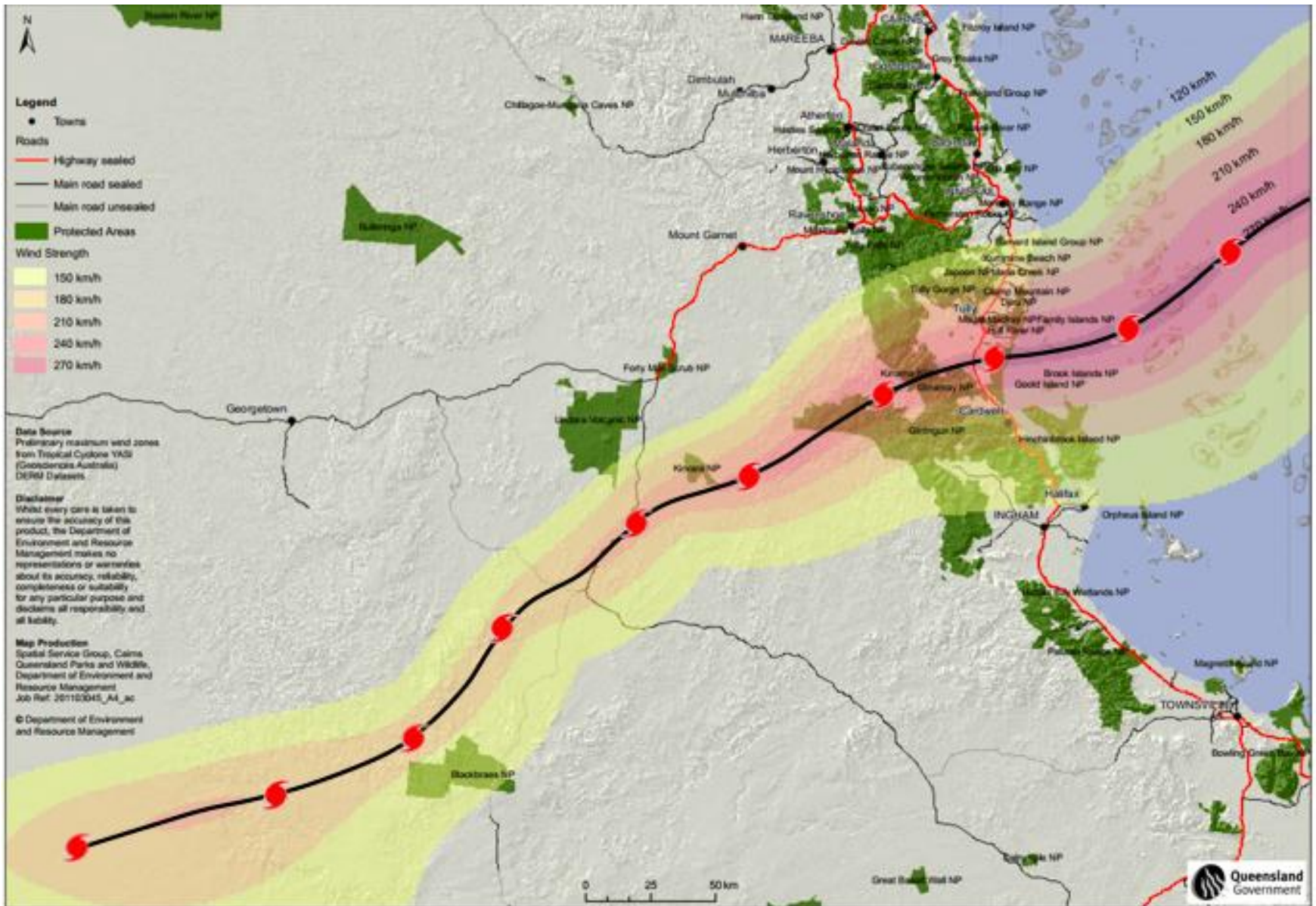


**intersection result**

# From Polygons to Point-Based Information

- Points likely to complement Polygons as the organizing framework for data integration, providing location-specific Information;
- The dynamic movement from Point to Point will pull out packets of Point-of-Relevance information on a string;
- Point-based information will be able to facilitate the convergence of information from multiple sources for a particular location;
- Points identified by Geocodes or Addresses.

# Line Trajectory of Tropical Cyclone Yasi





Address	12 BROSANAN STREET, TULLY
Type	Residential
Condition	Severe
Details	roof missing
Photos	1
Source	Y:TCYasi070211RDA1230FinalTC Yasi RDA

Directions: [To here](#) - [From here](#)



# Matrix: Data Structure for Statistics



Data Glossary Metadata More



## Statistics

Gender Inequality Index and related indicators [Search glossaries](#)

Source: Human Development Indices: A statistical update 2011 | [United Nations Development Programme](#)

[Download](#) [Explore](#)

HDI rank	Gender Inequality Index		Maternal mortality ratio	Adolescent fertility rate	Seats in national parliament (% Female)	Population with at least secondary education (% ages 25 and older)		Labour force participation rate (%)		Reproductive Health			Total fertility rate	
	Rank	Value				Female	Male	Female	Male	Contraceptive prevalence rate, any method (% of married women ages 15-49)	At least one antenatal visit (%)	Births attended by skilled health personnel (%)		
	2011	2011												2010
<b>VERY HIGH HUMAN DEVELOPMENT</b>														
1	Norway	6	0.075	7	9.0	39.6	99.3	99.1	63.0	71.0	88.0	..	..	2.0
2	Australia	18	0.136	8	16.5	28.3	95.1	97.2	58.4	72.2	71.0	100.0	100.0	2.0
3	Netherlands	2	0.052	9	5.1	37.8	86.3	89.2	59.5	72.9	69.0	..	100.0	1.8
4	United States	47	0.299	24	41.2	16.8	95.3	94.5	58.4	71.9	73.0	..	99.0	2.1
5	New Zealand	32	0.195	14	30.9	33.6	71.6	73.5	61.8	75.7	75.0	95.0	100.0	2.1
6	Canada	20	0.140	12	14.0	24.9	92.3	92.7	62.7	73.0	74.0	..	98.0	1.7
7	Ireland	33	0.203	3	17.5	11.1	82.3	81.5	54.4	73.0	89.0	..	100.0	2.1
8	Liechtenstein	..	..	..	7.0	24.0	..	..	..	..	..	..	..	..
9	Germany	7	0.085	7	7.9	31.7	91.3	92.8	53.1	66.8	75.0	..	..	1.5

# Unit Observation in Statistical Collection

- Individual entity as basic data unit (person, household, housing unit, enterprise, community, country);
- 'Location' information of limited interest or focus;
- Data Matrix structure designed for statistical computations, but not for spatial analysis;
- But individual entities can be LINKED through Geocodes

# Building Location-Based Data Structure

- No consistent Geocode to link statistical data to Location;
- Many countries working on National Address Management Framework to define an unique geocode data structure;
- Urgently need location-based data management practices with multiple databases linked through geocode;
- Statistical-Spatial Metadata Interoperability, Integrating SDMX/DDI (statistics) with ISO-19115;
- Need enabling policies and protocols.



# Lessons Learnt from Spatial Data Integration Project, Australia

- Pilot project 2009-2010.
- Integrating statistical population data with geographic information.
- Unit level geocoded (address) data integrated with unit level social data.
- A number of Implementation Problems:
  - Data Formats;
  - Coherence in Geocoding,
  - Integration of Multiple Data Sources.

# Location Analytics: Pulling the Information Together

- Greater, better use of information at specific location helps promote further integration;
- Confidentiality a major issue. Countries need to define clear boundaries. Crowd Sourcing, VGI and mobile device will push this boundary;
- Location Analytics provide location-based evidence to solve problems and gain insights;
- Many organizations actively developing Location Analytics.

# Migration Analytics

## Internal Migration in England & Wales, year ending June 2010



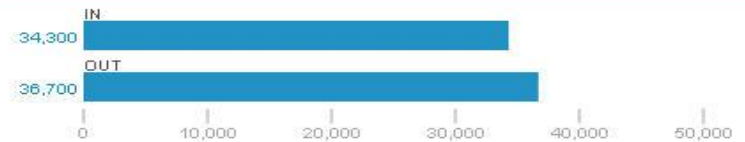
mouseover the map or the graph to see details of migration flows.  
click the 'clear' button to reset the map or use the list to select a different area

To From

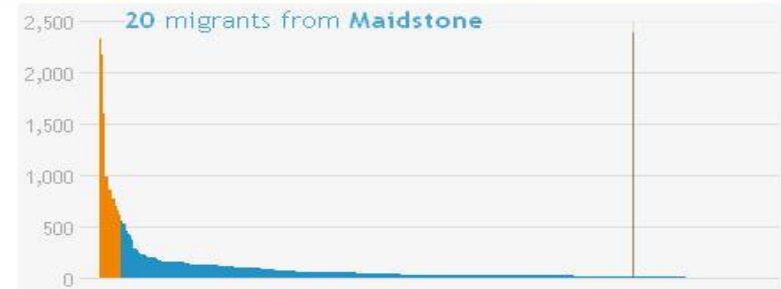
Manchester

### Manchester

inward and outward migration estimates



### inward migration, ordered by total number of migrants



Significant flows highlighted using a method adapted from [Holmes and Haggett \(1977\)](#).

Graphic by [ONS Data Visualisation Centre](#)  
Data source: [ONS Migration Estimates \(published 18th October 2011\)](#)

# An Action Agenda for Information Integration

- Information integration will continue to evolve at a fast pace, pushed by commercial interest and user demand
- Need the United Nations to facilitate collaboration of the two communities globally and nationally in:
  - the promotion and standardization of Geocoding process
  - the development of data management practices enhancing interface of location-based datasets from multiple sources
  - the development of Location Analytics
  - the promotion and sharing of best practices

THANK YOU

[paul.cheung@nus.edu.sg](mailto:paul.cheung@nus.edu.sg)