Is social data big data? Challenges for global social models

In the great chain of causes and effects no thing and no activity should be regarded in isolation

Alexander van Humboldt

Big Data

Big Data

Data is not information
Information is not knowledge
Knowledge is not wisdom
Clifford Stoll

Big Data

Information is not knowledge
Knowledge is not Wisdom
Wisdom is not Truth
Truth is not Beauty
Beauty is not Love
Love is not Music
Music is the best
Frank Zappa, Joe's Garage

Is social data big data?
Challenges for global social models

Big Data + Big {Modelling/ Simulation}

Is social data big data?
Challenges for global social models

Big Data + Big { Modelling/ Simulation }

“Big Understanding”

The best available data compression!

Is social data big data?
Challenges for global social models

Big { **Modelling/ Simulation** }

Process-based global social models are needed

Is social data big data?
Challenges for global social models

Big [**Modelling/ Simulation**]

Statistics are typically non-stationary –
+ non-gaussian, generated by non-linear dynamical systems

→ **Process-based global social models are needed**

Is social data big data?
Challenges for global social models

Big [**Modelling/ Simulation**]

Statistics are typically non-stationary –
+ non-gaussian, generated by non-linear dynamical systems

Everything is connected

→ **Process-based global social models are needed**

Is social data big data?
Challenges for global social models

Big [Modelling/ Simulation]

Statistics are typically non-stationary –
+ non-gaussian, generated by non-linear dynamical systems

Everything is connected

Process-based global *social* models are needed

Economics is not enough!

Is social data big data?
Challenges for global social models

The need for global models:

Globalization

Its a small world!

Global problems

- Poverty, development and urbanization
- Causes and impacts of climate change
- Ecosystems , their services, management and conservation
- Disease and pandemics
- Food Security, land-use change and deforestation
- Warfare and conflict
- Water, energy and other resource use (“peak oil”, “planetary boundaries”...)
- Finance and wealth distributions

Global models

We need to model the “Anthropocene”,

But current “Earth System Models” do not include people

Is social data big data?
Challenges for global social models

The need for global models:

Completeness
- **No magic!**

Capability
- **Ability to represent processes at all scales**

Boundaries
- **Artificial model boundaries may break dynamics**

Time and space scales
- **Larger systems act at longer times**

Ignorance
- **We don't know how social dynamics works**

Is social data big data? Challenges for global social models

Variety and Complexity

What is "social" data? Not just twitter and facebook...

What is needed for faithful dynamical models?

- **Personal** age, sex, height, weight, health...
- **Populational** births,death,family,demographic transition..
- **Material** food, shelter,clothing,furniture,
- **Infrastructural** roads,water,energy,housing,
- **Spatial/placial** home, work, schools, hospitals,pubs...
- **Financial** businesses, stock markets, banks, taxes, poverty...
- **Political** voting,lobbying, power and influence, governance...
- **Industrial** agriculture, engineering,mining,services...
- **Militarial** war,conflict,revolution,
- **Intellectual** books,newspapers,the net...
- **Connectival** family, friends, work colleagues, hierarchies, hard networks,
- **Psychological/ Behavioural** decision making, social interaction...

Dimensionality is very high

Interlinkages are largely unknown

Fragmentation is a problem

Where is the data? Censuses,companies, NGOs,households,governments...

Not longitudinal, non transparent, expensive, not available, not collected...

Is social data big data?
Challenges for global social models

Volume and Velocity

Potentially much bigger than the internet of things!

Supermarket purchase data: At least 1PB

Traffic flow data: 300PB/day

Variability and Veracity

Problematic for psychological/ behavioural information

Visualisation

How to represent high dimensional data at global scale?

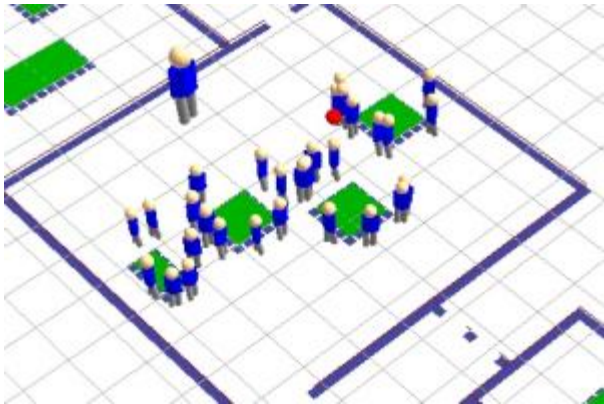
Is social data big data?
Challenges for global social models

Social systems are generally

- Composed of many *interacting* individuals
- Heterogeneous
- Spatially distributed
- Dynamics are generally complex (not just complicated!)
  - Sensitive to initial/boundary conditions
  - Path-dependent/contingent/adaptive
  - Non-decomposable
  - Tipping points/Phase changes
- Multiple interacting system types at different space and time scales

Agent based models

Deal with situations where we lack analytic power
Emergent properties arise from collective interactions
Multiple coupled systems can be dealt with
Test policy options where experiment is not possible
Very visual – good for policy communication
Social processes and networks in real-world situations



Is social data big data?
Challenges for global social models

Challenges I

Models need big memory

How to prioritise what to keep?

How to parallelise efficiently?

Multiple overlapping time-dependent non-local networks?

How to analyse the output?

Large complex datasets need to be challenged against incomplete and uncertain measurement

How to communicate the results?

How to give robust reliable advice when models are not well understood and incomplete?

Summary equations and emulation?

How to model models?

Challenges II

Model ownership
Democratization of knowledge
Policy assessment
Risk and environmental change

Model coupling
Cross-disciplinarity
Sharing and reproducing models/results
Joining complex dynamical models

Problem framing
What should be modelled?
Who for?

Vizualization
System size
Spatial extent
Complex interacting dynamical systems

Scaling
System size
Parameter space exploration
Processes at different scales
Micro-macro links

Validation
Reflexivity
Causality
Data integrity
Handling uncertainty

Complexity
What can be simulated?
How much complexity is “enough”?
How intelligent do agents need to be?