EMERGING METHODOLOGIES IN SOCIAL SCIENCE RESEARCH

- **Clifford Odimegwu**

Demography and Population Studies Programme

University of the Witwatersrand, South Africa

10/5/2021

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

WITS DEMOGRAPHY AND POPULATION STUDIES

school of SOCIAL SCIENCES

WITS School of Public Health

WELCOME ALL

- Social Science has a long history which includes such areas as Anthropology, Economics, Pol. Science, Psychology, Sociology etc.
- Social Science investigates human and social dynamics and subjects at all levels of analysis: decision-making, behaviour, groups, organizations, societies and the world system.

- Social Science investigates in the above disciplines or inter-discipline by using theoretical analysis and experimental results.
- Scale of problems, Social Science study range from micro to macro level
- Eg. Economists can study individual behaviours or predict global demographic changes by using demographic data or data collected by mobile phone.
- Traditional methods are founded as the principles of social investigations. These have their challenges – inaccuracies.

- Therefore, CSS emerges to take advantage of mathematical theories together with data processing and analyzing technologies from computational science to tackle those social issues.
- CSS is the integrated, interdisciplinary pursuit of social inquiry with emphasis on information processing through the medium of advanced computation.
- It is a fledging interdisciplinary field at the intersection of social sciences, computational science and complexity science.

- Statistics as a scientific method plays a prominent role across all the Social Sciences & their specialities.
- Over the past 2 centuries, Social Science (i.e. the study of society) have acquired 3 methodologies: Statistics, Mathematics & Computation.
  - These methods are acquired for purposes of description and induction (Statistics); analytic theoretical development (maths) and simulation of complex systems (Computation).

- CSS is a recent development – second half of the 20th Century and the invention of electronic computers. During the 1960s, Social Scientists began using computers to conduct statistical data analysis – days of SPSS, SAS, punched-card jobs.

- CSS is the integrated interdisciplinary investigation of social systems as information processing organizations and through the medium of advanced computational systems.

- So, computational paradigm in SS has dual foundations: substantive (theoretical) and instrumental (methodological).
- So nowadays, with the help of the above, accessibility of diverse data on human begins, traditional social issues can be investigated from a new perspective.
- With the emergence of CSS, we can look into research issues on Human Dynamics from a data-driven aspect.

- Human Dynamics refer to a branch of complex systems research in statistical analysis.
  - Examples: movement of crowds, queries & other systems of complex human interactions over communication networks including statistical modelling of human networks.

DATA-DRIVEN COMPUTATIONAL SOCIAL SCIENCE

- Individualistic analysis mainly focuses on individual attributes (e.g. basic human features and personal influence) and individual behaviours (e.g. basic human actions and behaviour predictions).
- Collective analysis focuses on community detection, evaluation & collective/group behaviours.

DATA-DRIVEN CSS

Research on relationship involves a discussion on the identification of different types of relationships between individuals.

- Based on individual's predisposition and their historical interactions, predictions of newly appeared relationships and vanished interactions can be made.
- Besides the topic, the procedure to solve the issues include:
  - Data collection
  - Data pre-processing
  - Data analysis (Statistics & ML)
  - Validation.

COMPUTATIONAL SOCIAL SCIENCE METHODS

Automated Information Extraction

- Content analysis
- Method of parsing and coding documents to extract information from data.
- Primary use in the production of events data which can be analysed through various methodologies (time series)
- Also used for anomaly detection and early warning, for monitoring trends & evaluating intervention or program performance.
- Can be used to mine real-time data streams such as news broadcasts or other electronic reports
- Can be used in Ethnography, relying on the production of written records through qualitative research.

CSS METHODS

Social Network Analysis (SNA)

- SNA provides insightful information and inference on the functionality of an organization, given its structural pattern of models and relatives.
- Can be used to design more robust and sustainable networks relevant to public policy (e.g. transportation, public health).
- SNA has numerous applications across Social Sciences by providing a deeper understanding of:
  - Belief Systems
  - Alliance & Treaty Systems
  - International and Transnational Organizations.

CSS METHODS

Social GIS

- Related to the field of spatial statistical analysis, but with a greater emphasis on visualisation of layers of Social data.
- Simulation Models
- Systems Dynamics
- Agent-Based Modeling

- While traditional methods of social sciences are scratching the surface of human society, the new tools of CSS can shed light on social issues from totally different angles.
- Computational social scientists try to exploit the advanced and powerful computation instrument to re-analyse the traditional social science discipline.
- The capacity to collect and analyse massive amounts of data has promoted the evolution of CSS and expedites the emergence of data-driven CSS.
- We now live life with many digital traces: we communicate with our friends through online social media, make mobile phone calls from anywhere, use public transportation with a transit card, and make purchases with a digital account

- We may go out by taxi or train, from which our moving traces are recorded by GPS devices.
- CSS could take full advantage of all these digital traces to better understand the individual and collective behaviors and furthermore to better understand our society and solve social problems

APPLICATION AREAS OF CSS

Application areas of CSS.

Application areas	Goals	Methods
Economics	Understand people behavior in economic activities; personal recommendation service	Agent-based model; recommendation technology
Sociology	Analyzing human interaction; social behavior and social systems	Social simulation modeling; social network analysis
Geography	Human mobility analysis; urban sensing; location-based recommendation service	Complexity modeling; automated information extraction
Public health	Anticipate and track disease outbreak; understand the spread of disease and prevention	Statistic; social simulation modeling
Environment	Nature preservation; pollution measurement	Data mining; big data analysis
Politics	Understand human voting behavior; party opinion diffusion analysis	Computational social choice

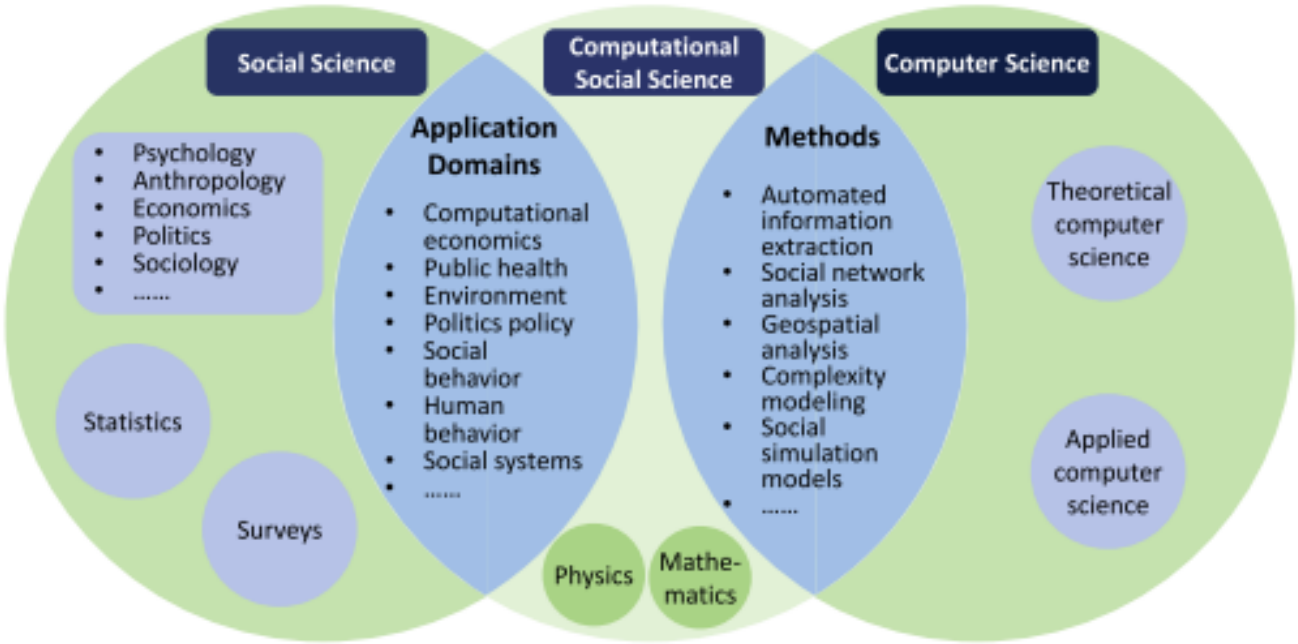
APPLICATION AREAS OF CSS

- CSS has huge potentials for applications in understanding human behaviors and social dynamics from various aspects. Examples
  - group interactions could be examined through online social media
  - CSS can address questions about temporal dynamics of human behaviour: what is individual mobility pattern? How does human relationship evolve? How to detect human community from social media data.
  - Scholars in this emerging area can use related data to solve these social problems with advanced computing technologies
  - As in above Table, CSS has many application fields such as economics, sociology/demography, geography, public health, environment, political science.

INTERDISCIPLINARY

- Important feature of CSS is the interdisciplinary nature. CSS is an integrated and interdisciplinary new area which aims to analyse previous and present social issues and human behaviors with an emphasis on information processing through the methods of advanced computation. Figure below:
- CSS mainly focuses on analysing the issues studied by social scientists with the computational methods which have been developed by computational scientists
- It is interdisciplinary: Social scientists provide the insight into such research issues as economics, politics and environment while the computer scientists contribute expertise in developing mathematical methods such as SNA, Data mining and machine learning
- CSS is a fundamental interdiscipline including all social science disciplines, applied computer social science and other related disciplines

INTERDISCIPLINARY FOCUS OF CSS



RESEARCH TOPICS IN CSS

1. Individuals: -
    1. Individual attributes
    2. Individual influence
    3. Individual behaviors
        1. Human actions
        2. Influence factors
        3. Behavior prediction
2. Relationships: individuals are embedded in social networks and therefore are connected to a neighbouring group of people

RESEARCH TOPICS IN CSS

1. Relationship identification
2. Relationship prediction
3. Relationship evoluation

RESEARCH TOPICS …

1. Collectives
    1. Community detection
    2. Community evolution
    3. Community behaviors
        1. collective decision-making
        2. Cooperation
        3. Contagion
        4. Human mobility

RESEARCH METHODOLOGY

- Data collection
  - Web resources
  - sensing data, via sensors, mobile devices,wireless etc can be used to collect data
  - Data from self-designed experiments

RESEARCH METHODOLOGY

- Data Pre-processing
  - Data cleaning
  - Data reduction – data reduced representation of original data sets
  - Data reformation
  - Data analysis

RESEARCH METHODOLOGY- DATA ANALYSIS

- Large data set calls for automated methods of data analysis, done by what is called Machine Learning.
- Machine learning is defined as a set of methods that can automatically detect rules in data and then to predict future data or perform other kinds of decision-making under uncertainty
- Later this week, w will cover Machine Learning…

CONCLUSION

- The availability of unprecedented amounts of data on human interactions in different social spheres opens the possibility of transceding existing social behaviour boundaries.
- These data can be used to validate the results of simulation models and socioeconomic theories.
- We should try to read and understand more about CSS and begin to apply them in our social science research across Africa!

THANK YOU