

COMPLEX-IT:

An AM-SMART Platform for non-expert
use of computational methods for data
exploration

Brian Castellani, PhD. FAcSS

Professor of Sociology

Director, Durham Research Methods Centre

Director, Wolfson Research Institute for Health & Wellbeing

Durham University, UK



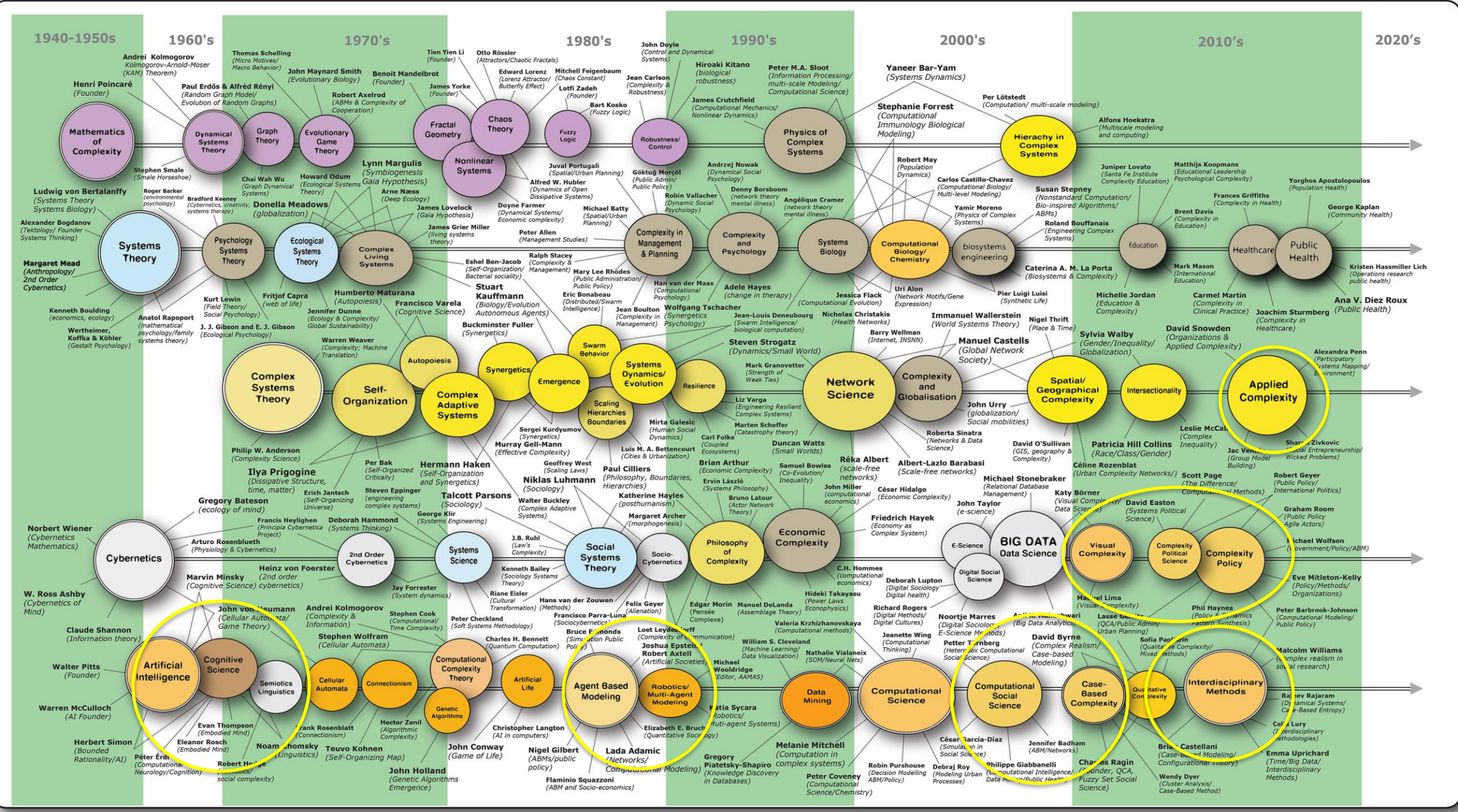
The Field of Case-Based Complexity

Case-based complexity is a suite of interdisciplinary methods first advanced by David Byrne and colleagues as an improvement on the conventions of case-based qualitative configurational analysis (QCA).

While marginal to mainstream computational methods, case-based complexity is an established field of study, particularly in sociology, policy studies, political science, governance, urban planning and public administration.

Case-based complexity is grounded on Byrne's novel insight that cases meet the definitional critique of complex systems.

An excellent introduction to this methodology is Byrne and Ragin's *Sage handbook of case-based methods*.



Case-based complexity can be divided into a more **computational** versus **qualitative** approach.

On the computational side are techniques such as cluster analysis and dynamic pattern synthesis.

On the qualitative side is process tracing and trajectory-based qualitative comparative analysis (TJ-QCA).

See, for example:

Cooper, B., & Glaesser, J. Using case-based approaches to analyse large datasets: a comparison of Ragin's fsQCA and fuzzy cluster analysis. *International Journal of Social Research Methodology*. 2011: 14(1), 31-48.

Haynes, P. *Social synthesis: Finding dynamic patterns in complex social systems*. Routledge, 2017.

Krueger, K., & Wright, M. Theory amidst complexity—using process tracing in ex-post evaluations. *New Directions for Evaluation*, 2022 (176), 119-128.

Case-based complexity is anchored in **four core epistemological arguments**:

1. Cases are the methodological equivalent of complex systems – that is, they are emergent, self-organizing, nonlinear, dynamic, etc – and therefore should be studied as such.
2. The case and its trajectory across time/space are the focus of study, not the individual variables or attributes of which it is comprised.
3. Cases and their trajectories are best treated as composites (profiles), comprised of an interdependent, interconnected sets of causal conditions, factors or attributes.
4. The wider social contexts/systems in which cases are situated needs to be considered.

CASE-BASED MODELLING AS CONFIGURATIONAL THINKING

Configurational theorising pushes the researcher to engage in three distinct ways of thinking about complex social causality that are, in combination, theoretically innovative.

The first, and perhaps most original, is *causal asymmetry*: the idea that the configuration of causal conditions that lead to some outcome may be very different from the configuration of conditions that leads to the absence of that outcome.

For example, the causal conditions that account for high performing, affluent schools can differ from those that explain the absence of high performance in economically deprived schools.

CASE-BASED MODELLING AS CONFIGURATIONAL THINKING

The second is that of the pair of *equifinality* and *multifinality*.

Equifinality concerns those instances where different configurations of conditions co-occur with similar outcomes.

Multifinality is the opposite of equifinality. It expresses that similar configurations of causal conditions can co-occur with the outcomes.

CASE-BASED MODELLING AS CONFIGURATIONAL THINKING

The third is *conjunctural causation*: the idea that a single condition impacts an outcome through its qualitative causal linkages with the other conditions in a configuration.

This way of thinking about causality is similar to what Warren Weaver definition of *organised complexity*, where the factors in a configuration are deeply interrelated forming an organic whole, such that any one condition's impact on an outcome requires an understanding of the others.

Configurations as Complex Systems

Here we will develop the idea of cases as complex systems a bit more into a formal outline:

- Cases are complex systems and complex systems are cases.
- Cases can be all types of social actors, from governments and organisations to school systems and cities to existing group of people positioned at the intersection of multiple systems of oppression.
- As a complex system, some cases, such as a school system or a poor urban community, can be comprised of a set of cases, acting as agents. Other cases, such as a person, will be singular in their social agency. Depending upon the desired level of granularity, a case, such as a government or country, can be treated as a singular entity with a singular configuration – as is done in economic complexity, for example, or diplomacy studies.

Configurations as Complex Systems

Here we will develop the idea of cases as complex systems a bit more into a formal outline:

- The construction of the case and its categories (causal conditions) are open to interrogation and are recognised as necessary traces of the system of study.
- Cases have boundaries, even if structurally or functionally open-ended, but again there is always a concern for the boundaries that categories create.
- Defining the boundaries is therefore a matter of focus, scale and agenda relative to some given outcome.
- Cases, as systems, are self-organising and emergent, where the whole is more than the sum of its parts; and yet the parts are important, as they help to increase the complexity of social inquiry and our understanding of causal complexity and to bring attention to intersections (nexus points) that are otherwise erased, marginalised or ignored.
- Causal conditions are the intersecting structures of a case, as a system. They may appear as intersecting forces of oppression or social factors.

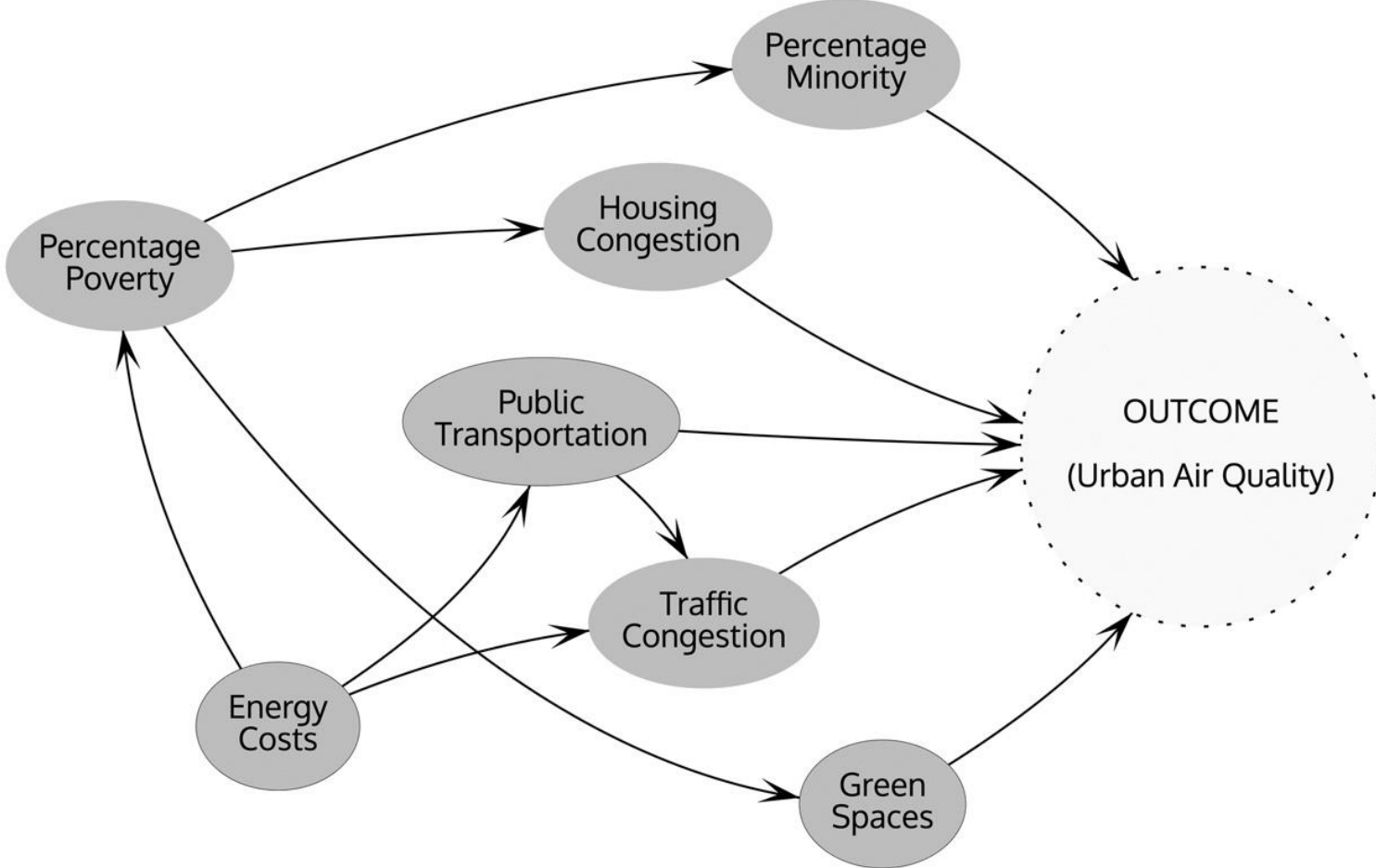
Configurations as Complex Systems

Here we will develop the idea of cases as complex systems a bit more into a formal outline:

- Configurations and their intersectional arrangements are the organisational patterns of a case, which can take the form of complex networks.
- The qualitative interactions amongst the configurations making up a case, relative to some outcome, constitute its dynamics, including issues of agency and power relations.
- Qualitative interactions are nonlinear, comprised of feedback loops, and constitute the case's complex causality.
- In terms of complex causality, both systems and cases are best understood in terms of causal asymmetry, equifinality, multifinality, conjunctural causation, and necessary and sufficient conditions.

The System in the Configuration

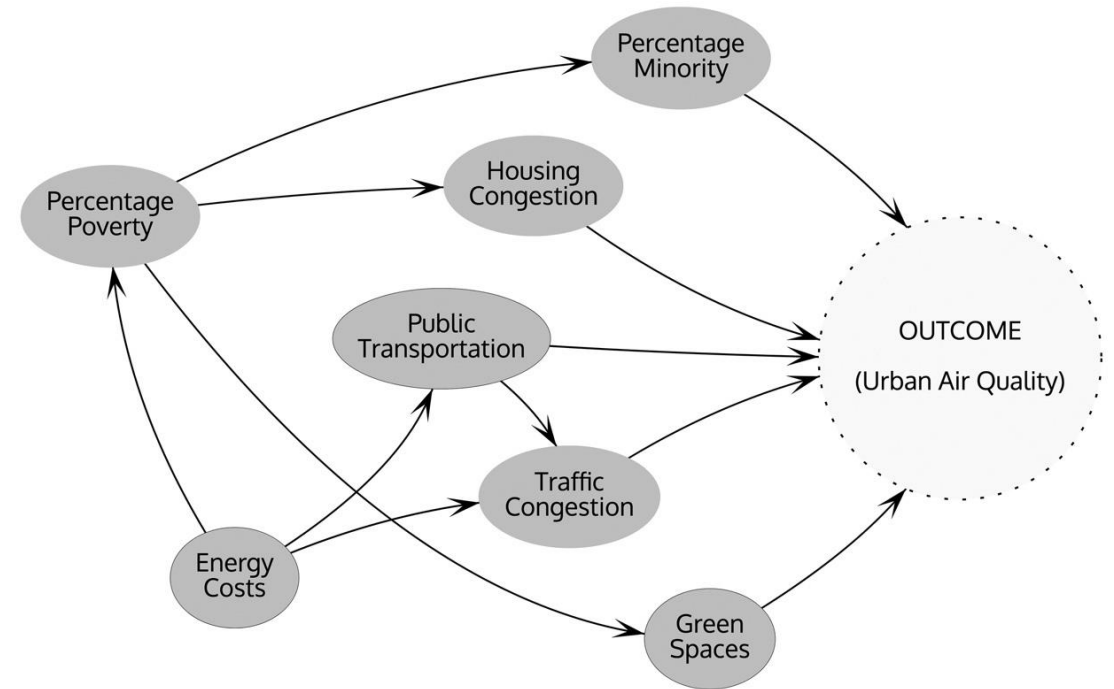
To help demonstrate how one can think about case-based configurations as systems (and vice versa), consider the following example. Suppose you had a systems map, like the one shown here.



The System in the Configuration

While the map is useful for making better sense of the policy system in which this public health issue is situated, the ultimate question for your urban community stakeholders is how this relates to them?

- How can they make use of the systems map?
- What does it tell them about which social determinants to address to improve air quality for their community?
- In other words, how do you put the case and its agency, identity, and relations of power back into the complex system of study?
- That is where configurational theory comes into play.



	ENERGY COSTS	HOUSING CONGESTION	PUBLIC TRANSPORT	TRAFFIC CONGESTION	PERCENTAGE IN POVERTY	PERCENTAGE MINORITIES	GREEN SPACE	OUTCOME (Urban Air Quality)
COMMUNITY A	High	Medium	Extensive	Severe	High	High	Poor	Poor
COMMUNITY B	Medium	High	Moderate	Moderate	Moderate	High	Poor	Poor
COMMUNITY C	Low	Low	Extensive	Low	Low	Low	Excellent	Excellent

To make use of the map you shift from constructing your system to understanding life within it – in this case, the air quality of a given urban community as brought forward through conjunctions of conditions, as shown in our map.

The result are shown in tablature form in the Table here.

- It features the three communities (the cases), each with its own vector configuration.
- The configuration is comprised of our seven conditions, ranging from such social determinants as energy costs to traffic congestion to access to green space.
- The outcome (as in the map) is urban air quality.
- Each community now has a range of scores on these seven conditions.

	ENERGY COSTS	HOUSING CONGESTION	PUBLIC TRANSPORT	TRAFFIC CONGESTION	PERCENTAGE IN POVERTY	PERCENTAGE MINORITIES	GREEN SPACE	OUTCOME (Urban Air Quality)
COMMUNITY A	High	Medium	Extensive	Severe	High	High	Poor	Poor
COMMUNITY B	Medium	High	Moderate	Moderate	Moderate	High	Poor	Poor
COMMUNITY C	Low	Low	Extensive	Low	Low	Low	Excellent	Excellent

Looking at the scores, Communities A and B struggle with issues of poverty and deprivation, sitting at the intersection of various interlocking systems of inequality, as in the case of medium to high levels of housing congestion, moderate to severe traffic congestion and limited access to green space.

Minorities are also disproportionately living in these two urban communities.

In switching to the table, you are effectively turning the systems map on its side to create a seven-factor configuration

Case-Based Modelling

Working within the general framework of case-based complexity, Castellani and colleagues have developed *case-based modelling*.

While primarily computational, case-based modelling is an interdisciplinary methods platform that employs a variety of techniques.

SEE:

Castellani, B & R. Rajaram (2023) *Big Data Mining and Complexity*. Volume 11 of the SAGE Quantitative Research Kit. For more on this approach, see our website: <https://www.art-sciencefactory.com/cases.html>

In terms of the core assumptions of case-based complexity listed earlier, case-based modelling positions itself as follows:

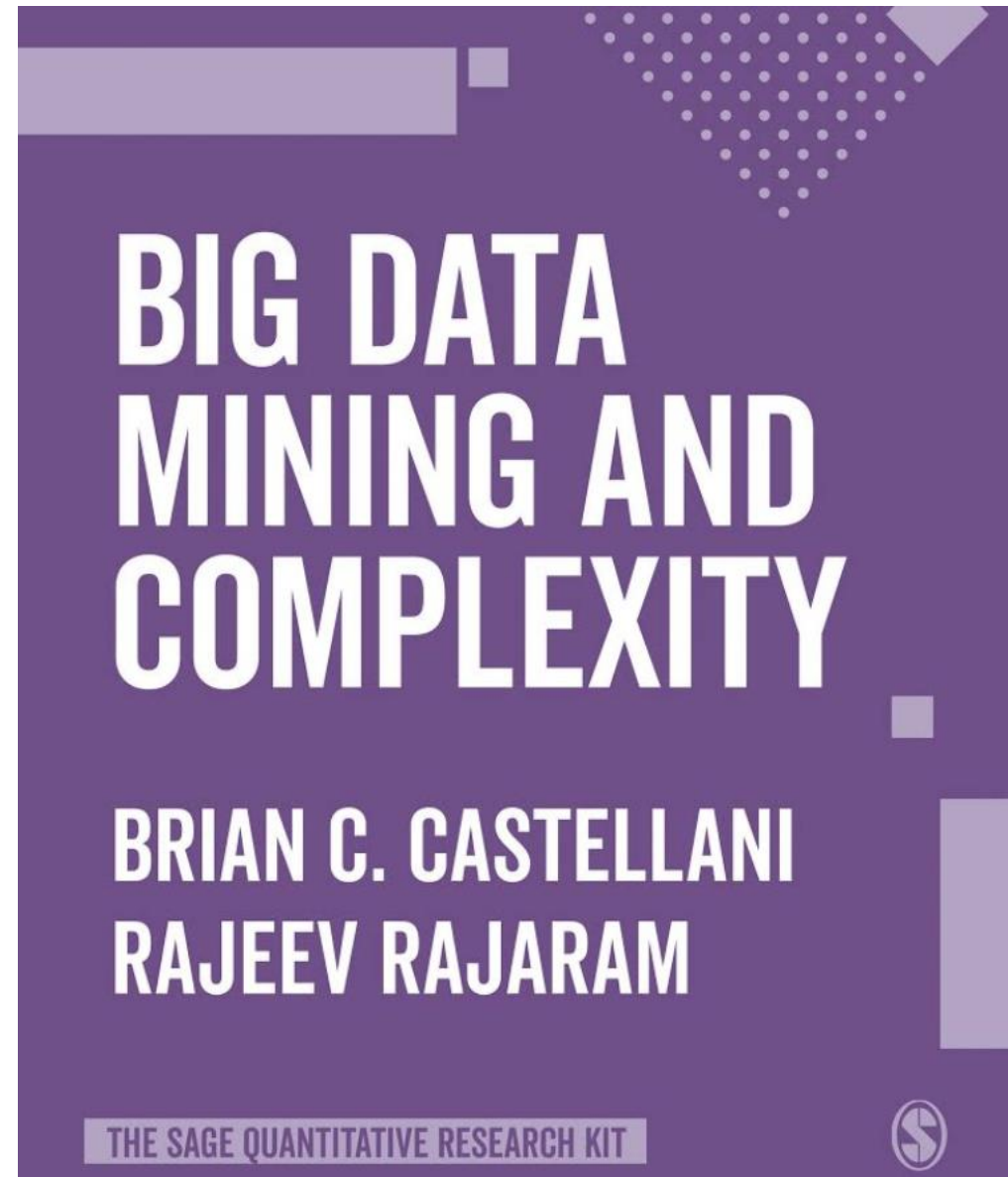
- Cases are best viewed as complex systems. In some instances, this means exploring a dataset as a complex system comprised of interdependent cases; in other instances, given an absence of relationships, it means exploring the cases in a dataset as independent complex systems.
- Cases and their multiple trajectories are dynamically evolving across time/space and, therefore, should be explored to identify their major and minor trend, as well as issues of equifinality and multifinality.
- These trends also should be explored in the aggregate for key global–temporal patterns, as in the case of spiralling sources and saddles.
- The social interactions amongst cases are also important, as are the multi-level social contexts and complex systems in which these relationships take place.

A distinct advantage of case-based modelling is the set of mathematical formalisms it has developed, which allow users to treat the compendium of computational modelling techniques – from network analysis to machine learning – as case-based.

In making this mathematical move, case-based modelling creates a platform for combining computational and qualitative methods, including QCA, systems mapping and AM-Smart methods.

For an overview of these mathematical formalisms, see *Big Data Mining and Complexity*.

SEE Castellani, B & R. Rajaram. Big Data Mining and Complexity.





COMPLEX-IT

Run Online or Download for R-Studio



WEB
VERSION



DOWNLOAD
VERSION

<https://www.complex-it-data.org/>



Build Your Model

1. Build database and import your cases

2. Cluster your cases

Confirm & Explore Your Model

3. Use AI to confirm your cluster solution

4. Compare and visualize your results

Run Scenario Simulations

5. Simulate your scenarios, policies, and interventions

Run Data-forecasting/classification

6. Use AI to predict the cluster membership of new cases

Systems Mapping Tab

7. Using Systems Mapping To Explore Cluster Variables

Export Your Results

8. Generate your report

Help

Select the number of clusters

3

Do you want to set a seed for reproducible results?

Fuel.poverty	Social.Isolation	Get.NHS.Health.check
10.091	45.233	32.759
11.614	48.155	30.53
10.646	45.562	37.842
10.567	45.854	34.402

(x = km@uclusters, dist = dissM)

3 clusters C_j
j : n_j | ave_{i(c_j)} s_i

1 : 57 | 0.41

2 : 22 | 0.41

COMPLEX-IT

Team

As a team we are committed to advancing a **case-based complexity approach** to research, policy and practice in an effort to advance the study of social complexity and to support decision making. We each bring to the team a wide range of methodological and programming expertise and are proud of the truly transdisciplinary and international makeup of our work.



Brian Castellani
Durham Research Methods Centre
Durham University



Corey Schimpf
Department of Engineering Education
University at Buffalo, SUNY



Peter Barbrook-Johnson
School of Geography and the Environment
Oxford University



Michael Ball
Computer Services
Kent State University, Ashtabula



Christopher Caden
Durham Research Methods Centre
Durham University



COMPLEX-IT is part of the AM-Smart movement in methods

- Advances in the integration of smart technology with interdisciplinary methods has created a new genre, **approachable modelling and smart methods** – AM-Smart for short.
- AM-Smart platforms address a major challenge for applied and public sector analysts, educators and those trained in traditional methods: accessing the latest advances in interdisciplinary (particularly computational) methods.
- AM-Smart platforms do so through nine design features. They are
 - (1) bespoke tools that
 - (2) involve a single or small network of interrelated (mostly computational) methods
 - (3) they also embed distributed expertise
 - (4) scaffold methods use
 - (5) provide rapid and formative feedback
 - (6) leverage visual reasoning
 - (7) enable productive failure
 - (8) promote user-driven inquiry
 - (9) while counting as rigorous and reliable tools

COMPLEX-IT is a case-based, mixed-methods platform for applied social inquiry to complex data/systems, designed to increase non-expert access to the tools of computational social science.

Presently, the platform is comprised of a bespoke suite of techniques, including:

1. cluster analysis
2. artificial intelligence
3. data visualization
4. data forecasting
5. case-based systems mapping
6. case-based scenario simulation

COMPLEX-IT supports applied social inquiry through a design-based emphasis on learning about the complex data/system under study. It does by

(a) identifying and forecasting major and minor clusters/trends

(b) visualizing their complex causality

(c) mapping and simulating scenarios for potential interventions.

COMPLEX-IT is that it is accessible through the web or can be run locally and is powered by R and the Shiny web framework and includes written and video tutorials.



COMPLEX-IT

Run Online or Download for R-Studio



WEB
VERSION



DOWNLOAD
VERSION

<https://www.complex-it-data.org/>



Build Your Model

- 1. Build database and import your cases
- 2. Cluster your cases

Confirm & Explore Your Model

- 3. Use AI to confirm your cluster solution
- 4. Compare and visualize your results

Run Scenario Simulations

- 5. Simulate your scenarios, policies, and interventions

Run Data-forecasting/classification

- 6. Use AI to predict the cluster membership of new cases

Systems Mapping Tab

- 7. Using Systems Mapping To Explore Cluster Variables

Export Your Results

- 8. Generate your report

Help

Select the number of clusters

Do you want to set a seed for reproducible results?

Fuel.poverty	Social.Isolation	Get.NHS.Health.check
10.091	45.233	32.759
11.614	48.155	30.53
10.646	45.562	37.842
10.567	45.854	34.402

(x = km@uclusters, dist = dissM)
 3 clusters C_j
 j: n_j | ave_{i∈C_j} s_i

