

Exploratory Models in a time of Big Data

Maja Mazur¹ and Ed Manley²

¹ Arts and Sciences, University College London, Gower Street, London WC1E 6BT, United Kingdom

² Centre for Advanced Spatial Analysis, University College London, Gower Street, London WC1E 6BT, United Kingdom

Acknowledgements

We would like to thank Emma Tobin, Carl Gombrich, Ben Kingsley and Tim Knipe who have contributed to this paper through various discussions and intellectual inspiration along the way.

Abstract

This paper aims to trigger discourse about the emergence of a new type of social scientific model – Exploratory Models – models which draw on Big Data, computer modelling methods and interdisciplinary research to tackle increasingly complex social scientific processes. First, define Exploratory Models referring to Batty and Morgan and Morrison. We then present major changes introduced by Big Data and computer modelling to the traditional modelling paradigm. We show how Exploratory Models circumvent challenges related to social phenomena’s idiosyncrasy, complexity, self-reflectivity and acceleration, which restrain predictive effectiveness of traditional social scientific models. We then show that Exploratory Models are better equipped to tackling interdisciplinary complex problems due to their capacity to process heterogeneous, even contradictory datasets. Having established that Exploratory Models are predominantly problem- and data-driven, we emphasize that scientific theory is indispensable to the progress of Exploratory Models. Finally, the development of an integrative platform for Exploratory Models is suggested as a way of maximizing the benefits of this approach, acknowledging the challenges to the proposed platform and to Exploratory Models in general. Discussion concludes by flagging areas for further research.

Introduction

Despite the velocity of changes in corresponding scientific methods, little holistic assessment of issues posed by Big Data and computer modelling in the context of social science has been done so far (Kitchin 2014; Crawford and Boyd 2012). Philosophy of science, academia and business need to reflect on the impact of models utilizing these technologies on society, not only on the generation of scientific knowledge. Otherwise, we risk falling into the trap of skewed data insights, rushed decision making (Crawford 2015) and unsustainable, yet highly profitable short-term solutions (Townsend 2013). Moreover, increased volumes of data on *what*, presented without the necessary context of *why* (Mayer-Schonberger and Cukier 2013) increases the risk of exploitation of a data illiterate population reliant on secondary interpretation (Townsend 2013).

With these issues in mind, we aim to trigger further discourse about the emergence of new types of social scientific model – Exploratory Models – that draw on Big Data and interdisciplinary research to tackle increasingly complex social scientific issues. We devote considerable amount of space to showing differences between the traditional and Exploratory Models and emphasize throughout, in contrast to authors such as Anderson (2008), the role of scientific theory in this data-driven modelling paradigm. Finally, given that our aim is to provoke further discussion, we flag challenges to Exploratory Models and provide our replies to some of them, the rest leaving exposed to debate.

Definition of Big Data and Exploratory Models

We define Big Data after Kitchin (2014 p1) as: ‘huge in volume, high in velocity, diverse in variety, exhaustive in scope, fine-grained in resolution, containing common fields that enable the conjoining of different data sets and enabling adding new fields or expanding their size rapidly’. For reasons shown below, we believe that this relatively new phenomenon has a significant impact on the traditional modeling paradigm and fuels the potential success of Exploratory Models.

We define 'Exploratory Models' as new types of models in social sciences, which emerged due to the ability of researchers to use Big Data and computer modelling to analyze social scientific phenomena at previously unattainable levels of granularity and complexity. Moreover, these models use Big Data to become epistemologically similar to experiments and to allow higher flexibility in data control and analysis. They may thus inform the debate even if sometimes incapable of making predictions (Banks 1993 cited in Batty 2013 p273). Batty contrasts Exploratory Models with the consolidative style, which focuses on the models that might ultimately provide accurate or focused predictions (ibid.). We discuss 'consolidative' models as 'traditional' or 'theor-driven' modeling.

The first tier of our definition draws on Kitchin's (2014) terms 'exploratory science' and 'exploratory model' referring to models using Big Data and computer modelling. Such models may be considered in parallel with data-driven science (Leonelli 2014), the new empiricism (Kitchin 2014) or computational social science (Lazer 2009; Chang et al. 2014 et al. 2013; Blok and Pedersen 2014). However, in our view the key characteristic of Exploratory Models is the ability to not only confirm the existing theory but to provide theory-independent insights. Hence, the second tier of our definition sees Exploratory Models as models that surpass the levels of epistemic autonomy of models described by Morgan and Morrison (1999).

Morgan and Morrison (1999) perceive models as autonomous because certain elements in their construction, rules or mathematical forms, are independent of researchers. Autonomous models can refine the theory used to design them and provide novel scientific insights (Morgan and Morrison 1999; Morgan 2013; Massimi 2011). Autonomous representation allows surprising the researcher (Morgan and Morrison 1999; Morgan 2013, Morgan 2005). They are thus epistemologically closer to experiments than to 'consolidative models', aiming to solely confirm the theory.

First, we argue that Exploratory Models use Big Data to replicate the elements of the real world in a quest to achieve the correspondence levels of experiments (Winsberg 2009). Experiments capture or replicate the world inside the laboratory by using the same kinds of forms as 'those existing in the wild' (Morgan 2005 p320). Models that rely on artificial mathematics or alternative

physical domain materials are less able to provide independent insights about the system (ibid.). However, experiments reach perfect correspondence of idiosyncrasy and complexity of social phenomena only in real settings, i.e. inside a city on unaware subjects. Ethical and practical problems make such experiments unacceptable (Yasserli 2013).

In our view, Exploratory Models may mitigate this problem. Exploratory Models enjoy previously unseen granularity (Blok and Pedersen 2014) and autonomy from the theoretical 'artificial setting' (Morgan and Morrison 1999 p176). This is often helped by using Big Data simulation. As a result, according to Morgan, the model-experiment divide is slowly disappearing, allowing most stabilized visual constructions to be accepted as knowledge and 'perceived as real' (2013 p246). Moreover, 'computer simulation may be applied to contradictory theories, i.e. quantum mechanics and classical molecular mechanics' (Winsberg 2009 p843). Hence, some disciplines, i.e. urban planning (DIY City; Townsend 2013) or virtual ethnography (Morgan 2013), embrace Exploratory Models to investigate and 'experience' (ibid.) otherwise unobservable or too complex phenomena. Thus, Exploratory Models aim to replicate the target system with the accuracy similar or aspiring to that of experiments. The challenges of achieving such correspondence are discussed later, using the interdisciplinary hybrid model developed by Manley et al. (2014) as an example.

Nonetheless, Exploratory Models are still far from fully delivering on the promise first made when they appeared in the literature. There are significant challenges to how much data can be modeled at what granularity. For example, there is no current model able to simulate driver behavior in a detailed fashion at a highly scalable level (Manley et al. 2014). There are also significant challenges related to the use of simulation in scientific investigation. Firstly, there are no comprehensive methods of verification and validation of simulations (Batty 2001 and 2013; Batty and Torrens 2005). Secondly, there are risks of data manipulation (Leonelli 2014) or rejection of simulations by the scientific community as not conforming with accepted ways of understanding (Lehtinen and Kuorikoski 2007; Morgan 2013). Finally, the comparative epistemological power of a simulation (-model) and experiment depends on the quality of

background knowledge (Winsberg 2009 p842). Some of these problems are further addressed in our discussion of the role of theory in building Exploratory Models.

Another characteristic of Exploratory Models contributing to their autonomy from theory is the increased flexibility in data control available to researchers (Chang et al. 2014; Wierzbicki 2007). Data flows are produced almost continuously and researchers control their timing (Ruppert, Law and Savage 2013). In comparison with censuses giving a snapshot of system's dynamic (Ruppert, Law and Savage 2013; Cukier and Mayer Schonberger 2013; Kitchin 2014), Exploratory Models allow data generation akin to data monitoring. Big Data does not only reject or minimize the spatial closure of data but also extends the time interval over which dynamics are captured. For instance, Exploratory Models of urban growth use satellite imagery applied to topographic and thematic mapping to obtain nearly real-time high-resolution data, which allows identifying the spatial and temporal evolution of problem areas within the city (DIY City). Continuously monitored real-time data may also help preventing or mitigating risks before they escalate (Daqing et al. 2014; DIY City). In the past, data suffered from a considerable discrepancy between the state of the system at a time of harvesting data and the state of the system at the time of model building. Exploratory Models, however, achieve both spatial and temporal correspondence of data, the resolution of which may now be calibrated by researchers. Nonetheless, there are important computational and technological limits, as well as costs, which all increase for large-scale data (Tong 2015). Before we address them, a discussion of major differences between traditional social scientific modeling and Exploratory Models follows.

Differences between traditional and exploratory models

Exploratory Models embrace the scientific view that acknowledges omnipresent uncertainty. Traditional modeling perceived social systems as clock-like mechanisms, subject to the inevitability of causal relations (Wierzbicki 2007 p617). Models served to verify analytically derived predictions about isolated systems (Morgan and Morrison 1999). As recently as in 2001, the global economy was also seen as a clock-like mechanism, 'governed by the laws of

supply and demand in a similar way as the laws of gravity govern the Universe' (Taleb 2012 p459). These reductionist views were revised in the light of complexity theory (Batty and Torrens 2005; Batty 2008), as well as events such as the 2008 financial crisis disturbing financial, political and social systems. The increasingly connected world (Mayer-Schonberger and Cukier 2013) is perceived more as a universe of information flows rather than a mechanism of law-like regularities (Batty 2013 p48). For instance, equilibrium is no longer considered paramount (Batty 2008; Townsend 2013; Mayer-Schonberger and Cukier 2013) and 'imbalance potentially becomes the new paradigm' (Verebes 2014 p38). As a consequence, models must not be isolated and 'reduced to only one variety of variables when several are present' (Batty and Torrens 2005 p754). Thus, determinism born out of the mechanistic view of systems gives way to the uncertainty towards theory-born predictions (Taleb 2012; Wierzbicki 2007; Verebes 2014). Abduction becomes an increasingly popular mode of inferential reasoning (Kitchin 2014).

Exploratory Models are 'good' science where traditional models had to use overidealization of assumptions or encountered paradox. Researchers increasingly emphasize operational success of models and their relationality (Kitchin 2014; Mayer-Schonberger and Cukier 2013), instead of theory-guided analytic confirmation (Winsberg 2009; Lazer 2009; Crawford and Boyd 2012). Although explaining phenomena accordingly to the parsimony rule (Crooks et al. 2008 p419) has long been associated with good science, it is nowadays acknowledged that 'the whole is more than the sum of its parts' (Verebes 2014; Batty 2013). Correlation is employed to scrutinize micro patterns with the intention of capturing emergent properties (Manley et al. 2014; Batty 2013; Batty and Torrens 2005). The operational success of correlation may elevate it above analytical explanation in research contexts that try to avoid commitment to one level of the phenomenon at a time, for instance where the two levels are contradictory (Winsberg 2009 p843). Furthermore, correlation used by Exploratory Models may also offer immediate benefits without understanding of the pattern (Cukier and Mayer-Schonberger 2013). For example, correlation analysis of moods on Twitter could predict 'daily up and down in Dow Jones Industrial Average with an accuracy of 87.6%' (Mestyan, Yasseri and Kertesz

2013 p1). Nonetheless, later discussion acknowledges limitations to using correlation as a scientific tool. We still see theory and analytical explanation as indispensable to exploratory scientific research.

Finally, Exploratory Models allow new ways of understanding the data fed into the model, for instance by using Big Data to challenge the relation between model inputs and outputs (Batty and Torrens 2005 p752). Traditional models investigate logical connections between inputs and outputs, not the dynamic itself (ibid.). If outputs and inputs are not logically connected, traditional practice changes the inputs, not relations (ibid.). However, in Exploratory Models, further examination of data relations takes place. This allows, for instance, accounting for self-reflectivity of systems, particularly problematic in econophysics (Bookstaber 2010). Big Data may produce data on such phenomena in real time (Wierzbicki 2007; Chang et al. 2014; Kitchin 2014). Given these differences, we believe that Exploratory Models overcome the below challenges faced by their traditional predecessors.

Exploratory Models as more successful at modeling complex social phenomena than traditional models

First of all, idiosyncrasy and complexity of social phenomena obstruct theory-led analytical models from revealing underlying causalities (Massimi 2011; Bogen and Woodward 1988). Moreover, social phenomena exhibit high variability, which hinders their conceptualisation (Henrickson, McKelvey 2002). For instance, temporal persistence of a pattern in a given context has sometimes decided of it qualifying as a phenomenon (ibid. p7289), revealing the absence of universal criteria for pattern-to-phenomena transition. Idealized assumptions narrow down a model's accuracy to limited circumstances and may render them inadequate (Batty 2013). Big Data provides more inclusive data classification and modeling (Manfred and Reggiani 2015 p334) by giving models the capacity to collect and analyze uncertain and unstructured data (Lazer 2009). Furthermore, although not discussed here, machine learning data classification methods (Aggarwal 2015; Henrickson 2002; Townsend 2013) support more effective assumption making and hence partially overcome the idiosyncrasy challenge. Finally, Exploratory Models analyze varied contextual clues (Chang et

al. 2014, p77) and thus decrease the risk of researchers omitting real underlying mechanisms or overlooking emergent properties (Yasseri 2013).

Exploratory Models are also more responsive to the relation between the system, the model (Yasseri 2013; Ruppert, Law and Savage 2013) and the researcher (Morgan and Grune-Yanoff 2013). Cognitive agents learn models of social systems (Bookstaber 2010) and influence them (Ruppert, Law and Savage 2013; Batty et al. 2013; Lazer 2014). Digitization increases the profitability of extremely quick reactions (Lazer 2014). For instance, agents consciously manipulated Twitter entries to affect stock prices (ibid.). Thanks to being able to gather almost real-time Big Data inputs Exploratory Models may be more effective in understanding and predicting self-reflectivity of social models.

Finally, Exploratory Models can use voluminous datasets, often coming from different disciplines, to be more successful at predicting and understanding cascading social phenomena. This is crucial, because social systems exhibit cascading properties (Daqing et al. 2014; Henrickson 2002; Townsend 2013; Batty 2013). Effectiveness of prediction and understanding of cascading phenomena is crucial from the practical point of view, as the acceleration of phenomena threatens social systems' stability (Taleb 2012) and risks significant financial losses. In 2003 a cascading energy blackout in the US and Canada had an approximate cost of 10 billion dollars (Daqing et al. 2014 p1). Cascading phenomena increase rapidly in intensity as a result of disturbance (ibid. p2). The more connections a given node already holds, the more likely it is to expand its network (Katz and Shapiro 1994), allowing potential damages to spread directly via a network's structural dependencies (Daqing et al. 2014). Overload cascading failures propagate usually due to cooperative overload pressure, without visible or direct casual relation (ibid.). Therefore, given 'the many-body effects of such failure interactions and the lack of real failure data' (Daqing et al. 2014 p1), purely analytical models can hardly capture or predict such complex dynamic (ibid.). Although developing causal explanations of networked phenomena acceleration is a challenge to our current science (Batty 2013), Exploratory Models utilizing nearly N=all dataset (Cukier and Mayer-Schonberger 2013; Taleb 2012) may show multiple structural dependencies that accompany and

precede acceleration (Daqing et al. 2014 p5). Thus they offer operational value and actionable results, which were unavailable to traditional theory-led models.

Furthermore, Exploratory Models may be more successful at understanding and predicting interdisciplinary, complex phenomena. This is now discussed using the interdisciplinary hybrid agent-based model of urban traffic developed by Manley et al. (2014). Vital in demonstrating how Exploratory Models support successful interdisciplinary research, the model combats the three-horned dilemma between generalizability over populations, precision of variable control and realism of observational context (McGrath 1982). Due to the complexity of the modeled system (urban road network) it analyzes micro level behaviors of agents to understand the macro pattern and to capture emergent properties (Manley et al. 2014). Thus, it models under the assumption of the system's nearly 'boundless complexity' (Batty and Torrens 2005 p747). It also uses human expertise to make necessary idealizations, decreasing the required computational resources. By analyzing both micro and macro levels of psychological factors, which affect route choices of drivers, the model targets to a certain extent the problem of idiosyncrasy.

Although the above example does not tackle self-reflectivity, it demonstrates model's ability to combine knowledge from different fields, here cognitive psychology and urban planning, to meet requirements of the interdisciplinary inquiry into agents' behaviour. This shows that capacity of Exploratory Models to process voluminous, diverse and potentially contradictory data allows them to be problem-driven and interdisciplinary than their traditional predecessors. The scope of investigation of an Exploratory Model is defined by the available data, not disciplinary boundaries, even if data is not explicitly linked to the problem. Hence, an 'interdisciplinary convergence' (Chang et al. 2014 p70) takes place, allowing usage of all potentially relevant knowledge to deliver most accurate solutions to the problem. The time it took in the past to exchange interdisciplinary knowledge at that level made such problem-based research hardly efficient (Cukier and Mayer-Schonberger 2013). However, Big Data fed into Exploratory Models increases the profitability of interdisciplinarity. Moreover, Mansnerus suggests that Exploratory Models may serve as 'bridges

between disciplines' (2013 p269), further contributing to proliferation and growth of interdisciplinary research.

To conclude, Exploratory Models can more representative of the target system and less constrained by disciplinary boundaries. Despite their increased autonomy from theory, we now show that theory is essential to the development of Exploratory Models.

The role of scientific theory in Exploratory Models

Firstly, theory is crucial to effective data interpretation and presentation. Data manipulation may produce spurious correlations. Social media or search engines, which are most often data sets for Big Data correlation studies, are unreliable sources of information not intended for scientific purposes (Lazer 2009 and 2014; Kitchin 2014; Crawford and boyd 2012). Google Flu Trends (GFT), for long considered a success story of correlation analysis (Townsend 2013), consistently overestimated flu trends for 100 out of 108 weeks, due to changes in algorithm's parameters, without many people realizing it (Lazer 2014). Platforms such as Twitter and Facebook are constantly re-engineered, and it is increasingly questioned whether studies conducted even a year ago on data collected from these platforms can be replicated (Lazer 2014), and hence verified (Batty and Torrens 2005). Moreover, despite their size 'they are not representative of all people' but of very particular groups (Crawford and boyd 2012 p669). Thus, theory should ensure that adequate parameters calibrate the exploratory model (Leonelli 2014; Lazer 2014). The primary function of the theory is to choose the necessary levels of idealization, the achievable resolution of patterns spotted in data and the maximum possible data stability.

Secondly, theory calibrates risk indicators, which monitor pattern abnormalities and network instabilities. For instance, in High Frequency Trading (HFT) (ESMA 2014) algorithms analyze, model and predict the risk of trading a particular asset (ibid.). Given that the algorithm's decisions are too fast for the human to fully supervise, the parameters' quality is the major insurance against costly mistakes. Therefore, relevant financial theory is solely responsible for the results achieved by such algorithms. Ironically, currently such 'data-driven' models are the most dependent on theories used to create them. Using artificial

intelligence to calibrate algorithm's formulas may in the future decrease this dependency.

Finally, theory and vernacular knowledge incorporated into the process of Exploratory Model creation highlight soft factors important for model implementation (Morgan and Grune-Yanoff 2013; Wierzbicki 2007; Trutnevyte 2014). For instance, a qualitative story surrounding the Exploratory Model of energy consumption in a city ensures its compliance with its regulatory contexts at the time of implementation and in the future. This is crucial to avoid research and operational implementation delays (Batty 2013; Crawford, 2015). Furthermore, Exploratory Models must be correctly interpreted and organized, as otherwise they only provide information, not knowledge (Wierzbicki, 2007). Moreover, we now argue that theory enables interdisciplinary Exploratory Models to be integrated into platforms and therefore respond to problems of very complex systems, such as cities or wide-spanning phenomena, which would not be analyzed by a singular Exploratory Model due to computing limits or data diversity.

The role of theory in integrating Exploratory Models

Theory organizes information collected by Exploratory Models in order to turn it into knowledge (Wierzbicki 2007). The focus of Exploratory Models is often on showing how much can be modeled and with what granularity. This is scientifically unproductive, as rendering a problem visible is not solving it (Townsend 2013). Thus, increasing the models' computational power without developing frameworks for understanding the results produced might generate models as complex as target systems (Wierzbicki 2007; Leonelli 2014). For instance, the Protein Data Bank (Protein Data Bank) is already too voluminous to be directly used in research and must be partitioned into secondary databases. We suggest developing a strategy for integrating Exploratory Models to avoid a similar outcome.

Integrating Exploratory Models should start with developing quantitative models and enabling exchange of information. Otherwise, the qualitative framework might be proven invalid by technical restrictions of the system. Simon's (1962) view of complex systems as hierarchical structures is thus

refined. Complex systems are hierarchic in nature, i.e. composed of interrelated sub-systems (ibid. 468). Contrary to Simon's assumption that internal dynamic of the system is denser than its external interactions (ibid. p477), urban science shows that social systems might exhibit as strong interactions with their external environment as between internal components (Batty 2013). Nevertheless, we take Simon's view that systems can be divided into sub-systems (1962 p468) until some elementary level is reached (ibid.), i.e. an individual agent. A broader definition of hierarchy does not necessitate power relations between systems (ibid.). Two systems might connect in a hierarchical, subordinate manner but they may also enter into horizontal relations, therefore rendering such hierarchy similar to the semi-lattice, also used in urban science conceptions of cities (Townsend 2013 p143; Batty 2013 p23). It is also possible that individual sub-systems interact with external sub-systems. Hence, the scope of the system should extend to horizontal, long-range spheres of influence or even 'sequences of networks' (Batty 2013 p106). Moreover, to fully utilize the potential of Exploratory Models it is worth exploring possible algorithms, which would incrementally update the data classification results for a database whose values keep changing, as opposed to the more traditional approach, where 'algorithms dealing with uncertain data only capture a data snapshot' (Cheng and Fang 2015 p441). Finally, hierarchy is not measured in 'spatial terms' (Simon 1962 p468) or according to 'size ordering' (Batty 2013 p152). Overlapping hierarchies of intensity of external and internal interactions are more fit to social systems (Simon 1962; Batty 2013 p175), as they may be linked to the intensity of potential cascading failures. Hierarchy is therefore linked to the volume of interactions between systems (Simon 1962) but also to the impact which one system's disturbance may have across network. It is important to note that due to horizontal connections between sub-systems, there may be strong long-range correlations (Daqing et al. 2014) between sub-systems. Hence, the question arises whether one measures a system's position in a hierarchy also by accounting for its long-range influence, or whether only close-range impacts are prioritized.

Despite its complexity, such a platform would enable quick and efficient exchange of information, essential to interdisciplinary convergence between

Exploratory Models. However, the platform dominated by quantitative models risks that policymakers or researchers from outside disciplines directly involved in model building would not understand complex results. Moreover, we cannot expect models of multiple systems to share spatial and temporal foci, disciplinary perspective, objectives, system boundaries or format of the inputs and outputs (Trutnevyte et al. 2014). Thus, qualitative storylines should put exploratory data analyses into context of existing models, scientific theories and operational policies, therefore turning them into knowledge. Furthermore, ‘interrelated ideas could be structured as overlapping hierarchies’ (Batty 2013 p175), thus enhancing their methodological conformance with quantitative representations.

Combining qualitative and quantitative representations of the system offers a possible methodology for verification of integrated Exploratory Models. Quantitative models may be validated against qualitative storylines and expertise of researchers (Trutnevyte et al. 2014; Wierzbicki 2007; Morgan and Grune-Yanoff 2013). Furthermore, theory allows testing models against other existing models in the field. First, theory locates a model’s area of expertise (Trutnevyte et al. 2014), which give priority of insights delivered by “expert models” over insights delivered by “less specialized” models (ibid.). For instance, in Trutnevyte three different models deliver two different values for the necessary back-up capacity of energy (2014). The prediction given by the model with the highest data resolution in this field, i.e. expertise, is given priority over the two others, which do not focus to that extent on the variable of required back-up capacity (ibid.). Although these methods of verification still need further research into their limits and risks, the suggested framework sets the ground.

What is more, combining theoretical and quantitative frameworks may improve the capacity to anticipate and mitigate the models’ self-reflectivity. None of individual Exploratory Models we evoked in this discussion managed to predict its own impact on the system. However, incorporating a theoretical framework, which accounts for not only formal soft knowledge, such as legal and institutional issues, but also vernacular knowledge concerning the behavior of individual humans, increases the probability of anticipating self-reflectivity (Ruppert, Law, Savage 2013; Mansnerus 2013). It might also discover ways in

which models would affect one another, as well as flag further emergent qualities. If successful, such integration considerably extends the correspondence of Exploratory Models, as it renders them more flexible with regards to the target system.

Finally, the inclusion of a variety of different models in an integrative platform largely circumvents the concern that models are only relevant to a limited number of contexts (Trutnevyte et al. 2014). It is possible that a model with multiple expertise areas may be developed in the future. However, due to our current technological constraints, as demonstrated by the hybrid agent-based model (Manley et al. 2014), lowering computational requirements without considerable trade-offs in accuracy is problematic (McGrath 1982; Chang et al. 2014). Integrating a variety of models with separate expertise fields circumvents this challenge to some extent. Due to the scope of this discussion, structure of an integrative platform is not further discussed. Major related challenges are now presented.

Scientific and philosophical challenges to Exploratory Models and integrative platforms

The first critique is analytical impenetrability of Exploratory Models. As a complex network of models, the platform exhibits few explicit casual relations between data or models. Thus, sources of distortions in results are analytically intractable (Winsberg 2009; Lazer 2009; Batty and Torrens 2005). Furthermore, even if a method for verification of individual models is found, the platform as a whole might require different verification criteria and show emergent properties. The question of the allowed degree of impenetrability of platforms and inability to track the exact ways different outcomes emerge may become as much of a political as a scientific question. It is possible that human expertise will remain a decisive factor in assessing the implementation feasibility of any platform.

Furthermore, some claim that data classification should precede model integration. Initiatives, such as the Chicago Data Dictionary (Chapin Hall; Townsend 2013; Glodsmith and Crawford 2014) have approached modeling complex systems from this angle. Integrating models prior to classifying data

implies accepting many idealisations for data presentation and selection. Moreover, uncovering techniques for minimum mathematical or modeling interference in data classification could increase the correspondence of Exploratory Models. Nevertheless, a strong competition on 'global data markets' (Townsend 2013 p301) between research centres, businesses and individuals forces researchers to adapt to competitive environments (Glaeser 2011). If they fail to deliver responses and instead only flag existing problems, the data will be 'most likely used for individual gain, instead of towards holistic improvement to our society' (Townsend 2013 p165). Moreover, simply revealing complexity and uncertainty may prompt rushed decision-making on operational and research levels. It is therefore secondary to delivering timely response mechanisms, which are facilitated by integrating models rather than data.

Finally, the scope of a suggested platform is unclear. Should a platform's boundaries be mapped according to structural similarity of models included or of intensity of their external interactions? There is a dilemma between automated and human-led partitions, already faced by secondary databases of the Protein Data Bank. Delineating boundaries for platforms of social systems, more complex and varied in nature, poses even more challenges. Furthermore, even if platforms would be problem-based, as Exploratory Models often are, questions arise about the boundaries of a problem. Hence, partitioning the system into platforms leads to self-contradictory reasoning: first, Exploratory Models assume systems to exhibit boundless complexity (Batty and Torrens 2005 p747); then an integrative platform restricts the systems' scope. Thus, platforms may be criticized as a mechanistic, albeit more complex, view of the world. The shift away from thinking about the world as a clock to conceiving it as a clock with more functions is not the goal of Exploratory Models.

Operational challenges

The first operational challenge concerns control. Data is the currency of information economy (Taleb 2012; Townsend 2013; Verebes 2014) and the proposed platform accumulates high quality scientific data. Many see such grand-scale civic and scientific surveillance as authoritarian (Verebes 2014 p14). Furthermore, the wider and more transparent platforms become, the higher a

platform's influence on the system and the higher profitability of manipulation. Data-intensive platforms bringing quick and accurate solutions increase the divide between data poor and data rich population (Crawford and boyd 2012; Townsend 2013). Due to costs of data curation (Leonelli 2014), the divide may correspond to financial inequality. Necessary regulatory measures should be implemented to avoid this negative influence of exploratory modelling on target systems.

Moreover, levels of desired democratization of data are not clear (Townsend 2013). Falling costs of research due to open source and Big Data seem temporal and soon costs of data curation may exceed profits from its availability (Leonelli 2014). Furthermore, Townsend worries that most smart apps and "civic hacks" serve their own profit rather than common good (2013 p230). Decentralizing science (ibid.) may allow technologies such as chemical weapons, TDCs kits (Fields 2011) or statistical sources to spread beyond control. It would be a very undesirable effect, if the good of the individual were gradually favoured over the good of the society (Townsend 2013).

What is more, concerns regarding data privacy appear (Yasseri 2013; Crawford and Boyd 2012; Blok and Pedersen 2014). Firstly, the suggested platform requires excessive access to personal data, often in unstructured form, which obstructs effective legal restrictions on the collection of such information (Townsend 2013 p293). Despite efforts such as Data Protection Reform (Slaughter and May 2013 and 2014; Norton Rose Fulbright 2014), the law lags behind technology. Moreover, recent studies show that healthcare providers forced to comply with Big Data technologies could not afford adequate cyber security measures (Maerian 2016). Furthermore, there is a trade-off between privacy and security (Townsend 2013; Rabari and Storper 2014). It is already possible to track individual Netflix users via "anonymous" data (Singel 2012). In addition, advances in visual machine learning techniques could lead to crimes being detected automatically by CCTV cameras (Li 2015). High numbers of "potential crimes" automatically reported to police stations will slow down police operations, every such "potential case" needing human interpretation. On the other hand, such techniques could effectively fight certain crimes, as illustrated by Polaris fight against human trafficking (Gesher 2013).

Furthermore, it is difficult to secure conscious consent from data producers. For example, it was shown that only about 7% of Britons actually read the Term and Conditions of social media sites or services bought (Smithers 2011). Regulating data collection for Exploratory Models is problematic, as traditionally the purpose of collection must be stated (Kitchin 2014; Crawford and boyd 2012), whereas Exploratory Models derive their efficiency and accuracy from analysing data not explicitly related to the problem. Hence, regulating justification processes for collecting such information poses a considerable challenge.

Exploratory Models also risk increasing a system's overall volatility. The human eye cannot control an algorithm's real-time reaction when response time is very short, i.e. nanoseconds (ESMA 2014; Taleb 2012). Therefore, results delivered by algorithms are not verifiable on a daily basis but only in the context of overall success or failure. Due to the acceleration of social scientific phenomena there is a risk of small technical flaws in one part of the network resulting in a general structural failure. Ideally, an algorithm should isolate models in the event of failure, as blockchain can isolate the sequence of the chain under attack (Szabo 1997). However, the complexity of social systems makes isolation of subsystems a yet unsolved task. Proponents of blockchain technology currently explore various solutions.

Moreover, a system's volatility is increased through the "Fat Finger" problem (Lewis 2015 p98; ESMA 2014; Hagstromer and Norden 2012): one mistaken data input into the parameters of the algorithm, for instance 1000 points instead of 100 may generate dire outcomes. In financial models, although Fat Finger risk brings temporary volatility, such as the Flash Crash¹ (The Economist 2010a and 2010b; ESMA 2014), the market quickly regains balance. However, financial models are unique in their capacity to regain balance with such speed. Other social models, i.e. of urban growth or military actions, do not have target systems with such rebounding mechanisms.

¹ On May 6th 2010 the Dow Jones Industrial Average fell by nearly 1,000 points in 30 minutes but rebounded quickly. Many blamed HFT firms, which fled the market in a 'cascading' run (The Economist 2010b).

Finally, due to increasing levels of self-reflectivity of social systems, the question of what new social groups the exploratory science may create should be investigated (Ruppert, Law and Savage 2013). Past social scientific measures, such as censuses and survey techniques generated new social entities, such as “the national economy” or caste groups (Ruppert, Law and Savage 2013). It may be observed that data overflow resulting from exploratory scientific methods and correlation techniques contributes to the increasing distrust of citizens towards the media and experts. The occurrence of events such as 2008 Financial Crisis despite the assurances from governments and experts of the soundness of our social system heavily undermined public confidence in the models and systems in place. Complexity of datasets used by Exploratory Models as well as their focus on solving the problem without necessarily providing a clear theoretical explanation thereof may thus deepen alienation of science, which previously relied on theory to establish a dialogue with the society.

Recommendations for further research

Firstly, visualization and simulation require further research in the context of exploratory science, particularly into relevant regulatory and verification methodologies. Interpreting data produced by Exploratory Models is a challenge, especially since there is a shortage of adequately qualified people (Chang et al. 2014). Interactive visual analytics (IVA) help interpreting complex patterns delivered by Exploratory Models and are designed to promote more transparent and creative decision-making (Cybulski 2015). Its potential should be therefore evaluated. What is more, philosophy of science should take a more operationalist, or even regulatory approach to investigating Exploratory Models. Modern philosophers of science should consider perspectives of both academia and business. Finally, the potential impact of Machine Learning on Exploratory Models and integrative platforms is a promising field for further investigation. Machine Learning may support both the construction of individual Exploratory Models, i.e. continuous improvement of the data production, and the construction of the platform suggested. Machine Learning could optimize information exchange, prevent structural failures and support human interpretation of the models’ results by suggesting novel insights, as done by the

IBM Watson Discovery system (IBM Watson). Its implications, potential and threats should be therefore evaluated.

Conclusion

To conclude, in this discussion we have aimed to present the most pressing and disruptive challenges posed by Big Data, computer modeling and Exploratory Models to the modeling paradigm in social sciences. In order to deliver a comprehensive and discourse-provoking discussion, we drew on a number of fields, such as philosophy of science and urban science. Discussion of literature on the topic and case studies of models utilizing Big Data and computer modelling revealed potential of Exploratory Models to overcome the challenges faced by traditional models. We therefore argue that Exploratory Models overcome challenges of idiosyncrasy and complexity, self-reflectivity and social systems entering into cascading effects, which obstruct success of traditional social scientific models. Furthermore, we show that Exploratory Models support interdisciplinary research and are well equipped to handle interdisciplinary complex problems. Finally, we evaluate challenges faced by exploratory science, such as integrating Exploratory Models. We reply to this challenge by suggesting an integrative platform for Exploratory Models. Throughout the discussion, we emphasise the indispensability of scientific theory for exploratory science. We hope that this discussion will inform further epistemological and empirical investigation of Big Data, computer modeling and exploratory science.

Word Count: 5982

References

- Aggarwal, C. C., 2015. *Data Classification. Algorithms and Applications*. London: Taylor & Francis Group / Chapman & Hall
- Anderson, C. 2008. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired* [online] 23 June. http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory [Accessed 20 April 2015]
- Batty, M. 2006. 'Model Cities'. [UCL working papers] 113. February 2007.
- Batty, M. 2008. Cities as Complex Systems: Scaling, Interactions, Networks, Dynamics and Urban Morphologies. *CASA UCL Working Papers Series*. 131. February.
- Batty, M. 2013. *The new science of cities*. London: The MIT Press
- Batty, M. et al. 2013. Smart London. In: Bell, S., Paskins, J. 2013. *Imagining the Future City: London 2062*. London: Ubiquity Press. pp 31 - 40
- Batty, M., 2001. Models in planning: technological imperatives and changing roles. *JAG*. 3(3) pp 252-266
- Batty, M., Torrens, P.M. 2005. Modelling and prediction in a complex world. *Futures*, 37. pp 745–766
- Bienn, A., 2015. Neural Networks: A Review. In: Aggarwal, C. C., 2015. *Data Classification. Algorithms and Applications*. London: Taylor & Francis Group / Chapman & Hall 2015
- Blok, A., Pedersen, M.A. 2014. Complementary social science? Quali-quantitative experiments in a Big Data world. *Big Data & Society*. pp 1–6
- Bogen, J., Woodward, J., 1988. Saving the Phenomena. *The Philosophical Review*, 97(3). pp 303-352
- Bookstaber, R., 2010. Here's Why Econophysics Will Never Work, Though Traders May Dream. *Business Insider*. [online] Available from: <http://www.businessinsider.com/heres-why-econophysics-will-never-work-though-traders-may-dream-2010-8#ixzz3Kr3LK0Pe> [Accessed 20 April 2015]

- Buckingham Shum, S., 2012. Towards a global participatory platform. Democratising open data, complexity science and collective intelligence. *European Physical Journal Special Topics*, 214. pp 109–152
- Canning, K., Rose, S., 2001. Gender, Citizenship and Subjectivity: Some Historical and Theoretical Considerations. *Gender & History*. 13 (3). pp 427–443
- Chang, R. M. et al., 2014. Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems*, 63. pp 67–80
- Chapin Hall. *Chicago City Data Dictionary*. [online] <http://datadictionary.cityofchicago.org> [Accessed 22 April 2015]
- Cheng, R., Fang, Y., 2015. Uncertain Data Classification. In: Aggarwal, C. C., 2015. *Data Classification. Algorithms and Applications*. London: Taylor & Francis Group / Chapman & Hall pp 417 - 441
- Crawford, K., boyd, D. 2012. Critical Questions for Big Data. *Information, Communication & Society* [online], 15(2). pp. 662 - 679
- Crawford, S., 2015. Responsive City. *MIT CMS/Writing* [online]. Available from: <https://vimeo.com/99133426> [Accessed 20 April 2015]
- Crooks, A., Castle, C. and Batty, M., 2008. Key challenges in agent-based modelling for geo-spatial simulation. *Computers, Environment and Urban Systems*, 32. pp 417–430
- Cybulski et al. 2015. Creative problem solving in digital space using visual analytics. *Computers in Human Behavior*, 42. pp 20–35
- Cybulski, J., L. et al. 2013. Creative problem solving in digital space using visual analytics. *Computers in Human Behavior. Elsevier*, 42. pp. 20–35
- Daqing, L. et al. 2014. Spatial correlation analysis of cascading failures: Congestions and Blackouts. *Scientific Reports*, 4 (5381). pp 1- 6
- DIY City. 2015. Managing mega-cities: spatial information techniques for better property management. *DIY City* [online]. <http://diycity.org> [Accessed 22 April 2015]
- Easterly, W., Kraay, A., 2000. Small States, Small Problems? Income, Growth, and

Volatility in Small States. *World Development*, 28 (11). pp 2013 - 2027

ESMA. 2014. High Frequency trading activity in EU equity markets. *ESMA Economic Report* [online] 1. Available from: http://www.esma.europa.eu/system/files/esma20141_-_hft_activity_in_eu_equity_markets.pdf [Accessed 22 April 2015]

Fields, D.R. 2011. Amping Up Brain Function: Transcranial Stimulation Shows Promise in Speeding Up Learning. *Scientific American* [online]. 25 November. <http://www.scientificamerican.com/article/amping-up-brain-function/> [Accessed 22 April 2015]

Fischer, L. 2013. Privacy Check: Social media changes highlight importance of terms and conditions. *The Pendulum* [online]. 24 February. <http://www.elonpendulum.com/2013/02/privacy-check-social-media-changes-highlight-importance-of-terms-and-conditions/> [Accessed 20 April 2015]

Fischer, M. M., Reggiani, A. 2015. Spatial Interaction Models: From the Gravity to the Neural Network Approach. *Urban Dynamics and Growth: Advances in Urban Economics* [online], 266. pp. 319-346. <http://www.emeraldinsight.com.libproxy.ucl.ac.uk/doi/pdfplus/10.1108/S0573-8555%282005%290000266012> [Accessed 20 April 2015]

Gesher, A., 2013. Huffpost Live: Fighting Human Trafficking Through Better Data Analysis. *The Palantir Blog* [online]. 11 September. <https://www.palantir.com/2013/09/huffpost-live-fighting-human-trafficking-through-better-data-analysis/> [Accessed 22 April 2015]

Giles, J. 2012. Making the Links. *Nature*. 23 August. 488. pp 488 - 450

Glaeser, E. 2011. *How urban spaces make us human. Triumph of the city*. London: Pan Books

Goldsmith, S., Crawford, S. 2014. *The Responsive City. Engaging communities through data-smart governance*. San Francisco: John Wiley & Sons <http://www.sciencedirect.com/science/article/pii/S0040162514002571> [Accessed 20 April 2015]

Hagstromer, B., Norden, L., 2012. The diversity of high frequency traders. *Stockholm University School of Business* [online]. 27 September. Available from: <https://secure.fia.org/ptg-downloads/The-Diversity-of-HFT.pdf> [Accessed 22 April 2015]

Henrickson, L., McKelvey, B., 2002. Foundations of “new” social science: Institutional legitimacy from philosophy, complexity science, postmodernism, and agent-based modeling. *Proceedings of the National Academy of Sciences*, 99 (3). pp 7288-7295

IBM Watson. *IBM Watson Discovery*. [online] Available from: <http://www.ibm.com/smarterplanet/us/en/ibmwatson/discovery-advisor.html> [Accessed 24 April 2015]

Katz, M. L., Shapiro, C., 1994. Systems Competition and Network Effects. *Journal of Economic Perspectives*, 8(2), pp 93-115

Kitchin, R. 2014. Big Data, new epistemologies and paradigm shifts. *Big Data and Society* [online] pp. 1 – 12 <http://bds.sagepub.com/content/1/1/2053951714528481> [Accessed 20 April 2015]

Lazer, D. et al. 2009. Computational Social Science. *Science*. 6 February. 323. pp 721 - 723

Lazer, D. et al. 2014. The Parable of Google Flu: Traps in Big Data Analysis. *Science*. 14 March. 343. pp 1203 -1205

Lehtinen, A., Kuorikoski, J. 2007. Computing the Perfect Model: Why Do Economists Shun Simulation? *Philosophy of Science*, 74(3). pp 304-329

Leonelli, S., 2014. Data Interpretation in the Digital Age. *Perspective on Science*,

Lewis, M. 2015. *Flash Boys. Cracking the Money Code*. UK: Penguin Books

Li, F.F. 2015. How we’re teaching computers to understand pictures. TED [online]. Available from: http://www.ted.com/talks/fei_fei_li_how_we_re_teaching_computers_to_understand_and_pictures [Accessed 22 April 2015]

- Mearian, L., 2016. 1 in 13 patients will have their records stolen after a healthcare provider data breach. *Computerworld* [online], 30th June <http://www.computerworld.com/article/3090566/healthcare-it/hackers-are-coming-for-your-healthcare-records-heres-why.html> [Accessed 2nd September 2016]
- Manley, E. et al., 2014. A framework for simulating large-scale complex urban traffic dynamics through hybrid agent-based modelling. *Computers, Environment and Urban Systems*, 44. pp 27–36
- Mansnerus, E. 2013. Modeling in the Social Sciences: Interdisciplinary Comparison. *Perspectives on Science*, 21(2). pp 267-272
- Massimi, M., 2011. From Data to Phenomena: A Kantian Stance. *Synthese*, 182 (1). pp 101-116
- Mayer-Schonberger, V., Cukier, K. 2013. *Big Data. A Revolution That Will Transform How We Live, Work and Think*. Great Britain: John Murray (Publishers)
- McAlister, J.W., 1997. Phenomena and Patterns in Data Sets. *Erkenntnis*, 47 (2). pp 217-228
- McGrath, J., 1982. Dilemmatics: The study of research choices and dilemmas. In: McGrath, J. *Judgment calls in research*. Beverly Hills: Sage Publications, pp 69-80.
- Mestyan, M., Yasseri, T., Kertesz, J. 2013. Early Prediction of Movie Box Office Success Based on Wikipedia Activity Big Data. *PLoS ONE*, 8(8). E71226
- Mitchell, S. D., 2004. Why integrative pluralism? *E:CO*, 6 (1-2). pp 81-91
- Morgan, M. S. 2013. Experiencing Life Through Modeling. *Perspectives on Science*, 21(2). pp 245-249
- Morgan, M. S., 2005. Experiments versus models: New phenomena, inference and surprise. *Journal of Economic Methodology*, 12(2). pp 317-329
- Morgan, M., Grune-Yanoff, T., 2013. Modelling Practices in the Social and Human Sciences. An Interdisciplinary Exchange. *Perspectives on Science*, 21(2). pp 143-156

Morgan, M., Morrison, M., 1999. Models as Mediating Instruments. In: *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.

Morgan, M., S. 2001. Models, Stories and the Economic World. *Journal of Economic Methodology*, 8(3). pp 361–84

Norton Rose Fulbright. 2014. Global data privacy. Directory. *A Norton Rose Fulbright directory* [online]. London: Norton Rose Fulbright, pp. 87 -88. Available from: <http://www.nortonrosefulbright.com/files/global-data-privacy-directory-52687.pdf> [Accessed 20 April 2015]

Norton Rose Fulbright. 2014b. High frequency and algorithmic trading obligations. *MiFID II / MiFIR series* [online] October. <http://www.nortonrosefulbright.com/knowledge/publications/115236/mifid-ii-mifir-series> [Accessed 23 April 2015]

Novak, E., Zhao, W., Reiser, R., A. 2014. Promoting Interdisciplinary Research Among Faculty. *Journal of Faculty Development* [online], 28 (1). pp. 1 – 6 http://www.academia.edu/6135044/Promoting_Interdisciplinary_Research_among_Faculty [Accessed 20 April 2015]

Protein Data Bank. [online] <http://www.rcsb.org/pdb/home/home.do> [Accessed 23 April 2015]

Rabari, C., Storper, M., 2014. The digital skin of cities: urban theory and research in the age of the sensed and metered city, ubiquitous computing and big data. *Cambridge Journal of Regions, Economy and Society*. pp 1 - 16

Resnick, J. C. 2011. Increasing Opportunity through Interdisciplinary Research: Climbing Down and Shattering a Tower of Babel. *Frontiers in Psychiatry* [online], 2. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3098683/> [Accessed 20 April 2015]

Rudin, C. 2012. Machine Learning for the New York City Power Grid. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [online], 34 (2). pp. 328-345 <http://hdl.handle.net/1721.1/68634> [Accessed 20 April 2015]

Ruppert, E., Law, J. and Savage, M. 2013. Reassembling Social Science Methods: The Challenge of Digital Devices. *Theory, Culture & Society*, 30(4). pp 22–46

Simon, H. A. 1962. The Architecture of Complexity. *Proceedings of the American Philosophical Society*, 106 (6). pp 467-482

Singel, R., 2012. Netflix Spilled Your *Brokeback Mountain* Secret, Lawsuit Claims. *Wired* [online], 17 September. <http://www.wired.com/2009/12/netflix-privacy-lawsuit/> [Accessed 22 April 2015]

Slaughter and May. 2013. Data Protection Reform. *Briefing*. [online]. London: Slaughter and May. Available From: <https://www.slaughterandmay.com/media/1957777/data-protection-reform.pdf> [Accessed 15 April 2015]

Slaughter and May. 2014. Selected legal and regulatory developments in data protection. *Data Protection* [online]. London: Slaughter and May. Available From: <http://www.slaughterandmay.com/media/2161364/data-protection-newsletter-may-2014.pdf> [Accessed 15 April 2015]

Smithers, R. 2011. Terms and conditions: not reading the small print can mean big problems. *The Guardian* [online] 11 May. <http://www.theguardian.com/money/2011/may/11/terms-conditions-small-print-big-problems> [Accessed 20 April 2015]

Spurious Correlations. [online] Available from: <http://www.tylervigen.com> [Accessed 20 April 2015]

Szabo, N. *Formalizing and Securing Relationships on Public Networks*. [online]. <http://szabo.best.vwh.net/formalize.html> [Accessed 22 April 2015]

Taleb, N., N. 2012. *Antifragile. Things that Gain from Disorder*. London: Penguin Books

Tgordon. 2013. Urban Prototyping London- Hackathon and Crackathon April 25 2013. *Imperial College London* [online]. Available from: <http://www.imperial.ac.uk/blog/digecon/2013/05/28/urban-prototyping-london---hackathon-and-crackathon-april-25-2013/> [Accessed 22 April 2015]

The Economist. 2010a. One Big Bad Trade. *The Economist* [online]. 1st October. http://www.economist.com/blogs/newsbook/2010/10/what_caused_flash_crash [Accessed 22 April 2015]

The Economist. 2010b. The flash crash: Autopsy. *The Economist* [online]. 7th October. <http://www.economist.com/node/17202255> [Accessed 22 April 2015]

The Economist. 2014. Russian money in Britain. Honey trapped. *The Economist* [online]. 22 March. <http://www.economist.com/news/briefing/21599408-london-has-more-lose-most-when-it-comes-scaring-oligarchs-honey-trapped> [Accessed 22 April 2015]

The Economist. 2015. Why Singapore became an economic success. *The Economist* [online], March 26. <http://www.economist.com/blogs/economist-explains/2015/03/economist-explains-23> [Accessed 22 April 2015]

Tong, H., 2015. Big Data Classification. In: Aggarwal, C. C., 2015. *Data Classification. Algorithms and Applications*. London: Taylor & Francis Group / Chapman & Hall 2015, pp 275 – 283

Townsend, A., M. 2013. *Smart Cities. Big data, civic hackers, and the quest for a new utopia*. London: W. W. Norton & Company

Trutnevyte, E. et al., 2014. Linking a storyline with multiple models: A cross-scale study of the UK power system transition. *Technological Forecasting and Social Change* [online], 89. pp 26 – 42

UCL Enterprise. 2015. Cisco [online] Available from: <https://www.ucl.ac.uk/enterprise/corporate-partnerships/partner-information/cisco> [Accessed 23 April 2015]

Verebes, T., 2014. *Masterplanning the adaptive city. Computational Urbanism in the Twenty-First Century*. London: Routledge

Wierzbicki, A. P., 2007. Modelling as a way of organising knowledge. *European Journal of Operational Research*, 176. pp 610–635

Winsberg , E., 2009. Computer Simulation and the Philosophy of Science. *Philosophy Compass*, 4(5). pp 835 – 845

Yasseri, T. 2013. The physics of social science: using big data for real-time predictive modelling. *The Policy and Internet Blog. Understanding Public Policy Online* [online]. November. <http://blogs.oii.ox.ac.uk/policy/physics-of-social-science-using-big-data-for-real-time-predictive-modelling/> [Accessed 22 April 2015]

