NATIONAL CENTRE FOR RESEARCH METHODS

Interdisciplinary report

October 2021

Computational Social Science: A Thematic Review

Robert Meckin, University of Manchester
Mark Elliot, University of Manchester

UKRI Economic and Social Research Council

Computational Social Science: A Thematic Review

Robert Meckin
Sociology
School of Social Sciences
University of Manchester
Robert.meckin@manchester.ac.uk

Mark Elliot
Social Statistics
School of Social Sciences
University of Manchester
Mark.elliot@manchester.ac.uk

Table of Contents

Abstract

The explosion of social digital data and the concomitant increases in computational capabilities along the data analytics pipeline (data acquisition, storage and analysis) impact upon the possibilities and choices for conducting social research. This report examines the emerging research field called computational social science (CSS). The aim of this review is to offer insight into the shape of CSS, its questions and methodologies, and how these relate to and interact with different social science disciplines. Two searches and hand sorting identified 41 of the most highly cited publications. The papers were initially categorised into two main groups of papers: substantive-technical contributions and critical-review contributions. The groups were thematically analysed. As a validation and refinement exercise, a further search identified thirty of the most recent CSS papers, which were also categorised and analysed. The review focuses on the first 41 articles as well as several other relevant articles are discussed that were identified through citations, additional *ad hoc* searches, and personal conversations. The substantive-technical literature and critical-review literature can each be sub-divided into three groups, and findings from these six groups are described. In the discussion, we draw out points related to interdisciplinarity and potential implications of the findings for engagement research communities.

Keywords: Computational social science, big data analytics, social media analysis, interdisciplinary research

Introduction

> *Our movements in public places may be captured by video cameras, and our medical records stored as digital files. We may post blog entries accessible to anyone, or maintain friendships through online social networks. Each of these transactions leaves digital traces that can be compiled into comprehensive pictures of both individual and group behavior, with the potential to transform our understanding of our lives, organizations, and societies. (Lazer et al. 2009, 721)*

The quotation above comes from a short and superficially unassuming paper that appeared in *Science* positing the development of a new approach to social research: computational social science or CSS (Lazer et al. 2009). The paper highlighted how access to new forms of data including transactional purchase data, public transportation data, mobile phone triangulation and file logs, records of social media usage, administrative data, and data from public and private surveillance, were creating large volumes of digital records of behaviour that, due to increases in computer memory and processing power, could be analysed and may offer new insights in social science.

Since the publication of the paper, computational approaches to social science have continued to be explored, perhaps increasingly, given the comparatively high number of publications that cite the paper. However, there are differing definitions for CSS and some contestation over the use of related terms, some which are discussed below. Taken collectively, though, the definitions we present highlight the social research possibilities and challenges created by the emergence of digital data storage and analysis including the potential to generate knowledge about many and arguably whole populations, the variety of digital methods at its disposal, and the ethical and epistemological issues at stake.

Some scholars refer to "the analysis of huge data sets as obtained, say, from mobile phone calls, social networks, or commercial activities" (Conte et al., 2012). The terms "big data", "BigData", "big data analysis", and "big data analytics" are frequently used in literature regarding CSS (e.g. Conte et al., 2012: 331; Lazer et al., 2014; Siverajah et al., 2017; Vargo et al., 2018).[1] Rather than focusing on the size of data, Cioffi Revilla (2010) classifies CSS into five areas, each of which, he argues, make their own contributions to analyses of social phenomena:

- *Automated information extraction, which can enable real time analysis of news and other reports.*
- *Social network analysis (SNA), which can help understand belief systems, organisations and network games.*
- *Geospatial analysis using social geographic information systems (socio-GIS).*

[1] The authors of the report feel it is important to note that "big data" is used in other science and research contexts including, for example, astronomy, biology and medicine, and that authors in the CSS contexts imply "social big data" and that, from our perspective, the term "big data" may not be a useful concept in capturing some of the more distinctive qualities of the contemporary social data environment, which arguably include the capabilities to link large data sets and analyse data as it is produced ie. in "real time".

- *Complexity modelling, which can be used to explore markets, international aid programmes and natural disasters.*
- *Social simulations modelling, where system dynamics and agent-based models can be used to understand issues like national and international exchanges and can be used to inform state domestic and foreign policy.*

(Cioffi Revilla, 2010)

Thus, a little over a decade ago, researchers gathered together different computational techniques, argued they were related to social research questions as situated in contemporary social, economic, technological and educational conditions.

Inevitably, perhaps, a different take on CSS emerged from the perspective of critical commentators. For example, boyd and Crawford (2012) define big data analyses as:

> *A cultural, technological, and scholarly phenomenon that rests on the interplay of:*
>
> *(1) Technology: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets.*
>
> *(2) Analysis: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims.*
>
> *(3) Mythology: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy.*
>
> *(boyd and Crawford 2012: 663)*

This definition casts big data analytics as a socio-technical phenomenon rather than as a neutral technology such that its context, use, and connections to other domains in society affect the particular affordances and challenges that big data analytics create. Importantly, as well as including the technological context and goals of such analyses, their definition draws attention to the epistemic promises that accompany increases in data size and computational power – promises and claims that they critique throughout their article. We revisit these critiques later in the report.

In order to better understand the shape and direction of the CSS field now that it is over a decade from its initial labelling, this report presents a literature review that sketches out the field, charting its methodological commitments and key debates. Next, we describe our review methods (for data generation and analysis). Then we discuss the findings of the review and close by highlighting some of the issues regarding interdisciplinarity and suggest potential directions for engagement and research.

Methods

Background

NCRM was funded for a fourth 5-year phase, running from 2020 to 2024. NCRM's remit is to identify and engage in methodological development in particular areas in the social sciences. A focus on CSS was identified in discussions that led the production of the National Centre for Research Methods (NCRM) strategic engagement plan (Elliot 2020). NCRM identified CSS because is an emerging field in social research with potential to influence projects in many sectors, and so gaining an overview of its methodological dimensions would support strategic engagement with practitioners in the field. CSS was also identified partly because of one author's (Elliot) data science expertise and knowledge and partly because the other author (Meckin) is interested in interdisciplinary collaboration and in the interplay between technologies and research methods from the perspective of science and technology studies (STS).

The authors agreed that Meckin should lead a literature review with the aims of identifying methodological approaches in CSS and to attempt to identify key debates in the field. The authors decided to begin with the Lazer et al. (2009) paper quoted in the introduction and treat it as the "source paper" of CSS. From there, Meckin was to conduct a narrative literature review that would meet the aims stated above, discussing the search strategy, readings and findings with Elliot every two weeks. Meckin would then draft a literature review and the authors would use that initial draft to develop the report, iteratively discussing the findings, claims and points to include, both in online meetings, in the margin using word-processor application "comment" features, and by writing and editing sections of text.

A literature review of one field (CSS) led by a scholar (Meckin) from another field of research (STS) is perhaps a strange thing. It can be difficult to know what is important, or what seemingly insignificant things turn out to be crucial. The initial approach, therefore, was more that of a cultural anthropologist. Albeit, partly given the restrictions in place because of the coronavirus pandemic, a deskbound one. It is often the case that initial acquaintance with an object of study can be rich and informative because everything is new. At the same time, it is not yet clear what is significant and what is not. Meckin had some prior knowledge of the intersection of computational methods and social data because he had read some academic commentary (e.g. boyd and Crawford 2012), media reports and social media discussions. Although he was coming to CSS with little interactional or analytical experience, Meckin had studied other emerging technosciences. Particularly relevant to this study was work he had conducted on digitalisation and automation in the field of synthetic biology. Also, in developing another project, he had also read some of the "politics of method" literature (e.g. Savage 2010, Savage & Burrows, 2007; Ruppert et al., 2013). Thus, while Meckin had no expertise in CSS, he was aware of some critiques of data science and its methods via other reading.

Selecting and categorising the papers

Overall, the literature review proceeded in three main steps:

1. Two different searches identifying 40 papers (top 20 from each search).
2. Analysis of these 40 papers.
3. Validation of analysis with an additional 30 papers.

In terms of the details of the literature review process, Meckin proceeded as follows. First, he read the Lazer et al. (2009) paper and then used Google Scholar (following a social media discussion regarding narrative reviews initiated by Deborah Lupton) to find the top most cited papers that also cited the Lazer et al. (2009) piece. This strategy created two problems. First, Google Scholar returned two entries for the Lazer et al. (2009) paper and neither were the original. Second, Google Scholar incorporates previous search histories and, probably, previous links that the searcher has clicked, into the results the search engine reports. But, because the algorithms are secret, biases based on search histories are introduced in an opaque way. One possible remedy is to continually set up new accounts that do not have search histories. Instead, Meckin switched to Web of Science which returned the original paper and which has the further advantage in that searches are potentially more replicable, meaning that one can rerun searches without the browser cache making a difference barring the addition of articles published in between searches or moving geographical or institutional location (Pozsgai 2020). The search in Web of Science returned 1477 publications that had cited Lazer et al. (2009).

The Web of Science search had good face validity – boyd and Crawford (2012) was the top hit, for example. The authors agreed to select the top twenty relevant articles for close reading, initially. Meckin decided relevance by reading the title and, if unsure, the abstract, to check that both computational methods and social questions were under discussion. He excluded one entry from the top twenty on the grounds it was primarily concerned with biology: *Network Neuroscience*. He downloaded twenty papers as .pdfs and four chapters from the book *Code/Space* (Kitchin & Dodge, 2011) and also downloaded the top 500 citations from the search. He analysed the twenty papers in three ways. He recorded features of the papers, such a year, authors, title, etc. in Excel. He loaded the papers into NVivo 12 to code[2] the texts in an open way by highlighting sections of text and moving them into new or existing codes. The coding structure was open and dynamic. It was open in that he started with no codes and generated them according to content with no limits on new codes. It was dynamic in that he sometimes combined codes, or moved them into hierarchies or groups. After reading each paper he also made notes in Scrivener 3 recording his reflections that, in some cases, included emotional responses.

As Meckin began classifying the first twenty papers – there appeared to be two main sets of papers. He considered them as *critical* papers and *substantive-technical* papers. Critical papers were commentaries and reviews of CSS and related topics that did not include primary research. Substantive-technical papers were contributions that were primarily technical and more focused on presenting methods and findings. He was able to further divide the substantive-technical literature into three categories that he called, at this point, *social cartography*, *transfer modelling* and *trace correlating*. As Meckin and Elliot discussed findings

[2] The authors have discussed the meanings of code. In the qualitative analysis sense intended here by Meckin, code means assigning symbols or categories for the purposes of classification and identification, and is a way of drawing out themes in qualitative data across multiple sources. Elliot, as data scientist, was most familiar with code in the software programming sense, where symbols provide instructions for computational processes. Kitchin & Dodge (2011), reviewed later in the report, play on the multiple meanings of code in their topological analysis.

and added more literature, the names changed to be conceptually more appropriate, but are broadly consistent with this initial classification.

The final category names for the substantive-technical literature, which we cover in detail later in the findings sections, are *network cartography, tracking influence and transference, and categorising and correlating digital traces*. We should say explicitly here, and it will be clear in the discussion, that the typology does not aim to be a comprehensive or complete set of categories but indicative of particular assumptions, interests and philosophies in play in CSS.

The second phase of the review involved retrieving another selection of twenty papers. Meckin searched the exact phrase "computational social science" in Web of Science. This returned 493 results (far fewer than Google Scholar). In this phase, he excluded papers that were repeats, so Lazer et al. (2009) and Kosinski et al. (2013) were removed. Again, he loaded the papers into NVivo 12 to thematically analyse them, made field notes in Scrivener 3, and recorded their features in Excel.

One paper presented an query in terms of inclusion: after reading the full text of Luke et al.'s (2005) article, the publication turned out to be a description of an agent-based model (ABM) called MASON. Although the publication predates the publication of the source paper, ABMs are an important method for CSS research (Cioffi Revilla, 2010). In Luke et al.'s (2005) paper, there is some mention of social questions and the later work that was done to add modules that could support social network modelling and system dynamic modelling. For these reasons, we retained this paper the review, even though it predates the Lazer et al. (2009) paper and is focused on the computation rather than social science, as it shows the ways that computational methods intersect with other fields of study.

Overall, this meant 41 individual papers were included in the analysis (the source paper and two searches each of 20 papers).

Thematic analysis

The main organisation of the review is from the article typology generated by reflecting on the papers' contents and features in Scrivener 3 and Excel. Meckin and Elliot discussed and refined the categories, sometimes because papers were hard to place. For instance, Farrell's (2016) paper on mapping online the discourse of climate scepticism appeared at first glance to be *network cartography*, but closer reading meant it was more closely affiliated with the final category *tracking influence & transference*.

After reading the 41 papers Meckin generated fourteen top level codes, seven of which contained further codes. In total, there were 103 codes. The top-level[3] codes were:

- Critique responses
- Data*
- Epistemology*
- Ethics*
- Findings
- Future

[3] * in the list means they contain other codes

- Impact (hoped hype)
- Innovation
- Interdisciplinarity*
- Method*
- Ontology*
- Secrecy
- Social questions*
- Software

The codes provided a subject reference for writing up the analysis. An overview of the substantive-technical literature is presented in Table 1 below.

Table 1: A grouping of 26 substantive-technical articles relating to CSS (see Appendix for a copy of the table with the references to indicate how the categories were generated).

Area (no. papers)	**Network cartography (3)**	**Tracking Influence & transference (13)**	**Categorising and correlating digital traces (10)**
Aims	Map patterns of social relations; improve on other (social network analysis) methods.	Chart spread or transfer of entities through groups or populations and effect on behaviour.	Classify digital traces and show that categories correlate with other phenomena (e.g. data from personality tests, stock markets or geographic location and therefore may predict individuals and populations)
Prominent methodological principles	Focus on spatial/network location and transactional connections. Relation derived from node activity. Description, accuracy,	Focus on network location, transactional connections and activity. Identification of causal/directional relations. Description, prediction, causality,	Focus on activity and/or cultural outputs. Assembly of aggregates for comparison. Extraction, classification, correlation, interpretation, inference,

	comparison, generalisation.	contextualised findings.	prediction, revelation.
Prominent concepts	Networks	Networks, contagion, influence, spread, behaviour, ties.	Language, personality, mood, traits, attributes.
Example Technique	Social Network Analysis	Agent based models	Factor Analysis
Data exemplars	Mobile phone location data, Bluetooth data, online game data	Mobile phone application download data; Combining data e.g. Bluetooth records with health service data	Social media posts, messages, photographs, “likes”, etc; Wikipedia editing activity
Disciplinary interactions and sources	Sociology, social-psychology, psychology	Management, politics, epidemiology, health, innovation, physics, media studies	Psychology, management

The method of sampling used creates sampling biases. First, some computer science and data science fields value conference proceedings as research outputs and so there is a risk that some aspects of CSS may advance through the channel of conference proceedings. However, both Google Scholar and Web of Science index conference proceedings and neither returned highly cited papers from proceedings. This may be because the norms of social science publishing may be dominant in CSS. This issue is not explored further in this report, but is an area that could be examined in future work.

The second bias is that searching for highly citing papers privileges older papers that have had time to be cited. This means that many of those included in this review were published between 2009 and 2012, with no papers identified after 2018. This latter issue presents a problem as key advances in CSS may be underway, and this review may therefore reflect an earlier state of the art. The validation step in the analysis, described below, goes some way to addressing this bias, as that included most recent papers that had not had time to be cited.

Other lines of inquiry

Sometimes, Meckin responded to emerging findings by diverging from the formal, structured mode of collection and analysis. This happened notably in three ways:

1. After Meckin noticed that many authors featured twice (or more) among the corpus of articles, he sometimes searched for them in Google, to understand the backgrounds to some of the scholars. For instance, he recognised the name “C. Cioffi-Revilla” from one of the earlier searches and who was listed as an author on Luke et al. (2005). He downloaded another paper *Computational Social Science* (Cioffi-Revilla, 2010), which was useful for characterizing the initial emergence of CSS. Meckin also used this method to help classify some of the papers – checking the disciplinary backgrounds and affiliations of authors to get a sense of contributors to the field. For instance, he searched for Nathan Eagle (Eagle et al., 2009), who was a PhD student of David Lazer, and turned out to be a successful technology entrepreneur, and Michal Kosinksi. Meckin recognized Kosinski’s name as he been interviewed on a documentary reporting on the Cambridge Analytica scandal among other things (Bartlett, 2017). These *ad hoc* searches helped develop background and contextual knowledge of connections across CSS, as well as links to popular media discussions.
2. Wang & Kosinksi’s (2018) article provoked a somatic response of adrenaline rushing through his legs. It’s not that normal to get a flight/fight/flock/freeze response to an academic paper, and he immediately searched for commentary and coverage in Google. This generated a couple of non-academic commentaries that are cited later.
3. Searching for citations that appeared to theoretically important to literature on network structure, for instance: Centola & Macy (2007) and Granovetter (1973).

So, while we systematically searched for literature to include, previous knowledge, experience and responsivity all played a part in creating a hinterland that affects the review, the extra literature we include, and the interpretations we present.

Validation

In order to test the categories of literature, Meckin searched “computational social science” in Web of Science, which returned 504 results this time (this search, on 23rd May 2021, was about two months after the first search) and chose the thirty most recent publications to see if he could position them within the emerging typology. It was fairly straightforward to classify papers with their primary affiliation, while several bridged two categories. In terms of numbers, they roughly reflect the first 41 papers in terms of very few critical external reviews (just one). However, *categorising and correlating digital traces* is by far the dominant category in this corpus, with far fewer papers being in either of the other two categories. The categorisation of the thirty papers is presented in table 2 below.

Two articles, apparently emerging from a single project, presented a problem in the validation step. Ramit Debnath and colleagues publishe the two papers in *Energy Research & Social Science* (Debnath et al., 2020; Debnath et al., 2021). The papers draw on an innovative methodological approach, integrating focus groups and a CSS method of text analysis, in a way that none of the other papers do. This suggests CSS methods can be used with other methods within methodological frameworks. The main issue for us, though, was that the identification of topic modelling did not fit within the three categories of literature – initially, it did not appear to fit with mapping, tracking or, notably, correlating. The authors decided to adapt the category name of “*Correlating digital traces”* to “*Categorising and correlating digital traces”* since addition of *“categorising”* seemed to better capture the methodological interests in the space, and to have the benefit of making more conceptual space for the social psychological papers that emphasise classification work.

Table 2. Categorising 30 most recent papers found with the term "Computational Social Science" in Web of Science.

	Authors	Article Title	Category	Subcategory
1	Tuninetti, M; Aleta, A; Paolotti, D; Moreno, Y; Starnini, M	Prediction of new scientific collaborations through multiplex networks	Substantive-technical	Cartography
2	Perez-Verdejo, JM; Pina-Garcia, CA; Ojeda, MM; Rivera-Lara, A; Mendez-Morales, L	The rhythm of Mexico: an exploratory data analysis of Spotify's top 50	Substantive-technical	Categorising
3	Guess, AM; Barbera, P; Munzert, S; Yang, JW	The consequences of online partisan media	Substantive-technical	Tracking
4	Taghikhah, F; Filatova, T; Voinov, A	Where Does Theory Have It Right? A Comparison of Theory-Driven and Empirical Agent Based Models	Critical	Internal method review
5	Kejriwal, M	On using centrality to understand importance of entities in the Panama Papers	Substantive-technical	Cartography (with some tracking)
6	Lockhart, JW	Paradigms of Sex Research and Women in Stem	Substantive-technical	Categorising
7	Gonzalez-Bailon, S; De Domenico, M	Bots are less central than verified accounts during contentious political events	Substantive-technical	Tracking
8	Tomaselli, V; Battiato, S; Ortis, A; Cantone, GG; Urso, S; Polosa, R	Methods, Developments, and Technological Innovations for Population Surveys	Critical	Internal tech review
9	Lam, JCK; Li, VOK; Han, Y; Zhang, Q; Lu, ZY; Gilani, Z	In search of bluer skies: Would people move to places of better air qualities?	Substantive-technical	Categorising
10	Olson, AW; Calderon-Figueroa, F; Bidian, O; Silver, D; Sanner, S	Reading the city through its neighbourhoods: Deep text embeddings of Yelp reviews as a basis for determining similarity and change	Substantive-technical	Categorising

11	Fu, Q; Zhuang, YF; Gu, JX; Zhu, YS; Guo, X	Agreeing to Disagree: Choosing Among Eight Topic-Modeling Methods	Critical	Internal method review
12	Debnath, R; Bardhan, R; Darby, S; Mohaddes, K; Sunikka-Blank, M; Coelho, ACV; Isa, A	Words against injustices: A deep narrative analysis of energy cultures in poverty of Abuja, Mumbai and Rio de Janeiro	Substantive-technical	Categorising
13	Gillani, N; Chu, E; Beeferman, D; Eynon, R; Roy, D	Parents' Online School Reviews Reflect Several Racial and Socioeconomic Disparities in K-12 Education	Substantive-technical	Categorising
14	Gallagher, RJ; Frank, MR; Mitchell, L; Schwartz, AJ; Reagan, AJ; Danforth, CM; Dodds, PS	Generalized word shift graphs: a method for visualizing and explaining pairwise comparisons between texts	Substantive-technical	Categorising & novel method
15	Schimpf, C; Barbrook-Johnson, P; Castellani, B	Cased-based modelling and scenario simulation for ex-post evaluation	Internal	Review
16	Peris, A; Meijers, E; van Ham, M	Information diffusion between Dutch cities: Revisiting Zipf and Pred using a computational social science approach	Substantive-technical	Tracking
17	Alassad, M; Spann, B; Agarwal, N	Combining advanced computational social science and graph theoretic techniques to reveal adversarial information operations	Substantive-technical	Tracking and categorising
18	Fuchs, C	Engels@200: Friedrich Engels and Digital Capitalism. How Relevant Are Engels's Works 200 Years After His Birth?	Critical	External
19	Pierri, F; Piccardi, C; Ceri, S	A multi-layer approach to disinformation detection in US and Italian news spreading on Twitter	Substantive-technical	Tracking
20	Carter, EB; Carter, BL	Propaganda and Protest in Autocracies	Substantive-technical	Categorising & a bit of tracking
21	Wu, AX; Taneja, H; Boyd, D; Donato, P; Hindman, M;	Computational social science: On measurement	Critical	Internal commentary

	Napoli, P; Webster, J			
22	Yang, T; Majo-Vazquez, S; Nielsen, RK; Gonzalez-Bailon, S	Exposure to news grows less fragmented with an increase in mobile access	Substantive-technical	Categorising & a bit of tracking
23	Botta, F; Moat, HS; Preis, T	Measuring the size of a crowd using Instagram	Substantive-technical	Categorising
24	Debnath, R; Darby, S; Bardhan, R; Mohaddes, K; Sunikka-Blank, M	Grounded reality meets machine learning: A deep-narrative analysis framework for energy policy research	Substantive-technical	Categorising
25	Ma, J	Automated Coding Using Machine Learning and Remapping the US Nonprofit Sector: A Guide and Benchmark	Substantive-technical	Categorising
26	Boyd, RL; Schwartz, HA	Natural Language Analysis and the Psychology of Verbal Behavior: The Past, Present, and Future States of the Field	Critical	Internal review
27	Theocharis, Y; Jungherr, A	Computational Social Science and the Study of Political Communication	Critical	Internal review
28	Young, JA	#SocialWorkEducation: A Computational Analysis of Social Work Programs on Twitter	Substantive-technical	Cartography & categorising
29	Feinberg, F; Bruch, E; Braun, M; Falk, BH; Fefferman, N; Feit, EM; Helveston, J; Larremore, D; McShane, BB; Patania, A; Small, ML	Choices in networks: a research framework	Substantive-technical	Tracking
30	Yantseva, V	Migration Discourse in Sweden: Frames and Sentiments in Mainstream and Social Media	Substantive-technical	Categorising

The final categories

Following the validation exercise, we settled on six areas of literature to discuss. In light of this, the review is presented in the following sections:

Substantive-technical literature:

1. *Network Cartography (“Cartography” for short)*;
2. *Tracking Influence & transference (“tracking” for short)*; *Categorising and*
3. *correlating digital traces (“Categorising”)*, and

Critical literature:

1. *Internal reviews*;
2. *External commentaries*; and
3. *Tangential contributions*.

Substantive-technical literature

Papers in this category contribute methods and knowledge to CSS and the implicit audience is typically other CSS scholars, as well as demonstrating the capabilities of CSS analyses to scholars within particular disciplines. The substantive-technical literature is divided into three categories. This is not an absolute distinction but is meant to be indicative of the analytical interests, aims, methods, philosophies and research areas that contribute to CSS. They are best regarded as poles of orientation and many papers, while orientated towards one pole, sit in the methodological space between them. Each of the three areas are described in terms of:

- Examples of studies
- Theory and concepts
- Interdisciplinary connections
- Aims

We have also added an additional subsection within our description of *categorising and correlating digital traces* to explore further one particular strand of the literature. This is because we feel that Michal Kosinski's papers that were selected in our review process, and that we mentioned in the methods section, merit additional reflection in light of the cultural connections and social claims they make. What follows in this section is an overview of the three CSS research approaches we identified.

Network cartography

This strand of literature deals with mapping social networks[4] using digital methods and digital traces. This strand contains the fewest number of articles in the review – just three. However, network cartography represents a distinct subsection of our substantive-technical literature.

Examples of studies

Two of the studies recruited university students as their participants and tracked their whereabouts and interactions using data from mobile phones (Eagle et al., 2009; Sekara et al., 2016). The other study used data from an online game, which comprised all the game data of approximately 20,000 players in the first 445 days of one "universe" in the online game "Pardus" (Szell et al., 2010). Eagle et al.'s (2009) study compared mobile phone data to survey data and claims that data from mobile phones (cell tower ID, Bluetooth, applications, phone status, and call log data) can accurately map onto self-reported data, meaning that mobile phone data could be used where self-report data is absent, incomplete, or otherwise unreliable. The Eagle et al (2009) study draws on data generated through the "reality mining" project[5]. They claim it is possible to infer the satisfaction participants feel with respect to particular groups within their networks. Sekara et al. (2016) used Bluetooth data to track individual phones' proximity to one another and were able to describe the differences between social patterns. They identify instances of single-event 'gatherings' and infer longer-term networks, such as friendship or

[4] These is an obvious interface here to the long standing interdisciplinary field of *social network analysis*. Many social network analysts would not identify themselves as computational social scientists – and our analysis could undoubtedly be enriched by considering this relationship in more detail. However, in the interests of parsimony we just note it for now and place it firmly in the "future work" tray!

[5] Details available here: http://realitycommons.media.mit.edu/realitymining.html.

class groups, and the relationship between these phenomena. Szell et al. (2010) use gaming data to map six different kinds of collegial or aggressive relationships between gamers and show that the multidimensionality of relationships is important to understand the “structure and stability” of social networks (Szell et al., 2010: 13640).

Theories and concepts

The studies use the concepts of individuals (nodes) and networks. They are primarily interested in the description of networks which they achieve by identifying individuals as phones or players and constructing different dyadic relations between nodes predominantly based on proximity, communication and interaction. The strengths of different associations are interpreted by the volume of interaction (proximity, calls, messages, transactions etc). In Szell et al.’s (2010) paper, the online game players can also attribute simple qualities to other players – whether the player classes another player as a friend or enemy – and they are able to integrate that into their six-part typology of relationships. That aside, the studies generally use activity, as the volume of interactivity, to infer qualities of relationships such as friendship or dissatisfaction. Eagle et al. (2009) also compare their results to a survey, although their survey and classifications do not reference existing sociological knowledge about friendships and acquaintances. Szell et al. (2010) use the idea of “ties” and are specifically interested in measuring the strength of weak ties.

The overall collection of relationships suggests an aggregate ontology in which discreet entities of nodes, agents or individuals are collected into larger assemblies of clusters, networks and groups. Perhaps, at its most fundamental, the network ontology starts with the assumption that all the nodes are the same, but relationships can vary in strength.

Interdisciplinarity

All three studies are heavily influenced by scholars in computer science, complex systems science and mathematics. They are primarily concerned with demonstrating the application of computational methods of digital data and their relevance for inferring social stuctures. Thus, from this small collection of papers, computational mathematics is the primary tool to describe the organisation of networks.

Aims

The *network cartography* studies aim to show that particular forms of digital data “have the potential to provide insight into the relational dynamics of individuals” (Eagle et al. 2009, 15274). This is orientated to showing the patterns and organisation of social groups, and individuals across social groups, and could be closely associated with social network analysis (SNA). One of the key claims is that mobile phone data and other digital trace data provide reliable data if self-reported data is unavailable or unreliable:

> *The field devoted to the study of the system of human interactions - social network analysis - has been constrained in accuracy, breadth, and depth because of its reliance on self-report data. (Eagle et al. 2009, 15274)*

A potential impact is to demonstrate the power of digital data because of the increase in available data points for analysis. For instance, by having the greater temporal resolution afforded by mobile phone data (e.g. Bluetooth connection data updated every 5 minutes), complex mathematics are apparently no longer required to show community structure meaning it is possible to understand better how people group together through time:

> *When single time slices are shorter than the rate at which social gatherings change, communities of individuals can be observed directly and with little ambiguity… Using a simple matching between time slices, we can infer temporal communities. These dynamic communities offer a powerful simplification of the complex system of social interactions as it develops over time. (Sekara et al. 2016, 9977)*

Accordingly, the claimed power of digital data lies in its comprehensiveness and resolution, and suggests that theory is less important given the data is more "telling" (see boyd and Crawford (2012: 665-666) for a critical discussion of this claim).

A way that the authors show the improved data is by attempting to show that digital data can be used to accurately infer networks and, from there, to infer the feelings of individuals in regards to particular networks:

> *The relationship between satisfaction and interactions patterns… was exactly as predicted, that is, having friends - especially ones to whom you were near at work predicted satisfaction with the work group, and calling friends while at work was associated with lack of satisfaction with the work group.*
>
> *What is important, from the perspective of this paper, is that the inferred friendship network … produced substantive-technically identical results to the self-report model, with a slightly improved fit. These nearly identical results suggest that it is possible to accurately infer subjective job satisfaction based solely on behavioral data, validating the inferred measure of friendship. (Eagle et al. 2009: 15277)*

This quotation argues that it is possible to gauge people's feelings about groups in their networks by examining their contact patterns with members of different groups at different times. The claim here is that it is possible to move from mobile phone location and connection records, to behaviour, to affective dispositions without needing to know the content of messages or calls.

Researchers set up *network cartography* approaches using digital data against other methods, such as survey and self-report data, with the promise of changing how social science understands social interaction patterns. There are caveats, though, as suggested by Szell et al:

> *Traditional methods of social science, such as small-scale questionnaire-based approaches, get more and more replaced by automated methods of data collection which allow for entirely different scales of analysis… This change of scale has opened new perspectives and has the potential to radically transform our understanding of social dynamics and organization… However, this large-scale perspective suffers from the drawback of a relatively coarse-grained representation of social processes taking place between individuals and of blindness in respect to the existence of different types of social interactions. (Szell et al., 2010, 13636).*

The authors indicate the need for *network cartography* to attend to different types of relations, rather than treating them in a flat, undifferentiated way. They go on to describe types of relations within a multiplayer online game, arguing that it is possible to characterise relationships in a multifaceted way, rather than assuming they are one-dimensional. Ultimately, a central aim of *network cartography* appears to be to demonstrate that digital data can produce high quality knowledge about the spatiotemporal dynamics of human relations and the structures of human networks.

Tracking Influence & Transference

This strand of the research contains the largest number of substantive-technical literature references of the initial corpus (13 papers). In some ways, it is closely related to *network cartography* but is distinguished, as discussed below, by its interest in entities that move between nodes in a network.

Examples of studies

We have selected several examples to give a sense of different aspects of this literature as it features in studies of management and innovation, politics, and studies of social media. First, Sinan Aral's work (Aral et al., 2009; Aral & Walker, 2012) is concerned with the adoption of a product (a mobile app) in networks of mobile phones. The earlier paper set up the problem of being able to tell whether people adopt an innovation because the people in networks are similar to one another, known as homophily, or because they influence one another to do so. The paper develops a method that can separate the effects of influence and homophily in a network. Their second paper in the review is aimed at developing a method that can identify "influential" and "susceptible" people within networks (Aral & Walker, 2012).

A second strand in our literature is that an interest in the spread of political sentiment online is high, possibly because of the aftermath of the 2016 US presidential election, the UK Brexit referendum, and the Cambridge Analytica scandal. Bail et al. (2018) report a US-based "field experiment" where participants signed up on social media and were regularly sent automatic messages on Twitter that opposed their claimed political views. The study found that those claiming to be Republicans posted more conservative views after exposure, while those claiming to be Democrats did not significantly increase their posts of liberal views. These first two examples – Aral's and Bail et al.'s work – use randomised experimental method designs and track the effect of "interventions" on activity.

The third example by Procter et al. (2013) presents an analysis of 2.6 million Tweets related to the August 2011 riots in England. They show that although misinformation does spread on Twitter and that social media can be used to incite illegal acts, "Twitter was used overwhelmingly for more positive ends" such as "the organisation of the riot cleanup [sic]" (Procter et al., 2013, 206). Here, the study aimed to nuance public debate, challenging idea that social media is necessarily harmful by showing how different messages spread through the networks.

Theories and concepts

There is a similar aggregate ontology to *network cartography*, with the added feature of elements that move through networks to change activity or the behaviour of nodes. The main concepts in the literature are related to contagion, influence, spreading, agents and networks.

Thus, the ontology includes transferable entities, like viruses or memes, that can move between agents or nodes within a network consequently changing their activity, or external conditions that can influence activity. Secondly, this attention to causality affords a potential for intervention by predicting diffusions that influence decision-making or activity. The literature reviewed was related to the spread of several types of entities that include:

- Viruses (Funk et al., 2010; Wang et al., 2016)
- Informational contagion (for a review see Vespignani, 2012)
  - Uptake of products (Aral et al., 2009; Aral & Walker, 2012)
  - Political messages (Bail et al., 2018; Bond et al., 2012; Farrell, 2016; Vargo et al., 2018)
  - News and memes (Procter et al., 2013; Weng et al., 2013)
  - Unknown influences (Christakis & Fowler, 2011)

The literature appears to be influenced by previous work on the “strength of weak ties” (Granovetter, 1973). The main insight is that there are two kinds of relationships between nodes: “strong ties” between like those between kin and friends, and “weak ties”, which connect nodes more distant from one another and thus link different groups together in a network. Granovetter shows that weak ties mean that particular kinds of contagion, in his case job information, can diffuse through networks because of the way that weak ties link together different groups. Indeed, strong ties tend not to be significant in their connections across groups. More recent developments have nuanced this insight to show to how transmissible entities can be understood: “simple contagion” can be to spread with just one exposure (a virus is the paradigm example), whereas “complex contagion” require more than one so that “successful transmission depends upon interaction with multiple carriers” (Centola & Macy, 2007: 703). This implies that strong ties might be more important for the spread of complex contagion.

This corpus of literature is concerned with measuring the strength of ties is of interest as it may affect how information flows (Bond et al., 2012; Weng et al., 2013). In the literature we identified, there is interest in the relationship between the organisation or shape of networks and transmission of different kinds of contagion. Weng et al. (2013) explore how memes move through social media networks. They find that it is possible to predict which memes might go viral on social media by examining their early patterns of diffusion. They also say:

> *Our method does not exploit message content, and can be easily applied to any socio-technical network from a small sample of data. This result can be relevant for online marketing and other social media applications. (Weng et al., 2013: 4)*

From this quotation is is possible to infer at least two ways to treat contagion – whether the content of the message or the structure of the contagion matters, or not. The advantage claimed above is that a method excluding content can be applied to any contagion in any network. However, this content-agnosticism is identified as a potential issue by some writers because interactions between a transmissible entity and a node are not known: “recipient selection and message content may be important aspects of influence and should therefore be estimated in future experiments” (Aral and Walker, 341). The other way of treating contagion, focusing on the content of messages, appears to be more connected to media studies and political studies. Here, the content of information and messages is treated important and may change the behaviour people, which may be detectable digitally as well as in non-digital domains (examples are Bail et al., 2018; Farrell, 2016; Procter et al., 2013; Vargo et al., 2018).

Interdisciplinarity

In contrast to the mathematical-computational disciplinary connections to *network cartography*, the disciplinary connections and influences in *tracking influence & transference* are more varied[6]. The studies of viruses tend to be influenced by biological and health sciences, and statistical physics (Wang et al., 2016). Innovation and business management are also important – there is a notable interest in marketing messages and product diffusion/adoption through networks (Aral et al., 2009; Aral & Walker, 2012). Finally, political science and media studies are also influential, in understanding the spread of certain messages, particularly through social media, as well as their potential effects on behaviours such as voting.

Aims

Tracking influence and transference is primarily concerned with showing how entities spread through networks. Authors claim that "peer effects are empirically elusive in the social sciences" (Aral & Walker, 2012, 337) and that many disciplines cannot address whether peers influence one another in terms of education, health, financial position, and so on. The two main aims are to: (i) track how an entity (of influence) spreads through a network and (ii) to predict how it might spread. This is claimed to have implications for interventions in public health, marketing, business and economics, politics, and media (Funk et al., 2010; Wang et al., 2016). Researchers may also seek to encourage the spread of entities, such as products or messages, and thus understanding different contagion in different kinds of networks have the potential to be used to develop marketing or political messaging strategies.

Categorising and correlating digital traces

The final theme in the literature appears to be less similar to *network cartography* and *tracking influence & transference* than they are to each other. The first two categories are interested in identifying topographical and/or causal descriptions of networks. In contrast, this strand of literature is more interested in how digital data in one domain corresponds to that in another domain. It may include causal claims, but may be about indicators for change, as well as about demonstrating the special qualities of computational approaches. It is primarily interested in statistical relationships across domains, usually between social media and another domain like markets, health or individual characteristics. It is therefore interested in relating features across different contexts.

Examples of research

We offer three kinds of examples to give a sense of the range of contributions to this category. Our intention is not to be comprehensive, but rather indicate the variability within the category.

- The first class of examples uses digital data to infer emotions and correlates public mood with data about other phenomena.
- The second class uses digital data to quantify activity and uses that information to correlate with other phenomena.

[6] This may, of course, be a feature of sample size.

- The third class is interested in classifying aggregate population data and using that to infer information about individual cases, particularly in relation to psychological and personal traits. (We discuss this further in an additional section.)

The first class is exemplified by *Twitter mood predicts stock market* (Bollen et al., 2011). The authors analysed approximately ten million tweets from 2.7 million users, including only those tweets that were explicitly related to emotions e.g. including the phrases “I feel” or “I am”, among others. They used two methods to ascertain mood. One, they used OpinionFinder to measures positive and negative sentiment from tweet text. Two, they used GPOMs (Google Profile of Mood States) which makes a classification of mood along six dimensions: calm, alert, sure, vital, kind and happy. The authors find that the *happy* dimension of GPOMs most closely correlates with *positive* on OpinonFinder. They correlate these dimensions with the Dow Jones Industrial Average. They claim that changes in mood happen 3 or 4 days before changes in the stock market. They claim that the “rather simple” analysis of textual sentiments can predict the stock market (Bollen et al., 211: 7). However, the predictive dimension is the GPOMs measure of *calm*, and not any of the other dimensions under study.

The second class of literature quantifies internet activity to predict, for example, the stock market through search behaviour (Curme et al., 2014) and box office receipts using editing activity on movie pages in Wikipedia (Mestyán et al., 2013). They claim that “since the methods presented here are independent of the language of the medium, they can be easily generalized to other languages and local markets” (Mestyán et al., 2013, 4). Curme et al. (2014) use 100 semantic topics and 30 words in each topic, then select 55 topics for study. They find that certain topics of internet searches related to US politics and business relate to stock market moves:

> *Our results provide evidence that for complex events such as large financial market moves, valuable information may be contained in search engine data for keywords with less-obvious semantic connections to the event in question. Overall, we find that increases in searches for information about political issues and business tended to be followed by stock market falls. (Curme et al., 2014: 11604)*

The third class relates large-scale digital data to infer information about individuals. Reece & Danforth (2017) use Instagram photos to claim they reveal predictive markers of depression. They extracted features (e.g. pixel averages; whether photographs contained human faces etc) of over 40 000 photographs from 166 individuals, 71 of whom were diagnosed with depression. About 13 000 randomly selected photos were used to develop the machine learning classifier. The selection of photos was manually rated on a 0-5 scale by workers on Amazon Mechanical Turk according to whether photos were interesting, likable, happy or sad. Then, 100 photos were analysed for each participant. For those who had been diagnosed with depression, the most recent photos preceding diagnosis were used, and the most recent 100 photographs for healthy participants. They claim their machine-learning classifier showed that it could positively identify photos taken by people diagnosed with depression with a higher rate of accuracy than clinicians. However, their pre-diagnosis classifier found only about a third of people subsequently diagnosed, which was not as good as clinicians. Only happiness and sadness, not likeability or interestingness, were predictors, such that “depressed participants’ photos were more likely to be sadder and less happy” (Reece & Danforth, 2017, 8). However, they note that the Amazon Mechanical Turk workers did not appear to associate darker, bluer, greyer pictures with depression, while academic research has previously noted the link. They also hedge by saying that depression is complex and that common and clinical meanings, as well as interactions with other conditions, are diverse and widespread.

Theory and concepts

Some of the central concepts in this category are personality, language, traits, attributes, moods and public mood, and it would seem that most of these are interpreted using a psychological perspective that focuses on the relation between individuals and aggregate populations. This is distinct from the *cartography* and *tracking* approaches that define individual nodes in terms of their relationship activity realised predominantly as communicative-transactional connections.

All three classes of examples show some evidence of an interest in understanding meaning as well as approaches that focus more on measures of activity. *Categorising and correlating digital traces* is often about analysing and classifying the content of messages e.g. Twitter, Facebook messages and Instagram photos. However, some papers are more interested in other digital markers, such as the number of revisions on a Wikipedia movie entry or quantifying search behaviour (Curme et al., 2014; Mestyán et al., 2013).

Interdisciplinarity

(Social) Psychology is the dominant discipline in this literature. Most of the contributions reference and use psychological concepts. Michael Macy, however, is a computational sociologist and studies moods in different global zones or cultures (Golder & Macy, 2011). Physics, business and economics are also important connections (e.g. Curme et al., 2014).

Curme et al. (2014) and Reece and Danforth (2017) make use of Amazon Mechanical Turk to pay workers a small sum to do the classification work. This does raise questions of ethics in collaborative research, but also of the relation between human meaning making and algorithmic training in analysing social data.

Aims

Categorising and correlating digital traces approaches demonstrate interest in categorising digital data from one domain often followed by showing correspondence to another domain. Typically, the methods involve automated feature extraction, algorithmic model building and correlation. Broadly, the latter fit into two categories. One of which is correlating general-to-general features across two domains, such as Twitter and stock markets (Bollen et al., 2011), and the other explores the relation between populations and individuals, such as links between language usage and personality traits (Schwartz et al., 2013).

A significant aim is to demonstrate the power of computational analytics when applied to social, economic and psychological phenomena. This can be in terms of using publicly available digital data to detect personality traits or to indicate future (predict) changes in other phenomena like economic trends. Titular claims include publications like *Private traits and attributes are predictable from digital records of human behaviour* (Kosinksi et al., 2013), *Twitter mood predicts the stock market* (Bollen et al., 2011) and *Early Prediction of Movie Box Office Success Based on Wikipedia Activity Big Data* (Mestyán et al., 2013). The possibility of generating predictions is therefore a key contribution that approaches in this category seek to make. There is evidence of attempting to contribute to various social science fields, notably economics and psychology.

Another temporal aspect to prediction in that several papers are interested in "nowcasting" as they search for real-time indicators of change in other volatile domains.

> *Researchers have acknowledged the limitations of this methodology [retrospective reports from university students] but have had no practical means for in situ real-time hourly observation of individual behavior in large and culturally diverse populations over many weeks… That is now changing. Data from increasingly popular online social media allow social scientists to study individual behavior in real time in a way that is both fine-grained and massively global in scale (11)[7], making it possible to obtain precise real-time measurements across large and diverse populations (Golder & Macy, 2011, 1879)*

This point speaks to our first footnote, on page 2, where we note our feeling that real-time analysis of social phenomena was arguably a distinctive feature of CSS. Finally, several articles make a claim that computational analysis is better than a human at making inferences or predictions in a particular domain (Reece & Danforth, 2017; Wang & Kosinksi, 2018). In other words, in this space, some researchers set up a competitive relationship between humans and computers, rather than setting up a comparison between research methods that we found in *network cartography*. Thus, the claimed potential impacts of CSS are bolder here than in *network cartography*.

Controversy and claims: A comment regarding the work of Michal Kosinski

In making sense of this space, we feel we need to make a note concerning one of the prominent contributors as they have been controversial both within and beyond academia. Of the ten papers in this section, Michal Kosinski is an author on four of them, suggesting he's an influential scholar in the field (Kosinski et al., 2013; Schwartz et al., 2013; Wang & Kosinksi, 2018; Wu et al., 2015). The works typically fit into the third class mentioned above, where *categorising and correlating* relates to inferring individual characteristics and features from digital data. We include the following reflections because they highlight how CSS methods and approaches are connected to debates in media, and indicate the ethical and epistemological challenges that may emerge with CSS work.

CSS methods have featured in popular media. As we explained in the methods section, Meckin recognised Kosinksi's name from the documentary *Secrets of Silicon Valley* (Bartlett, 2017). Bartlett was mainly investigating the promissory rhetoric of US technology firms and the reality of their operations. Bartlett interviewed Michal Kosinski and subjects his own Facebook profile likes to analysis. Kosinski's computer-based prediction is that Bartlett is 84% likely to be open-minded, liberal and artistic and 40% likely to have no religion. But, if he were religious, there's a 38% chance he would be Catholic. Barlett responds with incredulity, demonstrating disbelief as he says he was raised a Catholic. The program only briefly flashes the probability values on screen, and does not discuss them. The program does, then, contribute to the "mythology"

[7] This citation is Lazer et al. (2009).

(boyd and Crawford, 2012) that computation of social media data is somehow special has a capacity to “see into” private lives.

Indeed, privacy is one of Kosinski’s main stated concerns (in his writing). His work claims he is trying to highlight the exposure people subject themselves to as they create digital footprints:

> *On the other hand, the predictability of individual attributes from digital records of behavior may have considerable negative implications, because it can easily be applied to large numbers of people without obtaining their individual consent and without them noticing. Commercial companies, governmental institutions, or even one's Facebook friends could use software to infer attributes such as intelligence, sexual orientation, or political views that an individual may not have intended to share. One can imagine situations in which such predictions, even if incorrect, could pose a threat to an individual's well-being, freedom, or even life. Importantly, given the ever-increasing amount of digital traces people leave behind, it becomes difficult for individuals to control which of their attributes are being revealed. For example, merely avoiding explicitly homosexual content may be insufficient to prevent others from discovering one's sexual orientation. (Kosinski et al., 2013, 5805).*

However, this position is complex given the publications he’s involved in, his links to Cambridge Analytica and, just a paragraph earlier, the fact the authors list a set of “positive” uses, predominantly related to upselling and marketing strategies. They close by saying:

> *There is a risk that the growing awareness of digital exposure may negatively affect people's experience of digital technologies, decrease their trust in online services, or even completely deter them from using digital technology. It is our hope, however, that the trust and goodwill among parties interacting in the digital environment can be maintained by providing users with transparency and control over their information, leading to an individually controlled balance between the promises and perils of the Digital Age. (Kosinski et al., 2013: 5805).*

They invoke psychological concepts of trust and goodwill in their hope for a utopian digital age, suggesting it’s about a relation between individuals and digital technology. They leave out legal, institutional or socio-economic concepts, like regulation.

A more recent paper specifically addresses the issue of sexuality and makes that claim that *Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation From Facial Images* (Wang & Kosinksi, 2018). This paper describes how facial images on Facebook have been analysed and claims the authors can predict whether a Facebook user is hetero- or homosexual by analysing their facial structure. Sexuality had been inferred from profile statuses where users state the partners they are looking for. Facial structure was determined using feature extraction of users’ faces from images. The authors claim their work supports the view that pre-natal hormone exposure is a link between facial morphology and sexual orientation.

This study has been criticised from different angles, both online and in academic work:

1. The use of images without specific consent is unethical and may expose participants and other social media users to danger[8].
2. Another commentary[9] acknowledges the possible ethical and methodological and even assuming those issues are in fact unproblematic, they still find issues with the research logic:
    a. Claiming software could detect things humans could not was an unfair test – the software was specifically trained whereas the humans were not.
    b. Claiming the results back up "pre-natal hormonal exposure" is also not supported by evidence as it is not certain the algorithm uses "facial structure" in its analysis as opposed to cues like lighting, styling and so on. A slightly different take on this critique was made by blurring the faces (Leuner, 2019) which showed arguably little difference in accuracy.

Kosinski's work therefore highlights an ethical dimension to CSS, which may need more consideration. On the one hand, Kosinski claims to be highlighting issues of privacy and security, yet simultaneously making unsupported arguments and over-claims about the capabilities of the technology. We return to these issues in the discussion section.

Critical literature

The critical literature is mostly different from the substantive-technical literature in that it is generally concerned with describing and commenting on the field of the CSS. The papers contain varying levels of technicality, but do not tend to present new research or methods in and of themselves. We present three groups of literature: *critical-internal, critical-external* and *tangential*.[10]

One paper stands out as not fitting neatly into our categories. Borra and Rieder (2014) present a method for collecting and analysing tweets and discuss the contextual epistemological considerations of their technology development. The paper presents a technical method for data extraction and analysis but, importantly, discusses key implications for knowledge production in terms of the ontology produced by Twitter. The acknowledgements mention that scholars like Noortje Marres, David Moats and Emma Uprichard have commented on the work, which suggests the argument has been developed in dialogue with members of the Science and Technology Studies community. Due to its epistemological content, we have therefore decided to include it as a critical paper, from the inside, rather than as a technical paper.

Nine papers offer internal critique, and together offer a mix of critical methods evaluations, promissory and revolutionary rhetoric, and application of CSS to specific areas. Just two papers are in our critical-external category. Lastly, three papers that came up in the searches turned

[8] https://www.theregister.com/2019/03/05/ai_gaydar/
[9] https://www.callingbullshit.org/case_studies/case_study_ml_sexual_orientation_original_version.html
[10] The "internal" and "external" categories, which were generated by the content of the articles as well as the affiliations of the authors, might be regarded as problematic because they could be interpreted as a reification of a boundary between CSS practitioners' concerns and may, therefore, be interpreted to imply more or less legitimacy in terms of the arguments being advanced. Neither of these interpretations are intended here. The terms "internal" and "external" are used in this report because of their convenience in highlighting the different strands of literature by author affiliations and that the two categories of literature demonstrate different concerns.

out, on close reading, to contain points that intersected with CSS, but the central point of interest was different. These are described below as they show important connections between CSS and other discussions, debates and developments in the academe. These papers help inform a future exploration of interdisciplinarity.

Critical-internal

The nine papers reviewed here contain at least two themes; the publications (i) review and discuss the challenges of CSS and (ii) discuss the methods used within CSS. Sometimes these two themes are discussed in parallel in the same sections, and sometimes they are treated separately.

There are different ways that scholars divide up the methods space in CSS. Batrinca & Treleaven, divide them into two: 1) "computational science techniques", by which include machine learning and "complexity science", which can be deployed for data mining and simulation and 2) sentiment analysis, which is the identification of particular meanings in data (Batrinca & Treleaven, 2015; 103-105). Alternatively, in the case of methods for decision support tools, Wang et al. (2016) discuss five different areas of techniques, including mathematical and statistical approaches, data analysis (regression, clustering etc. using machine learning, data mining and other "artificial intelligence"), visualisation methods, cloud computing, and fuzzy sets and fuzzy systems (Wang et al., 2016: 753-755). And then, there are also different enabling technologies that allow batch, stream and hybrid (a combination of both) data processing (Wang et al., 2016: 756). Although authors analyse the methods space in different ways, there appear to be themes in terms of the issues they discuss, which we present below.

Challenges

As part of the internal critique literature there is frequently reference to the potential of CSS and big data analytics. Discussions of big data analytics in the context of CSS highlight features such as velocity, volume and variety but also including complementary characteristics such as veracity, variability and value (Wang et al., 2016). There appears to be attempts to capture something new about emerging computational possibilities, and that they present different kinds of methodological issues when it comes to collating and analysing data.

Lazer et al.'s (2009) central concern was in the potential for the social sciences that was afforded by the growing volumes of data and increases in analytic power. They present a conundrum that, on the one hand, CSS is potentially transformative of our understanding of society but that, on the other hand, social science is somewhat behind other sciences (such as biology and physics) in its adoption and support of high-power computational methods. They argue further that, due to the nature of data production and the ownership of computational infrastructure, private companies may keep data and analyses secret and universities are not set up to support infrastructure or collaboration, nor to train people with skills needed in this research space. Thus, in its formation, CSS was presented in terms of opportunities and challenges.

It is often in this context that authors present challenges as barriers or impediments to the emergence of the field. For instance, Siverajah et al. (2017) categorise the challenges into three areas in which:

- *Data challenges are the group of challenges related to the characteristics of the data itself*
- *Process challenges include all those challenges encountered while processing the data*
- *Management challenges tackle e.g. the privacy, security, governance and lack of skills related to understanding and analysing the data* [sic]

(Siverajah et al., 2017: 265)

Comparably, Batrinca and Treleaven (2015) categorise considerations as those relating to “data”, those relating to “process”, and those relating to “facilities” (by which they mean the organisational, material and knowledge production infrastructures). Thus, analysis of large social data sets generates challenges in terms how to access, handle, wrangle, store and process data to make knowledge and we discuss some of these below.

Access and security. One of the key concerns to realising the potential of CSS regards enabling access to data (Batrinca & Treleaven, 2015; Lazer et al., 2009; Chang et al., 2014). This is discussed in relation to data production being owned by private companies, for whom licencing the data may be source of revenue or give competitive advantage. This means that access can be prohibitively expensive (Borra & Rieder, 2014). Access to quality data can also be problematic, as there is a high level of production of data from “uncontrolled sources” that it may be difficult to check quality (Conte et al., 2012: 332). There are also concerns that data is often produced by private companies with biases towards particular uses, such as informing customer insight and marketing (Borra & Rieder, 2014).

Organisational. Several papers highlight concerns around institutional support and collaboration. For instance, the importance of developing inter-organisational relationships for value-co-creation (Chang et al., 2014) and the need for interdisciplinary engagement is mentioned, to ensure good quality research and access to data, by which the authors typically mean cooperation between academics and data-producing businesses (Chang et al., 2014; Lazer et al., 2009). Technical infrastructure, collaboration and support for training are all organisationally related issues.

Process. Chang et al. (2014) outline “practical considerations” for analysing social big data, which include data collection, reliability and cleaning; data acquisition and security. Validation of massive datasets can also be difficult as they may come from different sources and have inconstencies (Conte et al., 2012). There are significant concerns of the use and understanding of statistics to make claims (Conte et al., 2012; Lyons, 2011; Siverajah et al., 2017). Jungherr et al. (2012) show that Tumasjan et al. (2010) are opaque about their methods as they ran a comparable data collection and argue the authors of the earlier paper should have included the their rationale for the exclusion of a political party and their rationale for the date range. Jungherr et al. (2012) show variability across different date ranges, meaning the findings of the original paper are not generalizable beyond the apparently arbitrary dates chosen.

Borra and Rieder (2014) claim an analysis of up to 100 million tweets can be conducted on a single Linux machine, but larger datasets would need more infrastructure, perhaps through distributed computing. They provide a method for capturing and analysing tweets and offer 3 different modes of sampling for social researchers: 1% sample of tweets, keyword sampling, and a representative panel of 5000 accounts. Platforms, such as Twitter, structure the organisation and social possibilities, making possible certain kinds of entities and interactions (e.g. tweet length, hashtags, retweets), which means that these need to be taken into account when analysing the data and making claims.

Disposition of the researchers. There is an acknowledgement of "big data hubris" (Lazer et al., 2014). The authors point out, with respect to Google Flu Trends (GFT), that:

> *the odds of finding search terms that match the propensity of the flu but are structurally unrelated, and so do not predict the future, were quite high… in short, the initial version of GFT was part flu detector, part winter detector."*
>
> *(Lazer et al., 2014: 1203).*

The challenge here, then, is belief among researchers that CSS methods are superior to existing methods and that reflection about what new analytics can actually add is warranted.

Directions

The critical internal literature contains proposals for future research in CSS. Using the case of flu detection, Lazer et al. (2014) suggest that the improvement gains to be made on an already well-developed system are minimal and that it may instead be prudent to focus on areas where greater impact could be made. As Chang et al. (2014) discuss, this might include taking more account of contextual awareness of consumer decisions, increasing personalisation in marketing, focus on events in which decisions change, and exploring societal level analysis. Chang et al. (2014) also suggest "complementary research on data privacy" could be conducted. Finally, Conte et al. (2012) focus on emergence, suggesting emergent behaviour, social groupings and networks, and institutions, could all be studied.

This selection of papers therefore tends to deal with methodological problems within CSS, and conceptualises many of the challenges as separate from the analytic technology and thus more social in nature (e.g. to do with people, organisations, skills, etc).

Critical-external

Just two papers fit in the critical-external category. These cover a range of epistemic and ethical issues related to uses of social big data, particularly regarding knowledge claims and social use. boyd and Crawford (2012) is one of the most highly cited papers regarding CSS and big data analytics. They point out epistemological issues, arguing that big data changes the definition of knowledge. They suggest the need for epistemological modesty as claims to accuracy, objectivity, the importance of size, and the generalisability of findings, all need nuancing. Barocas and Selbst (2016) offer related points, from a legal perspective. They point out the processes of discrimination in what they call data mining, at first showing this in the nonpejorative sense, but going on to develop this argument to show that the discriminatory practices of data mining can produce illegal discriminations against people and communities, both inadvertently and deliberately. Thus, claims about the representativeness and generalisability of findings need to be understood in context.

The use of data presents ethical challenges – just because it is accessible does not mean that one should use the data for certain ends (boyd & Crawford, 2012). Data may also be generated unevenly across society, so that those who produce more data may reap more rewards of automated tracking. Barocas and Selbst (2016) use the example of an ingenious mobile phone app that can detect potholes in roads by sensing driving motion bumps, and then shares the data with highway maintenance. However, it demonstrates that where mobile phone ownership is lower (e.g. more deprived areas) potholes will be detected less frequently and the roads generally in a worse state of repair (Barocas & Selbst, 2016). Claims to comprehensiveness need to be understood in terms of who has access to technology, what they use it for, and how discrimination in data mining can reproduce or worsen societal inequalities. Indeed, the idea that CSS offers a more fine-grained and complete picture of social interactions is likely only related to those who regularly access digital devices and, perhaps, use them in particular ways.

The two external critiques, therefore, show how CSS data analytics are connected to other domains in society and show how CSS results may miss, skew or be more contextualised than claimed. What is also important is that they argue CSS methods should not be treated as just an academic endeavour and that academic techniques can have implications when tools are applied and developed elsewhere.

Tangential critical papers

Three publications indicate the intersections of CSS with other scholarship that relates to the digitalisation of society. All three papers emerged in the search of papers citing Lazer et al. (2009).

Kitchin and Dodge (2011) argue that, to date, studies of software and code have explored temporal issues and the ‘outputs’ of code e.g. technology. Their argument is that there is much to be gained from a spatial analysis of code, such that software codes for things beyond software. They argue that codes produces, in increasing complexity, coded objects, coded infrastructures, coded processes and coded assemblages. These coded entities afford particular actions which help sustain the entities. They show, for instance, that some spaces are dependent on code to function, such as airport check-in areas, while others are augmented by code, such as the use of a presentation in a lecture hall. The point here is that code, seemingly an immaterial entity, can be considered in relation to the way code makes particular spaces. Taking their argument specifically to CSS, we would need to interrogate what analytical spaces are produced through particular instantiations of CSS code and to explore the interdependency of those analytical spaces with formal computer code.

One paper in this category was an introduction to a special issue of *Organisation Science* (Yoo et al., 2012). This offers a critical introduction to scholarship on digital innovation from an organisational perspective. Their analysis is predominantly concerned with the features of digital technology. Their thematic introduction covers the generativity of digital technologies in social and organisational settings (media platforms and automobile manufacturing), the complexities and risks arising from digital convergence, and serendipity in spaces of digital innovation. They argue that new entities are emerging and producing a digital materiality e.g. microchipped running shoes (Yoo et al., 2012). The special issue therefore relates code and innovation, connecting management, organisation, and innovation studies; this indicates the broader context of CSS, building further on the points made briefly by Lazer et al. (2009). This is related to the arguments made be Kitchen and Dodge (2011) that software code, and the technologies that run on it, are productive of other entities.

Finally, Helbing (2013) is predominantly interested in advocating for complexity science in understanding the contemporary nature of risks. He argues that the world is increasingly networked and interdependent. He cites Lazer et al. (2009) when discussing the use of ABMs in simulating learning in large social groups. Helbing is also co-author of Conte et al.'s (2012) *Manifesto of computational social science* and indicates the role of CSS in informing understandings of the complexities of global risks.

Concluding Discussion

This review identified 41 articles related to computational social science (CSS) and validated findings with a further 30 papers. The review has also mentioned several other articles and commentaries in making sense of the literature, and the ways these were included is described in the methods. We will discuss briefly some of the issues raised in the review. First, we recap the top-level findings and revisit some of the definitions in the introduction. Second, we suggest some possibilities and motivations for engaging with CSS practitioners. Last, we outline some potential themes around which engagement could be organised or could explore.

First, our review suggests three broad categories of approach within CSS: *network cartography, tracking influence & transference,* and *categorising and correlating digital traces*. This is a different set of categories to the five methods outlined in the introduction (Cioffi Revilla, 2010). A rough resketch, though, might see the five methods aligned to particular approaches. Thus,

- Automated information extraction is used in all three approaches;
- SNA is associated with *network cartography* and *tracking influence & transference*;
- complexity modelling and social simulations modelling (ABMs) are also mostly affiliated with *tracking influence & transference;*
- Finally, geospatial analysis is not found in our review.

However, we would suggest that a rigorous analysis of the different techniques and methods might identify other methods that have come to the fore and help explain why geospatial analysis is underrepresented. Further work would be needed to explain whether this is because (i) geospatial analytics has not been developed as part of CSS, (ii) practitioners do not affiliate/identify with CSS, or (iii) other reasons or combinations thereof.

Secondly, the outcome of this review broadly aligns with boyd and Crawford's tripartite definition of big data: computing technology, developing patterns for social claims, and epistemic mythology (boyd and Crawford, 2012). There is evidence of the idea that increased computational power, working on digital social data, can generate patterns and infer knowledge about social phenomena. Furthermore, many of the papers claim improved scale, comprehensiveness or generalisability with CSS analyses. The promissory aspects of technosciences have been discussed elsewhere (e.g. Borup et al., 2006), but the specific promises of CSS and what they enable may warrant investigation.

Finally, the literature seems to suggest that CSS is not a fully-fledged and recognised approach. Many of the papers can be read as attempts to prove the use of computational analysis to address social science questions, rather than as reports on an accepted technology. Thus, the repeated comparisons to other methods, or to human capacity, seem to be there to convince readers that CSS has something distinct, and better, to offer. Again, further work, including engagement with practitioners, would be needed to explore these points further. This brings us to one of the central aims of the report, which was to explore CSS as a potential space for NCRM (and UK social science in general) to engage in.

Interdisciplinary engagement

There seem to a range of possibilities for following up this review, some of which are outlined below. From the analysis so far, it is possible to draw out several points for further exploration:

1. **What are the current and emerging methods, and what training is needed?**
   The review has suggested that there are identifiable approaches in CSS making use of multiple methods, techniques and processes. These include data extraction, feature extraction, network analysis and so on. Engagement (and specifically training needs analyses) might aim to and identify the methods that are challenging to learn and/or correctly deploy. Engagement might also explore the "fates" of methods, including why geospatial analysis was not found in our review and the hint from our validation step that there may be a shift of interest towards *categorising and correlating* approaches over *cartography* and *tracking.* Answers here have potential implications for future training and infrastructure provision.
2. **In what areas might NCRM help form knowledge exchange networks?**
   NCRM has a range of possibilities for structuring engagement. One of these is facilitating the formation of networks around methodological themes or areas of interest. These might result in, for example, virtual meetings or a series of workshops. There are different substantive-technical approaches and disciplinary interests across CSS. Network formation may therefore be around particular areas of interest, such as moods, news, markets, innovation or organisations. The review also indicates that the three substantive technical areas: network cartography, spreading/contagion and correlating and categorising traces may form useful points of interest for the research community.
3. **What about theory?**
   CSS is, to some degree, lacking theory (see boyd & Crawford, 2012). However, others show that the philosophy of CSS or Big Data is under consideration (Fuchs, 2020; Siverajah et al., 2017). The review has shown the theory of weak ties to be important in studies of networks and contagion (see *tracking influence and transference*).Furthermore, the *categorising and correlating* aspects make use of psychological theory. However, perhaps formal engagement in the philosophy, history and sociology of science might support the production of socially attuned claims?
4. **Ethics and Responsible innovation**
   There is a literature on the governance of research and innovation. Recent efforts (e.g. *Journal of responsible innovation*) have sought, in part, to challenge the "modest witness"

argument put forward by some scientists that they are not responsible for the findings and technologies they discover and that others then put to questionable uses. One responsible innovation (RI) framework arising from this literature encourages techno-scientific and innovation communities to adopt the AREA (*anticipate, engage, reflect and act)* Framework[11] (Stilgoe et al., 2013).

In the CSS literature, there is clear evidence of practitioners anticipating particular challenges and issues. However, the other dimensions of the AREA framework (engage, reflect and act) could be explored by facilitating engagement between practitioners, stakeholders, members of the public, and so on, to discuss about how best to develop CSS methods with the aim of making them socially robust and beneficial to society. There is scope to explore RI thinking with the CSS community, and with wider data analytics communities, to ensure that CSS research and innovation embeds understandings from other disciplines and other sectors of society.

Themes for consideration

There are several important epistemological themes that we want to comment on further and that may warrant further exploration as to their implications.

Context

The different approaches and literatures provide differing perspectives on the importance of context in analyses and what is valued in CSS. The *network cartography* and *tracking influence and transference* approaches are more orientated towards description of social phenomena and the production of tools, techniques and methods for those descriptions. Their primary aim is to offer a novel analysis of social organisation and social networks. The *categorising and correlating digital traces* approach is more orientated to correspondence in that it deals primarily with the degree of matching of phenomena in different spaces. In these terms, the *network cartography* and *tracking influence and transference* appear to produce their own context and concentrate on findings in that space. It is, then, analytical in nature and seeks to identify component parts and their relations. The *categorising and correlating* approach, in contrast, compares digital data across contexts. Understanding the production of analytical contexts, that is how they are shaped through code and computation, would be an interesting space to explore further. A thematic engagement here would contribute to some of the work already begun that explores how platform code produces particular contexts and social ontologies (Borra & Rieder, 2014; Kitchen & Dodge, 2011) and would address some of the epistemic issues, such as how comprehensiveness is shaped, identified as a problem by the critical-external literature we reviewed (boyd & Crawford, 2012; Barocas & Selbst, 2016).

Meaning

The extent to which approaches take meaning into account varies throughout the literature. *Network cartography* appears to be the least interested in meaning, although, Szell et al. (2010, 13637) suggest their approach does not treat relationships as the “volume of information exchanged”, but includes analysis of relationship content. *Influence & Transference* shows some interest in language and images, particularly in relation to political messages, but many

[11] https://epsrc.ukri.org/research/framework/area/

studies also find quantitative digital proxies for phenomena, such as the frequency of webpage edits rather than the content. The *Categorising and correlating* approach appears most interested in meaning and interpretation, with a major focus on mood and personalities. There is significant opportunity to explore issues and decision making in regard to what needs to be, or can be, included in particular analyses.

Relations to and with human capabilities

A significant framing of the way CSS works is to compare directly to humans, either within the study (Eagle et al, 2009; Wang & Kosinksi, 2018) or in other literature (Reece & Danforth, 2017). Wang et al. (2016) discuss the role and relation of humans in the interpretation and decision-making process. Furthermore, several studies recruit workers through Amazon Mechanical Turk (Curme et al., 2014; Reece & Danforth, 2017; Wang & Kosinksi, 2018). This use of low paid workers raises questions about ethics, equity and ownership in research. Exploring the relations between computational and human capabilities and capacities in social research may be a generative line of inquiry.

Combining methods

Another engagement theme could be the extent to which methods are mixed or combined. Eagle et al. (2009), for instance, benchmark their computationally-derived findings with a survey. Most of the studies that we have reviewed combine different computational techniques into analytical workflows, but do much less to clearly think about how CSS methods might be integrated with other methods. This speaks partly to the point made above, where humans are recruited to do upstream classification work – a process that needs much more exploration in terms of how people are actually doing this work on platforms like Amazon Mechanical Turk.

On this issue, though, we want to again mention two papers from the validation exercise. These are papers led by Ramit Debnath that contribute to the field of energy studies (Debnath et al., 2020; Debnath et al., 2021). The papers are particularly interesting because of the combination of methods from both computational and interpretative qualitative analyses. The projects combine CSS methods and focus groups in an integrated methodology oriented to topic modelling. This perhaps indicates the possibility of combining multiple methods within one investigation and thus further engagement to understand how methods might be combined and integrated may be profitable for creating projects that address interdisciplinary or societal problems.

Explanatory resources

The relation between psychological conceptual resources and digital traces is an important dimension that could be explored further. The *approaches* tend focus on analysing digital traces of people's activity[12] (for a discussion of what this produces, see Ruppert et al., 2013). The *Cartography* and *tracking* approaches focus on movement or transactional connections, which many of the authors refer to as "behaviour", and the *tracking* and *categorising and correlating* approach can focus on the products of cultural practices (photography, social media messages, etc). The *categorising and correlating* approach, in particular, demonstrates more interest in concepts like mood, features and traits of individuals. However, there is evidence of individualistic interpretations of mood in the other approaches, too. In *cartography* there was

[12] "Activity" is our term; the literature often refers to "behaviour"

discussion of satisfaction while in *tracking* there was discussion susceptibility to influence. Thus, psychology can be used as an explanatory resource for describing results. Psychology, in terms of processes in people's heads, can also be seen as a limitation or confounding factor (e.g. Aral and Walker, 2012). The potential ability to identify, statistically categorise and make claims about individuals has many ethical concerns but, crucially, means that limited explanatory resources are used to interpret results.

Exploring different conceptual underpinnings may change the shape of projects and the explanations arising from them. For example, thinking of the social as intersecting practices instead of aggregates of human individuals, may be one avenue of inquiry. Such a framework has a potential confluence with CSS since analyses influenced by theories of practice also focus on activity and action. Another related alternative may be to consider how different moods and traits are a consequence of the affordances of techniques and tools used in analytical procedures. In other words, this interpretative inversion emphasises that these phenomena are the outputs of analyses rather than the causes, which means they could not be used to explain phenomena because that would introduce circularity. Thinking through these alternative frameworks may be a way to develop CSS in different directions and integrate CSS with other debates and developments in the social sciences.

In meeting its aims, this report has generated an overview of CSS that can be used as provocations in conversations about the future training, methodological development and overall directions of CSS. This report has sought to offer insight into the emerging field of CSS by exploring the kinds of studies that scholars conduct, the methods they use, and the debates they engage in. CSS is a multifaceted research area, with contributions to scholarship across the social research spectrum. There are various challenges and questions that emerge with CSS methods, and these vary across the CSS space. The review suggests that there are methodological dimensions and that further analysis, engagement with scholars, stakeholders and with publics, might be useful to support the development of CSS and its integration in social research.

References

Alassad, M., Spann, B. & Agarwal, N. (2021) Combining advanced computational social science and graph theoretic techniques to reveal adversarial information operations. *Information Processing and Management,* 58(1): 102385.

Aral, S. & Walker, D. (2012) Identifying Influential and Susceptible Members of Social Networks. *Science*, 337 (6092), 337-341. DOI: 10.1126/science.1215842

Aral, S., Muchnik, L., Sundararajan, A. & Jackson, M. (2009) Distinguishing Influence-Based Contagion from Homophily-Driven Diffusion in Dynamic Networks. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (51), 21544-21549. https://doi.org/10.1073/pnas.0908800106.

Bail, C.A., Argyle, L.P., Brown, T.W., Bumpus, J.P., Chen, H., Fallin Hunzaker, M.B., Lee, J., Mann, M., Merhout, F. & Volfovsky A. (2018) Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* Sep 2018, 115 (37) 9216-9221; DOI: 10.1073/pnas.1804840115

Barocas, S. & Selbst, A. (2016) Big Data’s Disparate Impact. *California Law Review*, 104 (3), 671-732. DOI: http://dx.doi.org/10.15779/Z38BG31.

Batrinca, B. & Treleaven, P.C. (2015) Social media analytics: a survey of techniques, tools and platforms. *AI & Society* 30, 89–116. DOI: https://doi.org/10.1007/s00146-014-0549-4

Bollen, J., Mao, H. & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science,* 2 (2011), 1-8. DOI: 10.1016/j.jocs.2010.12.007.

Bond, R., Fariss, C., Jones, J., Kramer, A., Marlow, C., Settle, J. & Fowler, J. (2012). A 61-million-person experiment in social influence and political mobilization. [Research Letter]. *Nature*, 489, 295–298. DOI: https://doi.org/10.1038/nature11421.

Borra, E. and Rieder, B. (2014). Programmed method: Developing a toolset for capturing and analyzing tweets*. Aslib Journal of Information Management*, 66(3): 262–78. DOI: https://doi.org/10.1108/AJIM-09-2013-0094

Borup, M., Brown, N., Konrad, K., & van Lente, H. (2006). The Sociology of Expectations in Science and Technology. *Technology Analysis & Strategic Management* 18 (3–4): 285–298.

Botta F, Moat, HS, & Preis, T. (2020). Measuring the size of a crowd using Instagram. *Environment and Planning B: Urban Analytics and City Science*. 47(9):1690-1703. doi: https://doi.org/10.1177/2399808319841615.

boyd, d. & Crawford, K. (2012) Critical questions for big data. *Information, Communication & Society*, 15:5, 662-679, DOI: 10.1080/1369118X.2012.678878.

Boyd, RL, Schwartz, HA. (2021) Natural Language Analysis and the Psychology of Verbal Behavior: The Past, Present, and Future States of the Field. *Journal of Language and Social Psychology*. 40(1):21-41. DOI: https://doi.org/10.1177/0261927X20967028.

Carter EB, Carter BL. Propaganda and Protest in Autocracies. *Journal of Conflict Resolution.* 65(5): 919-949. DOI: https://doi.org/10.1177/0022002720975090.

Centola, D and Macy, M. (2007) Complex contagion and the weakness of long ties. *American Journal of Sociology*, 113(3): 702-734. DOI: https://doi.org/10.1086/521848.

Chang, R.M., Kauffman, R.J. & Kwon, Y.O. (2014) Understanding the paradigm shift to computational social science in the presence of big data. Decision Support Systems, 63 (2014), 67-80. DOI: https://doi.org/10.1016/j.dss.2013.08.008

Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J. -P., Sanchez, A., Nowak, A., Flache, A., San Miguel M. & Helbing D. (2012). Manifesto of computational social science. Eur. Phys. J. Spec. Top. 214, 325–346. DOI: https://doi.org/10.1140/epjst/e2012-01697-8

Curme, C., Preis, T., Stanley, H.E., & Moat H.S. (2014) Quantifying the semantics of search behavior before stock market moves. Proceedings of National Academy of Sciences. 111 (32), 11600–11605. DOI: https://doi.org/10.1073/pnas.1324054111

Christakis, N.A. & Fowler, J.H. (2013), Social contagion theory: examining dynamic social networks and human behavior. *Statistics in Medicine*, 32: 556-577. https://doi.org/10.1002/sim.5408.

Debnath, R., Darby, S., Bardhan, R., Mohaddes, K. & Sunikka-Blank, M. (2020) Grounded reality meets machine learning: A deep-narrative analysis framework for energy policy research, *Energy Research & Social Science,* 69: 101704. DOI: https://doi.org/10.1016/j.erss.2020.101704.

Debnath, R., Bardhan, R., Darby, S., Mohaddes M., Sunikka-Blank, M., Villaça Coelho, A. & Isa, A. (2021) Words against injustices: A deep narrative analysis of energy cultures in poverty of Abuja, Mumbai and Rio de Janeiro. *Energy Research & Social Science,* 72: 101892. DOI: https://doi.org/10.1016/j.erss.2020.101892.

Eagle, N., Pentland, A. & Lazer, D. (2009) Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences* 106 (36), 15274-15278; DOI: https://doi.org/10.1073/pnas.0900282106.

Elliot, M. (2020) Strategic Engagement Plan NCRM Phase IV. *NCRM unpublished internal document.*

Epstein, B. (2018) "Social Ontology", *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2018/entries/social-ontology/>. Last accessed: 7th May 2021.

Farrell, J. (2016) Corporate funding and ideological polarization. Proceedings of the National Academy of Sciences, 113 (1), 92-97. DOI: https://doi.org/10.1073/pnas.1509433112.

Feinberg, F., Bruch, E., Braun, M., Hemenway Falk, B., Fefferman, N., McDonnell Feit, E., Helveston, J., Larremore, D., McShane, B., Patania, A. & Small, M. (2020) Choices in networks: a research framework. *Marketing Letters* 31, 349–359. DOI: https://doi.org/10.1007/s11002-020-09541-9.

Fu, Q., Zhuang, Y., Gu, J., Zhu, Y. & Guo X. (2021) Agreeing to Disagree: Choosing Among Eight Topic-Modeling Methods. *Big Data Research*. 23: 100173. DOI: https://doi.org/10.1016/j.bdr.2020.100173.

Fuchs, C. (2021) Engels@200: Friedrich Engels and Digital Capitalism. How Relevant Are Engels's Works 200 Years After His Birth? *TripleC: Communication, Capitalism & Critique*, 19(1): 15-51. DOI: https://doi.org/10.31269/triplec.v19i1.1228.

Funk, S., Salathé, M. & Jansen, V.A.A. (2010) Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J. R. Soc. Interface*, 7: 1247–1256. DOI: http://doi.org/10.1098/rsif.2010.0142.

Gallagher, R., Frank, M., Mitchell, L., Scwartz, A., Reagan, A., Danforth, C. & Dodds, P. (2021). Generalized word shift graphs: a method for visualizing and explaining pairwise comparisons between texts. *EPJ Data Science*, 10: 4. DOI: https://doi.org/10.1140/epjds/s13688-021-00260-3.

Gillani, N., Chu, E., Beeferman, D., Eynon, R., & Roy, D. (2021). Parents' Online School Reviews Reflect Several Racial and Socioeconomic Disparities in K–12 Education. *AERA Open*, 7(1): 1-16. DOI: https://doi.org/10.1177/2332858421992344.

Golder, S.A. & Macy, M.W. (2011) Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science*, 333 (6051), 1878-1881. DOI: 10.1126/science.1202775.

González-Bailón, S. & De Domenico, M. (2021) Bots are less central than verified accounts during contentious political events. *Proceedings of the National Academy of Sciences*. 118 (11): e2013443118; DOI: https://doi.org/10.1073/pnas.2013443118.

Granovetter, M. (1973) The Strength of Weak Ties. *American Journal of Sociology*, 78:1360–80.

Guess, A., Barberá, P., Munzert, S. & Yang, J. (2021) The consequences of online partisan media. *Proceedings of the National Academy of Sciences* 118 (14): e2013464118; DOI: https://doi.org/10.1073/pnas.2013464118.

Helbing, D. (2013) Globally networked risks and how to respond. *Nature,* 497, 51–59. DOI: https://doi.org/10.1038/nature12047.

Jungherr, A., Jürgens, P. & Schoen, H. (2012) Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. "Predicting Elections With Twitter: What 140 Characters

Reveal About Political Sentiment." Social Science Computer Review. 2012;30 (2):229-234. doi: https://doi.org/10.1177/0894439311404119

Kejriwal M (2021) On using centrality to understand importance of entities in the Panama Papers. *PLoS ONE* 16(3): e0248573. DOI: https://doi.org/10.1371/journal.pone.0248573.

Kitchin, R., & Dodge, M. (2011). Code/space: Software and everyday life. Cambridge, MA: MIT Press.

Kosinski, M., Stillwell, D. & Graepel, T. (2013) Digital records of behavior expose personal traits. *Proceedings of the National Academy of Sciences*, 110 (15), 5802-5805; DOI: https://doi.org/10.1073/pnas.1218772110.

Lam, J., Li, V., Han, Y., Zhang, Q., Lu, Z. & Gilani, Z. (2021) In search of bluer skies: Would people move to places of better air qualities? *Environmental Science and Policy*, 117: 8-15. DOI: https://doi.org/10.1016/j.envsci.2020.12.012.

Lazer, D., Kennedy, R., King, G. & Vespignani, A. (2014) The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343 (6176), 1203-1205. DOI: 10.1126/science.1248506.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D. & Van Alstyne, M. (2009) Computational social science. *Science*, 323, (5915), 721–723. DOI: 10.1126/science.1167742.

Leuner, J. (2018) A Replication Study: Machine Learning Models Are Capable of Predicting Sexual Orientation From Facial Images. Masters in Information Technology (Big Data Science) Dissertation, University of Pretoria, Department of Computer Science, Pretoria, South Africa.

Lockhart, J. (2021) Paradigms of sex research and Women in stem. *Gender & Society*, 35(3): 449-475. DOI: https://doi.org/10.1177/08912432211001384.

Luke, S., Cioffi-Revilla, C., Panait, L ., Sullivan, K. & Balan, G. (2005) MASON: A multiagent simulation environment. *Simulation*, 81 (7), 517-527.

Lyons, Russell (2011) The Spread of Evidence-Poor Medicine via Flawed Social-Network Analysis. *Statistics, Politics, and Policy*, 2 (1): Article 2. DOI: 10.2202/2151-7509.1024.

Ma, J. (2021) Automated Coding Using Machine Learning and Remapping the U.S. Nonprofit Sector: A Guide and Benchmark. *Nonprofit and Voluntary Sector Quarterly*, 50(3): 662-687. DOI: https://doi.org/10.1177/0899764020968153.

Meckin, R. (2016) *Making research translatable: Articulating and shaping synthetic biology in the UK* [PhD Thesis]. Sociological Studies, University of Sheffield.

Meckin, R. (2020) Changing Infrastructural Practices: Routine and Reproducibility in Automated Interdisciplinary Bioscience. *Science, Technology, & Human Values*. 45(6):1220-1241. DOI:10.1177/0162243919893757

Mestyán, M., Yasseri, T. & Kertész, J. 2013. Early prediction of movie box office success based on Wikipedia activity big data. PloS One, 8 (8) (2013), Article e71226. DOI: https://doi.org/10.1371/journal.pone.0071226

O'Brien, S.O. (2010) Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research. International Studies Review, 12 (1), 87–104. DOI: https://doi.org/10.1111/j.1468-2486.2009.00914.x

Olsen, A., Calderón-Figueroa, F., Bidian, O., Silver, D. & Sanner, S. (2021). Reading the city through its neighbourhoods: Deep text embeddings of Yelp reviews as a basis for determining similarity and change. *Cities*, 110: 103045. DOI: https://doi.org/10.1016/j.cities.2020.103045.

Pérez-Verdejo, J.M., Piña-García, C.A., Ojeda, M.M., Rivera-Lara, A. & Méndez-Morales, L. (2020) The rhythm of Mexico: an exploratory data analysis of Spotify's top 50. *Journal of Computational Social Science,* 4: 147–161. DOI: https://doi.org/10.1007/s42001-020-00070-z.

Peris, A., Meijers, E. & van Ham, M. (2021) Information diffusion between Dutch cities: Revisiting Zipf and Pred using a computational social science approach. *Computers, Envrionment and Urban Systems*, 85: 101565. DOI: https://doi.org/10.1016/j.compenvurbsys.2020.101565.

Pierri, F., Piccardi, C. & Ceri, S. (2020) A multi-layer approach to disinformation detection in US and Italian news spreading on Twitter. *EPJ Data Sci*ence, 9: 35. DOI: https://doi.org/10.1140/epjds/s13688-020-00253-8.

Pozsgai, G., Lövei, G.L., Vasseur, L., Gurr, G., Batáry, P., Korponai, J., Littlewood, N.A., Liu, J., Móra, A., Obrycki, J., Reynolds, O., Stockan, J.A, VanVolkenburg, H., Zhang, J., Zhou, W. & You M. (2020) A comparative analysis reveals irreproducibility in searches of scientific literature. bioRxiv. Preprint. doi: https://doi.org/10.1101/2020.03.20.997783.

Procter, R., Vis, F. & Voss, A. (2013) Reading the riots on Twitter: methodological innovation for the analysis of big data, *International Journal of Social Research Methodology*, 16:3, 197-214, DOI: https://doi.org/10.1080/13645579.2013.774172.

Reece, A.G. & Danforth, C.M. (2017) Instagram photos reveal predictive markers of depression. EPJ Data Science, 6 (15), DOI: 10.1140/epjds/s13688-017-0110-z.

Ruppert, E., Law, J. & Savage, M. (2013) Reassembling Social Science Methods: The Challenge of Digital Devices. *Theory, Culture & Society*. 30(4) 22–46. DOI: https://doi.org/10.1177/0263276413484941.

Savage, M. (2010) *Identities and Social Change in Britain since 1940: The Politics of Method.* Oxford: Oxford University Press.

Savage, M. and Burrows, R. (2007) The coming crisis of empirical sociology. *Sociology*, 45(5): 885–889. DOI: https://doi.org/10.1177/0038038507080443.

Schimpf, C., Barbrook-Johnson, P. & Castellani, B. (2021). Cased-based modelling and scenario simulation for ex-post evaluation. *Evaluation.* 27(1):116-137. doi: https://doi.org/10.1177/1356389020978490.

Schoen, H. , Gayo-Avello, D. , Metaxas, P.T. , Mustafaraj, E. , Strohmaier, M. and Gloor, P. (2013), “The power of prediction with social media”, Internet Research , Vol. 23 No. 5, pp. 528-543. DOI: https://doi.org/10.1108/IntR-06-2013-0115

Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Dziurzynski, L., Ramones, S.M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M.E.P. & Ungar, L.H. (2013) Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLOS ONE,* 8(9): e73791. https://doi.org/10.1371/journal.pone.0073791.

Sivarajah, U., Kamal, M.M., Irani, Z. & Weerakkody, V.(2017) Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263-286. DOI: https://doi.org/10.1016/j.jbusres.2016.08.001.

Stilgoe, J., Owen, R. & Macnaghten, P. (2013) Developing a framework for responsible innovation. *Research Policy*, 42(9): 1568-1580. DOI: https://doi.org/10.1016/j.respol.2013.05.008.

Szell, M., Lambiotte, R. & Thurner, S. (2010) Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences,* 107 (31), 13636-13641; DOI: https://doi.org/10.1073/pnas.1004008107.

Taghikhah, F., Filatova, T. & and Voinov, A. (2021) Where Does Theory Have It Right? A Comparison of Theory-Driven and Empirical Agent Based Models. *Journal of Artificial Societies and Social Simulation* 24 (2): 4. DOI: 10.18564/jasss.4573.

Theocharis, Y. & Jungherr, A. (2021) Computational Social Science and the Study of Political Communication. *Political Communication*, 38(1-2): 1-22. DOI: https://doi.org/10.1080/10584609.2020.1833121.

Tomaselli, V., Battiato, S., Ortis, A., Cantone, G.G., Urso, S. & Polosa, R. (2021) Methods, Developments, and Technological Innovations for Population Surveys. *Social Science Computer Review.* March 2021. DOI: https://doi.org/10.1177/0894439321994218.

Tuninetti, M., Aleta, A., Paolotti, D., Moreno, Y. & Starnini, M. (2021) Prediction of new scientific collaborations through multiplex networks. *EPJ Data Science.* 10: 25. DOI: https://doi.org/10.1140/epjds/s13688-021-00282-x.

Vargo, C.J., Guo, L. & Amazeen, M.A. (2018) The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. New Media & Society, 20 (5), 2028-2049. DOI: https://doi.org/10.1177/1461444817712086

Vespignani, A. (2012) Modelling dynamical processes in complex socio-technical systems. *Nature Physics*, 8, 32–39. DOI: https://doi.org/10.1038/nphys2160.

Wang, H., Xu, Z., Fujita, H., & Liu, S. (2016). Towards felicitous decision making: An overview on challenges and trends of big data. *Information Sciences*, 367, 747–765. DOI: https://doi.org/10.1016/j.ins.2016.07.007.

Wang, Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*, 114(2), 246–257. DOI: https://doi.org/10.1037/pspa0000098.

Wang Z., Bauch C.T., Bhattacharyya S., d'Onofrio A., Manfredi P., Perc M., Perra N., Salathé, M. & Zhao D. (2016) Statistical physics of vaccination. *Physics Reports*, 664, 1-113. DOI: https://doi.org/10.1016/j.physrep.2016.10.006.

Weng, L., Menczer, F. & Ahn, YY. (2013) Virality Prediction and Community Structure in Social Networks. *Sci Rep* 3, 2522, 1-6. DOI: https://doi.org/10.1038/srep02522

Wu, A., Taneja, H., boyd, d., Donato, P., Hindman, M., Napoli, P. & Webster, J. (2020) Computational Social Science: On measurement. *Science*, 370 (6521): 1174-1175. DOI: 10.1126/science.abe8308.

Wu, Y., Kosinski, M. & Stillwell, D. (2015) Computers judge personalities better than humans. *Proceedings of the National Academy of Sciences*, 112 (4), 1036-1040. DOI: 10.1073/pnas.1418680112.

Yang, T., Majó-Vázquez, S., Nielsen, R.K. & González-Bailón, S. (2020) Exposure to news grows less fragmented with an increase in mobile access. *Proceedings of the National Academy of Sciences.* 117 (46): 28678-28683. DOI: https://doi.org/10.1073/pnas.2006089117.

Yantseva, V. (2020). Migration Discourse in Sweden: Frames and Sentiments in Mainstream and Social Media. *Social Media + Society*. https://doi.org/10.1177/2056305120981059.

Yoo, Y., Boland Jr. R.R.J., Lyytinen, K. & Majchrzak, A. (2012) Organizing for Innovation in the Digitized World. *Organization Science,* 23(5):1398-1408. DOI: https://doi.org/10.1287/orsc.1120.0771.

Young, J.A. (2020) #SocialWorkEducation: A Computational Analysis of Social Work Programs on Twitter. *Journal of Social Work Education*. DOI: https://doi.org/10.1080/10437797.2020.1817821.

Appendix: Three categories of substantive-technical literature

This table represents a phase in the analysis of the substantive-technical literature in which 26 papers were categorised with their dominant research approach.

Area	**Network cartography (3)**	**Influence & transference (13)**	**Categorising and correlating (10)**
Citation	Eagle et al. 2009 Szell et al. 2010 Sekara et al. 2016	Bond et al. 2012 Aral and Walker 2012 Aral et al. 2009 Funk et al. 2010 Christakis and Fowler 2011 Vespignani 2012 Wang et al. 2016 (physics of vaccination) Luke et al. 2005 Weng et al. 2013 Bail et al. 2018 (hard - has some correlating aspects, too. Political polar) Vargo et al. 2018 Procter et al. 2013 Farrell 2016 (some influence, political polar)	Kosinski et al. 2013 Bollen et al. 2011 Golder and Macy 2011 Schwartz et al. 2013 Wu et al. 2015 Mestyan et al. 2013 Wang and Kosinski 2018 O-Brien 2010?? (Not sure about this one; sits between) Reece & Danforth 2017 Curme et al. 2014

Aims	Map patterns of social relations; improve on other (social network analysis) methods.	Chart spread or transfer of entities through groups or populations and effect on behaviour.	Classify digital traces and show that categories correlate with other phenomena (e.g. data from personality tests, stock markets or geographic location and therefore may predict individuals and populations)
Prominent methodological principles	Focus on spatial/network location and transactional connections. Relation derived from node activity. Description, accuracy, comparison, generalisation.	Focus on network location, transactional connections and activity. Identification of causal/directional relations. Description, prediction, causality, contextualised findings.	Focus on activity and/or cultural outputs. Assembly of aggregates for comparison. Extraction, classification, correlation, interpretation, inference, prediction, revelation.
Prominent concepts	Networks	Networks, contagion, influence, spread, behaviour, ties.	Language, personality, mood, traits, attributes.
Example Technique	Social Network Analysis	Agent based models	Factor Analysis
Data exemplars	Mobile phone location data, Bluetooth data, online game data	Mobile phone application download data; Combining data e.g. Bluetooth records with health service data	Social media posts, messages, photographs, “likes”, etc; Wikipedia editing activity