

Towards an understanding of research ethics for computational, e-social science: a geographical perspective

Richard Harris

Spatial Modelling Research Group, School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1SS.

rich.harris@bris.ac.uk

Abstract. There has been a rapid growth of geographical information technologies and the amount of personal data routinely collected about individuals and their lifestyles. The geographical dimension of these technologies – notably geographical information systems (GIS) – has sparked considerable discussion on whether they be welcomed in human geography, with consideration to: whether to research individuals as amassed bits of data; the tensions between explaining and understanding social processes, structures, human agency and behaviour, and the merits of qualitative and quantitative approaches for doing-so; reconciling post-modern and post-structuralist concerns for difference with empirically-driven search for pattern and order; issues of geosurveillance; what it means to have personal data privacy, freedom of information and control of one's digital persona; and the processes by which software routines create or sustain social-economic divisions in terms of digital divides. Here I reflect on these debates to argue that whilst innovations such as ‘the Grid’ theoretically make it possible for new computational tools to be brought to bear on a diverse range of social science research problems, the long-term sustainability of innovative e-social science research, its methods, practices and technologies, is best ensured by an early dialogue with those who might otherwise wish to contest and debunk it.

Introduction

This paper focuses on grid computing as a practice of e-social science. It argues that the potential for e-social science as an explicitly computational project “characterized by massive distributed datasets, distributed computing and the need for complex visualizations” (Martin 2005 p.2) will only be realized if critiques of a ‘databased’ approach to social research are incorporated and considered at the project’s inception.

For this paper it is imagined that users of ‘the Grid’ will wish to harness the collaborative power of distributed computing to identify, explore, verify or explain patterns and trends in extensive, though not necessarily complex, digital datasets of a size that prohibits effective analysis on a

single ('desktop') machine. Their research may be abductive, inductive or deductive in nature, where the differences in these forms of reasoning are taken to lie on a broad continuum from using the computer to generate new ideas and knowledge from the 'raw' data alone (in principle free from the researcher's *a priori* ideas and assumptions – pure abduction); through using the computer to create or infer information from data collected about assumed socio-economic phenomena or processes (induction, e.g. by algorithmic classification); to using or generating data to confirm or falsify existing theory, laws or models (deduction). As such, the data processing undertaken on the grid could involve any one of a swathe of techniques that include: generating simple statistics summarizing class distributions; multivariate reduction of the data into a small number of groups or clusters; iterative repeat testing for patterns of significance within subsets of data; mapping, visualization, interactive or exploratory analysis; the application of artificial intelligence; and simulation and forecasting. Succinctly stated, this perspective on e-social science is similar to Couclelis' (1998 p.17) 'working definition' of geocomputation, though without her geographical emphasis: "the eclectic application of computational methods and techniques to portray spatial properties, to explain geographical phenomena, and to solve geographical problems."

Whilst there is no assumption that either the patterns or trends examined using the Grid or the researcher's interests necessarily will be spatial, drawing on the adage that 'everything which happens, happens somewhere' it is likely that at least some users will be interested in issues of socio-economic differentiation that do indeed have geographical dimensions. To such researchers, the application of grid computing may represent only an evolution of existing practices of geocomputation (Longley *et al.* 1998), Geographical Information Systems (GIS) (Longley *et al.* 2005) and/or geosimulation (Benenson & Torrens 2004) that already lie on the abduction-induction-deduction continuum for computer-dependent geographical research (together with methods of geovisualization; Unwin & Fisher 2001). Martin (2005 *ibid.* p.3) makes a linkage, stating that

many have questioned whether there is really any difference between GIS and geocomputation, but ... a real difference may be about to emerge with respect to their ability to contribute to the new e-science environments. Some of the enduring fundamentals of socioeconomic GIS research agendas remain to be addressed, but to these should be added the development of appropriate spatial methodologies for e-social science.

If so, then it is prudent to remember that the practice of GIS and of its associated technologies have been sharply debated in (human) geography where their place in social research has been, at times, strongly contested and resisted (Pickles 1999; Schuurman 2000). With the centrality of computing in the emerging vision for e-social science, it would be surprising if similar concerns and hostilities were not encountered there, too.

For the rest of this paper, then, I focus on enduring critiques of GIS and quantitative methods, especially, and of 'scientific' practices more generally within human geography; and on the responses to those critiques. The purpose is more than disciplinary navel-gazing or, worse, triumphalism. The critiques are not exclusively geographical, usually originate from outside the discipline and have wider relevance. That the arguments have resonated so strongly in human geography is in part due to the hybrid nature of a subject which can be found in faculties of arts, humanities, science and social science, and in part due to the rapid proliferation in the routine and increasingly 'invisible' collection of geographically-referenced digital data and of

technologies to process them, used for the everyday functioning (and structuring) of society. Yet because the debates are not uniquely geographical so sensitivity to them should be central to the shaping and prospects of a computational e-social science. By anticipating and reflecting on potential critiques, a geographical perspective is offered towards a research ethic for computational, e-social science that gives specific consideration to broader political motivations for collecting data about people and places, issues of representation, explanation and understanding, and concerns about privacy and surveillance.

Turning Spaces

Alongside the evolving sophistication of quantitative analysis, there has come a proselytizing zeal to present enumeration as *the* way of doing things, matched only by the 'reverse swing' of those who would dismissively regard qualitative analyses as *the* way (Cloeke et al. 2004 p.248, original emphasis).

Antagonisms between the quantitative and qualitative 'camps' in human geography are easily exaggerated – a process which leads too readily to setting them in opposition and the construction of false dichotomies: those who generalize and those who do not; those who are rigorous in their methods and those who are not; those who are sensitive to diversity and those who are not; those who seek to explain and those who seek to understand; those who valorize objectivity and those who consider the contingencies of knowledge; those who are policy relevant and those who 'fiddle whilst Rome burns; ...and so forth.

Unfortunately, from inception such dichotomies tend to be self-preserving as researchers, their methods and their practices are sifted and sorted under different 'labels', caricatured and exposed to critique and counter-critique. Such debates are 'headline grabbing' and eminently publishable, though whether they reflect the day-to-day sentiments, motivations, interests and practices of the wider research community is questionable. A focus on antagonism will distract from the more widespread tolerances and mutual respect that may actually characterize a traditionally diverse but holistic discipline.

Nevertheless, published narratives of the modern history of (human) geography certainly identify a tugging match between quantitative and qualitative approaches (amongst others). Many of the origins for this lie in the period from about the mid 1960s to the mid 1970s, an epoch sometimes described as 'the quantitative revolution' and exemplified by texts such as Haggett (1965) and Abler *et al.* (1971). The epistemological basis of this period is usually taken to have been set out by David Harvey (1969) in *Explanation in Geography*, wherein a realist, positivistic and hypothetico-deductive framework for geographic enquiry was constructed. To what extent this theoretical scaffolding actually supported or held the quantitative practices of the period – and even less so today – is difficult to know, given the temptation to retrofit particular philosophical perspectives upon the research of the time. Although some scholars clearly were inspired by classical science and its emphasis on laws, general theory, mathematical theorems, axioms, and experimental proof (transplanting physical laws such as the gravity model into the social realm), others adopted a less deterministic approach, focusing on statistical associations, incorporating randomness and uncertainty, and avoiding simplistic notions of cause and effect. Nevertheless, there clearly was an increased emphasis on numeric and formulaic abstraction, on mathematical

and computational analysis, and an appeal to rational behaviour, the principle of least effort, utility maximization, the efficiency of industrial production and other economic theorizations which could be modelled to determine their socio-spatial consequences, or to identify localities where theory and reality didn't mesh, thence raising the question 'why?'

Famously (within geography) David Harvey – the author of *Explanation in Geography* – was also one of the first antagonists of the model-based approach to geographical enquiry. His move towards Marxian perspectives and a politicized research agenda accompanied a wider unease about the culpability models emphasizing the efficient and rational working of an industrial society might have in maintaining, even legitimizing, structural inequalities – the capitalist *status quo*. Others sought to re-establish the human in human geography, rejecting notions of rationality and the artificial construction of the statistically average person. Instead, humanistic geography offered "a very different conception of humanity, a vision which recognized humans in all their flawed ambiguities as experiencing, perceiving, feeling, thinking and acting beings" (Cloke *et al.* 2004 p.22) – character traits that are poorly accommodated by logical positivism and an emphasis on empirical verification that validates knowledge by prediction and observation. Humanistic geography, in turn, evolved with critiques from feminist geography, for example (initially due to its lack of appreciation for gender issues), opening-up increased awareness of the politics of representation; of the embedded and unequal power relationships which structure discourse. New ways of reading and writing 'the world' have followed, with a current interest in 'non-representational' understanding of how social phenomena are performed and embodied, using perspectives that tend to reject essentialist categorizations such as human and nature, or the living and the technological, and which instead have looked towards 'post-human' and hybrid geographies (to understand, for example, that human life and technology, including the Grid, are produced through each other in complex ways; Dodge & Kitchin, 2005a).

Against either the emerging 'cultural turn' or the critiques of capitalist practise, the development of GIS as a maturing, pervasive and highly profitable technology could be regarded as a retrograde step if incorporated into human geography. With a focus on vector and raster data structures and the geometric manipulation of data, GIS could be portrayed as entailing a narrow, empirical and positivistic view of geographic space, seeing it as absolute and a container in which objects and phenomena have their place. For example, a technical conceptualization of a neighbourhood or community – one that can be encoded for computer reading – is to define it solely by the innermost surface of a polygonal object said to represent its boundary. That the place is surrounded by other neighbourhood objects is irrelevant to its own intrinsic definition, whilst the study and manipulation of the objects may encourage a false sense of separation between the users of a GIS and the world contained and naturalized on their computer screens. The separation in turn plays to a belief that the user is independent of and external to the apparently naturally forming object and able, therefore, to gain an impartial and objective understanding of the objects, primarily by collecting data about them (Curry 1998; Raper 2000).

Such reasoning is, of course, cumulative and each piece of the critical jigsaw does not necessarily fit together. For example, the limits of an absolutist notion of space were already well understood by quantitative geographers, many of whom *also* saw GIS as a rather limited technology, albeit with potential (e.g. Gatrell 1983 had demonstrated that there are many different concepts of distance and a variety of spaces to be explored and mapped). Nonetheless,

given the growing disciplinary interest in deconstructing ‘grand narratives’ by exposing the partisan underpinnings of taken-for-granted geographical representation and discourse (such as those provided by maps – digital or otherwise), and by drawing on philosophical and social constructivist perspectives on science – including the logical limits of the hypothetico-deductive method of knowledge construction and the ways in which scientific enquiry are socially constituted by the networks in which scientists operate as opposed to any objective and teleologically driven progress – so

human geographers were not immune to the growing body of work on the interdigitization [interdigitation] of science and culture. Articles discussing the merits of cultural studies of science made their way into geography... They were, in a sense, the continuation of a tradition of criticism focusing on positivism in geography. It was only a matter of time before GIS, at first glance a rapidly growing positivist technology, attracted the attention of human geographers (Schuurman 2000 p.571).

Schuurman (*op. Cit.*, after Pickles 1999) argues that the “debates over GIS in geography were overt and apparently motivated by a genuine desire on both sides to steer the discipline in an appropriate and responsible direction, despite occasional outbreaks of hostility” – and a provocative rhetoric. Taylor (1990), for example, describes GIS as the ‘positivist’s revenge’, links the technology to imperialism and suggests it is only equipped for what Cloke *et al.* (2004 p.283, after Geertz 1973 and Denzin 1978) describe as thin description: it “focuses more on facts about, and characteristics of, the research subject.”

In a strongly worded reply to Taylor, Openshaw (1991) dismisses the ‘pseudophilosophical’ objections of the ‘technical cripples in geography’ and instead views GIS as the unifying technology which can put ‘Humpty-Dumpty [geography] back together again’ (see Schuurman 2000). Openshaw (1997) also produced an impassioned rebuttal of *Ground Truth: the social implications of geographic information systems*, a volume edited by Pickles (1995) who states in the Preface that:

Ground Truth is first, a book about the transformation of data handling and mapping capabilities that have emerged in the past two decades, and the impact they have had within the discipline of geography. Second, it is a book about the constellation of ideas, ideologies, and social practices that have emerged within the development of new forms of data handling and spatial representation. Third, it situates GIS as a tool to protect disciplinary power and access to funding; as a way of organizing more efficient systems of production; and as a reworking (and rewriting) of a broader history of cultural codes – the creation of new visual imageries, new conceptions of earth, new modalities of commodity and consumer, and new visions of what constitutes market, territory, and empire.

Although Openshaw’s is an overreaction to what is generally a useful and thoughtful volume, it is perhaps understandable given the various provocations which appear in the book together with a tendency to exaggerate or see only the worst in geographical information technologies – one contributor links GIS and, by implication its practitioners, to those experts who are said to disenfranchise the victims of the bombings of Hiroshima and Nagasaki (For a calmer review of *Ground Truth* see Flowerdew 1998). And, despite the rather dismissive way in which they are dealt with, Openshaw (1997) usefully identifies thirteen lines of critique that are raised against GIS and which have wider relevance to a computationally-based e-social science: a GIS geography can be condemned as a handmaiden of the state; GIS is a revival of discredited positivist epistemologies; GIS emphasizes maps and spaces; GIS is an ordered representation of space within a Cartesian framework; GIS is a discipline dominated by facts rather than

knowledge and understanding; a GIS geography implies a neglect of themes not included in the data; GIS continues and reinforces a structural distortion of knowledge and uneven data access; GIS is not a value-free technology; GIS has a propensity for evil intent; GIS is the technology of surveillance; Geodemographic Information Systems [see below] are an apotheosis of instrumental reason; grandiose claims and hype; and the inevitability of ethical inconsistency.

In fact, as Schuurman (*ibid.*) notes, the timing of publication of *Ground Truth* obscures the meetings that were taking place between ‘the protagonists’ and ‘the antagonists’ of GIS that directly or otherwise led to: the National Center for Geographic Information and Analysis Initiative 19 (NCGIA i19: www.geo.wvu.edu/i19/) – looking at the social implications of how people, space and environment are represented in GIS (see Pickles 1999); greater awareness of privacy issues and the limits of representation (Curry 1998, 1999); increased awareness of how maps can ‘lie’ or be used politically (Monmonier 1996); issues of geosurveillance (Monmonier 2002); and to an increased shaping of an ethics of GIS use.

GIS, Ethics and Codes

This increasing use by cities of advanced technologies, in general, and of GIS, in particular, has provided unprecedented opportunities for improved efficiency in data access, management and dissemination. However, these technologies can also aggravate and exacerbate problems when they are misused (Esnard 1998 p.33).

Partly drawing on personal experience, Esnard (1998) documents some ‘abuses and questionable uses’ of GIS for local governance. The examples are intended to demonstrate how these computer-based information systems can be used to ‘mislead the public, marginalize community groups, violate the open public review process, create problems with data quality, and drive specific political agendas’ (*ibid.* p.36). From these Esnard offers what are described as ‘Portable, Provisional Codes’ (PPCs) that can serve as a simple common language of ethical behaviour amongst local government officials using GIS, and as a set of ethical guidelines, at least until relevant professional codes of conduct are updated. The five PPCs are: 1. at all times the process should be given precedence over the final map or analytical project (p.38); 2. data quality and data description should [in the US] satisfy standards established by the Federal Geographic Data Committee to ensure data uniformity and to make sharing possible at the local, state, and federal levels (p.39); 3. community needs, local knowledge, and social and cultural contexts must be represented when applicable (*op. Cit.*); 4. an individual or group’s right to confidentiality and privacy should not be violated, nor should there be intentional misuse, misrepresentation, or falsification of data to place an individual or group at an unfair advantage (p.41); and 5. in compliance with the Open Records [Freedom of Information] Act, the request for public access data by the public should be accommodated (*op. Cit.*).

Overall, the focus of the PPCs “is on process, data quality, community needs, liability, and data access as the framework for *reflective* use of GIS technology” (*ibid.* p.37, emphasis added), avoiding the mis- or non-representation of individuals, groups or communities, documenting the preferences, choices, selections and biases that condition an analytical outcome, and paying regard to issues of data protection, access and confidentiality.

Similar concerns are taken-up by the GIS Certification Institute (GISCI) which requires that its Code of Ethics be signed as a condition of certification for a 'GIS professional'. That Code (which may be viewed at www.gisci.org/code_of_ethics.htm) is described as deontological:

based on the ethical principle of always treating others with respect and never merely as means to an end... It requires us to consider the impact of our actions on other persons and to modify our actions to reflect the respect and concern we have for them. It emphasizes our obligations to other persons, to our colleagues and the profession, to our employers, and to society as a whole.

Whilst the Code deliberately is similar to that of other professional societies, it has some guidelines that, though described as 'unique to the GIS profession', actually could apply to any data-based research. These include

the encouragement to make data and findings widely available, to document data and products, to be actively involved in data retention and security, to show respect for copyright and other intellectual property rights, and to display concern for the sensitive data about individuals discovered through geospatial or database manipulations.

The Code of Ethics recognizes GIS as a social practice where

new technological capacities and an expansion of the scope of their application in many areas of social life have made it increasingly important to think about the ways in which the logics, systems, and representations deployed by contemporary GIS support particular types of social practice and inhibit others (Pickles, 1999 p.56)

And both sets of codes (the PPCs and GISCI's) recognize that which is legal is not equivalent to that which is ethical.

In philosophy, ethics defines what is good for the individual and for society and establishes the nature of duties that people owe themselves and one another. While law often embodies ethical principles, law and ethics are far from co-extensive. The law does not prohibit many acts that would be widely condemned as unethical, for example, lying or betraying the confidence of a friend ... much of the law is far from simply codifying ethical norms (Cho 2005 p.214).

This distinction between law and ethics especially is pertinent when considering empirical and data driven representations of peoples and places. With consideration to 'obligations to individuals in society' GISCI's Code of Ethics states that

the GIS professional will 1. Respect Privacy: protect individual privacy, especially about sensitive information; be especially careful with new information discovered about an individual through GIS-based manipulations (such as geocoding) or the combination of two or more databases [and] 2. Respect individuals: encourage individual autonomy – for example, allow individuals to withhold consent from being added to a database, correct information about themselves in a database, and remove themselves from a database; avoid undue intrusions in the lives of individuals; be careful when disclosing information about an individual; treat all individuals equally, without regard to race, gender, or other personal characteristic not related to the task at hand.

Within the European Union, these respects are also partially legal obligations defined under Directive 95/46/EC (dated 24 October 1995) and by Regulation Number 45/2001 (dated 18 December 2001) which together harmonize national laws across EU member states on the processing of personal data, with the intention of protecting the rights and freedoms of the persons concerned, in particular their right to privacy. Within the UK the Directive is enforced under the 1998 Data Protection Act, the eight statutory principles of which require that personal data must be: fairly and lawfully processed; processed for limited purposes; adequate, relevant

and not excessive; accurate; not kept for longer than is necessary; processed in line with people's rights; secure; and not transferred to countries without adequate protection. In addition, the UK Freedom of Information Act (2000) gives people a general right of access to information held by or on behalf of public authorities and promotes a culture of openness and accountability of public sector bodies. Both Acts apply to academic research (see www.informationcommissioner.gov.uk).

Legally the Data Protection Act imposes restrictions on the storage and manipulations of personal data in a distributed computing environment. But these should not overshadow – or be said to cover – the broader ethical issues at stake. As Cho (2005) notes, data protection precisely is that – the protection of data (against misrepresentation, against misuse, against mistake) not the data subject. The two overlap, of course, but it would be wrong to suppose, for example, that Data Protection prevents what the subject may consider to be personal data from being routinely collected and stored in neither an anonymous nor confidential state.

Moreover, informational privacy – here to be understood as the undertakings made by companies to limit the dissemination, manipulation and application of personal data – should not be synonymized with personal privacy. Personal data is that which relate 'to a living individual who can be identified from those data or from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller.' Personal privacy is the desire of an individual to be left alone but which rarely is a legal right. Personal privacy may be infringed by representations formed and decisions taken on the basis of aggregate (that is, 'not personal') datasets (Curry, 1997).

The *bête noire* of social theorists are geodemographic information systems (Goss 1995a, 1995b; Curry, 1998; Philips & Curry 2002; Monmonier 2004) which are used in marketing but also increasingly in public policy applications to make operational decisions that affect people on the basis of the externally defined 'type' of area in which the people live. They often are used to simplify and make sense of otherwise complex information sets and their effects can be fairly benign (receipt of particular 'junk mail') or positive (they have been used to adopt more sensitive policing of communities, Ashby & Longley 2005 and to monitor socio-economic participation in UK Universities, see Harris *et al.*, 2005). However, when decisions are made on the basis of at least partly constructed (partly artificial) differences between people and places there is a risk of stigmatization and of contributing to, not remedying, perceived or actual socio-economic divides. To some it is fair and makes sound economic sense to, say, adjust car insurance premiums to a geodemographic measure of risk associated with the owner's address. To others, it is profoundly unjust if the level of risk taken on is not a matter of individual choice but due to socio-economic determinants such as job or housing markets.

Geodemographics catches the social theorist's eye partly because of the nomenclature and language of labelling neighbourhood types (e.g. 'Industrial Grit' or 'Satellites and Scratchcards'), partly because of the excessive rhetoric of intimacy, power and weaponry propounded by early geodemographic advertising (see Goss 1995a, 1995b), and partly because it supports critical commentary on information economies, consumerism and the practices of advanced capitalism – most notably in the way populations are differentiated by place of residence and accorded different levels of service (e.g. response time when telephoning a bank's

call centre), access to facilities and so forth. But, as an example of social sorting, impinging on personal (private?) lives, geodemographics is an exceptionally visible and well advertised case of the software or code that are “increasingly central to the spatial formation of collective life”; that “mediates, supplements, augments, monitors, regulates, and produces collective life” (Dodge & Kitchin 2005a pp. 162 & 169).

Critiques of the pervasiveness of such code often focus on the proliferation of surveillance technologies such as CCTV, high resolution Remote Sensing, mobile phone tracking or the Radio Frequency Identification (RFID) chips embedded into consumer products; and on concerns about digital discrimination and digital divides. These cannot be ignored by an e-social science that explicitly incorporates algorithmic formulations – code – to sift and sort through large or complex datasets that directly or otherwise create visual, numeric, statistical or other representations of societies, of people, of places. One positive response to these critiques would be to place an ethos of openness at the heart of grid technologies, using open source programming languages, public access (control?) over the data processed and well documented, non-technical ‘meta-data’ explaining how the software are processing, selecting and categorizing data and therefore the data subjects (though there are clearly issues of IPR and copyright to be negotiated). Yet, by itself, openness is not sufficient because

code does not simply exist, created and working in easily defined and examined ways... The technicity of codes [the productive power of codes to make things happen] is not... deterministic (i.e., code determines in absolute, nonnegotiable means, everyday practices) or universal (i.e., such determinations occur in all places and all times in a simple cause-and-effect manner). Rather, as noted, technicity is contingent, negotiated, and nuanced; it is realized through its practice by people in relation to historical and geographical context. As such, there is no neat marriage between coded objects, infrastructures, processes and assemblages, and [particular effects of code. Instead, technicity varies as function of the nature of code, people, and context (Dodge & Kitchin, *ibid.* p. 170).

To unpack this a little: developing and promoting grid technologies will have uneven social consequences that may or may not be intended or considered desirable. Consider the following contribution from a Group Leader of one of Europe’s largest providers of personal, consumer data to the letters page of the January 2005 issue of *Database Marketing*, a magazine for the managers and users of UK target marketing and customer management systems (www.dmarket.co.uk).

Grid technology does not only provide speed. It gives us the ability to store and process virtually unlimited volumes of data at a lower cost. Now we can keep customer data over the whole lifetime of interaction with a company. And, as the more data you put into a model the better it is, we will be able to build models which have the depth of data and learning that enables them to be truly predictive... I firmly believe that grid technology will be one of the biggest influences in driving direct marketing into a truly data-focused discipline.

Under this vision of the future I can confidently expect my digital persona – who I may not have met or who’s existence be aware of – to be given increased voice to talk on my behalf and help anticipate and order the types of good and services the ‘real me’ is presumed to want. Sensationalist, perhaps but less so to argue that a computational e-social science will be socially sterile if it is not sensitive or contributing to broader debates about the social effects of the technologies employed; if it is not conscious of the contested political environment in which the regular collection of socio-economic data as a management practice of governance is viewed by some as promoting accountability and choice but by others as part of a wider neo-liberal political

agenda that disenfranchises those who are less able to exercise their choice (consider, for example, the potential socio-spatial effects of publishing school performance ‘league tables’ on a market-driven system of home ownership); if it is not wise to the risk of marginalizing those ‘knowledge systems’ that either are not easily encoded into digital databases or which are not central to systems of governance or of commerce – e.g. emotional attachments and the sacredness of place, local knowledge, historical memory of past events (Pickles, 1999). All of which is to take seriously the obligations to individuals in society suggested by the GIS Code of Ethics.

There are further issues of what social science is or should be about that relate to the on-going critiques of quantification and computation in human geography. Provided it is understood then computation, modelling and quantification is not just about enumeration (counting), that qualitative practices can also be enumerative, that the term ‘thin description’ is neither a derogatory nor pejorative term and as meaning neither theory nor explanatory ‘lite’, and that most, if not all, academic practices involve some level of abstraction or generalization, then there seems no reason to take umbrage at Cloke *et al.*’s (2004 p.283) suggestion that enumeration (in human geography, notably quantitative geography and GIS) is a form of thin description “capable of identifying certain characteristics and patterns of data, but incapable of describing or explicating the meaningful nature of social life” – that is, of ‘thick description’ that includes “information about the context of an act, the intentions and meanings that organise actions and its subsequent evolution.” The same may also be true of a computational e-social science, although the distinction is challenged by current commercial research into digital life-logs “as a form of pervasive computing consisting of a unified, digital record of the *totality* of an individual’s experiences, captured multi-modality through digital sensors and stored permanently as a personal multi-media archive” (Dodge & Kitchin, 2005b p.2, original emphasis) – itself raising ethical issues of ownership and whether “a record of the past that includes every action, every event, every conversation, every material expression of an individual’s life” (*op. Cit.*) should, like the human memory, be able to forget.

Conclusion

In this paper I have drawn on debates in human geography to consider how the prospect of a computational e-social science might be received in it and allied disciplines. In writing, the hope was to be neither polemical nor imbalanced but to consider the issues with mind to personal research experience of analysing and assessing the significance of patterns and trends in large, consumer databases originally compiled for the marketing of goods and services. Yet, perhaps arising from this background is an ulterior motivation – to suggest the recurring caricature of quantitative or computational social scientists (or human geographers anyway) as necessarily lacking in a self-critical reflexivity and research sensitivity is, at least in part, a rhetorical construction fabricated from the need to establish an ‘other’ to be critiqued from alternative epistemological and ontological positions as from genuine understandings of the actual research practices of those who are critiqued. Those alternative perspectives are certainly important but their dichotomised expression is unhelpful, possibly self-defeating. As Schuurman (2000 p.587, after Haraway 1991) emphasizes with respect to GIS:

if social theorists want to influence GIS, then they must make their arguments relevant to the technology Current research on ontologies and interoperability demonstrates the viability of incorporating social, philosophical *and* technical considerations in GIS ... Social theorists [need] to engage in the construction of the cyborg rather than critique it from afar ... Constructive critique requires a depth of understanding about the science or technology being investigated.

The dialogue does, however, need to be two-way, since “GIS scholars are apt to respond by discounting critical attention to *implications* of their work in favor of attention to theory and practice situated within the technology” (ibid. p.587, both original emphases).

For the future prospects of e-science, although innovations such as the Grid will theoretically make it possible for new computational tools to be brought to bear on a diverse range of social science research problems, if the possibilities are not to be eschewed and (perhaps, rightly) resisted by a broader and less technically driven research community, then it is vital to place wider, societal and representational concerns as central to the evolution of e-social science, helping to inform a research ethic that avoids an unnecessary schism between those who 'do computational, e-social science' and those who critique it. As Woolgar (not dated p.3) notes, “there is a tendency to distinguish between research which uses the grid and research which is about the grid.” Against this he asks:

is it possible to specify research which integrates an understanding of the social and economic conditions of uptake and use, with the actual development and design practices? In other words, what are the prospects for an interactive social science approach which articulates the advantages of a social shaping perspective for practice and policy?

At the time of writing there is no one clear definition of what e-social science is or what it entails, albeit that there is “reasonable consensus that, at minimum, e-social science includes High Performance Computing (HPC), Grid computing (where numbers of computers are linked to increase computing power), and associated technologies, to support operations that desktop computing cannot efficiently provide” (Fielding, not dated p.8). There is a risk that e-social science will be seen, as geocomputation has been, as a “haphazard collection of techniques developed for all sorts of different purposes in a variety of areas with sometime unproven intellectual credentials” (Couclelis 1998 p.22). Again, as with geocomputation, it is not yet clear whether e-social science actually is about computation or ‘just’ computing, where “as a first approximation this is the difference between approaching the computer as an integral component of the modeling of complex systems and processes, versus treating it as an equation-solving and data-handling device” (ibid. p.20). As Fielding (*op. Cit.*) notes

Generic e-Science makes a threefold distinction between the Computational Grid (primarily designed to support high performance computation), Access Grid (designed to support collaborative visualisation involving distant researchers) and Data Grid (designed for transporting large volumes of data).

It is the first of these ‘grids’ that is central to this paper and to Martin’s (2005 p.2) description of a Grid-enabled geocomputational toolkit offering “automated data mining, visualization of spatial data uncertainty, incorporation of an explicitly spatial dimension in simulation modeling and neighbourhood classification from multi-source distributed datasets.” However, all three need to be considered in terms of what Woolgar (ibid. pp. 1-2) describes as social shaping: “defined very broadly to include all social scientific aspects of the genesis, use, implementation and effects of the new e technologies” – not only “the likely effects of the widespread adoption

of grid technologies on [the construction of] (social) scientific knowledge” (*ibid.* p.3) but also the likely effects on those about whom social scientific knowledge is being generated. In other words an epistemological and ethical framework for e-social science needs to be developed and debated. To this end, the ESRC’s recent call for proposals to investigate the social shaping and the socio-economic impact of e-science warmly is welcomed.

References

- Abler, R., Adams, J.S. and Gould, P. (1971): *Spatial Organization: the Geographer’s View of the World*, Prentice-Hall Inc., New Jersey.
- Ashby, D.I. and Longley, P.A. (2005) ‘Geocomputation, Geodemographics and Resource Allocation for Local Policing’, *Transactions in GIS*, vol. 9, pp. 53-72.
- Benenson, I. and Torrens, P.M. (2004): *Geosimulation: Automata-based Modeling of Urban Phenomena*, Wiley, Chichester.
- Cho, G. (2005): *Geographic Information Science: mastering the legal issues*, Wiley, Chichester.
- Cloke, P., Cook, I., Crang, P., Goodwin, M., Painter, J. and Philo, C. (2004): *Practising Human Geography*, Sage, London.
- Couclelis, H. (1998): ‘Geocomputation in Context’, in P.A. Longley, S.M. Brooks, R. McDonnell and B. Macmillan (eds.): *Geocomputation: a primer*, Wiley, Chichester, pp. 17-29.
- Curry, M. (1997): ‘The Digital Individual and the Private Realm’, *Annals of the Association of American geographers*, vol. 87, pp. 681-99.
- Curry, M. (1998): *Digital Places: Living with Geographic Information Technologies*, Routledge, London.
- Curry, M. (1999): ‘Rethinking privacy in a geocoded world’ in P.A. Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind (eds.): *Geographic Information Systems: principles, techniques, management, applications*, Wiley, New York, pp. 757-66.
- Denzin, N. (1978): *The Research Act*, McGraw Hill, New York.
- Dodge, M. and Kitchin, R. (2005a) ‘Code and the Transduction of Space’, *Annals of the Association of American geographers*, vol. 95, pp. 162-80.
- Dodge, M. and Kitchin, R. (2005b) ‘The ethics of forgetting in an age of pervasive computing’, CASA Working Paper Series, UCL, London, No. 92 (www.casa.ucl.ac.uk)
- Esnard, A.-M. (1998): ‘Cities, GIS and Ethics’, *Journal of Urban Technology*, vol. 5, pp. 33-45.

- Fielding, N. (not dated): 'Qualitative research and E-Social Science: appraising the potential', NCESS Consultation Document (www.ncess.ac.uk).
- Flowerdew, R. (1998): 'Reacting to Ground Truth', *Environment and Planning A*, vol. 30, pp. 289-301.
- Gatrell, A. (1983): *Distance and Space: A Geographical Perspective*, OUP, Oxford.
- Geertz, C. (1973): *The Interpretation of Cultures*, Basic Books, New York.
- Goss, J. (1995a): 'Marketing the new marketing: the strategic discourse of geodemographic information systems' in J. Pickles (ed.): *Ground Truth: the Social Implications of Geographic Information Systems*, The Guilford Press, New York, pp. 130-70.
- Goss, J. (1995b): 'We Know Who You Are and We Know Where You Live: the Instrumental Rationality of Geodemographic Systems', *Economic Geography*, vol. 71, pp. 171-98.
- Haggett, P. (1965): *Locational Analysis in Human Geography*, Edward Arnold, London.
- Haraway, D. (1991): *Simians, cyborgs and women: the reinvention of nature*, Routledge, NY.
- Harris, R., Sleight, P. and Webber, R. (2005) *Geodemographics, GIS and Neighbourhood Targeting*, Wiley, Chichester.
- Harvey, D. (1969): *Explanation in Geography*, Edward Arnold, London.
- Longley, P.A., Goodchild, M.F., Maguire, D. J. and Rhind, D.W. (2005): *Geographic Information Systems and Science*, Wiley, Chichester.
- Longley, P.A., Brooks, S., McDonnell, R.A. and Macmillan, B. (1998): *Geocomputation: a primer*, Wiley, Chichester.
- Martin, D. (2005): 'Socioeconomic GeoComputation and E-Social Science', *Transactions in GIS*, vol. 9, no. 1, pp. 1-3.
- Monmonier, M. (1996): *Bushmanders and Bullwinkles: How Politicians Manipulate Electronic Maps and Census Data to Win Elections*, The University of Chicago Press, Chicago.
- Monmonier, M. (2002): *Spying with Maps: Surveillance Technologies and the Future of Privacy*, The University of Chicago Press, Chicago.
- Openshaw, S. (1991): 'A View on the GIS Crisis in Geography, or, using GIS to put Humpty-Dumpty back together again', *Environment and Planning A*, vol. 23, pp. 621-28.
- Openshaw, S. (1997): 'The truth about Ground Truth', *Transactions in GIS*, vol. 2, pp. 7-24.

- Philips, D. and Curry, M. (2002): 'Privacy and the phenetic urge: geodemographics and the changing spatiality of local practice', in D. Lyon (ed.): *Surveillance as Social Sorting: Privacy, Risk and Automated Discrimination*, Routledge, London, pp. 137-51.
- Pickles, J. (1995): *Ground Truth: the Social Implications of Geographic Information Systems*, The Guilford Press, New York.
- Pickles, J. (1999) 'Arguments, debates and dialogues: the GIS-social theory debate and the concern for alternatives' in P. A. Longley, M. F. Goodchild, D. J. Maguire and D. W. Rhind (eds.): *Geographic Information Systems: principles, techniques, management, applications*, Wiley, New York, pp. 757-66.
- Raper, J. (2000): *Multidimensional Geographic Information Science*, Taylor & Francis, London.
- Schuurman, N. (2000): 'Trouble in the heartland: GIS and its critics in the 1990s', *Progress in Human Geography*, vol. 24, pp. 569-90.
- Taylor, P.J. (1990): 'GKS', *Political Geography Quarterly*, vol. 9, pp. 211-12.
- Unwin, D. and Fisher, P. (2001) *Virtual Reality in Geography*, Taylor & Francis, London.
- Woolgar, S. (not dated): 'Social Shaping Perspectives on e-Science and e-Social Science', NCESS Consultation Document (www.ncess.ac.uk).