In search of the data effect: When best available data overrides the real requirement

Nicholas Chrisman

GEOIDE Network et sciences géomatiques,
Université Laval, Québec CANADA
nicholas.chrisman@geoide.ulaval.ca

Critical approaches have a long history in cartography (Blakemore and Harley, 1980; Harley 1989). In the realm of GIS, a related, but not identical critical thread has emerged (Schuurman 1999). It has become accepted, in some measure, that geographic information technologies influence society and that they are influenced by society (Sheppard 1995, Chrisman 2005). In much of this work, the issues of society and GIS have been taken as theoretical, rather than examining the practical consequences. This paper attempts to fill a gap by providing some practical examples of decision making systems reliant on data sources. In specific, these examples will turn on what could be termed "the data effect" – when the existing data sources influence the kind of analysis and possibly the decisions. Such a data effect is important because it shows the path-specific behaviour wherein past decisions, made for different purposes, have an influence on a different sector. It shows also the role of non-human actors (such as databases) in hybrid systems.

This paper also has a methodological contribution. Much of the direct evidence about the social component in GIScience has come from participant observation and other forms of ethnographic research. These methods involve huge time commitments and consequently cover few cases. Ethnography is often practiced by PhD students, once in their career. Afterwards, as newly established scholars, they must resort to less time-intense methods under the pressure of academic jobs. Consequently, there is also a methodological question of how best to conduct research in this sphere.

This paper considers a specific question that deals with the social and institutional aspects of GIScience, and it concentrates on methodology applicable to address a specific research problem, the detection of a possible "data effect" – the influence of pre-existing data sources on subsequent analytical practice. This effect can arise since it is substantially easier to put an existing data resource to use than to generate a new one. As it is explained in many circumstances, databases are a form of capital, either as a corporate resource, or as a public good. Data access policy frequently assumes that the value is always positive. Existence of a certain data resource creates an opportunity effect, making subsequent transactions cheaper, if all is well managed. This would apply to decision making applications, for example. The downside is that less suitable parameters will be used if they come from existing sources compared to those specifically tailored to the question at hand. It is easy to imagine that stakeholders in decisionmaking might be tempted by some form of data effect; it can reduce budgets, shorten times, and build coalitions of data users. The reasons to align with existing data sources are so varied, the effect is hard to detect. In a single ethnographic setting, it would be hard to argue that a certain choice was a deflection from a more rational procedure.

This paper argues that any data effect will be best detected through comparative retrospective studies of multiple replicates, rather than participant observation in specific situations. The case of low-level radioactive waste (LLRW) disposal will suffice to introduce the practice (Chrisman, 2010). In the 1980s, the US government made the states responsible for low-level radioactive waste generated in their borders. States were encouraged to form "compacts" and to locate facilities according to common rules (10 CFR 61) (Nuclear Regulatory Commission 1988; 1991). The regulatory guide counseled the use of GIS to perform siting studies. The fascination with this application of GIS remains current with the National Geographic Society's (2006) curriculum for secondary schools, and in the specialists of multicriteria methods (Carver 1991; Evans and others 2004) who use LLRW as a case study. About fifteen states of the USA have conducted siting projects to some stage of completion over the past thirty years. Millions were spent. Organized opposition has driven most states back to the drawing board due to court challenges or denials at administrative hearings. Revised projects provide replications in the experimental design. The history of these actual projects challenges the glib assumption that GIS is 'simply a tool' in a rational decision-making process. GIS serves no privileged, 'objective' role. GIS proponents may shrug off these failures as the result of 'politics', as if some remote force of nature beyond control. On the contrary, 'politics' of various forms play a necessary and inescapable role in all social contexts. And politics is a multi-sided arena where human and non-human actors interact in interesting ways (Latour 1999). This search for the data effect is essentially to demonstrate the role of a non-human actor, the existing database.

On the subject of politics, the case of Connecticut can set the stage. The public authority charged with making the decision for siting a facility for LLRW disposal attempted to make their selection as objective as possible (Connecticut 1991). They applied their criteria in a blind manner, so that the commission members were discussing just the nature of the criteria, not where these phenomena existing across the landscape. The work was done totally in the attribute space, without maps. The idea was that this would eliminate any bias based on local preferences or affiliations. Eventually, the barrier was removed, and the result of their analysis appeared for public examination. The candidate sites were clustered in the district of a particularly powerful elected politician, the President of the State Senate. Rather quickly, the candidates were withdrawn from consideration. The result so discredited the technocratic approach that the commission had to change models completely from the GIS-based top-down approach to a bottom-up 'volunteer' scheme. This story does not establish the specific power of one particular database, but rather provides a caution that a firm barrier between the technical and the political is not productive. The non-human actor in this case was the map.

One core issue in siting LLRW repositories involves groundwater, since buried waste is most likely to be transported offsite by this route. Of course, groundwater is one of the hardest factors to measure, since by definition it is invisible from the surface in most settings. Many states used surrogate information to stand in for the harder-to-obtain groundwater layer. In some cases (Newberry 1994) it was geological information, contrasting carbonates with high porosity to shales and schists with lower transmisivity. Others used soil data, contrasting sands to clays and other materials. No matter how detailed and appropriate the geological data source, groundwater is a contentious topic. In the case of Martinsville, Illinois, this became the core of the incredibly lengthy public hearings (Illinois 1994). In ninety-two days of public hearings, there was no detail to small to be reviewed and reanalyzed by each stakeholder. No database was good enough. In the opposite extreme case of California, a simple precipitation map was used to distinguish dry from wet with a simple threshold. Every place with over ten inches of annual rainfall was judged to pose a risk of too much groundwater movement. This criterion applied in most other states (or in most of Europe) would have simply ruled out the whole political unit. Of course, groundwater transmission is much different in desert settings, but even in California, desert conditions vary according to latitude – not to mention the geological conditions of interest in other regions. This is an example of substituting something quickly available for something really intended in the analysis.

Perhaps the most direct example of the data effect comes from the case of Pennsylvania (Chem-Nuclear Systems 1992, 1993, 1994). The site selection process of this state was conducted in discrete stages, first screening at a gross state-wide level, then two stages of refined scale (termed regional and local). Most of the criteria covered in the first phase were bureaucratic entities, such as state parks and protected lands. It was in the next two stages that the physical issues were addressed, but only in areas that had survived the prior screening. This is totally logical, from a data processing viewpoint. If an area is going to be excluded for some simple big bureaucratic polygon, there is no point developing a finely detailed geographic database inside the excluded area. However, the order of operations was not entirely based on scale. Areas of past mining cover large regions of the state of Pennsylvania, largely coal mines in carboniferous geology that mixes limestone with coal seams. If limestone geology is a bad idea for LLRW, and old mine is even worse, in that there are pre-made openings covering a large area. The problem is that the documentation of old mines is highly suspect and antiquated. Each mine survey, duely deposited in the state regulatory office, was done on a local datum, disconnected from other surveys and frames of reference (horizontal or vertical). Many of these records date back over a century or more, recorded on paper that is deteriorating. The LLRW team examined these sources, and estimated it would cost millions of dollars to digitize the original surveys and transform them into a known coordinate system. It was not really the mandate of the LLRW contractor to go clean up the mine records for the State. This really critical criterion was not treated directly. The availability of general geological strata was used as a surrogate for having the actual mine surveys. Any area with coal near the surface was treated as having been mined, which substantially exaggerates the extent of this criterion. In this case, the contractor probably erred on the side of caution, since most of the error is a false positives (saying an area had mines when it just had unexploited coal seams). For the purposes of this paper, the direction of the error is not so important. The data effect can be documented, and becomes more clear due to comparative methodology.

The comparative approach can be carried to an international scale, where the constrasts get more marked. For example, Belgium classifies its LLRW according to the heat that will be generated (IAEA 2004 section 4.1), and therefore is less worried about groundwater (in what is admittedly a wet environment across the whole country). Parenthetically, Belgium is also applying a volunteer approach using local incentives. The current volunteers (Dessel, Flanders) are located 100 km from the site of AGILE 2011. One of the French LLRW depositories is built specifically around a groundwater monitoring design (IAEA 2004 Section 4.2). The facility, near Soulaines, departement de l'Aube, is located on top of an impervious clay layer, in a bed of quite pervious sand. The groundwater movement is directly into a local stream, thus the site is deemed to be suited to long-term monitoring. Here, the information sought is about the future. Any leak from the site is intended to be detected by a network of sensors included in the design. A subsequent facility (Morvilliers) was located 2km away, but this time in the clay layer. In both cases, the sites were characterized in great detail, not at the exploratory scale used in the USA for site screening studies. With such detail, the data effect is less important, though placing depositories in porous and impervious strata shows a bit of inconsistency in approach.

While there is much debate about methods to apply to science and technology studies, the specific setting of LLRW disposal siting falls under a regime of state regulation. Radioactive waste is inherently contentious; a hot potato that glows in the dark. The field of drug regulation is a close cognate. A recent study by sociologists at University of Sussex (Abraham and Davis, 2007) applied a methodology of retrospective analysis to drug regulation and the interactions of politics and science. Their comparison contrasted the regulatory history of the same drug between USA and UK. While it may be truism that regulation occurs in a particular historical and political setting, just how that setting influences decisions takes substantial untangling. In the case of drug regulation, the study was hampered by proprietary claims to the studies submitted to the regulatory authorities. Still, the authors were able to compare between jurisdictions, and determine which issues were due to the biological response of the drug and which issues were imposed by the social and political setting. The comparative setting is crucial for this, and having just two countries still permitted substantial conclusions.

In the case of LLRW, there are many more siting studies, and thus an even greater chance to draw conclusions. The examples provided above provide a solid basis to establish the phenomenon of the data effect. This short paper also serves to validate the use of a comparative, retrospective approach. At this point, the analysis is in process, but the work in progress may interest those working in critical GIS and GIScience more generally.

This retrospective study covers a particular point in history that is perhaps hard to reconstruct for those who live awash in piles of data, freely disseminated on the web. The 1980s were a period during which GIS labored hard to build data resources. It was expensive, and time-consuming. Choices had to be made, since resources were not infinite. The LLRW projects were the kind of one-off project use that was typical of the time before the infrastructures were established though long hard effort. In the epoch of data scarcity, perhaps the data effect may be stronger since the cost comparison would be more pronounced. Another study would have to be mounted to see how a plethora of sources changes the effect.

The author acknowledges the assistance of Francis Harvey and Chloé Griot in compiling data sources used in this paper, as well as the support of the Canadian Social Sciences and Humanities Council for this current work as well as past support from the US National Science Foundation.

REFERENCES

Abraham J. and Courtney Davis C. 2007 Deficits, Expectations and Paradigms in British and American Drug Safety Assessments: Prising Open the Black Box of Regulatory Science, Science, Technology & Human Values, 32: 399-431.

Blakemore, M.J. and Harley, J.B. 1980 Monograph 26, Cartographica 17(4): 5-120.

Carver, S. J., 1991 Integrating multi-criteria evaluation with geographic information systems, International Journal of Geographical Information Systems, 5, 321-339.

Chem-Nuclear Systems. 1992 Pennsylvania Low-Level Radioactive Waste Disposal Facility Phase I (Screening) Siting Plan, Report S80-PL-001 Rev. 1, Chem-Nuclear Systems Inc.

Chem-Nuclear Systems. 1993 Pennsylvania Low-Level Radioactive Waste Disposal Facility Site Screening Interim Report Stage Two - Regional Disqualification, Report CNSI-PA-212, Chem-Nuclear Systems Inc., in Partnership with Pennsylvanians for Responsible Low-Level Radioactive Waste Disposal.

Chem-Nuclear Systems. 1994 Pennsylvania Low-Level Radioactive Waste Disposal Facility Site Screening Interim Report Stage Three - Local Disqualification, Report CNSI-PA-217, Chem-Nuclear Systems Inc. in Partnership with Pennsylvanians for Responsible Low-Level Radioactive Waste Disposal.

Chrisman, N.R. 2005 Full Circle: More than just Social Implications of GIS, Cartographica 40(4): 23-35.

Chrisman, N.R. 2010 Detecting the social influences in GIS technology: Getting value from comparative retrospective studies, paper presented at GIScience 2010, Zurich.

Connecticut 1991 Low-Level Radioactive Waste Disposal Facility Site Screening Report, Draft Report Connecticut Hazardous Waste Management Service.

Connecticut 1994 CHWMS believes volunteer approach best way to site LLRW disposal facility, CT LLRW Program Report, Connecticut Hazardous Waste Management Service.

Evans AJ, Kingston RP and Carver SJ 2004 Democratic input into the nuclear waste disposal problem: the influence of geographical data on decision making examined through a web based GIS, Journal of Geographical Systems, 6(2), 117-132

Harley, J. B. 1989 Deconstructing the map, Cartographica 26 (2): 1--20.

IAEA 2004 Disposal of low activity radioactive waste, Proceedings of an International Symposium, Cordoba, Spain, 13-17 December 2004, International Atomic Energy Agency, 497 p.

Illinois 1992 Martinsville, Report of the Illinois Low-Level Radioactive Waste Disposal Facility Siting Commission on its Inquiry into the Martinsville Alternative Site, Report 1418-350-12/92 Illinois Low-Level Radioactive Waste Disposal Facility Siting Commission.

Latour, Bruno 1999 Pandora's hope, Cambridge MA, Harvard University Press.

Martin, E. W. 2000 Actor-networks and implementation: examples from conservation GIS in Ecuador, International Journal of Geographical Information Science 14: 715-738.

National Geographic Society 2006 Lesson Plans: Low-level radioactive waste, http://www.nationalgeographic.com/xpeditions/lessons/18/g912/nuclearwaste.html

Newberry, William F. 1994 Comparative Approaches to Siting Low-Level Radioactive Waste Disposal Facilities, Report DOE/LLW-199, National Low-Level Waste Management Program, EG&G Idaho Inc.

Nuclear Regulatory Commission 1988 Guidance for selecting sites for near-surface disposal of low-level radioactive waste. Regulatory Guide 4-19. Office of Nuclear Reg. Research.

Nuclear Regulatory Commission 1991 Licensing Requirements for Land Disposal of Radioactive Waste. In Code of Federal Regulations, 10 CFR 61, p. 124-151. Washington DC: Government Printing Office.

Poore, Barbara and Chrisman, N.R. 2006 Order from Noise: Towards a social theory of information, Annals AAG 96(3): 508–523.

Schuurman, N. 1999 Critical GIS: Theorizing an emerging discipline, Cartographica 36 (4): 1–108.

Schuurman, N. 2000 Trouble in the heartland: GIS and its critics in the 1990's, Progress in Human Geography 24(4): 560-572.

Schuurman, N. 2006 Formalization matters: Critical GIS and ontology research, Annals of the Association of American Geographers 96(4): 726–739.

Sheppard, E. 1995 GIS and society: towards a research agenda, Cartography and Geographic Information Systems 22: 5-16.

Sheppard, E. 2005 Knowledge production through Critical GIS: Genealogy and Prospects, Cartographica 40(4): 5-21.

Sieber, R. E. 2000 Conforming (to) the opposition: the social construction of geographical information systems in social movements, International Journal of Geographic Information Science 14(8): 775-793.