

GEOGRAPHICAL PRACTICE AND THE ONTOLOGIES OF GEOGRAPHIC INFORMATION SYSTEMS

Michael R. Curry
Department of Geography
University of California, Los Angeles
Los Angeles, CA 90095

Presented at the Annual Meeting
Association of American Geographers
Pittsburgh, PA
4-8 April 2000

How do we find the ontological commitments in geographic information systems? That is to say, how do we determine what is real within such a system. I shall suggest here that there are three ways, coinciding with three ontological orders. And as might be imagined, those ways are not equally useful.

The ontology of institutionalized information

The first approach takes the perhaps prosaic view that what is taken to be real within a geographic information system is just what you put into it. Here we look in a sense at the supply end, and develop an ontology of institutionalized information. What is taken to be real within the GIS is just some subset of what is taken to be real within the world.

Hence, we if collect information about point output of pollutants in a given area, that will be what is taken to be real. Or if we collect information about residents of an area and their socio-economic characteristics, then those kinds of individuals and those characteristics will be real.

This does seem a prosaic view. But at the same time, it has the substantial advantage that it appeals to the very things that one does in creating any database. And it points to the dangers involved there: one fails to create categories for particular character-

istics and needs to add them later. Or one aggregates data in a particular way, and later needs them to have been aggregated in a different way.

And this view has a second advantage; it may also appeal to a typical way, and a very old one, of thinking about the world more generally, as consisting of objects that are characterized in terms of sets of attributes.

At the same time, one can appeal to this way of thinking about the ontological commitments involved in a GIS without this substance-attribute model; one can have a system that consists of particulars or other kinds of objects or agents. Indeed, this is just the notion of ontology that would be involved in what I have termed “place information systems,” where the formal structure of the system may consist of nothing more than records, and each one may be free-form, including audio, video, narrative, and so on, and where the creator of the system makes no prior claims about the truth of any of the material within the system.

Whatever its intuitive appeal, though, there are some obvious difficulties with this approach. It is inelegant. And just to the extent that it accepts as given a particular set of elements or types of elements, it is uncritical, and may be analytically useless. Still, I mention it because for a large number of situations such a system may be very useful indeed.

Normative ontologies

For various reasons, it generally seems to make sense to go beyond the bare-bones, first-order ontology and to construct a GIS using what amounts to a shop-built ontology. There, one decides at the outset what kinds of things are going to be allowed to exist, and constructs a data structure composed of those elements. And this, then, is a normative ontology, rather than a descriptive one. Here again, this might be an ontology that allows objects or agents or substances-with-attributes.

What would this be like? In fact, if we look to history as a model, we need only look to the medieval annals to begin to see what it might be like. In the *Annals of Saint Gall*, found in the *Monumenta Germaniae Historica*, we find the following:

709 Hard winter. Duke Gottfried died.
710 Hard year and deficient in crops
711
712 Flood everywhere
713
714 Pippin, Mayor of the Palace, died
715 716. 717.

.....

And so on (White 1981, p. 7)

Here there is a normative division of data, by year. But this points to something more. This excerpt is from Hayden White, who points out that it was only after the Middle Ages that people began to believe that there was a shape and direction to history, and to believe that it therefore made sense to think of history as a story. Today it is hard for us to read even this sketchy annals and not see some sort of story. But White is insistent; those who wrote this piece would not have seen it this way.

In fact, we know that in this second-order ontology something essential is missing. What is missing is the possibility of emergent properties and relationships. Indeed, and in fact, as soon as one begins to create categories of representations one opens the way for these elements. This is clear in one way in the move from the annals to the narrative, and in another as we develop visual representations of spatially arrayed data. And it seems clear that one today wants these sorts of emergent properties, although it should

also be clear that this need not be the case; if we wish to use a database simply as a catalog of elements, then such emergent objects become nuisances.

The ontology of GIS in practice

This, I would argue, should make it clear that we really need to move away from the view that we can usefully talk about the ontological commitments that are built into a GIS, without looking to the users of the systems. This, though, turns the question into one that requires the use of the various tools of the social sciences and humanities.

And here it becomes helpful to rethink the ontology in the GIS, to imagine a third-order ontology consisting of whatever one can read off of the representations that one can create. So if I create a visual representation, then I may have not only individuals but also clusters; I may have not only individuals with characteristics, but also patterns.

Now, if it were not already obvious, this should make it obvious to us that when we begin to talk about emergent objects as elements of an ontology we immediately move away from solid ground. Can we talk about a triangle as part of the ontology “built in” to a GIS? What of a crowd? Or a hole (Casati and Varzi 1994)?

But that doesn’t mean that the project can’t be done; rather, it only means that in carrying out such an ontology we need to move away from the view that we can achieve the level of certainty that seems to attend other forms of inquiry.

Let me give a couple of examples. First, take a geodemographic system. These systems typically begin with rather a wide range of data. There are census data, that introduce individuals, but where those individuals are characterized only in terms of characteristics--income between this and that, such and a such a race, and so on. And there are in fact only aggregates of individuals, where they are aggregated by location. In the end, we simply have a table of individuals with certain characteristics at a given place. Added to these data are households, each with a spatially defined location. And then

there may be names of individuals, associated with a location, then a household, and then the characteristics delineated in the census data. And finally, there may be information about the individual, including information about the objects that that person owns, as well as spending habits, travel, communications, and so on. More recently available may be information about the sites that that person visits on the internet, the amount of time spent at particular sites, and the files that are viewed or downloaded.

We might want to describe what I have just described as the first order ontology, an ontology this is simply a catalog of the stuff in the system. But of course the creators of geodemographic systems also create--this is the whole idea--a second order ontology, which creates emergent elements. So we have Margot and Elliott, from NDL:

Margot & Elliott

Married Households
Homeowners
Median Age = 56
Median Income = \$73,366
% U.S. Households = 8.1

The second largest group of the population, this segment is comprised of **established, affluent, empty-nest** households. Over 40% of all U.S. households reporting an income over \$75,000 are found in this segment. The vast majority are still in the work force (almost exclusively in **white-collar occupations**), and over half are **dual income households**. Their lifestyles have a distinctly **urban** character (over 80% of these households appear in large metropolitan areas). They are not averse to using credit to support their sophisticated interests, which include **foreign travel, wine connoisseurship, cultural activities, and gourmet cooking**. They are active investors in **stocks, bonds, and real estate**. While they are not especially active in the great outdoors, they do pursue **physical fitness**. Their sporting activities tend toward club sports like **golf and tennis**, and over half of them **walk for their health**, making them the top segment for that activity. **Avid book readers**, they also find time for **TV sports** and visiting with their **grandchildren**. They are the leading segment for both foreign and domestic **luxury car ownership**. They are most over-represented on the **west coast** and in the **northeast**.

Cohorts® Segment 1

Vignette

Margot checks the clasp on her heavy gold necklace one last time and smooths it so it lies flat. This is their third fund raiser this month, but for the first time she's ready before Elliott, thanks to his late tee-off. She glances at the clock. He'd better get out of that shower soon. It won't look good if the chairman of the board is late for the symphony benefit.

She reads the day's mail while she waits. A letter from their daughter, half crazed with midterms. Margot chuckles at her rantings—you'd think she was the first person to endure the rigors of med school. She puts the letter aside. Credit card bills, a catalog, a postcard from their son on winter break in Costa Rica. She checks the postmark. Only took three months to get here.

Elliott walks into the room, taking a final tug on a perfect bow tie. "Have you seen my cummerbund? Wow, you look great, hon." He bends to kiss her cheek.

Margot smiles. "The red one's on the bed, the black one's on the chair. Take your pick."

"Thanks."

Smoothing his wet hair, Elliott heads back to the bedroom. Margot picks up a magazine. In a minute he'll be back for help with his cuff links.

Note that they are first described as members of a particular group--married homeowners, affluent, empty-nesters. Then some of their purchasing habits are mentioned--foreign travel, wine, stocks and bonds. This is the second-order ontology that I mentioned above.

And then there is something else, a vignette, that lays out a brief narrative, a day in the life of Margot and Elliott. And what is interesting about this particular kind of sys-

tem is that the creator has in this way provided us with the third-order ontology. We are given a narrative account--and sometimes even a sketch--of the life of a person or family, but this is something imagined by someone on the basis of the second order ontology. (In the earliest geodemographic systems we had to do this step ourselves, but now it is done for us.)

There are several things to note about this third level ontology. One is that it constitutes what I would term a soft superorganicism. By that I mean the following. One longstanding way to think about culture imagines that people can be divided into groups, where a group has some sort of internal coherence. Racial and ethnic categories are examples of such groups, as are class categories, and so on. These are termed superorganic just because the category is imagined to transcend and to exist apart from any particular individual.

In geodemographic systems, though, we have what I term the soft superorganic just because the membership of a person in such a category--affluent homeowners--is contingent on a statistical aggregation of data, so that people are capable of easily moving from one group to another. Some might want to say, well of course, because that is exactly what the world is like. But in fact, that isn't universally the case.

Indeed, there are many groups for which membership is a very different matter, deriving from personal and family history, place of birth, and so on. And the point is, that if we imagine these geodemographic groups as arrayed along a continuum, with at one end groups whose membership is fluid, and at the other groups membership in which is very much fixed by the past, those who look at the systems will be unable on the basis of the data within the system to determine which group is which. So while the producers of the systems have tried to help us along by providing one version of a third-order ontology, users are forced to make various judgments about that ontology themselves.

And in fact, a lot is known about how these judgments are created. People pick certain elements from what is given to them, and take those elements as paradigmatic and symbolic, constructing narratives around those symbols. The narratives are constructed in part of shards from the user's own experience--personal, literary, cinematic, and so on. The important points here are two. First, this construction is inevitable. And second, just because the systems are designed for a certain body of users, these elements need to be seen as intrinsic elements of the systems. It will not do to see them as simply idiosyncratic, or beyond the control of the system designers.

Now let me turn to a second example. Consider a system that consists of a landscape and a set of agents, which behave in accordance with a set of rules. So a swarm-like model. We drop some sort of input into the system--an environmental change--and see how the agents respond, and how the landscape is changed.

In this case we might see the landscape and the agents as the elements of the first order ontology. Here the pattern that emerges as the model runs constitutes a second order ontology. What are we to say about of a third-order ontology? Here I would refer back to my earlier example, of the medieval annals, to note that a user of this system will inevitably create this third order ontology in terms of some kind of plot. We are just incapable today of coming up with an annals and leaving it at that.

There are a number of ways of characterizing these plots; one of them--in an earlier work White analyzed the writing of history in this way (White 1973)--is in terms of the standard literary tropes, of comedy, tragedy, irony, and so on. And White has shown for history in general, as others have shown for the history of science, that one can talk systematically about the propensity of individuals and members of certain groups--social, historical, and geographical--to organize their thinking about change in terms of one or

another of these tropes. So when we look to geographic information systems that incorporate some element of time, we need there too to look to the shape of that time.

Now one final point. One important way in which we judge anything is by its pedigree. In the case of a book, for example, it is of course commonplace, though that need not have been, that it has an author, about which things can be said. Just because of the partiality of that or any other any representation we use that pedigree as a kind of filler, to help us to decide whether a person's partial representation of the world is really one that can be taken seriously.

As we know, in the context of geographic information systems this pedigree function is in part taken up by metadata; ideally, at least, we use metadata to help us make sense of the quality of the data and the reliability of the person or organization involved in their collection or creation. That is, in describing the ontological commitments involved in a geographic information system one always makes some implicit or explicit reference to the provenance of those data. In an important sense, then, metadata operate at each of the three ontological levels that I have described.

But if we look back to the example of the book, or of social life more generally, we find that in important ways metadata fail to provide the information that one would like to have in making these judgments. This is in part, as Adam Seligman has shown, because the systems have emerged from within the context of a society in which the localized social practices that in one sense underpin the development of trust have disappeared (Seligman 1997). But it is in part because the format of metadata and their relative invisibility to users render them unable to offer the sorts of evidence--in the form of narratives, anecdotes, and the like--that are so important to the development of trust. And it is, finally, because those metadata almost inevitably refer not to the practices carried out by individuals, but rather to rules putatively appealed to by groups of individuals.

All of this means that each of the three ontological orders lacks a certain sort of clarity, just because the understanding of each relies on judgments that must themselves be made without a firm foundation. In the end, each requires an appeal as the basis of trust to extrinsic features, such as the institutional provenance of the system, and this tends, I would argue, to maintain existing systems of power and prejudice.

WORKS MENTIONED

Casati, R., and Varzi, A. C. 1994: *Holes and other superficialities*. Cambridge, Mass.: MIT Press.

Seligman, A. B. 1997: *The problem of trust*. Princeton: Princeton University Press.

White, H. 1973: *Metahistory: The historical imagination in nineteenth-century Europe*. Baltimore: Johns Hopkins University Press.

White, H. 1981: The value of narrativity in the representation of reality. In *On narrative* ed. W. J. T. Mitchell. Chicago: University of Chicago Press, pp. 1-23.