

Ethics of Network Data Collection: A Report on the Sunbelt Session

Charles Kadushin and Session participants

This is a summary of issues discussed in the February Cancun Sunbelt session on ethics and social networks. It is hoped that persons who were not able to attend will add further stories, issues, ideas and solutions. All this will be eventually written up for *Social Networks*.

The following is the nub of the issue:

Recent controversies over the role of social network analysis in counteracting terror have spurred an interest in the ethics of social network data collection and analysis. Unlike most social science observation, social network data in principle require tracking individual respondents or organizations (units) and their interrelationships. The main issue is under what circumstances can individual unit identities be revealed and when ought they not to be revealed. Some data are public data to begin with, but even here there is an issue over the adequacy of social network analyses to pinpoint the network location of individual units with sufficient accuracy to take actions that might have repercussions for those units. Do we really know enough accurately to pinpoint the “bad guys?” Moreno, who invented sociometry, insisted that it was not valid *unless* subjects knew that their answers would have consequences. Since Moreno, the general tendency has been not to reveal individual identities. The famous “Bank Wiring Room” omitted details that would have revealed the identity of the subjects. A practical issue is the reaction of literate “natives” to the revelation of the details of their interrelations. Finally, the shadow of the big gorilla in the room – institutional review boards with their concerns with privacy and possible harm – is ever present for academic researchers.

The session addressed the following questions that were put up as a slide presentation:

1. Personal experiences in data collection: what kinds of data have you collected, when, and under what circumstances?
2. Further:
 - Are there different consequences for different kinds of data collection?

- Has the climate and nature of network data collection changed over the past 20 years?
 - What differences have the context made: e.g. academic research, or government or corporate clients?
3. What have been your policies about confidentiality and disclosure under different circumstances and have they changed over the years?
 4. What has been the reaction of respondents to network questions?
 5. What has been the reaction of clients and people studied to the publication of your reports?
 - Do you prepare one kind of report for clients and another for general distribution?
 6. If members of an organization have been the subjects of the research, what kind of report should be made to them?
 - Is it different from the report given to those who paid for the research/
 7. What about data based on publicly available data?
 - Does it make a difference if the data are such that considerable processing is required, or in contrast, if the data are already in the form that they can be easily used for network analysis?
 8. Are there special considerations for defense or law enforcement analyses or consultation?
 9. What have been your experiences with Institutional Review Boards?
 10. Is there a bottom line to this discussion?
 - General principles?
 - Should INSNA take a formal stand on any of these issues?

To be sure, this was a long and complex agenda and even though the session ran overtime (the paper that was to follow had been cancelled), many of the issues were barely touched. Concerns about IRB reactions occurred at almost every point. On the other hand, although some persons present did do organizational consulting, no one spoke up who did mainly consulting, so the point of view of those not bound by IRB concerns was not represented.

The question was raised about the confidentiality of our own proceedings. I stated that we were operating under the premises of a focus group. Participants own positions would not be identified or attributed and confidentiality of what was said in the group precluded it being circulated outside of the group. Hence the remarks that follow, unless attributed to me, are not identified with any individual and details that might disclose the identity are omitted. By and large, comments are paraphrased, although the session was taped.

We began with a series of stories about experiences. I told about one of my first experiences with asking "Three best friends," in a study of Juilliard School of Music in the early 60's. This question appeared to be responsible for a rumor that I was trying to discover who was homosexual. Response rates in some departments suffered. In the late 60's I did studies of elites in Yugoslavia and later in Czechoslovakia and then on the American intellectual elite. They all contained data on who the respondent talked with about what that were used to construct very large sociograms. During the '68 student "revolution" on the Columbia Campus, I took all the data home for safekeeping, since I was fearful of the data on Yugoslavia, with names, being publicly circulated. When the Russians reentered Czechoslovakia in August after the famous democratic spring of 1968, the data were temporarily placed in a safe deposit box.

In discussion the issue was raised about roster style data collection (people's names are placed on a roster and the respondent is requested to identify the kind and extent of relationship with that person). If a person does not consent to participate in the network study as a respondent, what about having their name on a roster for people to respond to? In one case 90 percent of people who did not want their names on a roster themselves were willing to fill out the questionnaire. But symmetrizing a matrix can include people who otherwise did not respond or were not willing to be included.

What about gaining the consent of partners in an HIV study? In ethnographic field work with high risk populations what is one to do when one knows someone is HIV positive, has just injected himself and then another person wants to use the same needle who does not know that the first person is HIV?

What about ego network research? Do we have to ask permission of each person named in the network for him or her to be included in the research? One IRB insisted on this, thus effectively blocking the study. It was pointed out that one definition of a human subject is anyone who can be

identified, in which case consent may be required. Can a blanket exemption be obtained in cases in which the subject cannot be identified or the identification is locked up under special constraints as in survey research?

NSF in cultural anthropology is developing a document to deal with these complex issues because there are general problems with participant observation research. Meanwhile, IRB's are left in the lurch.

Most people agreed that their policies on collection and reporting have changed over the past several years.

Special problems: There are special issues for health providers. Special issues for government consultation. What do with people not capable of giving informed consent. These are all issues which need to be discussed and for which we do not now have clear answers.

Are there different standards under different conditions? Consulting may be different. Even if a university researcher but acting as a consultant, may have different standards. We may now be in the golden age in which respondents don't really understand the consequences of what they divulge. Once they do, we may have difficulty in data collection. There may be problems in the accuracy of reporting. My own policies (Charles Kadushin) are not to do network research if the results are to be used in personnel decisions. Further, I hold and own the data, not the group I am consulting for. If they are not comfortable with this, I don't accept the job.

In terms of publishing reports for whom, there is the possibility of giving each respondent personal feedback of where they stand in the group; they are giving you something and you are giving something in return. There is on the other hand no way of automating individual feedback of say, 1,000 node networks. [There are ways of automating 360 feedback, however]. In organizations, there may be consequences for refusing to participate either in 360 feedback or in network studies: those who do not participate are known and noted.

There are problems with minors, where rules are more strict, especially where alcohol and drugs are concerned. In detail, it comes down to whether the research is going to benefit people in general (risk vs benefit – a classic human subjects concern). There are problems in classroom network research where positive consent is now required. While many schools try to get blanket

consent at the beginning of the school year this may be satisfactory for much research, but in social network data collection, the systematic absence of some nodes can cause serious interpretation problems. Some schools did engage in a strong attempt to contact parents, but some parents felt safer in the default position of not giving consent. The result was that some students were singled out: when most of the class went to the computer lab to participate in the network study they were left behind. If schools allow the researcher to make the calls, then can use that opportunity to interview the parents and collect data on them to match with the students. But still may have some holes in the network.

Many IRB's (in the school context as well as other contexts) don't understand the network issues and have little experience with it and tend to view it negatively. This is one reason we (INSNA) may have to have some formal stands and inputs to IRB's. We have to educate IRB's on this. One participant has standard piece for IRB's explaining networks and what the legal issues are. Had sent this out to listserv a couple of years ago and got only two or three people who were interested.