



**Crime Mapping and
Data Confidentiality Roundtable
July 8-9, 1999**

**Sponsored by: National Institute of Justice,
Crime Mapping Research Center**

***"What is the appropriate model for partnerships between law enforcement agencies
and researchers with regard to data sharing?"***

by

Andreas M. Olligschlaeger,
System Scientist,

H. John Heinz III School of Public Policy and Management
Carnegie Mellon University

Just as researchers have a demonstrated need for access to as detailed crime data as possible, law enforcement agencies must not only protect the confidentiality of their information, but also the privacy of witnesses, victims, confidential informants and other innocent parties. The question is how to satisfy and protect the interests of all parties.

In the early days of crime mapping geocoded data most often consisted of flat files derived from 911 or record management systems, and contained attributes such as the type of crime, the time it occurred, police response time, etc. Most of this information is a matter of public record. Modern police information systems contain far more information than in the old days. They tend to be relational in nature and are often integrated with other systems, such as intelligence and case management. These systems have the potential to produce far more detailed (and, to researchers, extremely valuable) geocoded data that in the past might have been stored in the form of scraps of paper in a detective's notebook or desk drawer. Examples of such data would include the location of undercover sting operations, the home addresses of known gang members, and case management data on manpower deployment in an area. Clearly this type of geocoded information is far more sensitive than 911 or incident reports.

From past experience and conversations with other researchers it appears that historically, the sharing of geocoded data has occurred on an informal basis. The amount and detail of information given to researchers depends heavily on how well the law enforcement agency trusts the researcher in question. Indeed, sometimes law enforcement agencies find themselves in a position of having to give out more data than required because they do not have the resources or expertise to create data sets in the format that a research project requires. This definitely works in the favor of a

researcher that has a long standing relationship with a police department, but can make it difficult for young researchers or those new to the field to obtain quality data. A formal process, along with guidelines and appropriate legislation would certainly help, and would serve to protect both the disseminating agency as well as researchers.

The question is what should this process look like? Each research project has different data requirements, and the results of the research can be presented in varying degrees of detail (for example, pin maps vs. choropleth maps). For some data sets pin maps may be appropriate, for others not. For this reason it is important that any model can accommodate different types and sources of geocoded information on a case by case basis.

The following steps might form the basis of such a model:

Determine the type, format and nature of the data required to do the research.

The researcher(s) and the law enforcement agency should work together to ascertain what data will be needed. A part of this process should be determining the sensitivity of the data. While the results of the research might not be sensitive, the raw data might. If that is the case agreements (such as a non disclosure agreement) should be signed by the researcher guaranteeing that the confidentiality of the data will be upheld. If the research is federally funded, agreements could be backed up by civil and/or criminal penalties.

Decide how the results of the research will be presented.

While it might be appropriate to show address-level data for some data sets, it might not be appropriate for others. Determine which data, if any, can be displayed on a map (whether at the address level or aggregated) and published. Care should be taken to protect the privacy of any innocent individuals, as well as the integrity of sensitive law enforcement information. Law enforcement agencies should have a chance to review any research results before they are published.

Perform background checks on research personnel having access to confidential data.

This is already standard procedure for many law enforcement agencies, but is not always done.

Decide where confidential data is to be stored.

Sensitive or restricted data should always be kept on secure servers. Some agencies, while not averse to providing researchers with confidential data, do not want to release sensitive information outside the domain of their control. Often the open nature of university computing environments does not support secure storage. Indeed, universities are favorite targets for hackers. If secure storage is not possible in a university environment, perhaps the researchers could work in the law enforcement

agency on police equipment.

Destroy raw data once the research is completed.

This is a step often overlooked by researchers. Even if data is stored on a secure system, computers do get replaced, and it is easy to forget after a few years that the data is still there.