

**STATEMENT OF THE HONORABLE WM. LACY  
CLAY  
AT THE HEARING ON  
DATA MINING**

**MARCH 25, 2003**

Thank you Mr. Chairman. I would like to join you in welcoming the witnesses to today's hearing, and thank them for taking the time to share with us their knowledge on this subject. I am sorry that former Majority Leader Armey cannot be with us today. His defense of individual privacy during his career in the House is admirable. I am sure he would have added an important voice to this discussion.

I was pleased to read Mr. Rosen's testimony because it reflects by basic reaction to the issue -- data mining can be used well or badly, but it is all in how it is used. The more openness and oversight to the process, the less likely serious violation of citizen rights.

Let's be clear from the beginning. Data mining is profiling using computers and statistical models. We have all seen TV shows where the police have contacted the psychologists at Quantico and gotten a profile of the criminal that leads to his capture. We are also aware of individuals who have been wrongly arrested because they fit some profile. Indeed, innocent people arrested on a profile, have been wrongly convicted. One of the problems

with profiles is that they too often create the presumption of guilt. We have that same problem with data mining.

When credit card companies use data mining to track our purchases and then try to quickly stop fraudulent use of our cards, few people object. However, the government must be much more careful in using these techniques. First, much of what has been proposed for government use of data mining violates the basic principles of the Privacy Act. Second, when government uses these techniques, we have to be much more concerned with the cost of being wrong. If the credit card company is wrong, it often means nothing more than answering a phone call. When the government is wrong, the consequences are far greater.

Let me give you a simple example. One of the companies that produce face recognition software claims that they have an accuracy of 99.32%. Let's stop for a moment and think about what that means. About 20 million passengers pass through Dulles Airport each year. If we used this face recognition software to identify suspected terrorists, and no terrorists passed through Dulles at all, then 165,000 people would be stopped. Those people would be stopped and treated as terrorists, and the officials would be saying to themselves -- this guy has to be a bad guy. After all, this system is accurate more than 99% of the time. How would you feel if you were stopped as a terrorist, denied your rights, and subjected to the kind of interrogation we reserve for this kind of criminal?

The Transportation Security Administration (TSA) wants to create a system that uses data mining to give a terrorist score to every person who buys an airplane ticket. Those with high scores would be searched carefully. Those with low scores would go through the system with minimal screening.

To make matters worse, the TSA wants to keep the information on its data mining a secret. You won't know what information was used to create your terrorist score, and you will have no right to examine that information and correct errors. If you get a high score because of some mistake in the data or the computer program, you are stuck with it. If that makes traveling more difficult for you, you are out of luck.

Mr. Rosen proposes an oversight system for these kinds of security systems. I look forward to discussing that proposal. However, I would like to close with a thought from the world of cryptography -- the science of securing messages. In the 19<sup>th</sup> century, the cryptographer Auguste Kerckhoffs set down a principle that guide the most advanced work in cryptography today -- in good systems, the system should not depend on secrecy, and it should be able to fall into enemy's hands without disadvantage. In other words, the system should keep messages secret even if the enemy knows how the system works. That is the basic principle that underlies today's public key infrastructure. Unfortunately, that is not the principle that guides the systems being set up by agencies like TSA.