CS 325 - Reading Packet: "A few words on databases, society, and ethics"

A few words on databases, society, and ethics

This is a huge topic by itself, but I would feel remiss if we did not at least note that just because one *can* store something in a database, that does not mean that one *should*. Public policy, privacy concerns, efficiency, security concerns, commerce, and more all collide in inter-tangled ways when one starts really thinking about the issues in this area. When you start connecting databases to the Internet (or providing access to databases via applications available on the Web), these concerns are magnified; but even if a database is not connected to the Internet, there are still concerns. And you have seen, based on the power of relational algebra, how the possible questions you can ask about interrelated data stored in relations are essentially limitless. (And that's before one even adds in the concept of **data warehouses**, which may combine information from multiple databases as well as other sources!)

If you start thinking about just a FEW of the issues/questions related to this area, it doesn't take long to come up with a BIG list (and know that you've probably left out a lot!):

* Should there be limitations in storing some kinds of data?
* Should there be limitations in what data is accessible via the web, or via wide-area networks?

* Who should be able to have access to what data?

* What are the issues involved in a unified national database of citizens?

* Does an individual have the right to know what data is stored about himself/herself?
* Does an individual have the right to be able to challenge or correct inaccurate data stored about himself/herself?
    * Credit-reporting agencies store a great deal of data about individuals. Consumer advocates had to fight to win the right for individuals to receive a free credit report when they are turned down for credit based on the information in such databases, and to be able to contest inaccurate data then found -- before that, someone might have to pay just to find out what information is stored about himself/herself in such a database, and might not have recourse if one John Smith's default on a payment was recorded in another John Smith's record.

* Consider e-commerce web sites, or medical offices, or hospitals. What are "reasonable precautions" for a database administrator with regard to protecting data?
* How should financial data be protected?
* How should medical information be protected?
    * (At the same time, however, one can see that it might be very useful, for example, if an emergency room doctor could find out that an out-of-the-area patient who has just been rushed in has an allergy to latex, or has diabetes, etc.)

* Is the fact that someone was convicted of a crime public knowledge, or private knowledge?

    * (Consider Megan's Law, which provides for a publicly-accessible database of the locations of sex offenders -- http://www.meganslaw.ca.gov/)

* Is it OK for others to monitor, and compile databases of, your web shopping habits? what websites you visit? how long you spent at each of a list of web pages?
    * Consider the possibilities of Google Analytics (http://www.google.com/analytics/), or of RFID (radio-frequency identification) used in conjunction with databases

And that's just the tip of the iceberg.

Why mention this here (especially with so few answers to so many questions)? Because you may find yourself in the middle of some of these, given what you know about databases -- you could be asked to design a database or database application for someone, or serve as a database administrator for an organization. Perhaps you can be in a better position to affect public policy in these areas, given that you are now more aware of the possibilities of databases and database queries than the average person. At least, you should be aware that there are real concerns in this area, and will likely be lively debate on them for the foreseeable future.