

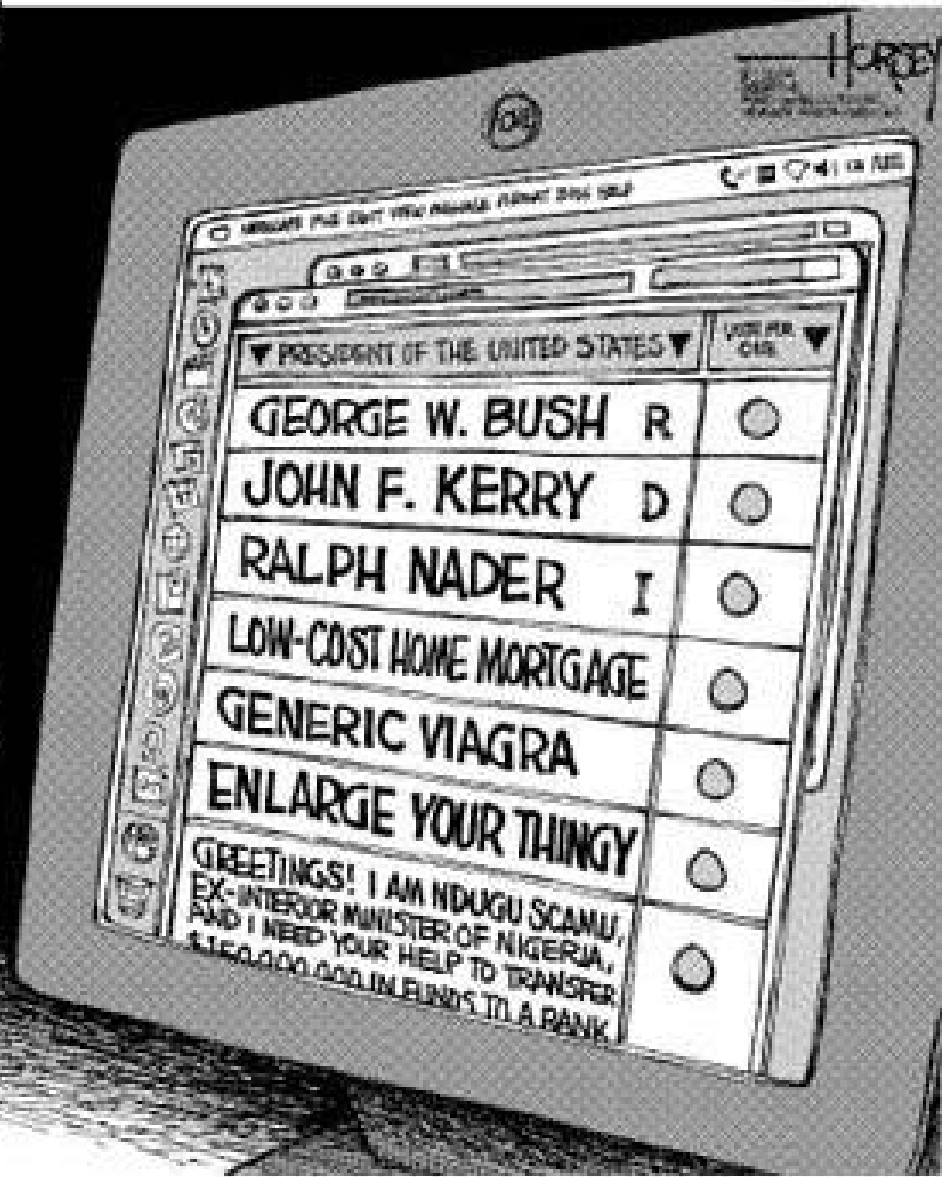
Privacy and Security

Barbara Simons

Outline

- Why we can't write correct software
- What is privacy?
 - Opposite of security? - tradeoff?
 - Fair Information Practices
 - Some early legal decisions
- Can privacy and surveillance coexist?
 - Total Information Awareness
- Voting - privacy & security issues

THE HAZARDS OF ONLINE VOTING...



Why we can't write correct software

- Halting Problem (Turing)
 - There exist problems with yes/no solutions for which we cannot compute the answer
 - It is impossible to determine for all programs whether or not they will halt on a particular input
- Cannot write a general purpose program to analyze correctness of all other programs
- Security implications

What is privacy?

- How to measure?
 - Percentage of privacy loss? - number of bits?
 - Different meanings in different cultures and even in same culture
 - e.g. positive pregnancy test
- Fair Information Practices
<http://www.privacyrights.org/ar/fairinfo.htm>

Fair Information Practices HEW '73

- Collection limitation
 - No secret collections of information
- Disclosure
 - You must be able to find out what information is being stored & how it is being used
- Secondary usage
 - You must be able to prevent information about you that was obtained for one purpose from being used for another without your consent

Fair Information Practices con't

- Record correction
 - You must be able to correct or amend a record of identifiable information about you
- Security
 - Any organization creating or using records of identifiable personal data must assure data reliability and must take precautions to prevent misuse of the data

Canadian Standards Association Model Code for the Protection of Personal Information '95

- Developed by consensus process that included Canadian Direct Marketing Association
- Became basis for Personal Information Protection and Electronic Documents Act
 - Approved 2000
 - Applies (among others) to businesses that trade data interprovincially & internationally starting Jan 2001

CSA Code

- **Accountability:** organization is responsible for personal information under its control
- **Identifying purposes:** must be identified by organization at or before information is collected.
- **Consent:** knowledge & consent required for collection, use or disclosure of personal information except where inappropriate.

CSA Code con't

- Limiting collection: to that which is necessary for the purposes identified by the organization - information collected by fair & lawful means.
- Limiting use, disclosure and retention: personal information to be used only for purposes for which it was collected unless consent for other uses is given - information retained only as long as necessary.

CSA Code con't

- Accuracy: personal information shall be as accurate, complete & up-to-date as necessary
- Safeguards: personal information shall be protected by appropriate security safeguards
- Openness: an organization shall make readily available to individuals specific information about its policies & practices relating to its handling of personal information.

CSA Code con't

- Individual access: upon request, an individual shall be informed of the existence, use & disclosure of personal info about the individual & given access to that info - an individual shall be able to challenge accuracy and completeness of info & have it amended as appropriate.

CSA Code con't

- Challenging compliance: an individual shall be able to challenge compliance with the above principles with the person who is accountable within the organization.

Some Early US Legal Decisions

- “Privacy” not in US Constitution
- Griswold v. Connecticut ‘65
 - State law restricting contraceptives (Hepburn)
 - Justice William O. Douglas: privacy found in “penumbra” of the Bill of Rights
- Olmstead v. United States ‘28
 - no protection against wiretap
 - Justice Brandeis: privacy the right to be left alone

Katz v. United States '67

- Olmstead reversed
- wiretap is search under 4th Amendment
- Justice John Harlan gave 2-part formula
 - Did person demonstrate expectation of privacy?
 - Would society consider expectation reasonable?
- 4th Amend. does not imply right to privacy

Whalen v. Roe '77

- NY State law required names/addresses of purchasers of some prescription drugs reported and stored in computer databases
 - Supreme Court did not find unconstitutional
 - Protection against unwarranted disclosures “arguably has its roots in the Constitution”
 - Justice William Brennan: might be future need to restrict technology used for computerized data
 - Justice Potter Stewart: no general Constitutional right to privacy, but some protections against gov't

Can Privacy and Surveillance co-exist?

DARPA Security with Privacy study

Total (Terrorism) Information
Awareness (TIA)



History of TIA

- Poindexter presentation to ISAT's (Inf Science & Technology) Security with Privacy panel 5/20/02
- Markoff NY Times article 11/9/02
- Safire NY Times article 11/14/02
- EPIC FOIA, Nov 2002
 - Privacy and civil liberties implications of TIA
 - DARPA study, Feb. 2003
 - Stated explicitly does NOT apply to TIA

History of TIA (con't)

- Internal and external (Newt Minow) oversight committees created, Feb 2003
- Congressional Research Service reports, March 2003
- Poindexter resigns, Aug 2003
- TIA “disbanded” Sept 2003
 - Some projects transferred to other agencies
 - Classified research?

What is data mining?

- Database queries?
- Attempt to determine what is relevant in a very large collection of data
 - If we know what we are looking for, don't need data mining - not targeted
 - Statistical analysis - hypothesis testing
 - Numerical tests?
 - Trying to predict human activities
 - Problems of modeling future attacks

USACM letter on TIA

<http://www.acm.org/usacm/>

Recommendation

- Rigorous, independent review that examines
 - Technical feasibility
 - Practical reality
- “Vast amount of information and misinformation ... likely to be misused to the detriment of many innocent American citizens.”

Database risks

- Immense databases are security & privacy risk
 - Compiled from financial, medical, educational, telephone, and travel records - many inaccuracies
 - Large quantity of sensitive info
 - Targets for malicious computer users, criminals, and terrorists
 - Problems of secure environment
 - Insider risks: domestic law enforcement, intelligence personnel, systems administrators, etc.

Security risks

- Identity theft risk from mega databases
- Examples of information theft
 - Help desk employee who used passwords from banks and credit companies to obtain PII of over 30,000 people in 3 years - sold info
 - Theft of computer disks from TriWest Healthcare
 - Info on over 500K military personnel and families
 - Soc sec numbers, medical claim history, etc
 - Blackmail or identity theft
 - Recent hacker break-in to U of CA computer

Privacy Risks

- Fair Information Practices
 - Prohibit secret databases and mandate fairness, accountability, and due process
 - Need for oversight and control especially great when collection of PII done without knowledge or consent
- “Privacy enhancing technologies” cannot protect privacy, since surveillance compromises privacy
 - Citizens could not verify that info about them correct
 - Lack of protection against harassment or blackmail

Economic risks

- E-commerce depends on privacy protection
- EU Data Privacy Directive could result in exclusion of US companies from EU based e-commerce
 - Companies may have to develop parallel systems to satisfy TIA in US and privacy regs elsewhere
- Identity theft costly to businesses and gov't

Personal risk - False Positives

- Incorrectly labeling someone a terrorist
 - Even exceptionally high accuracy rate would result in large number of false positives
 - Both terrorists and law abiding citizens would modify behavior
 - Would you purchase 1-way ticket with cash?
 - Lawful behavior might be avoided out of fear of being labeled terrorist
 - Waste resources checking out innocent suspects

False Positives

- Financial Services Technology Consortium credit card fraud analysis
 - 500,000 samples, 100,000 of them fraudulent
 - 20% false positive and 20% false negative rates
 - $.8 * 100,000 = 80,000$; $.2 * 400,000 = 80,000$
 - Therefore, half of all samples not fraud
 - "Credit card fraud detection using meta-learning: Issues and initial results" by Stolfo et al
- Suppose 500 terrorists out of 200,000,000
 - Same percentages 40,000,300 of which 400 terrorists
 - False positive 0.2%; false negative 5% = 400,474 of which 475 terrorists

Voting

Why is e-voting an issue now?

- Florida! - wrong conclusion reached by some
- Help America Vote Act (HAVA)
 - Almost \$4B for new voting equipment
 - Must replace punch card and lever machines by 2004 - can get waiver until 2006
 - National Institute of Standards and Technology (NIST) charged with setting standards
 - No money allocated

Computer based voting machines

Optical Scan

- Advantages
 - Cheaper than touch screen machines
 - Voter verifiable paper ballot
 - If precinct based scanner, can check ballot for overvote and undervote
- Disadvantages
 - Multi-lingual ballot can be a problem
 - Disabled voters?

Screen based systems

- Earphones for people with vision impairment and literacy problems
- Multiple language capability
- Avoids overvotes and warns of undervotes
- Voters can modify votes prior to finalizing their ballots
- Satisfy HAVA requirements for disabled

Ballot marking & generating Systems

- Optical scan machine used for tabulation
 - Could have earphones so blind voters could verify their ballots
- Vogue Election Systems
 - purchased by ES&S - now ES&S AutoMARK
 - Touch screen machine marks optical scan ballot
- Populex
 - Stylus used to mark screen
 - Prints optically scannable ballot

Direct Recording Electronic (DRE)

- Most have no voter verifiable audit trail
 - Ballot images printed at end of election!
- Because there is no audit trail, testing and security become incredibly important
 - Must be securely stored prior to, during, and after elections
 - Must be extensively tested
- Current testing and security grossly inadequate

Independent Testing Authorities (ITAs)

- Testing and results are secret
- Tests scripts - tests to inadequate standards
 - Does not do code review
 - Does not test Commercial off the shelf Software (COTS)
- Must test for likely bugs
 - Unlikely to detect clever Trojan Horse
 - If malicious code uses randomization, may not be able to determine if bug or intentional
 - May not be repeatable (because of randomization)

DREs

- Will be used by approx 30% of U.S. voters in Nov.
- Small number of vendors nationally
- Proprietary software (secret)
 - Independent computer security experts not allowed to view or test software
 - Code held in escrow not sufficient
 - Independent experts not allowed to examine code

DREs that produce paper ballots

- Avante: Vote-Trackker
 - Ballot under glass - cannot be touched by voter
- AccuPoll
- Both had paper ballots as part of initial design
- Sequoia retrofitted
 - Ballots stored sequentially on paper role
 - Privacy concerns

DREs that typically do not produce paper ballots

- Diebold
- ES&S
- Sequoia
- Hart Intercivic
- Some smaller companies

Horror stories

Boone County, IN 11/4/03

- Initial election results: > 144,000 ballots cast in a county with < 19,000 registered voters
 - Only 5,532 actually voted.
- County Clerk Lisa Garofolo: problem caused by a “glitch in the software”.
- Ballots are electronic; no paper ballots that could be used to verify the computerized election results accurate once the “software glitch” was fixed.

Broward County, 1/8/04

- Special election **only** to fill House seat 91
- 134 voters used ES&S DRE without casting any votes for any candidates
- 7 candidates, Ellyn Bogdanoff beat Oliver Parker by 12 votes
- Broward Mayor Ilene Lieberman calling for paper ballots

Voter Verifiable Paper Ballots

- Voter must be able to verify ballot
 - Ballot deposited in a secure ballot box.
 - Voter can't keep it because of possible vote selling.
 - Ballot handling and counting must be observable.
 - Manual recounts *must* be performed.
 - When elections are suspect.
 - When candidates challenge.
 - Randomly, even when elections go smoothly.