

VANDERBILT
School of Medicine

More than a Matter of Trust: Security & Privacy in Voter Registration Records

Presented at the National Academies Committee on State Voter
Registration Databases Workshop II

Bradley Malin, PhD (b.malin@vanderbilt.edu)
Assistant Professor of Biomedical Informatics, School of Medicine
Assistant Research Professor of Computer Science, School of Engineering
Vanderbilt University
November 30, 2007

VANDERBILT
School of Medicine

Questions

- *What principles should guide security decisions? How might these apply to voter registration databases?*
- *What privacy considerations need to be taken into account, especially with the impact of combining and linking data?*
- *What standard, adversarial test could be applied against each state's database? What would you include in such a test?*

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 2

VANDERBILT
School of Medicine

Question 1

- *What principles should guide security decisions? How might these apply to voter registration databases?*
- Principles of systems and data protection are dependent on who has access.
- There are multiple environments to consider
 - Private access
 - Collection and use of personal information to maintain accuracy and prevent fraud
 - Public access
 - political and related uses; or non-marketing

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 3

VANDERBILT
School of Medicine

Public Information Sharing

- Example: Washington
 - If you are a voter, your name, address, political jurisdiction, gender, date of birth, voting record, date of registration, and registration number are public information under state law. (RCW 29A.08.710)
- This is public record by law and does not violate security or privacy, however ...

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 4

VANDERBILT
School of Medicine

Behind Closed Doors

- More than public information for registration of voters
- NAS Committee: Recommend how to evolve and maintain voter registration databases in a way that enables sharing of information with other states securely and accurately in fulfillment of HAVA
- Example: Pennsylvania
 - "Statewide Uniform Registry of Electors", county election officials have direct access to the centralized statewide database
 - The state uses "identifying number, name, and date of birth" for linking to motor vehicle and/or Social Security records
 - Use a hybrid match: the number and first two characters of the last name must match exactly, with discretion left to the county commission to determine if the rest of the record is a match
 - Currently uses the AAMVA (American Association of Motor Vehicles Administrators) criteria to match information with SSN digits: exact match of the SSN-4, first name, last name, month of birth, and year of birth.

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 5

VANDERBILT
School of Medicine

Several Principles

- People are inherently curious.
- Account for the hacker, but don't forget about the insider

<ul style="list-style-type: none"> ■ "Easier" Goals <ul style="list-style-type: none"> □ Authenticate the users □ Access Control □ Logs and Audits □ Encrypt data in transmission □ Encrypt data at rest □ Try to minimize data on mobile devices <ul style="list-style-type: none"> ■ Passwords, encrypt, etc. 	<ul style="list-style-type: none"> ■ "Harder" Challenges <ul style="list-style-type: none"> □ Define use cases □ Specify workflows □ Define roles □ Develop Standards □ Learn or specify baselines of systems use
---	--

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 6

VANDERBILT School of Medicine

Secure Data Sharing

- NAS Committee: Recommend how to evolve and maintain voter registration databases in a way that enables sharing of information with other states securely and accurately in fulfillment of the Help America Vote Act of 2002
- Is data sharing centralized or distributed?
- What information is being shared?
 - Is it public information (names, dates of birth, etc.) or private information (Social Security Numbers)?

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 7

VANDERBILT School of Medicine

Private Information Sharing

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 8

VANDERBILT School of Medicine

Private Information Retrieval

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 9

VANDERBILT School of Medicine

What Data is Being Shared?

- Does voter registration, and other personal information, need to be compared in the clear?
- Computer science research has shown that comparison of encrypted records is possible.
 - Secure multiparty computation (e.g. Clifton; Wright; Nissim)
 - Need for standardized tools

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 10

VANDERBILT School of Medicine

Questions

- *What privacy considerations need to be taken into account, especially with the impact of combining and linking data?*
- Concerns are dependent on the realm
 - Information disclosed to the public is different than information used by the private data holders
- Threats exist primarily because we do not consider the availability of other databases

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 11

VANDERBILT School of Medicine

Privacy Concerns

- Privacy violations via inference: *voter registration history*
- Privacy violations via linkage: *voter registration lists*
- Untargeted Identification (UI): Identify any individual
 - Mass UI: Identify as many individuals as possible
- Targeted Identification (TI): Identify a specific individual
 - Mass TI: Identify as many pre-specified individuals as possible
- Ability to achieve attacks is dependent, in part, on the economics of data access

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 12

VANDERBILT School of Medicine

Information Database CD-ROM - Mac/PC Interface

Mac/PC State Secretary of State

Order Monthly Voter Registration Database CD-ROM

To order a monthly subscription of the Voter Registration Database CD-ROM you may either **download the order form** or enter the required information below and pay with your VISA or MasterCard.

Format Options

Choose from the following options: Diskettes CD-ROM Microsoft Access

Subscription Options

Subscription monthly price: \$30.00
 Subscription frequency: [1] times
 Subscription total: \$30.00

Billing Information

* First Name:
 * Last Name:
 * Email:
 * Address:
 * City:
 * State: Please Choose
 * Zip:
 * Phone:

Shipping Information

Nat. Acad. Voter Reg. Wksp. 13

Washington

VANDERBILT School of Medicine

Scott County, Tennessee Election Commission - Voter Registration Lists - Mac/PC Interface

Scott County Election Commission

VOTER REGISTRATION LISTS PURCHASE BY CITIZENS
 Phone: 423-663-2210

Voter registration records are updated daily. Voter registration lists may be purchased by any citizen who carries on a form provided by the State Election Commission that the list will be used for political purposes only. The forms are available from county election commission offices. Upon receipt of the completed form and payment of the cost, the election commission will prepare the information.

The Scott County Election Commission can provide lists of registered voters by district, precinct, ward, or a county-wide list can be prepared. Lists may contain voters' names, address, voting history etc. We do not include social security numbers or phone numbers with voter information. Information may be obtained in the following formats:

Printed lists - \$.25 per page
 Computer generated disc (floppy or CD) - Excel spreadsheet
 \$25.00 setup fee
 \$1.50 per floppy or CD
 \$1.50 administration fee per floppy or CD
 \$38.00 total
 Address labels - \$.20 per label

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 14

Tennessee

VANDERBILT School of Medicine

Privacy Registration Lists and Files - Mac/PC Interface

Water Registration Lists and Files

Water Registration Lists and Files

File Name	File Size	Cost of	Cost of
		Registration List	Registration File
Waterwide Voter File		\$500	
Congressional, State Senate, State House, or Judicial District	\$ 10,000 - \$50,000	\$50	\$250
	\$50,001 - \$100,000	\$75	\$400
	\$100,001 - \$200,000	\$100	\$600
	\$200,001 - \$500,000+	\$250	\$1,250
Geographic Unit of Interest	\$ 10,000 - \$50,000	\$50	\$250
	\$50,001 - \$100,000	\$75	\$400
	\$100,001 - \$200,000	\$100	\$600
	\$200,001 - \$500,000+	\$250	\$1,250
Municipal Unit of Interest	\$ 10,000 - \$50,000	\$50	\$250
	\$50,001 - \$100,000	\$75	\$400
	\$100,001 - \$200,000	\$100	\$600
	\$200,001 - \$500,000+	\$250	\$1,250
County, or Municipal Election Precinct	\$ 10,000 - \$50,000	\$50	\$250
	\$50,001 - \$100,000	\$75	\$400
	\$100,001 - \$200,000	\$100	\$600
	\$200,001 - \$500,000+	\$250	\$1,250
Special Suburban Districts	\$ 10,000 - \$50,000	\$50	\$250
	\$50,001 - \$100,000	\$75	\$400
	\$100,001 - \$500,000+	\$250	\$1,250

You will need a voter file in order to associate a voter with his/her voter history. The voter and history files both contain the voter registration number, and can be linked on this field.

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 17

Georgia

VANDERBILT School of Medicine

Privacy Violations By Inference

Name	Date	Number of Voters
Catoosa	9/18/2007	1
Cobb	9/18/2007	1
Clayton	9/18/2007	1
Lee	11/06/2007	1
Gwinnett	9/18/2007	1
Dekalb	9/18/2007	3
Chattahoochee	11/06/2007	3
Sumter	11/06/2007	3
Seminole	11/06/2007	4
Charlton	11/06/2007	5
Dodge	9/18/2007	6
Mitchell	3/20/2007	7
Dawson	9/18/2007	7

Registration number 76686

When it is revealed how Catoosa county voted in this election (aggregate results), then we uniquely link this voter to their vote.

Georgia Voting 2007 History

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 16

VANDERBILT School of Medicine

Principle

- Census and other government agencies have developed tools for analysis and protection of contingency tables for almost half a century
- Disclosure control
 - Statistical: suppress and/or permute voter histories with small cell counts (e.g. Fienberg; Duncan; Domingo-Ferrer)
 - Computational: generalize geocodings until cell counts are above minimal threshold (e.g. Sweeney)

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 17

VANDERBILT School of Medicine

Privacy Violations By Linkage

(Sweeney 1997, 1998)

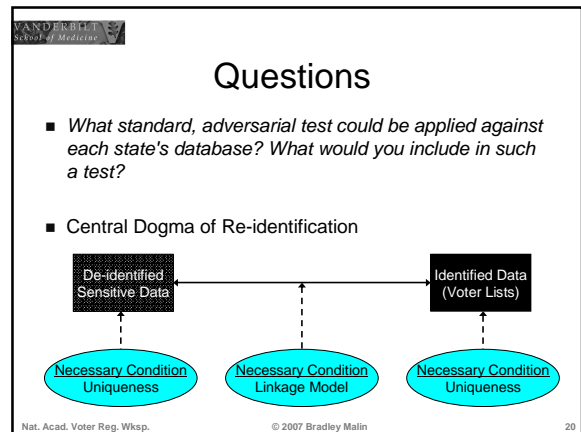
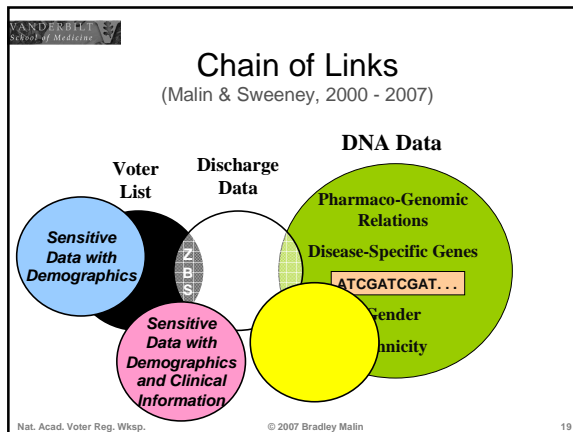
Discharge Data: Ethnicity, Visit date, Diagnosis, Procedure, Medication, Total charge

Voter List: Name, Address, Date registered, Party affiliation, Date last voted

Intersection: Zip, Birthdate, Sex

87% of the United States is RE-IDENTIFIABLE

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 18



Policy

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 21

Policy

RCW 29A.08.740
Violations of restricted use of registered voter data? Penalties? Liabilities. (Effective January 1, 2006.)

- (1) Any person who uses registered voter data furnished under RCW 29A.08.720 for the purpose of mailing or delivering any advertisement or offer for any property, establishment, organization, product, or service or for the purpose of mailing or delivering any solicitation for money, services, or anything of value is guilty of a class C felony punishable by imprisonment in a state correctional facility for a period of not more than five years or a fine of not more than ten thousand dollars or both such fine and imprisonment, and is liable to each person provided such advertisement or solicitation, without the person's consent, for the nuisance value of such person having to dispose of it, which value is herein established at five dollars for each item mailed or delivered to the person's residence. ...
- (2) Each person furnished data under RCW 29A.08.720 shall take reasonable precautions designed to assure that the data is not used for the purpose of mailing or delivering any advertisement or offer for any property, establishment, organization, product, or service or for the purpose of mailing or delivering any solicitation for money, services, or anything of value. However, the data may be used for any political purpose. Where failure to exercise due care in carrying out this responsibility results in the data being used for such purposes, then such person is jointly and severally liable for damages under subsection (1) of this section along with any other person liable under subsection (1) of this section for the misuse of such data.

[2005 c 246 § 19. Prior: 2003 c 111 § 249; 2003 c 53 § 176; 1999 c 298 § 2; 1992 c 7 § 32; 1974 ex.s. c 127 § 3; 1973 1st ex.s. c 111 § 4. Formerly RCW 29.04.120.]

Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 22

- ### Technology + Policy
- Must combine security and privacy risk evaluations with economic and policy deterrents
 - Current regulations are insufficient: nothing to prevent the linkage of this information with other records.
 - Legal protection from marketing, but does not prevent an adversary from linking a voter registration record to financial, health, and other types of information
- Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 23

- ### Principles
- What standard, adversarial test could be applied against each state's database? What would you include in such a test?
 - Develop a threat model
 - Model what other databases in the state are available?
 - Document what type of information the records contain
 - Be aware of who has access to these records
 - Rank risks by
 - Accessibility – How many people potentially have access to the data?
 - Sensitivity – What do the available records communicate about the individuals?
 - Cost – What are the economic barriers to access and linkage?
- Nat. Acad. Voter Reg. Wksp. © 2007 Bradley Malin 24