Might Privacy and Security Issues Frustrate National Health Information Technology Initiatives?

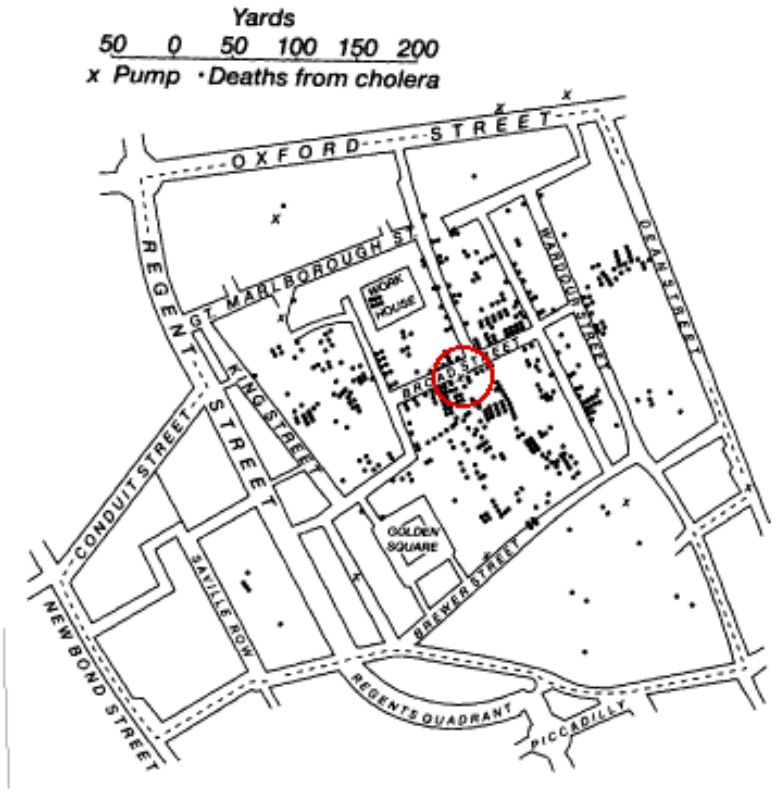
The Technology Perspective

Kenneth D. Mandl, MD, MPH

Harvard Medical School
Center for Biomedical Informatics

Children's Hospital Informatics Program at the Harvard-MIT
Division of Health Sciences and Technology

John Snow and the Broad Street pump

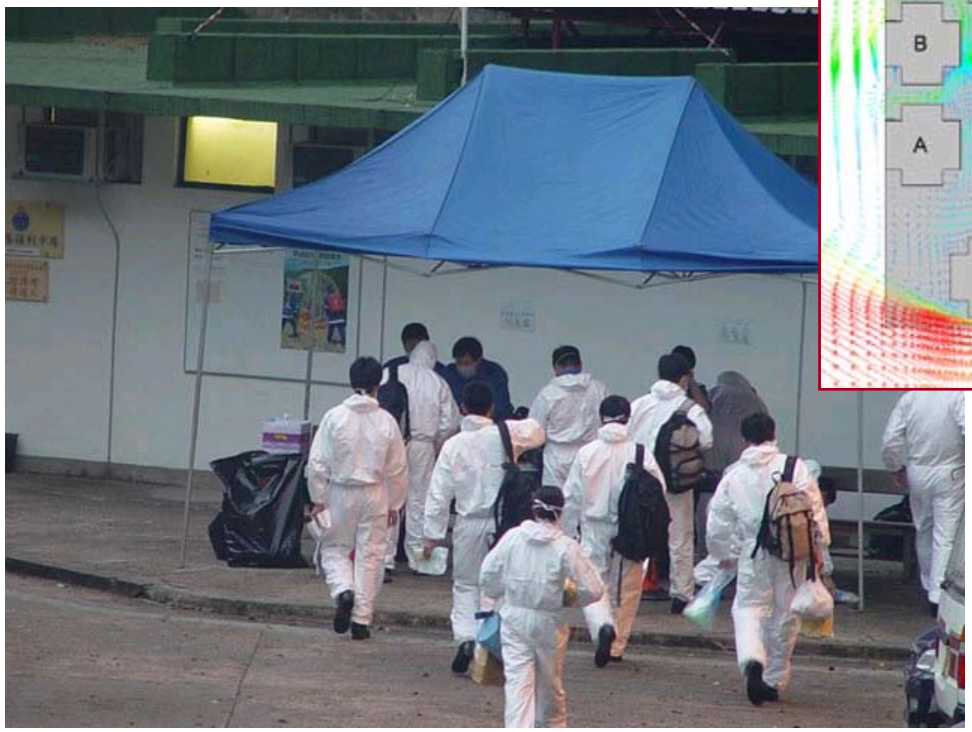


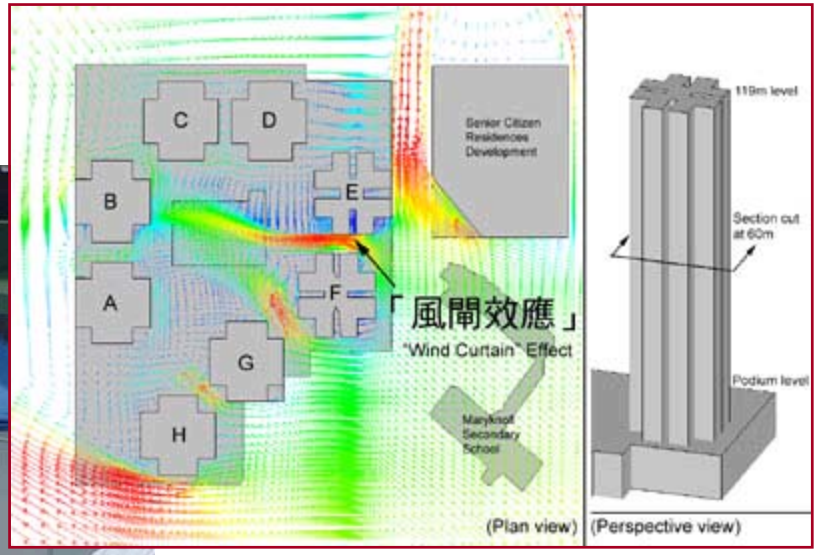
Tradition of mandatory reporting

- Some data should flow freely in the NHIN
  - ✓ E.g., data for mandatory infectious disease reporting

- Mandatory reporting of disease involves full identification of the individuals
- Little public debate about the mandatory reporting of
  - ✓ Cholera
  - ✓ Measles
  - ✓ Syphilis
  - ✓ Neisseria meningitidis

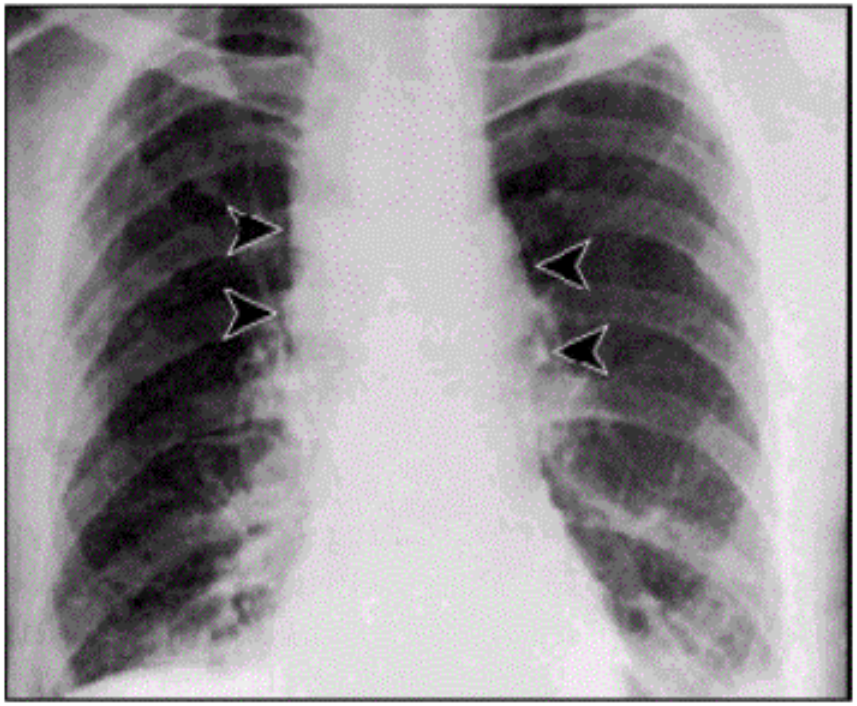
But, we want to find the next Amoy Gardens





This, however, requires a data-mining approach

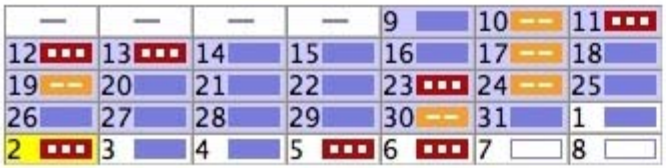
How Anthrax drove the technology



Early detection!!

Focus shifted to:

- Real time
  - $$$ Investment
- Data processing
  - New kinds of data
  - Monitoring many patients to detect patterns



SELECT HOSPITAL/SYNDROME

	BIDMC	CAMB	CHB	MGH	SOMER	WH
RESP						
GI						
RASH						
NEURO						
HEMOR						
ALL						
INJ						

SELECT ABBERRATION

Aberrations For Current Selection

BIDMC	RESP	S 4/2	details	export

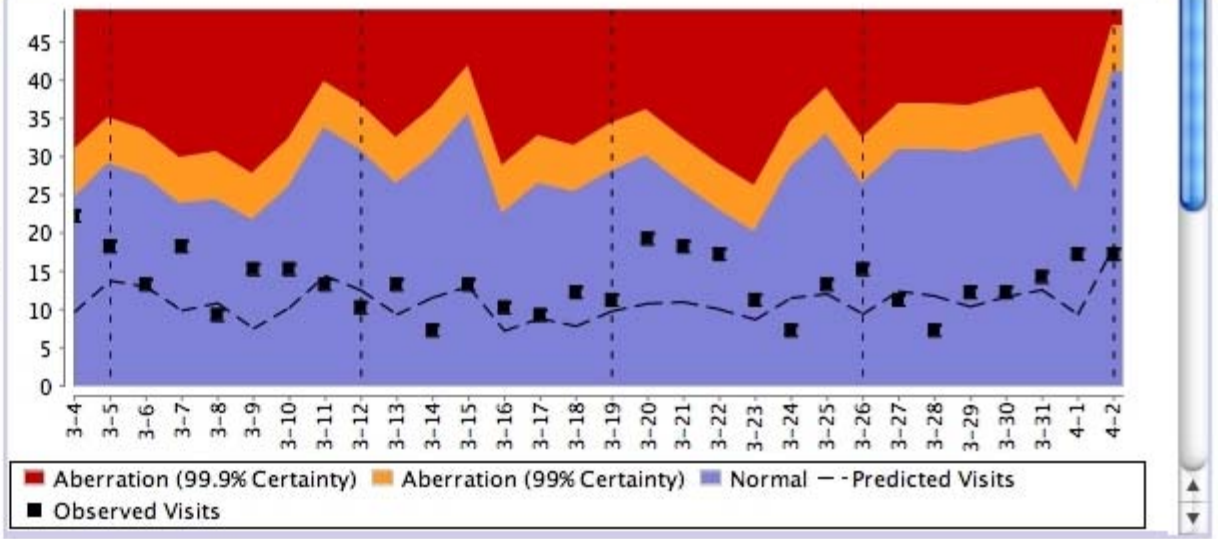
Other Aberrations

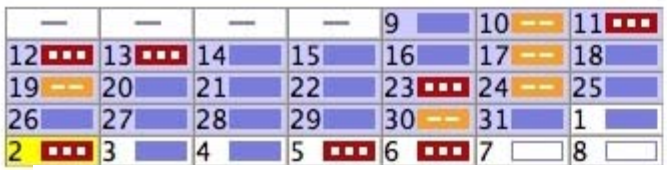
MGH	RESP	S 4/5	details	export
BIDMC	RESP	S 4/2	details	export
CHB	GI	S 3/11	details	export
CAMB	RESP	S 3/23	details	export

LEGEND

Active	Inactive	
		- Normal
		- Aberration (99% Certainty)
		- Aberration (99.9% Certainty)
		- Result not yet available
S/T		- Spatial/Temporal Aberration

TEMPORAL





So, how do we find disease outbreaks *and* protect privacy?

SELECT ABBERRATION

Aberrations For Current Selection

BIDMC	RESP	S 4/2	details	export

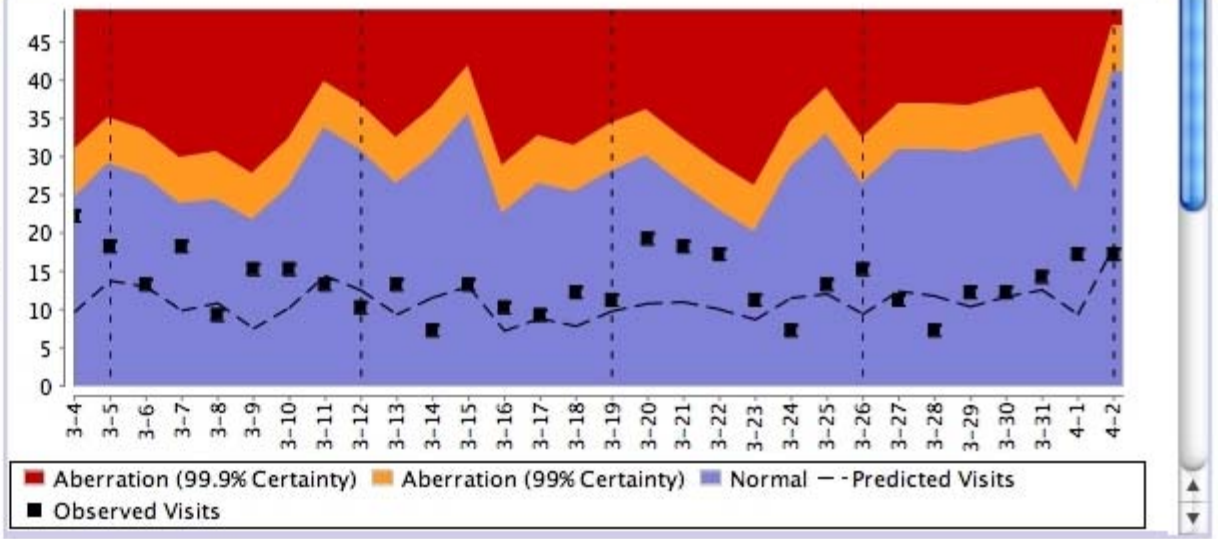
Other Aberrations

MGH	RESP	S 4/5	details	export
BIDMC	RESP	S 4/2	details	export
CHB	GI	S 3/11	details	export
CAMB	RESP	S 3/23	details	export

LEGEND

Active	Inactive	
		- Normal
		- Aberration (99% Certainty)
		- Aberration (99.9% Certainty)
		- Result not yet available
	S/T	- Spatial/Temporal Aberration

TEMPORAL



New imperatives and opportunities for data exchange

- Public health went from a data-poor enterprise, to one in which there is increasing data sharing with health care
  - ✓ This is important, because doctors and health care institutions (who have the data) do not focus on public health issues
  - ✓ So how do we handle this sharing?

- As the NHIN emerges, we have the opportunity to think carefully about preserving privacy

Why care about privacy?

- Health care data are very disclosing, e.g., a medication list
- Concern about linkage—employer-based health care, life insurance, stigmatizing conditions
- Secondary uses of healthcare data are often not restricted—e.g., pharmacy data
- Banks can put $ back into your account, and plan for fraud

Five principles

- Do not rely on technology alone—need rules, regulations, policies, legislation
- Allow strong institutional control
- Allow strong personal control
- Obscure the patient identity
- Err on the side of data security over efficiency

1. Policy

Critical to drive
and to complement technology

Policy

- Limit accesses to authorized individuals
- Educate those individuals about risks
- Implement regulations to enforce good behavior
- Strictly control on secondary uses of data
- Use IRB's whenever possible
- Consider a public health version of the IRB process
- Legislate to protect insurability—to reduce the overall privacy implications of disclosure

2. Institutional control

Follows from "policy" principle—health care institutions, heavily regulated, are enforcers of policies

Institutional control

- It is technically very difficult for each piece of information to travel with the policies around consent in perpetuity
- What leaves the institution is the institution's responsibility regardless of whether it going to
  - ✓ Public health
  - ✓ Personal health record
  - ✓ Research project (best developed framework)
- This approach leverages institutional control over employees, institutional enforcement of policies, implementation of audit trails etc.

Institutional control

- A corollary of "Institutional control" is to always share only the minimal dataset
- Technology must allow sharing of minimal data with reach back capability
  - ✓ This requires a distributed database with robust authorization and access controls

Institutional control

- e.g.—for biosurveillance, work with "de-identified data" to detect aberrations, and then dig back in—WITH PROPER AUTHORITY--when investigation is required
  - ✓ coming up—what does de-identified mean?

- For this, we use peer-to-peer architectures

3. Personal control

Models for allowing the patient to control access

Personal control

- Giving control to institutions can facilitate personal control—institutions can enforce the wishes of their patients
  - ✓ Simplest model is opt in and out at initial consent
- Another model is for institutions to release information to patients in *containers* called personally controlled health records. Then the patients can themselves handle consent and access.

Personal control

- The Indivo Health project, formerly PING, being rolled out in several test beds including
  - ✓ **MIT Medical**
  - ✓ Harvard University Health Services
  - ✓ HP
  - ✓ MA Share
  - ✓ Children's Hospital Boston
- E.g., a patient might make data available for
  - ✓ Public health
  - ✓ Research
  - ✓ Post-marketing surveillance (see web.mit.edu/cbi/)

4. Obscure the patient identity

Why take chances?

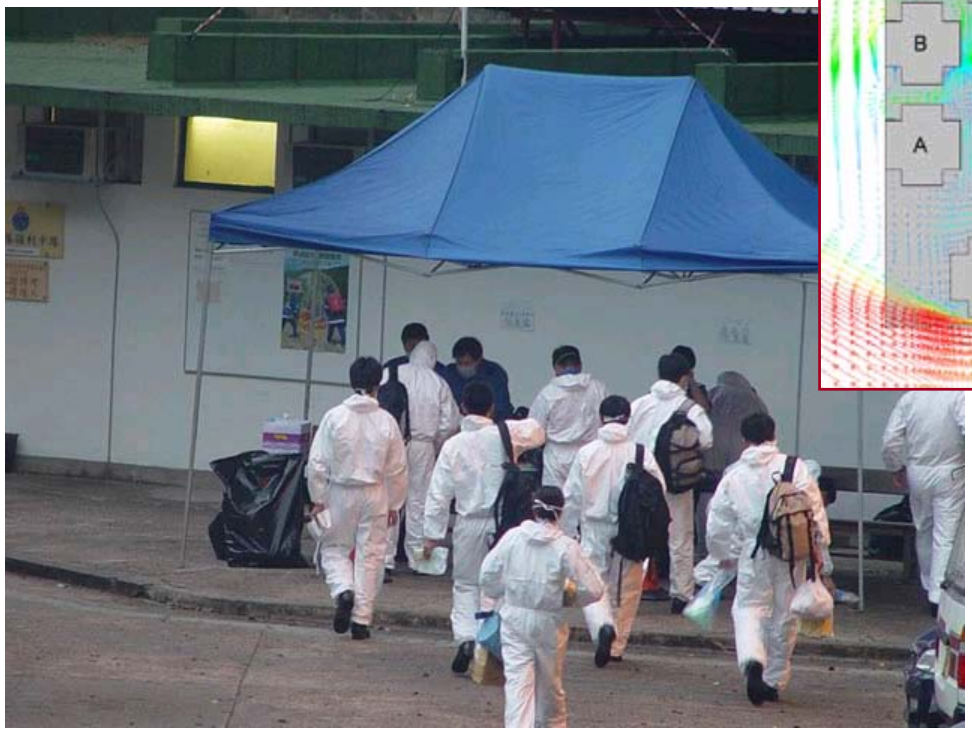
Obscure the patient identity

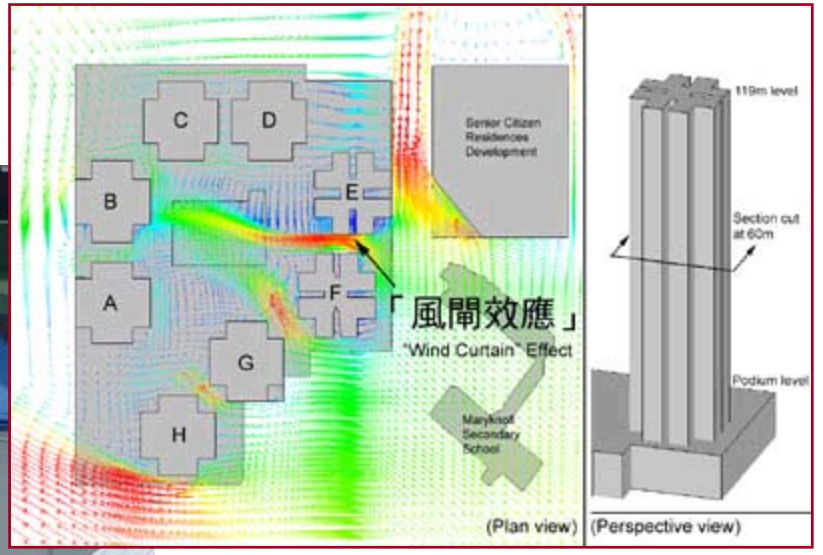
- Sweeney *--date of birth*, *gender*, *5-digit ZIP* combine to identify 87% of the US population

- Emerging issues--spatial data—a newer data type for the health care industry, increasingly used in surveillance

Obscure the patient identity

- We want to find the next Amoy Gardens

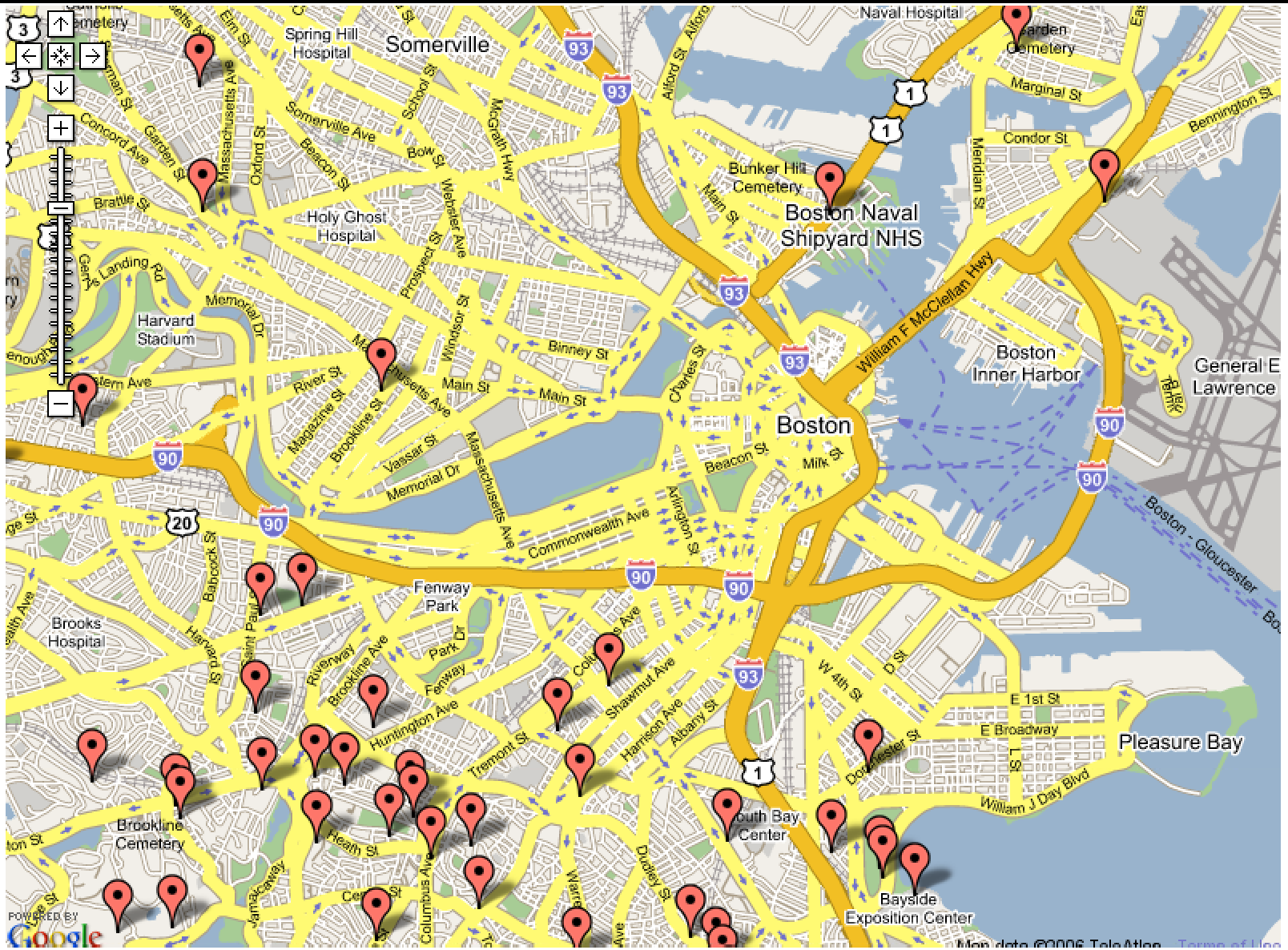




Most surveillance systems use zip codes—which lowers the resolution

Obscure the patient identity

- But point location data yield a superior spatial clustering detection
- Yet, point location data are very revealing of identity



Cassa et al *JAMIA* 2006

5. Encryption

Protect against failures of the first four approaches

Encryption

The New York Times
nytimes.com

May 31, 2006

Stolen VA Data Goes Beyond Initial Reports

By THE ASSOCIATED PRESS

Filed at 6:40 p.m. ET

WASHINGTON (AP) -- Personal information on 26.5 million veterans that was stolen from a Veterans Affairs employee this month not only included Social Security numbers and birthdates but in many cases phone numbers and addresses, internal documents show.

Meanwhile, VA Secretary Jim Nicholson said Wednesday that he had named a former Arizona prosecutor as a special adviser for information security, a new three-month post that will pinpoint security problems at the VA and develop recommendations for improvements.

The three pages of memos by the VA, written by privacy officer Mark Whitney and distributed to high-level officials shortly after the May 3 burglary, offer new details on the scope of one of the nation's largest security breaches. The memos were obtained Wednesday by The Associated Press.

They show that a file containing 6,744 records pertaining to "mustard gas veterans" -- or those who participated in chemical testing programs during World War II -- was breached, and that a "short file" with as many as 10 diagnostic codes indicating a veteran's disability also was stolen.

At the same time, however, the memos suggest that the data might be difficult to retrieve by thieves.

"Given the file format used to store the data, the data may not be easily accessible," stated one memo dated May 5 and distributed internally

- Here, encryption of data would have helped enormously
  - ✓ Ping model—individually encrypted records



Children's Hospital
Informatics Program

Harvard
Medical School