Safeguarding Confidentiality in GPS Travel Data Collection







Introduction



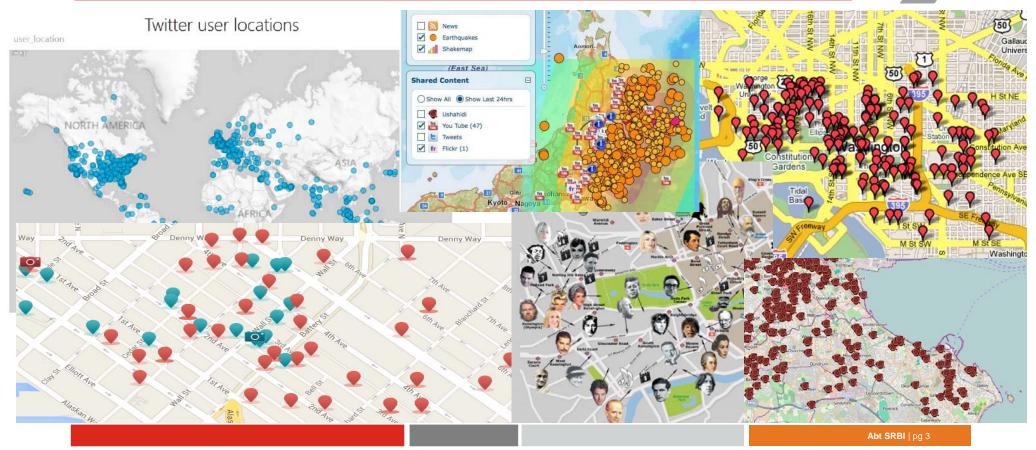
Abt SRBI

- National survey / data collection firm, founded in 1981.
 12th largest research firm in USA (by annual revenue Honomichl List)
- Headquarters in NYC, offices in DC, Chicago, Boston, Arizona, North Carolina, Florida.
- Transportation, Social Policy, Market Research, Health, Energy, Elections, etc.
- Local NYC Transportation clients: LIRR, MetroNorth, PANYNJ, and NYMTC.
 Onboard LIRR / MetroNorth surveys, Origin & Destination, Travel Surveys, etc.

Abt SRBI Geographic Information Systems (GIS) group

Timothy Michalowski, Senior Statistical GIS Analyst Dara Seidl, Statistical GIS Analyst







- The keeping private of one's geographic location.
- An individual's right to control access to his/her personal location data, including how and to what extent it is shared with others (Kwan et al. 2004; Kar et al. 2013).









Legal

- Privacy Act of 1974 –
 Protects personal records under control of federal agencies from being disclosed without written consent
- Health Insurance Portability Accountability Act 1996 (HIPAA) Protects confidentiality of health information collected in research, as well as informed consent



- Institutional Review Boards (IRBs)
- Disclosure Review Boards (DRBs)

Where there are no legal or institutional protections, it falls on the research organization to protect confidentiality.





Privacy

- The individual right to control one's personal information
- About the individuals



Confidentiality

- Researcher's agreement with participant about how information will be handled
- About the data



Potential Geoprivacy Violation - Ex #1



- eightmaps.com
- 2009 point mapping of political donors for California's Proposition 8 in opposition of same-sex marriage
- Location is strong personal identifier
- Mashups and linked databases can reveal sensitive personal information
- Donor address data were of public record



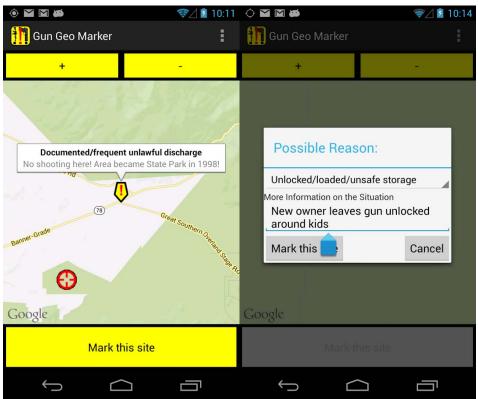
Potential Geoprivacy Violation - Ex #2



 2013: New York newspaper maps open database of gun ownership



- Gun Geo Marker
 - Public pinpoints homes/businesses of suspected unsafe gun owners
 - Implications for firearm theft, personal reputation

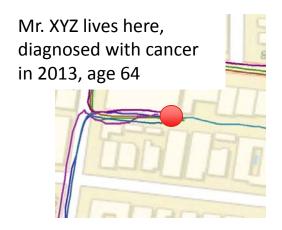


Geoprivacy Violations



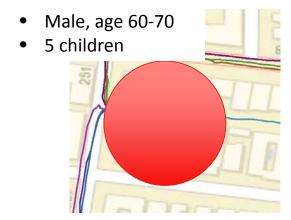
Overt

 Confidential data are explicitly disseminated



Deductive

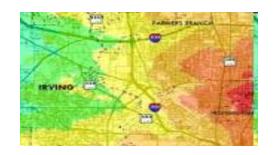
Details are released that can uniquely identify individuals through deduction



Travel Behavior Surveys



- Metropolitan and Regional Transportation Planning Organizations (MPOs / RTPOs) Department of Transportation (NYMTC, NYCDOT)
- Used for urban/regional planning
 - Travel demand models
 - Regional capacity and level of service planning
- Travel Survey participants
 - Recruited randomly from general population
 - Representative sample of entire population
 - Self report of travel behaviors
 - Incentives provided for completion





Traditional Travel Survey v. GPS



Travel: How did you get to Location 1?

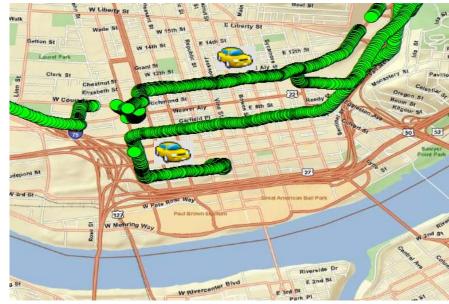
1. What type(s) of transportation did you use to go to Location 11?

				1			\rightarrow			
1 st	1 st			2 nd (if needed)				3 rd (if needed)		
1	Car, van, truck	4	Public Bus		7	Amtrak	10		Taxi/Shuttle	
2	Walk	5	Light Rail (Hiawatha)		8	Bicycle	11		Dial-A-Ride	
3	School Bus	6		mmuter Rail orthstar)	9	Motorcycle/ Moped	1	2	Other (specify)	

- 2. If you used a <u>bus/train</u> for this trip, did you use a pass? ☐ Yes ☐ No --> How much did you pay?
- 3. If you used car/van/truck or motorcycle/moped for this trip . . .
 - A. Were you the . . .? □ Driver □ Passenger
 - B. <u>Including yourself</u>, how many people were in the vehicle? 1 2 3 4+ <u>Including yourself</u>, how many are household members? 1 2 3 4+ Which household members were with you?
 - C. Was this vehicle from your household? ☐ Yes ☐ No
 - D. Did you pay a toll? ☐ Yes ☐ No
 - E. How much, in total, did you personally pay for parking?

 Nothing

Travel Diary Example



GPS Data Example

<u>Advantages of GPS</u>: Route information, lower respondent burden, reduced data entry, increased data quality and data volume

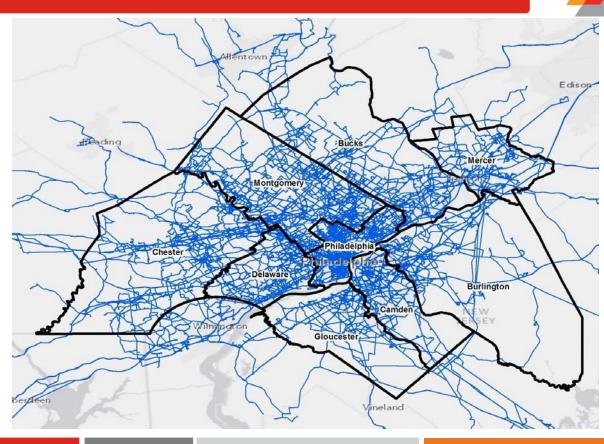
GPS Travel Surveys

- GPS data loggers & GPS smartphone apps for travel data collection
 - Behavioral data every 1 second of travel
 - Up to 3 meters of accuracy
 - Direction, speed, altitude
- Precision of data collection has great utility for travel models
- Increases participant vulnerability to identification



GPS Travel Survey - Philadelphia Region

- Delaware Valley Regional Planning Commission (DVRPC)
- August 2012 to April 2013
- ~750 households participated
- ~1,500 GPS units sent out



GPS Travel Survey - Los Angeles Region

- Southern California
 Association of
 Governments (SCAG)
- April 2012 to October 2012
- ~1,000 households participated
- ~2,000 total GPS units sent out



GPS Travel Surveys



- 1 second frequency GPS travel data capture
- 1 person, 1 day = ~10,000 data points
- 1,000 persons, 3 days = ~30,000,0000 data points

High Frequency +
High Accuracy =
LOW PRIVACY





GPS Travel Survey Considerations



- Higher precision of travel data
- Exact origins & destinations
- Routes, distances, times, speeds
- Trip segments
- Travel modes

- Higher risk of privacy violations
- GPS travel paths are an even stronger personal identifier
- Only 4 location points over 15 hours
 needed to uniquely identify 95% of individuals (de Montijoye et al. 2013)









GPS Survey Privacy Steps



Recruitment

Deployment

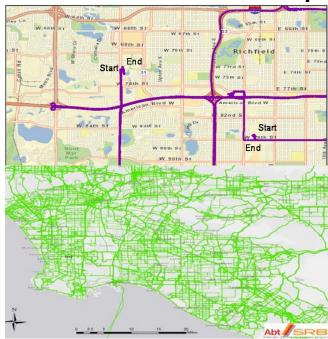
Data processing

Analysis

Data distribution

Data archiving

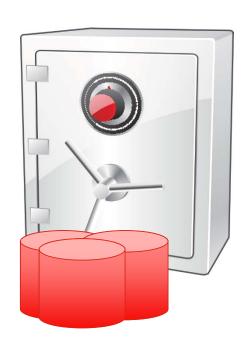
Collecting and managing GPS travel data is a complex process with numerous steps





Recruitment

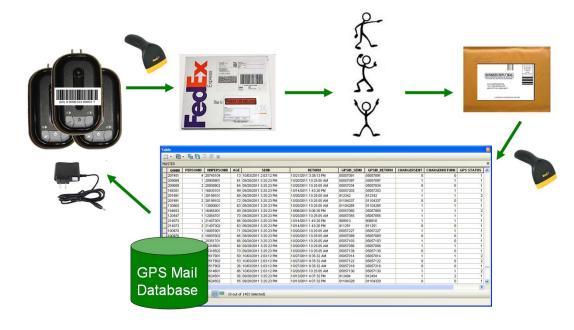
- Ensure understanding of the GPS project with the participant
- Confidentiality agreements, in accordance with your IRB
- Secure storage of participant names and addresses on restricted drives
- Use of encryption for databases
- 1-800 helpline for questions, feedback





Deployment

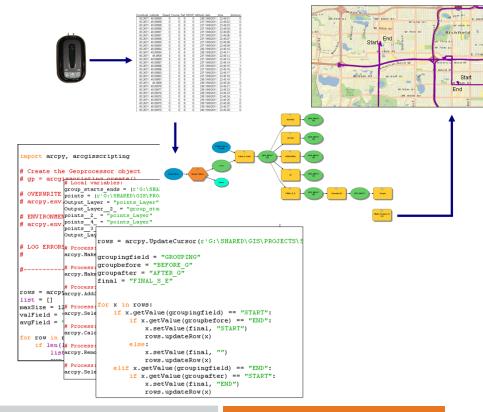
- No personal information on deployed GPS devices
- Use of barcode scanning to check in/out GPS devices
- GPS data cannot be extracted from device without specialized software and/or passwords
- Encrypted databases





Data Processing

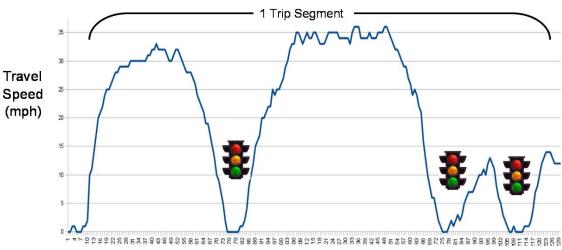
- All name and characteristic identifiers removed
- Anonymous records IDs
- Automated scripts to process data into trips limit manual handling of location data
- The more automated, the less personal



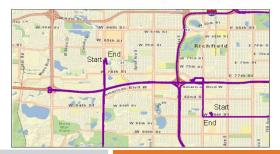


Analysis

- Most analyses do not require individual point data
- Aggregated analyses
 - Travel time, speeds
 - Highway traffic
- Conduct analysis with automated algorithms generating impersonal results



Travel Time (seconds)



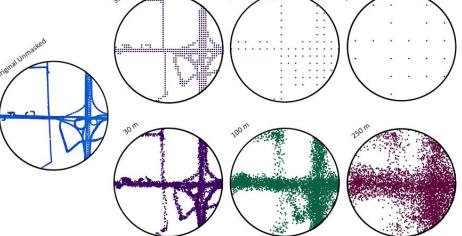


Grid masking

Data Distribution

- Aggregate (i.e. TAZ)
- Density maps
- Masking techniques
 - Grid masking
 - Random perturbation
 - Donut masking





Abt SRBI | pg 22

Random

perturbation



Data Archiving

- Ensure secure IT system
- Delete personal identifiers
- Delete original GPS data with precise locations
- Delete all data from GPS devices, use locks, limited access
- Understand threats to electronic information security





Data Archiving



Data Archiving

- Secure data portals
- http://www.nrel.gov/vehiclesandfuels/secure_transportation_data.html
- National Renewable Energies Laboratory (NREL)
 Transportation Secure Data Center (TSDC)
- Open public access to cleansed data with travel distance and speeds
- Apply for access to spatial GPS travel data in a secure remote environment with analysis tools
- Promotes access to GPS data results while maintaining participant confidentiality





Conclusions



GPS travel data collection offers great advantages but...



A CAUTION



- A breach in confidentiality can occur at any stage in the GPS process
- Plan at each stage to protect participants' personal location information
 - No personal identifiers on GPS devices
 - Encryption of data tables
 - Physical security of GPS equipment
 - Anonymous data processing

- Automation of tasks
- Aggregation for analysis
- Limited distribution
- Deletion of original GPS data

Conclusions



Privacy

The individual right



- About the individuals
- Personal privacy controls
- Participation in GPS Surveys
- Confidentiality agreements

Confidentiality

Researcher's agreement



- About the data
- Responsibility stay with the data
- After GPS travel data is collected the privacy responsibility and concerns remain

Conclusions



A CAUTION

GPS travel data privacy is not just the responsibility of and for the protection of the individual participant...

Responsibility & protection of the client (Ex: Regional Planning Agency)

Responsibility & protection of the researcher (Ex: Abt SRBI)

Responsibility & protection of the end user (Ex: Newspaper publishing a map)

GPS data collection privacy violations could halt funding, discredit the entire travel study, impact organizations, result in lawsuits, etc.

Thank You



Timothy Michalowski Dara Seidl

Abt SRBI - 275 Seventh Ave Suite 2700, New York, NY 10001 646-486-8404

t.michalowski@srbi.com d.seidl@srbi.com



References



- de Montjoye, Y.-A., Hidalgo, C.A., Verleysen, M., Blondel, V.D. 2013. Unique in the crowd: the privacy bounds of human mobility. Scientific Reports 3. http://www.nature.com/srep/2013/130325/srep01376/full/srep01376.html
- Kar, B., R.C. Crowsey, and J.J. Zale. 2013. The Myth of location privacy in the United States: surveyed attitude versus current practices. *The Professional Geographer* 65(1): 47-64.
- Kwan, M. P.; Casas, I.; Schmitz, B. C. 2004. *Protection of geoprivacy and accuracy of spatial information: how effective are geographical masks. Cartographica* 39: 15-28.
- U.S. Department of Health and Human Services http://www.hhs.gov/ocr/privacy/