

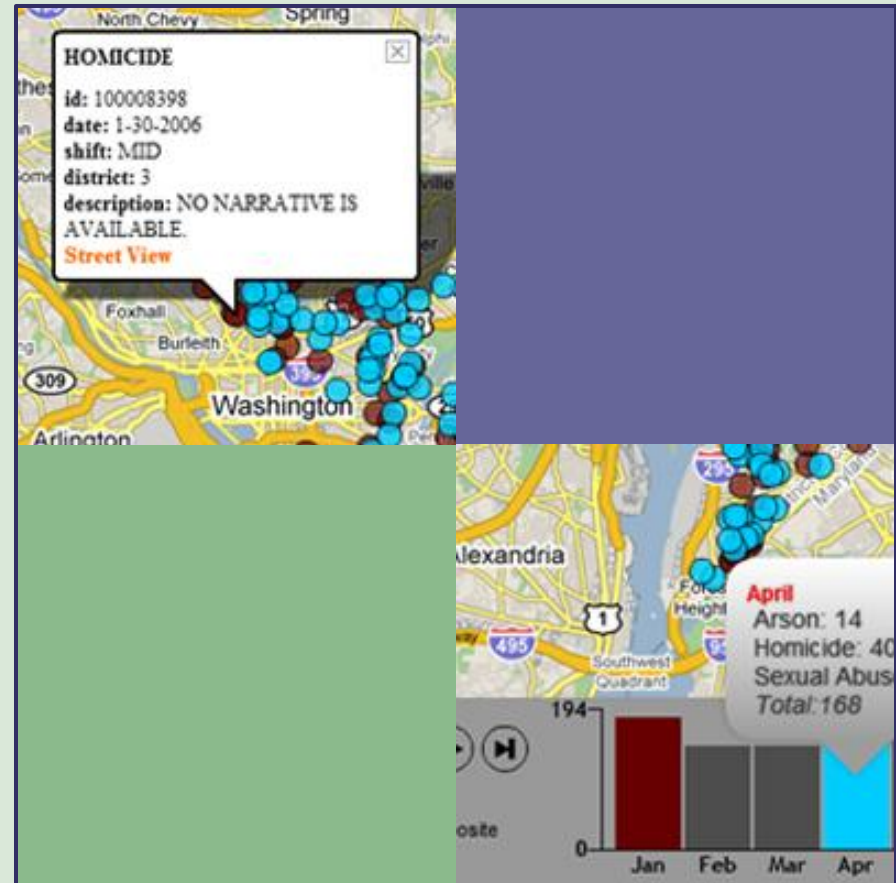
Web-based geovisualization and geovisual analytics to support crime analysis

DHS S&T Site Visit

21 January 2010

Project Team:

Robert E. Roth | reroth@psu.edu
Kevin Ross | kevin.ross@psu.edu
Benjamin G. Finch | bgf111@psu.edu
Wei Luo | wul132@psu.edu
Craig McCabe | cam509@psu.edu
Alan MacEachren | maceachren@psu.edu

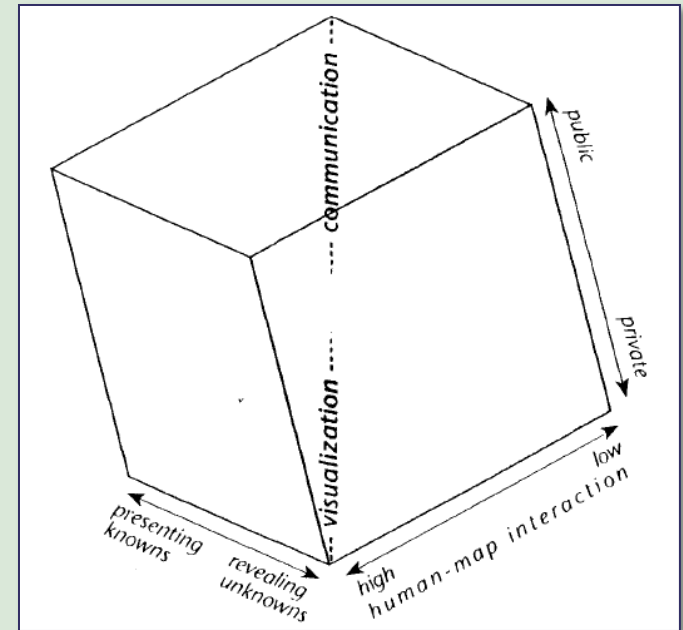
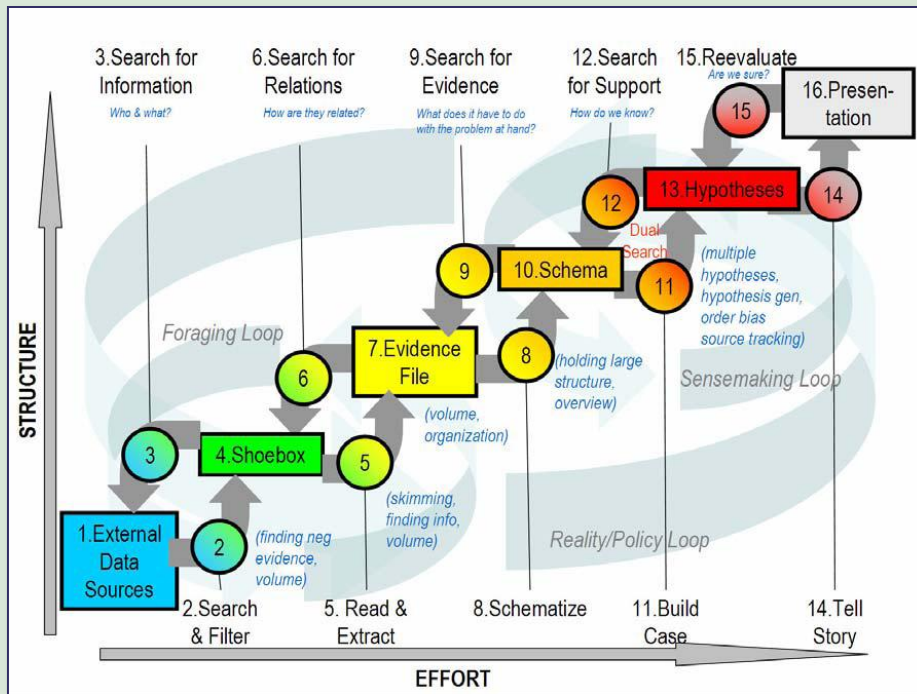


background

Geovisualization:

interactive maps to support exploration, hypothesis generation, and knowledge construction

Sensemaking Process – Pirolli & Card (2005)

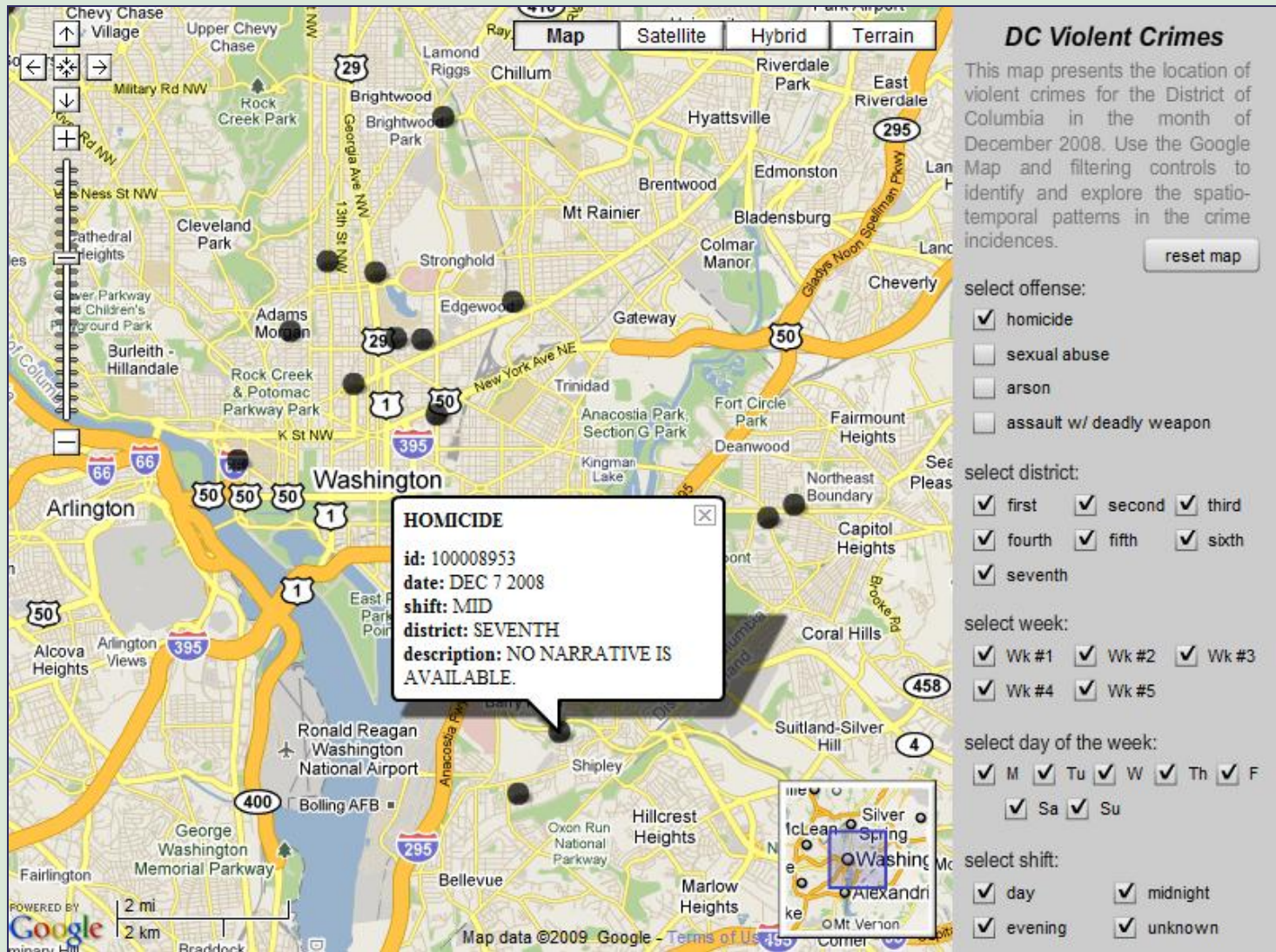


Cartography3 - MacEachren (1994)

Geovisual Analytics:

interactive maps to support the reasoning or sensemaking process

background



the DC CrimeViz prototype

NEVAC A near real-time visualization for understanding spatio-temporal patterns of violent crime in the District of Columbia

Kevin S. Ross, Craig A. McCabe, and Robert E. Roth
Faculty Advisor: Dr. Alan M. MacEachren, Director, North-East Visualization & Analytics Center
GeoVISTA Center, North-East Visualization & Analytics Center, Penn State University

Introduction

DC CrimeViz is a web-based map application for exploring spatio-temporal patterns of violent crime in the District of Columbia. Utilizing a suite of geovisual analytics tools, analysts can uncover detailed spatio-temporal patterns of DC crime, extending the value of the data derived from the existing website.

The client-server application plots up-to-date crime incidents on a basemap using the Google Maps API for Adobe Flash. The user can then filter the visualization by **time** (e.g. police districts or service areas), **type** (linear or composite aggregations), and **attributes** (e.g. type of crime), animating the selections via different temporal aggregations.

Using our application, analysts can generate hypotheses about the etiology of crime clusters [1, 2] to support policy decisions for the reduction and prevention of such crimes.

Data

The District of Columbia publishes violent crime incidents to their web-accessible Data Catalog site in near real-time. This dataset is of particular significance because of its fine spatial and temporal precision – crime incidents are geocoded to within one-half of a city block by eight-hour police work shifts.

An automated script processes the daily RSS feed broadcast by the DC Data Catalog and stores the cleaned data in a PostgreSQL database on a GeoVISTA Center server. These data are automatically retrieved by the DC CrimeViz client at run-time, ensuring the user always has access to the most up-to-date information.

Temporal Aggregation

DC crime data can be temporally aggregated on-the-fly using either linear or composite approaches [3, 4]. Linear aggregation chronologically steps through crime incidents using weekly, monthly, or yearly histogram bins. Alternatively, composite aggregation groups incidents cyclically to highlight recurring patterns, using day of the week (all Mondays, Tuesdays, etc.) day of the month (1st through 31st), or month of the year (January, February, etc.).

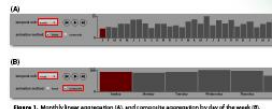


Figure 1. Monthly linear aggregation (A) and composite aggregation by day of the week (B).

Street View

Implementing the Street View capabilities of Google Maps, users can access a street-level perspective of crime scene locations. Employing the principles of 'overview first, zoom and filter, then details on demand' [5], DC CrimeViz provides the ability to drill-down through multiple levels of spatial and contextual detail.



Figure 4. A Google Street View window is opened by clicking on the crime description.

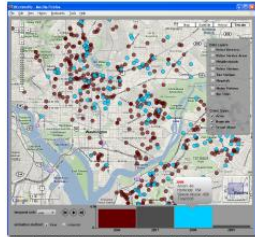


Figure 5. The DC CrimeViz application, displaying incidents for 2006 and 2007.

Animation

After applying space, resulting subset of data, one of six available temporal aggregation methods can be used to animate the data.



Figure 2. Screen progressive screen captures of an animation using weekly composite aggregation. Frames can be incremented manually or automatically.



Figure 3. Overlay of Police Districts boundaries.

Contextualizing Crime

DC CrimeViz contextualizes the crime data by providing access to additional data layers available on the DC Data Catalog website. These include emergency response locations (police/fire/hospital), transportation sites (Metro stations, bus stops), neighborhood composition, and socioeconomic statistics aggregated to police service areas. These supplementary data layers assist the analyst in drawing connections between crime 'hot spots' in space and time, and potential contributing factors.

The screenshot shows the DC CrimeViz prototype interface. The main map displays crime incidents in Washington, DC, with various data layers and crime types visible. A data panel on the right lists 'Data Layers' (Police Districts, Police Service Areas, Neighborhoods, Police Stations, Fire Stations, Hospitals, Metro Stations, Bus Stops) and 'Crime Types' (Arson, Homicide, Sexual Abuse). A 'temporal unit' dropdown is set to 'week', and an 'animation method' dropdown is set to 'composite'. A bar chart at the bottom shows crime counts by day of the week for Wednesday: Arson: 31, Homicide: 69, Sexual Abuse: 200, Total: 300.

Andrianti, 2005. "Evaluating crime visualization in a web-based study of DC crime." *International Journal of Geographical Information Science*, 19(12), 1527-1547.

De Weert, C., Pyle, D., Robinson, A., Hollenbeck, D., Pappas, D., & M. MacEachren. 2007. "Visual Analysis of Historical Violent Crime Patterns." *Information Visualization*, 4, pp. 89-101.

ESRI. 1996. "The area leave it A task by data aggregation for information visualization." *IEEE Workshop on Visual Language '96*, San Antonio, CA, September 1996, pp. 250-263.

Acknowledgements

This project was funded through the Regional Visualization & Analytics Center (RVAC) Center of Excellence by a grant from the Department of Homeland Security, Science and Technology Directorate, Office of University Programs. It was also supported by the National Visualization & Analytics Center, a U.S. Department of Homeland Security program supported by the Army Research Laboratory (ARL) (ARL/DC/06-01), a U.S. Department of Energy Office of Science laboratory.

Further Information

DC Data Catalog: <http://data.dc.gov/>
DC CrimeViz Application & Metadata: <http://www.geovista.psu.edu/DCCrimeViz/>
The GeoVISTA Center: <http://www.geovista.psu.edu/>
North-East Visualization & Analytics Center: <http://www.geovista.psu.edu/NEVAC/index.html>

**DHS Summit Student Poster Competition:
3rd Place (125+ entries)**

Study #1: Talk Aloud Discount Usability Study

***Goal:** identify usability errors in the DC CrimeViz prototype to improve/refine it for use in further testing

***Goal:** improve the practice of usability evaluation of geovisualization and geovisual analytics software

Study #2: Needs Assessment Interviews

***Goal:** improve the understanding of the current practice of crime mapping and analysis

***Goal:** identify the mapping and spatial analysis needs of crime analysts to determine how to best extend the prototype to a fully-featured crime mapping system

research: talk aloud study

Talk Aloud Discount Usability Study

Who: 5 participants (2 undergraduates, 1 graduate, 2 faculty)

Where: Human Factors Lab, 229A Walker Building

When: 11-13 November 2009

How: one administrator, two note takers recording critical incidents

Talk Aloud Protocol (60 minutes)

- Introduction (5 minutes)
- Opening (5 minutes)
- Tasks (35-40 minutes)
- Cognitive interview (10-15 minutes)

research: talk aloud study

Data Layers Panel (WHAT) & Data Issues

	P1	P2	P3	P4	P5	Freq	Fix?
Application breaks when viewing 'Bus Stops' data layer	x	x	x	x	x	5	Yes
Sexual abuse cases after 2006 not mapped	x	x	x	x	x	5	Yes
Loading screen does not provide feedback	x			x		2	Yes
Add a data layer with metro lines as well as stations			x			1	No
Data Layers panel overlaps the Google Maps inset	x					1	No
Add a layer reset feature			x			1	Yes
Crime layer check boxes low on the visual hierarchy				x		1	Yes
One misregistered data point		x				1	Yes

research: talk aloud study

Temporal Controls (WHEN) & Histogram Issues

	P1	P2	P3	P4	P5	Freq	Fix?
Ambiguity in the meaning of linear and composite aggregation	x	x	x	x	x	5	Yes
Add ability to select multiple histogram bars for persistent highlighting in the map	x	x	x	x	x	5	No
Lag in the animation and in histogram brushing when there are a small number of bins	x	x	x	x	x	5	No
Ambiguity in the meaning of temporal unit when composite is applied	x	x		x	x	4	No
Data filtering not reflected in the histogram tool tip		x	x		x	3	Yes
Add a clear division by year for the linear-month histogram	x		x			2	Yes
Add a scroll feature to the histogram so that the bins could be wider				x	x	2	No
Animations continued to play or stopped in unexpected ways when interacting with the histogram or map		x	x			2	Yes
Unclear labels on temporal legend	x				x	2	Yes
Unable to discriminate the different types of crime in the histogram without brushing				x		1	Yes
Add ability to customize the bin widths					x	1	No
Add a reset animation feature		x				1	No
Ambiguity in interpreting composite-month because of extra Jan and Feb from 2009		x				1	Yes

research: talk aloud study

Spatial Controls (WHERE) & Map Issues

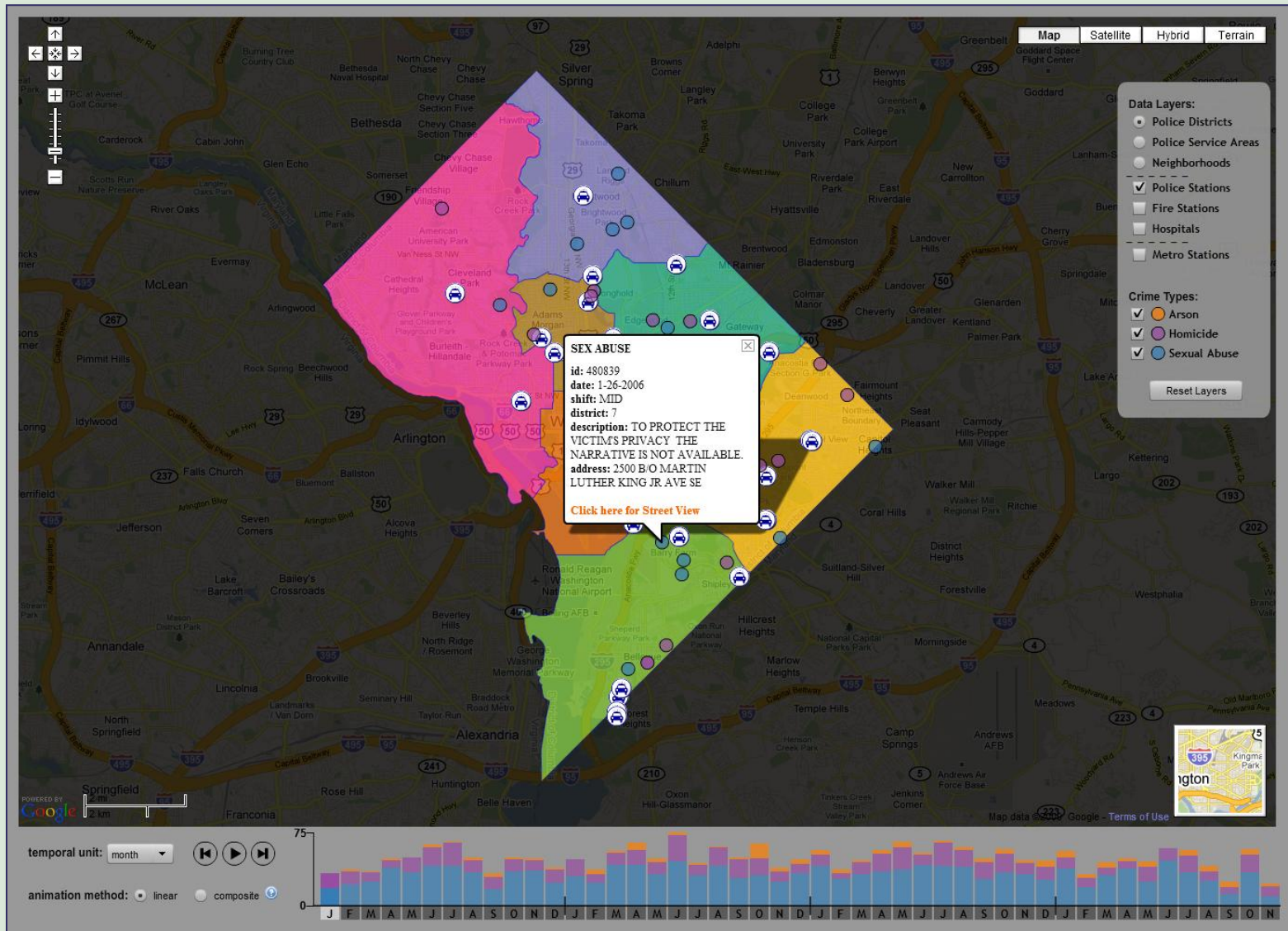
	P1	P2	P3	P4	P5	Freq	Fix?
Add a search feature by police record number	x	x	x	x	x	5	No
Add a search feature by address	x	x	x	x	x	5	No
Unable to discriminate the different areal boundary layers because they are the same color	x	x	x	x	x	5	Yes
Unable to retrieve information about districts (both IDs, population, and incidents) and POIs	x	x	x	x	x	5	Yes
Overlapping incident symbols / Too much data on the map / Add data aggregation option	x	x	x	x	x	5	No
Unable to discriminate the different types of crime without filtering/brushing them		x	x	x	x	4	Yes
Add scroll zooming using the mouse wheel	x	x			x	3	Yes
Add ability to zoom into a feature	x	x			x	3	Yes
Add a spatial extent reset feature		x	x	x		3	Yes
Add a measurement tool (linear) or distance query tool (circular from point)	x	x			x	3	No
Unable to discriminate the different Points of Interest		x		x	x	3	Yes
Add cluster analysis feature		x		x	x	3	No

research: talk aloud study

Spatial Controls (WHERE) & Map Issues

	P1	P2	P3	P4	P5	Freq	Fix?
Lag in panning and zooming when numerous points are shown	x					1	No
Add a search feature by Point of Interest					x	1	No
Add rubberband zoom using Shift+Drag		x				1	No
Unclear that Street View is available	x					1	Yes
Selection of 'fullscreen' instead of 'close' in Street View					x	1	No
Street View does not work in Internet Explorer					x	1	Yes
Information window should close when clicking outside of it		x				1	No
Information window should include the address		x				1	Yes
Add ability to show the case ID on mouse over of the point symbol					x	1	No
Add buffer feature				x		1	No

research: talk aloud study



Needs Assessment Interview Study

Who: 9 participants (5 civilians, 3 officers, 1 recently retired officer)

Where:

- New York Police Department
- Harrisburg City Bureau of Police (n=2)
- Akron Police Department
- Cleveland Police Department
- Department of Justice (Washington DC)
- Alexandria Police Department (n=2)
- Philadelphia Police Department

When: 30 November 2009 – 7 December 2009

How: one interviewer, audio recorded

research: interview study

Interview Protocol (60 minutes)

- Introduction (5 minutes)
- Biographic/Background (5 minutes)
- Data/Information Characteristics (5 minutes planned, usually lasted 15 minutes)
- Mapping and Analysis Practices (20 minutes)
- Use (10 minutes)

Transcription and Coding

- 39-code scheme (six high level categories)
- Two independent coders (inter-coder reliability=81.4%); interviewer reconciled differences

research: interview study

Key findings

Data: Maintain in-house voluminous (7k to 2.5mil records), multivariate datasets; rarely share with other agencies & rarely use federal data

Representation: Primarily make 'push-pin' (1-to-1 dot maps) using color to represent time and 'hot-spot' maps

Interaction: Filtering/focusing and sequencing are common, other map interaction operators rarely available

Spatial Analysis: Rarely apply transformations beyond basic filtering and KDE

Temporal Analysis: Extreme variation across agencies, with several agencies applying trend analysis/forecasting and automated spatio-temporal clustering algorithms

Use: Focus on tactical instead of strategic analysis; most agencies are underfunded and understaffed

future directions

Analyze/write-up an online survey study (n=9) evaluating the prototype

Continue development on the DC CrimeViz prototype

- this includes needs that are currently met by existing software and those that are not currently met, specifically focusing on features that support analytical reasoning

Complete an participant observation study of the CompStat process

- hope to contrast sessions using static versus interactive maps

Implement the CrimeViz concept for the Harrisburg, PA Police Department.

- hope to include follow-up study to learn about the insights generated by the tool

questions?

Check out the application at:

<http://www.geovista.psu.edu/DCCrimeViz/>



Thanks!

Rob, Kevin, Ben, Wei, Craig, & Alan