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

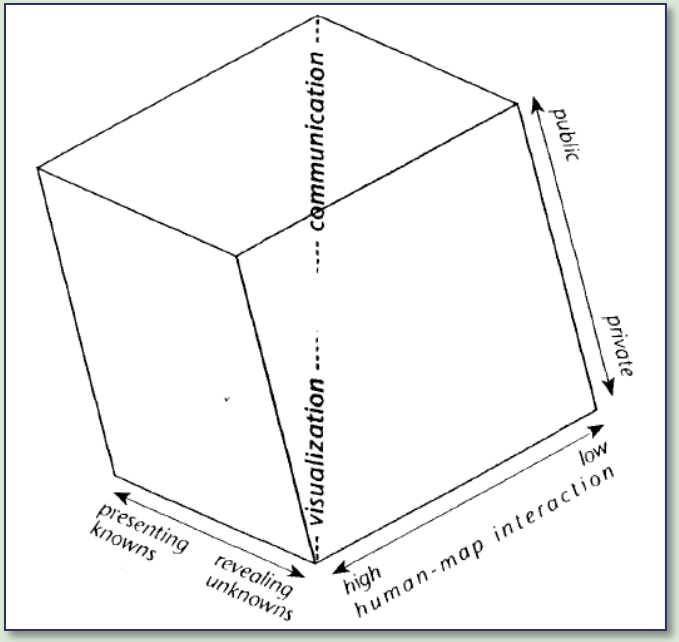
GeoVISTA Center

VACCINE
Visual Analytics for Command, Control, and Interoperability Environments
A U.S. Department of Homeland Security Center of Excellence

background

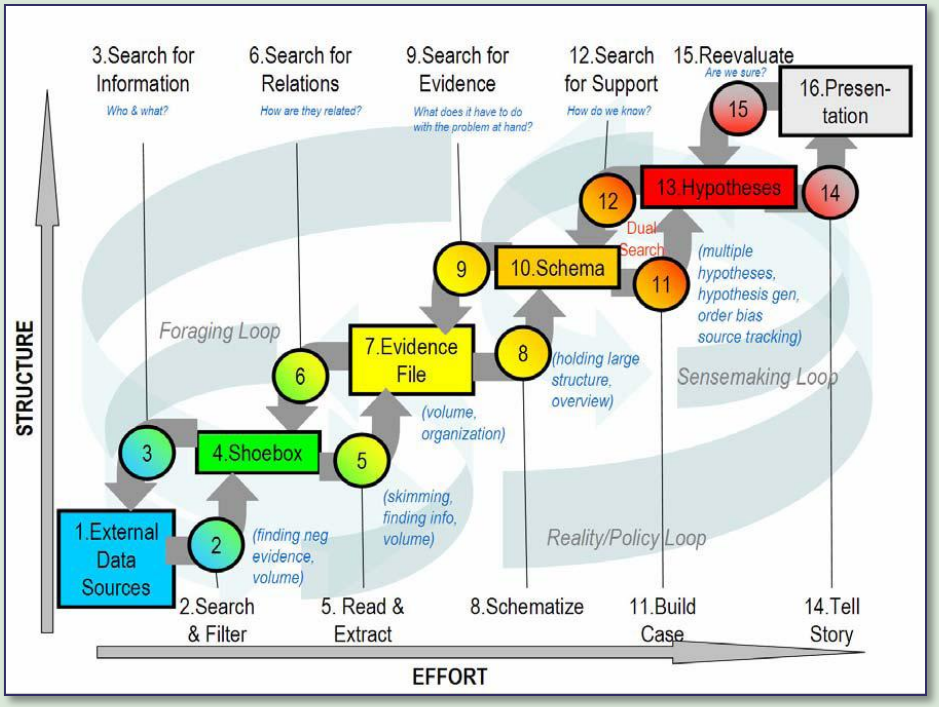
Geovisualization:

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



Cartography3 - MacEachren (1994)

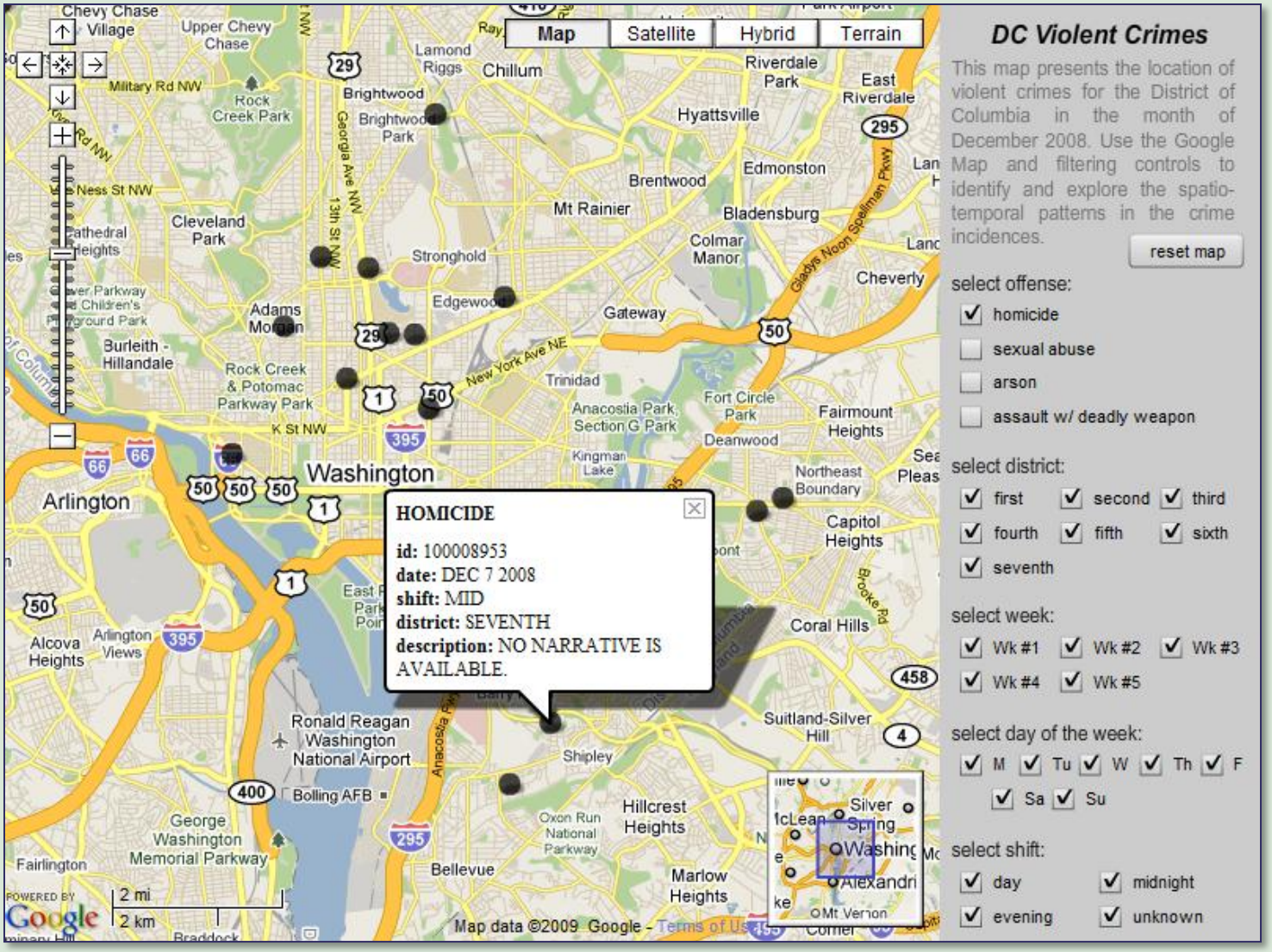
Sensemaking Process – Pirolli & Card (2005)



Geovisual Analytics:

interactive maps to support the reasoning or sensemaking process

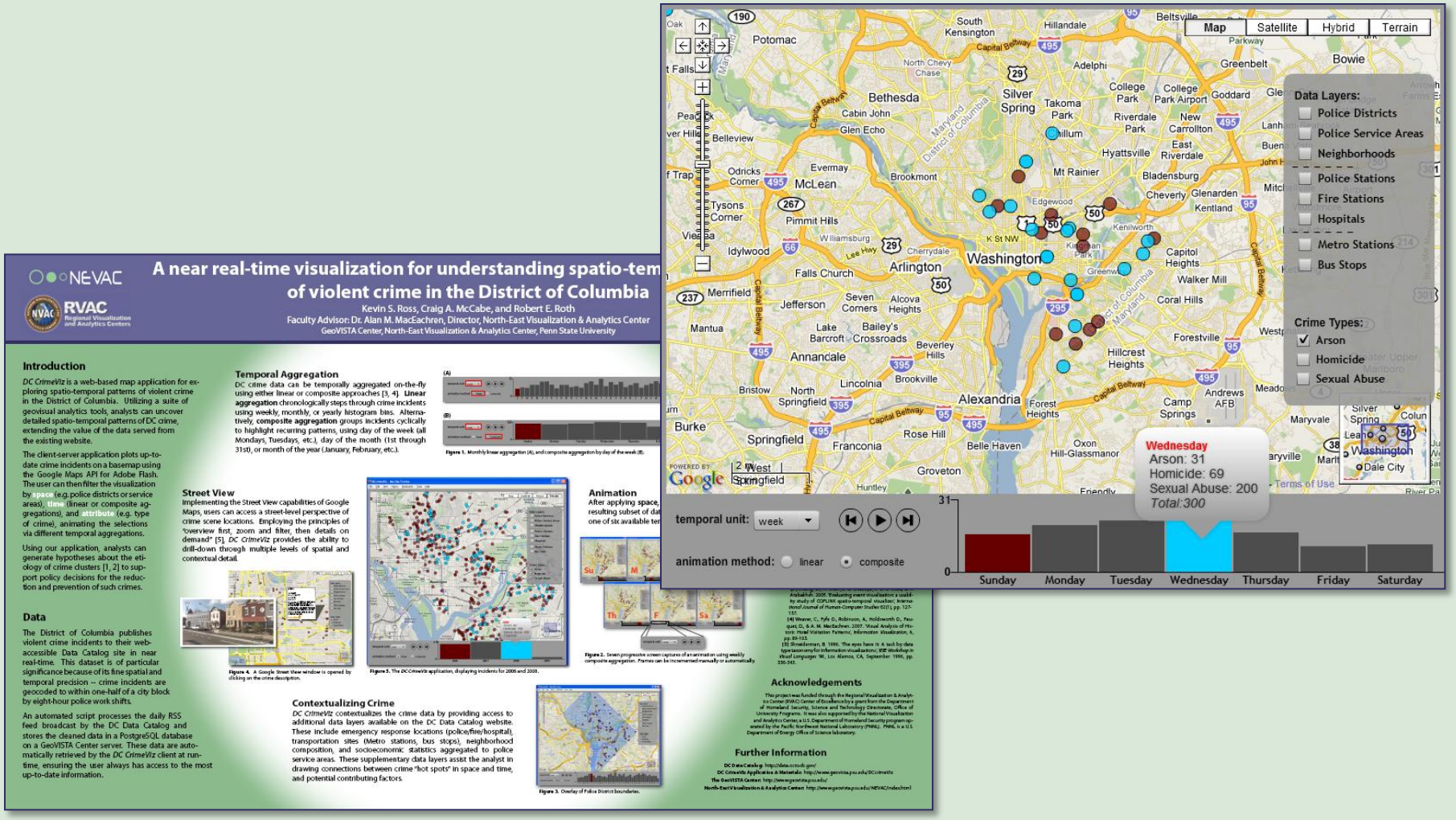
background



Geography 461 Lab #2

background

the DC CrimeViz prototype



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

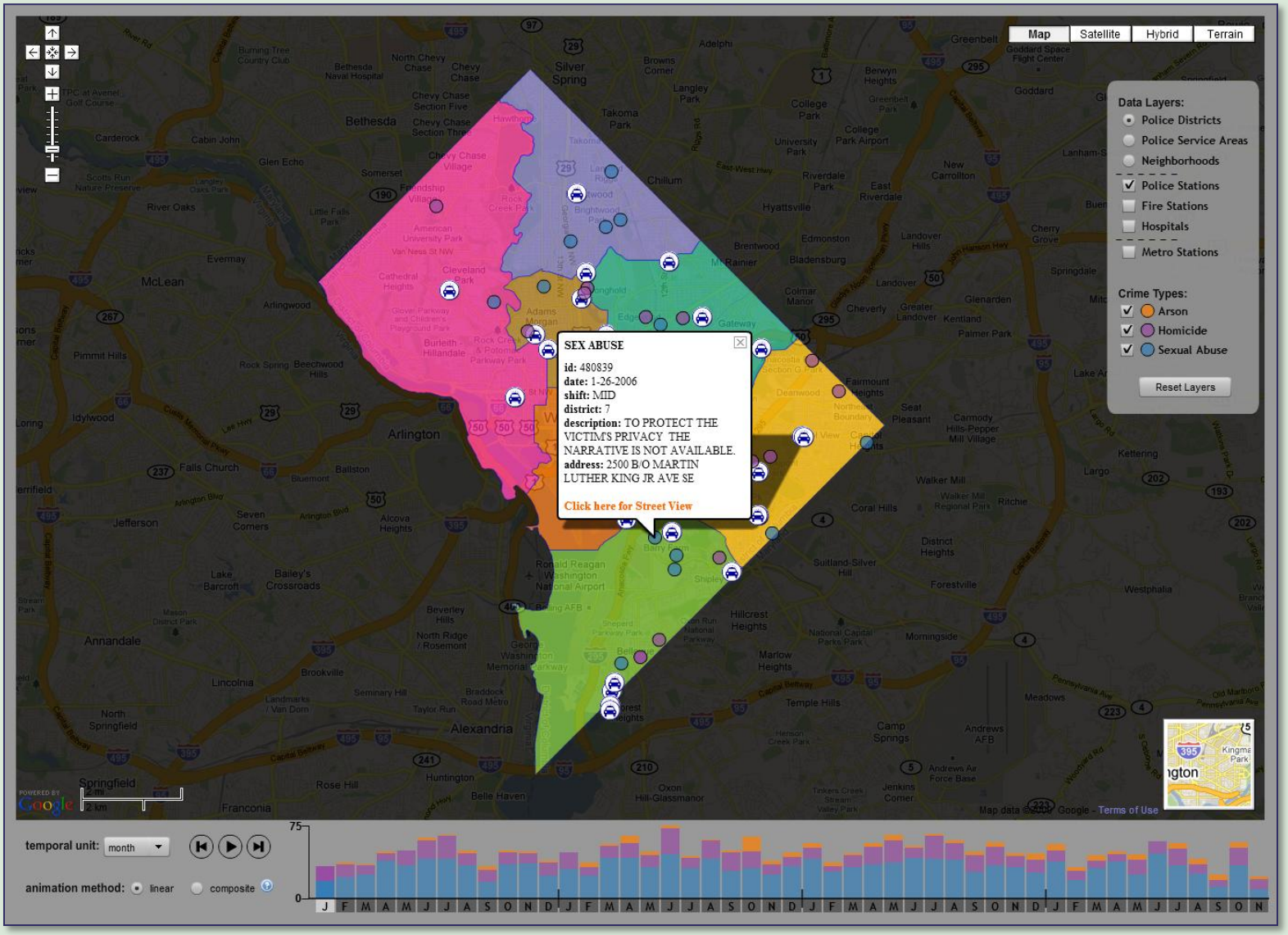
	P1	P2	P3	P4	P5	Freq	Fix?
Add a search feature by police record number	x	x	x	x	x	5	No
Add a search feature by address	x	x	x	x	x	5	No
Unable to discriminate the different areal boundary layers because they are the same color	x	x	x	x	x	5	**Yes**
Unable to retrieve information about districts (both IDs, 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



revised DC CrimeViz prototype

research: interview 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

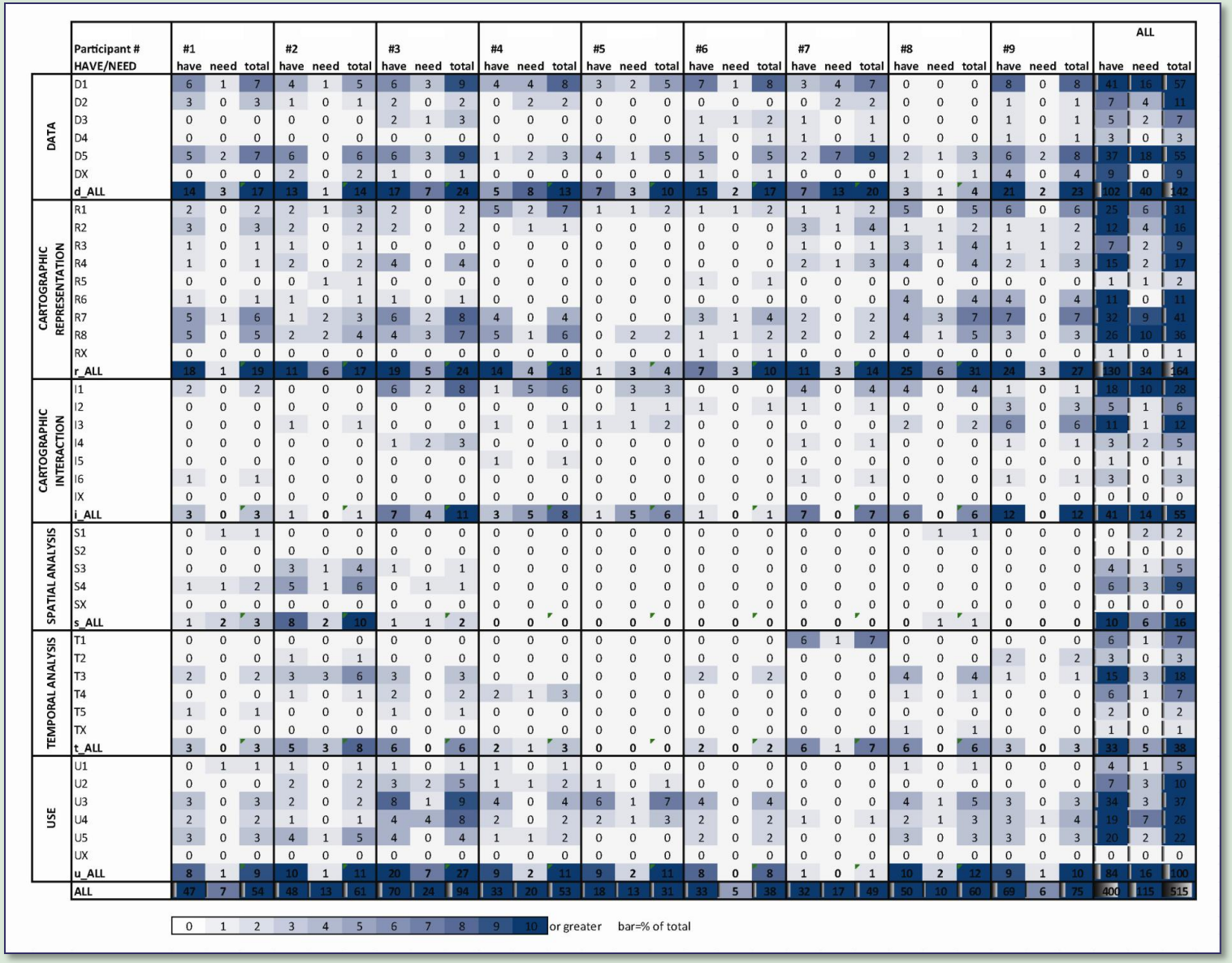
Interview Protocol (60 minutes)

- 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



	Participant #	#1			#2			#3			#4			#5			#6			#7			#8			#9			ALL		
	HAVE/NEED	have	need	total	have	need	total	have	need	total	have	need	total	have	need	total	have	need	total	have	need	total	have	need	total	have	need	total	have	need	total
DATA	D1	6	1	7	4	1	5	6	3	9	4	4	8	3	2	5	7	1	8	3	4	7	0	0	0	8	0	8	41	16	57
	D2	3	0	3	1	0	1	2	0	2	0	2	2	0	0	0	0	0	0	0	2	2	0	0	0	1	0	1	7	4	11
	D3	0	0	0	0	0	0	2	1	3	0	0	0	0	0	0	1	1	2	1	0	1	0	0	0	1	0	1	5	2	7
	D4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	1	0	1	3	0	3
	D5	5	2	7	6	0	6	6	3	9	1	2	3	4	1	5	5	0	5	2	7	9	2	1	3	6	2	8	37	18	55
	DX	0	0	0	2	0	2	1	0	1	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	4	0	4	9	0	9
	d_ALL	14	3	17	13	1	14	17	7	24	5	8	13	7	3	10	15	2	17	7	13	20	3	1	4	21	2	23	102	40	142
CARTOGRAPHIC REPRESENTATION	R1	2	0	2	2	1	3	2	0	2	5	2	7	1	1	2	1	1	2	1	1	2	5	0	5	6	0	6	25	6	31
	R2	3	0	3	2	0	2	2	0	2	0	1	1	0	0	0	0	0	0	3	1	4	1	1	2	1	1	2	12	4	16
	R3	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	3	1	4	1	1	2	7	2	9
	R4	1	0	1	2	0	2	4	0	4	0	0	0	0	0	0	0	0	0	2	1	3	4	0	4	2	1	3	15	2	17
	R5	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	1	2
	R6	1	0	1	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4	0	4	4	0	4	11	0	11
	R7	5	1	6	1	2	3	6	2	8	4	0	4	0	0	0	3	1	4	2	0	2	4	3	7	7	0	7	32	9	41
	R8	5	0	5	2	2	4	4	3	7	5	1	6	0	2	2	1	1	2	2	0	2	4	1	5	3	0	3	26	10	36
	RX	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1
	r_ALL	18	1	19	11	6	17	19	5	24	14	4	18	1	3	4	7	3	10	11	3	14	25	6	31	24	3	27	130	34	164
CARTOGRAPHIC INTERACTION	I1	2	0	2	0	0	0	6	2	8	1	5	6	0	3	3	0	0	0	4	0	4	4	0	4	1	0	1	18	10	28
	I2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0	1	0	0	0	3	0	3	5	1	6
	I3	0	0	0	1	0	1	0	0	0	1	0	1	1	1	2	0	0	0	0	0	0	2	0	2	6	0	6	11	1	12
	I4	0	0	0	0	0	0	1	2	3	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	3	2	5
	I5	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
	I6	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	3	0	3
	IX	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	i_ALL	3	0	3	1	0	1	7	4	11	3	5	8	1	5	6	1	0	1	7	0	7	6	0	6	12	0	12	41	14	55
SPATIAL ANALYSIS	S1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	2	2
	S2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	S3	0	0	0	3	1	4	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	5
	S4	1	1	2	5	1	6	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	3	9
	SX	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	s_ALL	1	2	3	8	2	10	1	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	10	6	16
TEMPORAL ANALYSIS	T1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	1	7	0	0	0	0	0	0	6	1	7
	T2	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	3	0	3
	T3	2	0	2	3	3	6	3	0	3	0	0	0	0	0	0	2	0	2	0	0	0	4	0	4	1	0	1	15	3	18
	T4	0	0	0	1	0	1	2	0	2	2	1	3	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	6	1	7
	T5	1	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	2
	TX	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1
	t_ALL	3	0	3	5	3	8	6	0	6	2	1	3	0	0	0	2	0	2	6	1	7	6	0	6	3	0	3	33	5	38
USE	U1	0	1	1	1	0	1	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	4	1	5
	U2	0	0	0	2	0	2	3	2	5	1	1	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	7	3	10
	U3	3	0	3	2	0	2	8	1	9	4	0	4	6	1	7	4	0	4	0	0	0	4	1	5	3	0	3	34	3	37
	U4	2	0	2	1	0	1	4	4	8	2	0	2	2	1	3	2	0	2	1	0	1	2	1	3	3	1	4	19	7	26
	U5	3	0	3	4	1	5	4	0	4	1	1	2	0	0	0	2	0	2	0	0	0	3	0	3	3	0	3	20	2	22
	UX	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	u_ALL	8	1	9	10	1	11	20	7	27	9	2	11	9	2	11	8	0	8	1	0	1	10	2	12	9	1	10	84	16	100
	ALL	47	7	54	48	13	61	70	24	94	33	20	53	18	13	31	33	5	38	32	17	49	50	10	60	69	6	75	400	115	515

0 1 2 3 4 5 6 7 8 9 10 or greater bar=% of total

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/

GeoVISTA Center

VACCINE
Visual Analytics for Command, Control, and Interoperability Environments
A U.S. Department of Homeland Security Center of Excellence

Thanks!

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