VACCINE

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

VACCINE ANNUAL REPORT – YEAR 4

Addendum A - Presentations

April 1, 2012 – June 30, 2013

Cooperative Agreement No. 2009-ST-061-CI0001

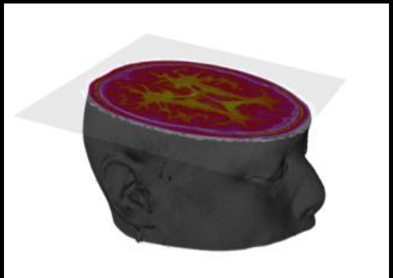
PURDUE UNIVERSITY

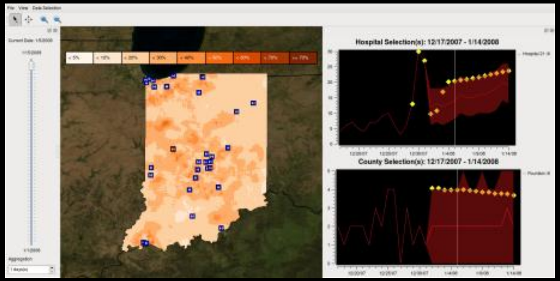
HOMELAND SECURITY UNIVERSITY PROGRAMS
TODAY'S RESEARCH & EDUCATION, TOMORROW'S SECURITY

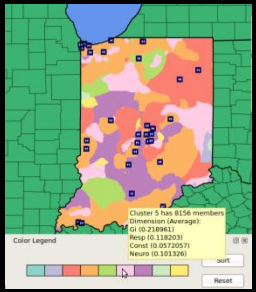
Addendum A – Presentations

Interactive Table of Contents

SIMON FRASER UNIVERSITY, CA







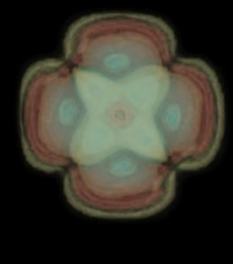


Table of Contents

Analytical Brushing in Spatiotemporal Analysis

Ross Maciejewski

Ross Maciejewski
E-mail: rmacieje@asu.edu
http://rmaciejewski.faculty.asu.edu/

Spatiotemporal Analysis

Much of the data being collected and analyzed today is spatiotemporal

- Twitter[1,2]
- Emergency Departments[3]
- Law Enforcement[4]
- Mobile Phones[5,6]
- Animal Migration[7]
- Flow Simulations[8]

1 - A. M. MacEachren, A. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, and J. Blanford. "Senseplace2: Geotwitter analytics support for situational awareness." In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, 2011.

2 - Chae, J., Thom, D., Bosch, H., Jang, Y., Maciejewski, R., Ebert, D. S., Ertl, T., "Spatiotemporal Social Media Analytics for Abnormal Event Detection Using Seasonal-Trend Decomposition," *IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2012.

3 - Maciejewski, R., Rudolph, S., Hafen, R., Mitchell, M. A., Larew, S. G., Cleveland, W., Ebert, D., "Forecasting Hotspot – A Predictive Analytics Approach," *IEEE Transactions on Visualization and Computer Graphics*, 17(4): 440-453, 2011.

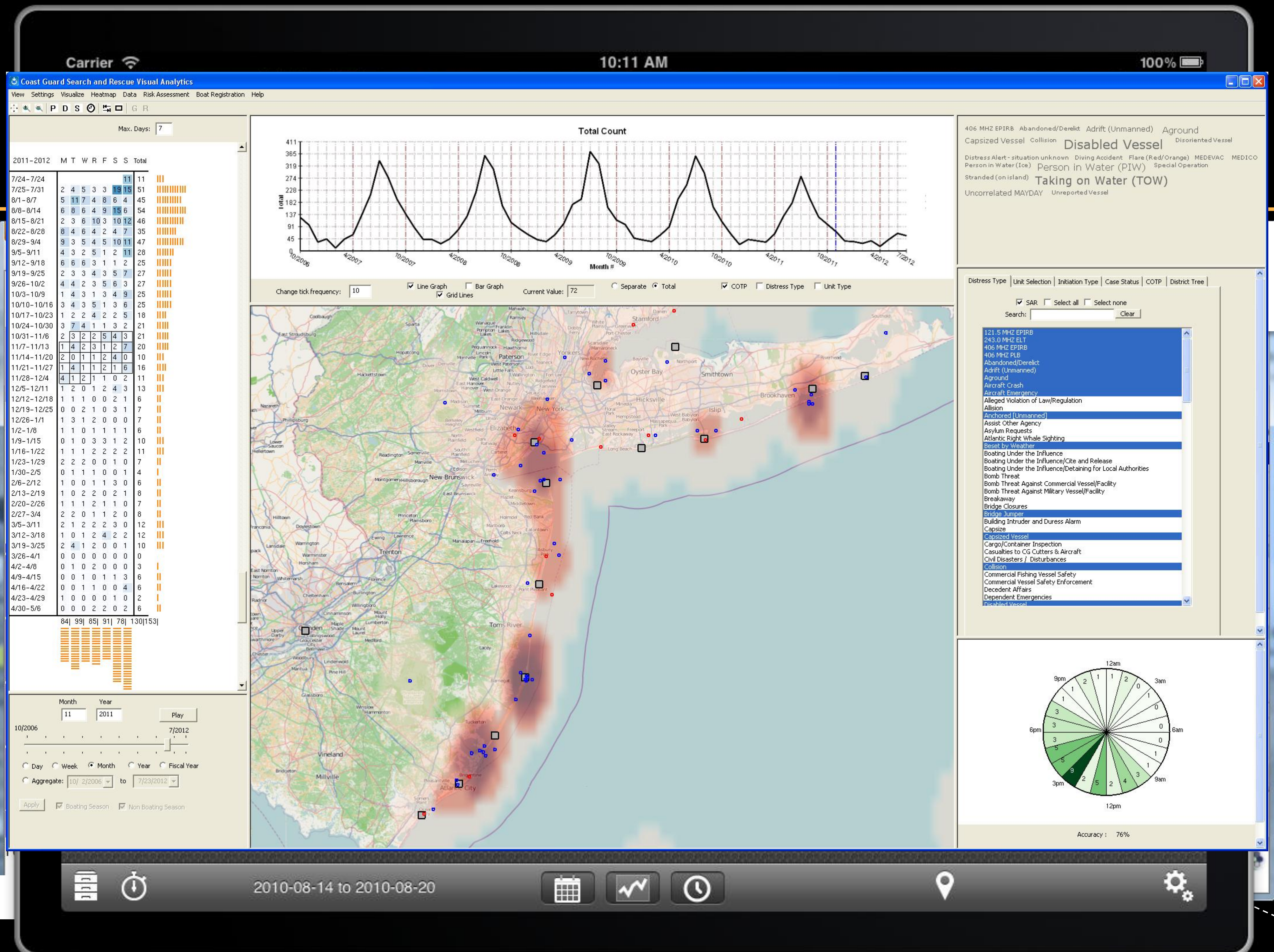
4 - Malik, A., Maciejewski, R., Collins, T., Ebert, D., T., "Visual Analytics Law Enforcement Toolkit," *IEEE International Conference on Technologies for Homeland Security*, 2010.

5 – Z. Shen and K.-L. Ma. Mobivis: A visualization system for exploring mobile data. In *IEEE Pacific Visualization*, pages 175–182, 2008.

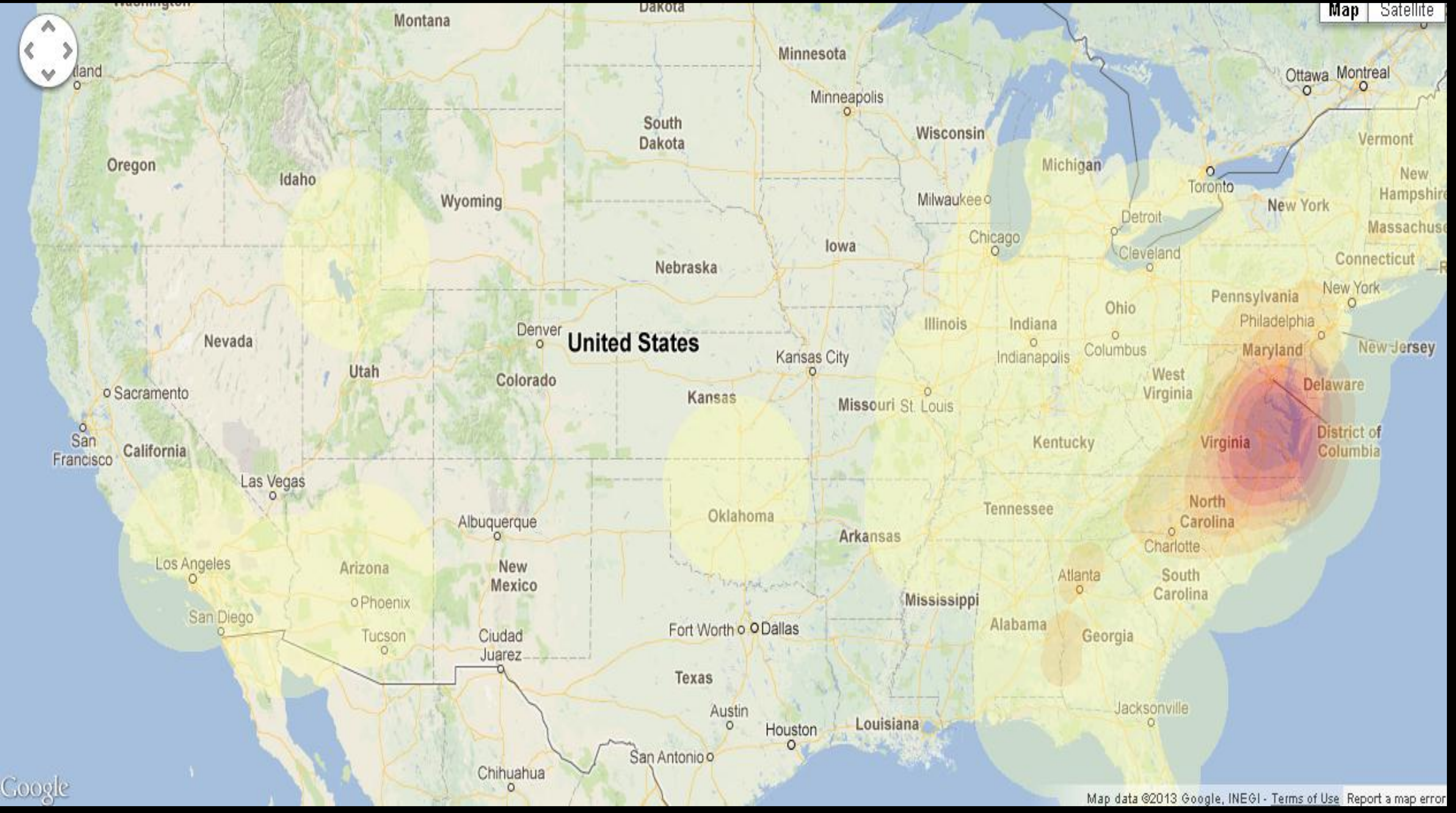
6 - G. L. Andrienko, N. V. Andrienko, C. Hurter, S. Rinzivillo, and S. Wrobel. "From movement tracks through events to places: Extracting and characterizing significant places from mobility data." In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, 2011.

7 – N. Ferreira, L. D. Lins, D. Fink, S. Kelling, C. Wood, J. Freire, C. T. Silva, "BirdVis: Visualizing and Understanding Bird Populations." IEEE Trans. Vis. Comput. Graph. 17(12): 2374-2383 (2011)

8 – J. Woodring, H.-W. Shen, "Semi-Automatic Time-Series Transfer Functions via Temporal Clustering and Sequencing." Computer Graphics Forum 28(3): 791-798 (2009)


Carrier
10:11 AM
100%
Coast Guard Search and Rescue Visual Analytics
View Settings Visualize Heatmap Data Risk Assessment Boat Registration Help
Max. Days: 7
2011-2012 M T W R F S S Total
Total Count
Month #
Change tick frequency: 10
Line Graph
Bar Graph
Grid Lines
Current Value: 72
Separate
Total
COTP
Distress Type
Unit Type
406 MHZ EPIRB Abandoned/Derelict Adrift (Unmanned) Aground
Capsized Vessel Collision Disabled Vessel Disoriented Vessel
Distress Alert - situation unknown Diving Accident Flare (Red/Orange) MEDEVAC MEDICO
Person in Water (Ice) Person in Water (PIW) Special Operation
Stranded (on island) Taking on Water (TOW)
Uncorrelated MAYDAY Unreported Vessel
Distress Type Unit Selection Initiation Type Case Status COTP District Tree
SAR Select all Select none
Search: Clear
121.5 MHZ EPIRB
243.0 MHZ ELT
406 MHZ EPIRB
406 MHZ PLB
Abandoned/Derelict
Adrift (Unmanned)
Aground
Aircraft Crash
Aircraft Emergency
Alleged Violation of Law/Regulation
Allision
Anchored (Unmanned)
Assist Other Agency
Asylum Requests
Atlantic Right Whale Sighting
Beset by Weather
Boating Under the Influence
Boating Under the Influence/Cite and Release
Boating Under the Influence/Detaining for Local Authorities
Bomb Threat
Bomb Threat Against Commercial Vessel/Facility
Bomb Threat Against Military Vessel/Facility
Breakaway
Bridge Closures
Bridge Jumper
Building Intruder and Duress Alarm
Capsize
Capsized Vessel
Cargo/Container Inspection
Casualties to CG Cutters & Aircraft
Civil Disasters / Disturbances
Collision
Commercial Fishing Vessel Safety
Commercial Vessel Safety Enforcement
Decedent Affairs
Dependent Emergencies
12am 3am 6am 9am 12pm 3pm 6pm 9pm
Accuracy : 76%
Month Year
11 2011
Play
10/2006 7/2012
Day Week Month Year Fiscal Year
Aggregate: 10/ 2/2006 to 7/23/2012
Apply Boating Season Non Boating Season
2010-08-14 to 2010-08-20

Distribution of Surname Cutchin

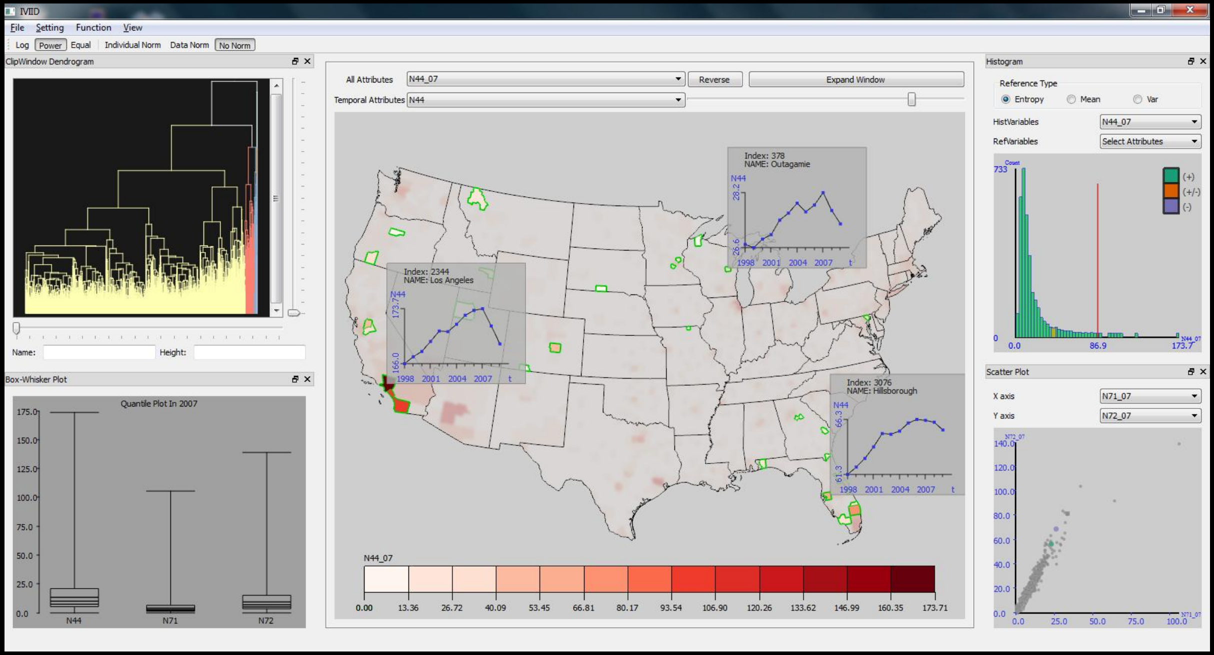


http://www.uncertaintyofidentity.com/SurnameSearch.html

Coordinated Multiple Views in Spatiotemporal Analysis

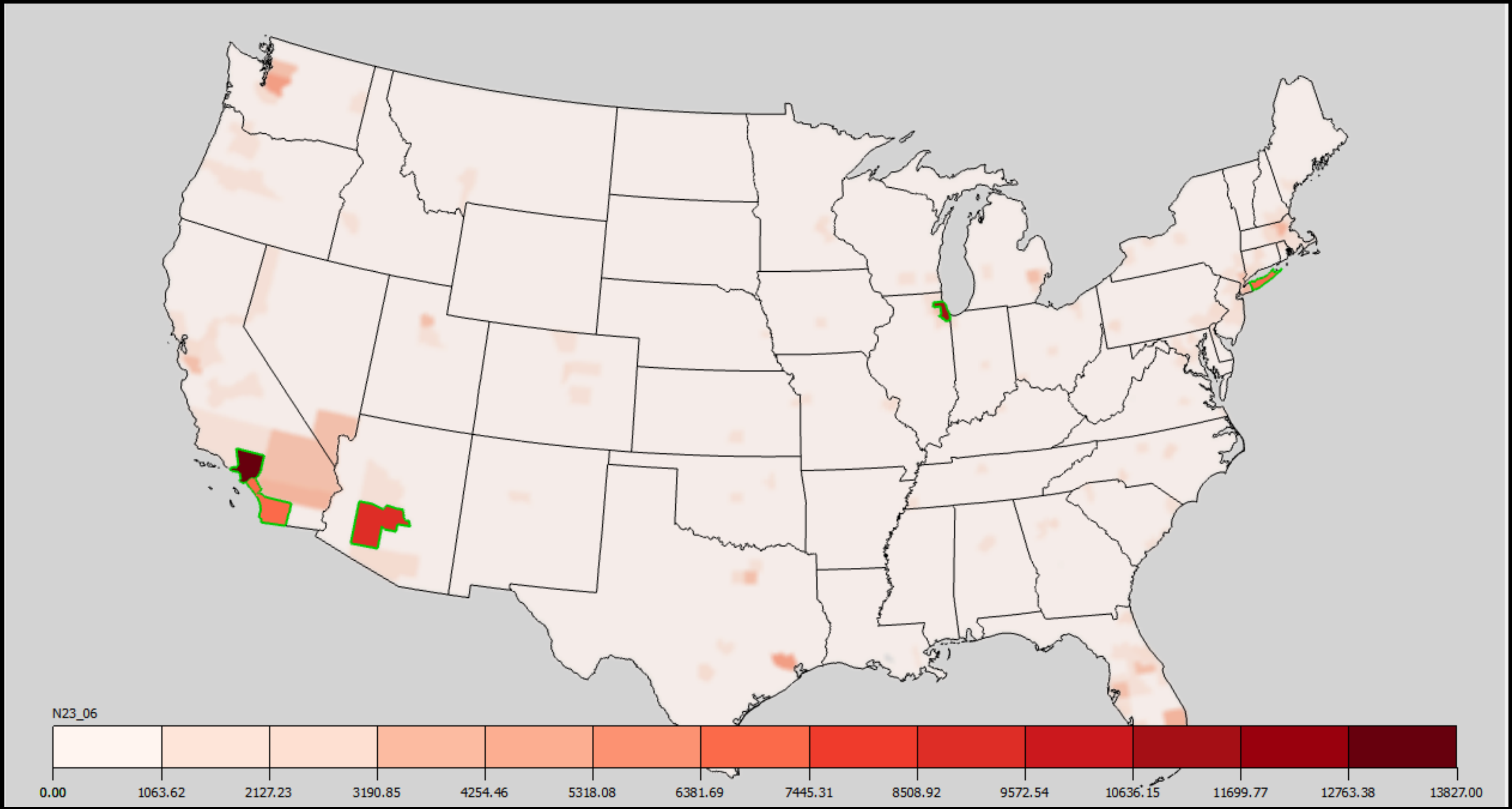
To facilitate answering these questions, we typically rely on interactive coordinated multiple views[1]

- Geographic Maps
- Scatterplots
- Parallel Coordinate Plots
- Dendogram
- Histogram
- Time Series Plots
- Tabular View
- Box-and-Whisker

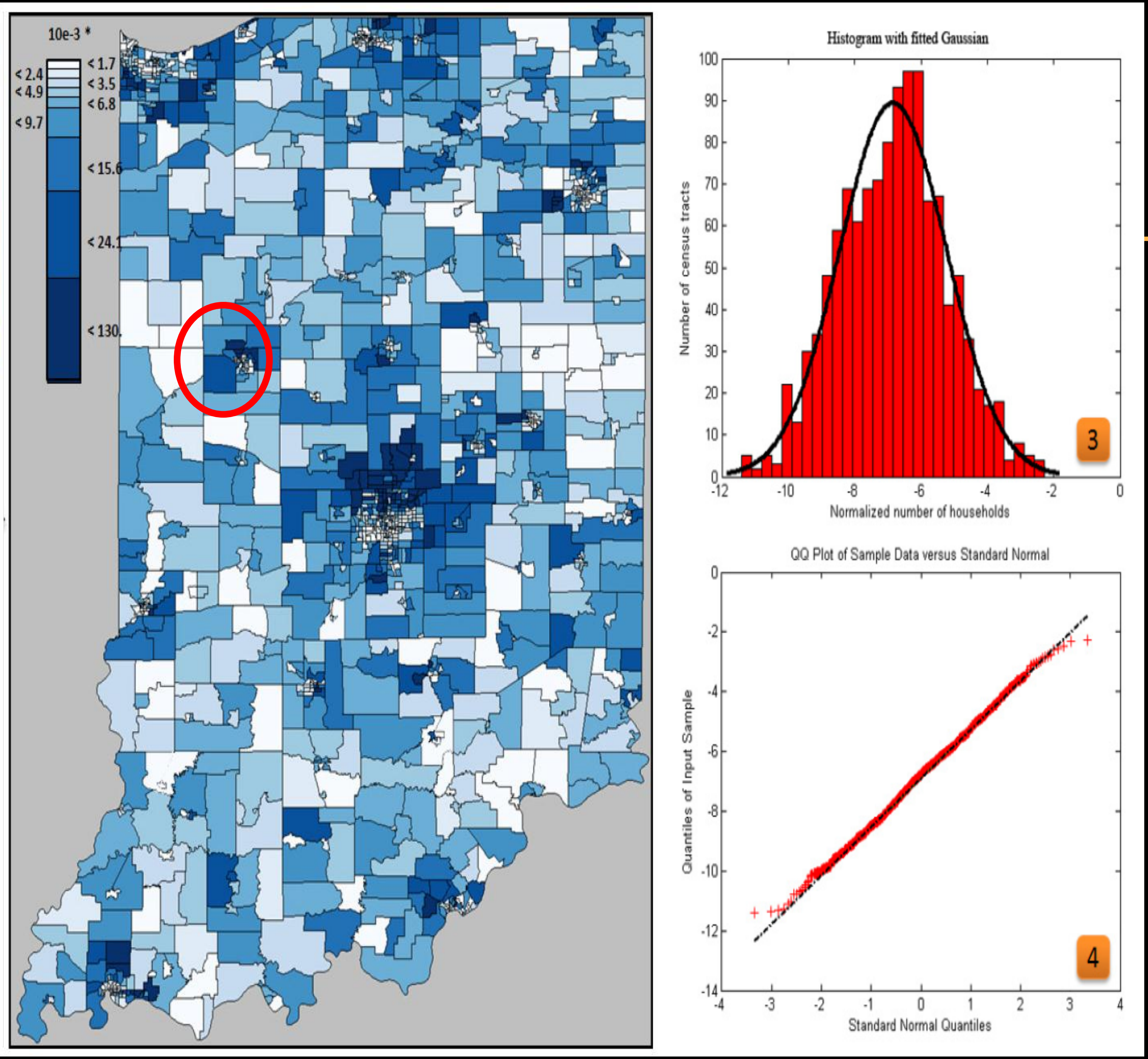


1 – M. Q. W. Baldonado, A. Woodruff, and A. Kuchinsky, “Guidelines for Using Multiple Views in Information Visualization,” in *Proc. of Advanced Visual Interfaces (AVI 2000)*, pages 110-119. ACM Press, 2000

Facilitating Spatial Analysis: Focus+Context Brushing



1 - J. W. Tukey. *Exploratory Data Analysis, 1977*. Addison-Wesley
2 – N. Andrienko, G.Andrienko, and P.Gatalsky, “Exploratory Spatio-Temporal Visualization: an Analytical Review *Journal of Visual Languages and Computing*,” special issue on Visual Data Mining December 2003, v.14 (6), pp. 5403-5413
3 - A. Robinson. Highlighting in geovisualization. *Cartography and Geographic Information Science*, In Press.



Maciejewski, R., Pattah, A., Ko., S., Hafen, R., Cleveland, W. S., Ebert, D. S., "Automated Box-Cox Transformations for Improved Visual Encoding," *IEEE Transactions on Visualization and Computer Graphics*, 19(1): 130-140, 2013.

Visual Analytics

Visualization is good for exploring data, but we can do more than just explore

"Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces"[1]

Visual analytics combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets.[2]

"A graphic display has many purposes but it achieves its highest value when it forces us to see what we were not expecting."[3]

1 - J. J. Thomas and K. A. Cook (Ed.). *Illuminating the Path: The R&D Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.
2 – Keim et al. chapter in *Information Visualization: Human-Centered Issues and Perspectives*, 2008.
3 – H. Wainer. *A Graphic Discovery: A Trout in the Milk and Other Visual Adventures*.

Spatiotemporal Analysis

Key spatial questions tend to be about causal relationships

- Do neighboring spatial regions have similar values?
- Do neighboring spatial regions have similar temporal trajectories?
- Are disparate spatial regions showing similar trajectories or values?
- Are temporal trends lagging or leading in neighboring spatial regions?

Key Temporal questions tend to focus on change[1]

- What is the magnitude of the change?
- What is the shape of the changes?
- How fast is the change occurring (what is the velocity of the change)?
- What is the direction of the change?

1 - S. Few. *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press, 2009.

Challenges in Spatiotemporal Analysis

Exploring trends and patterns in spatiotemporal data

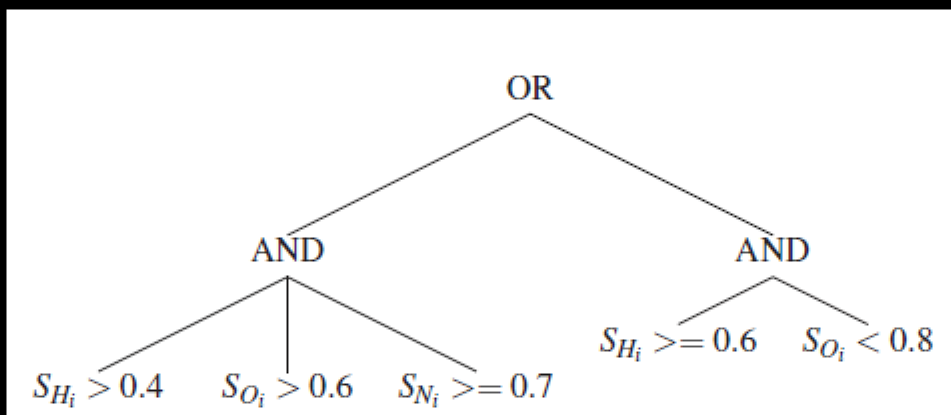
- Correlations among multi-variate data
- Explore possible predictive links among different variables

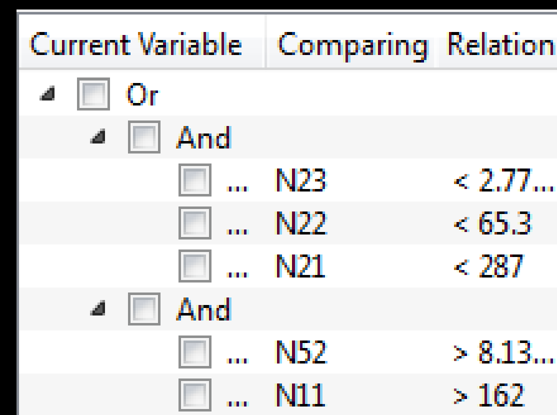
Challenges:

- End users understanding underlying statistical algorithms
- Applying algorithms at appropriate spatial and temporal scales
- Dealing with noisiness of the real-world data

Spatial Analysis: Linked Multivariate Filtering

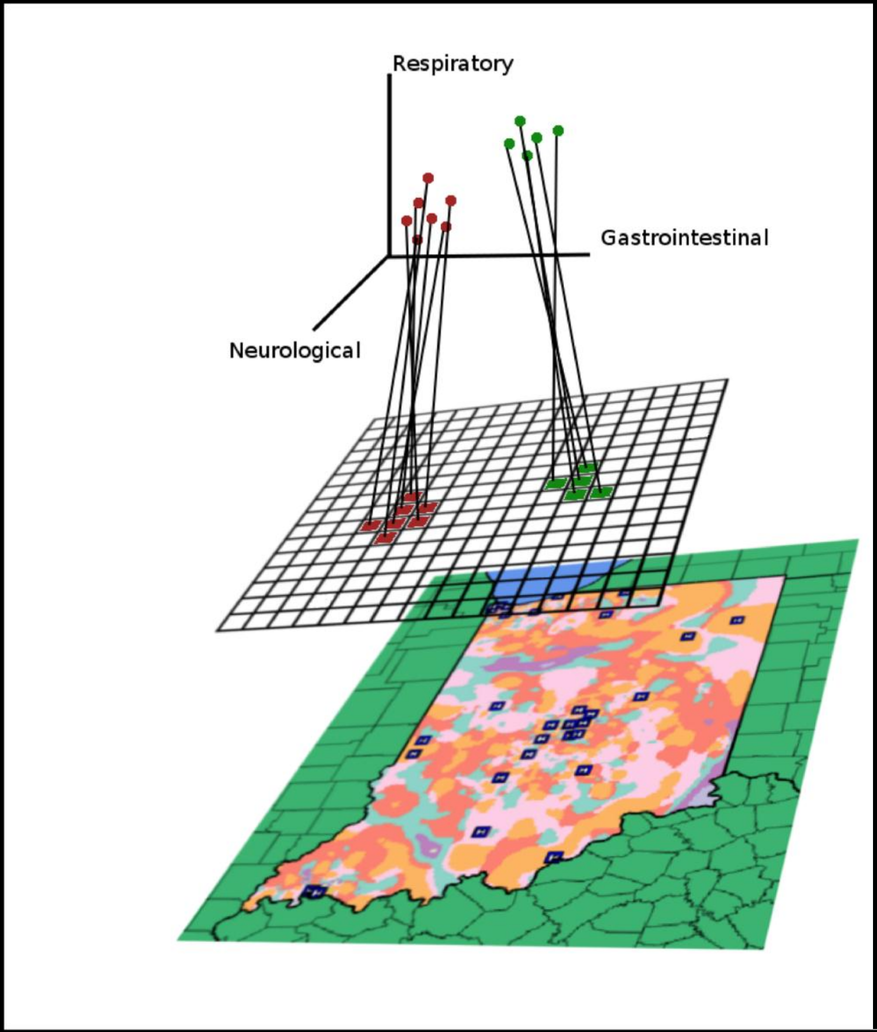
Link complex data queries directly to an interactive logic tree to facilitate multivariate queries

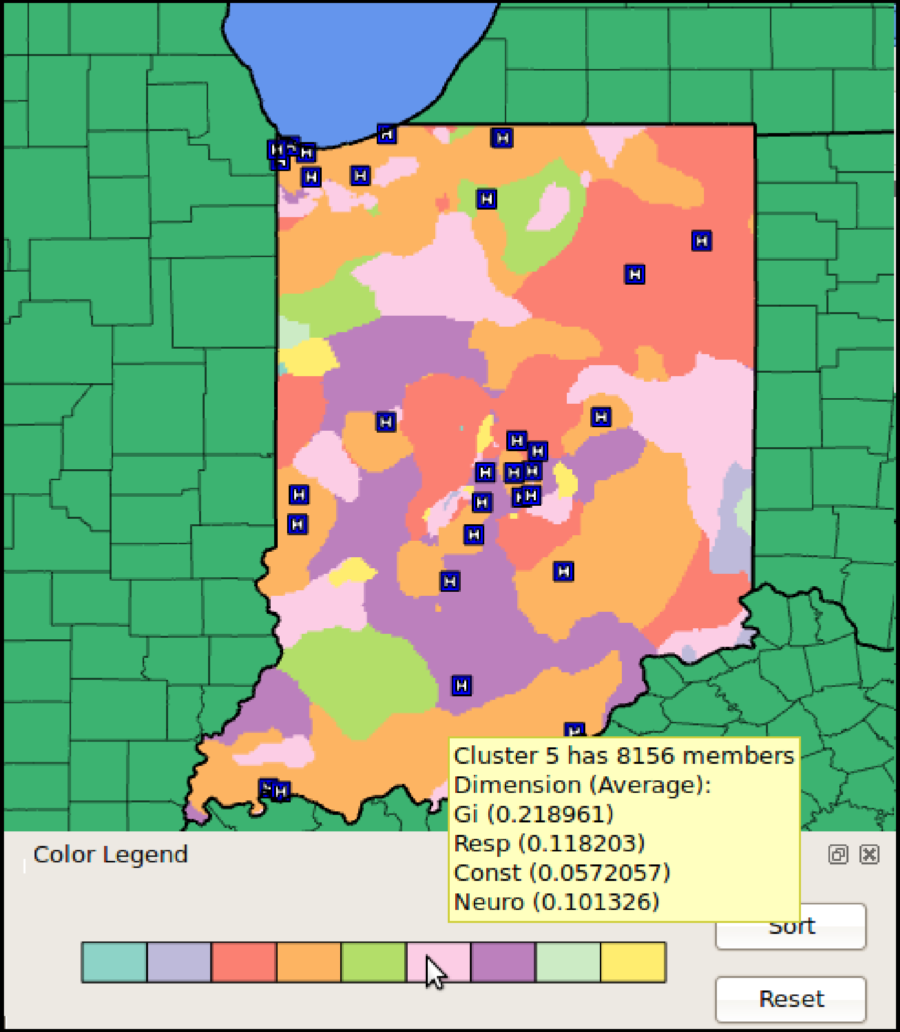




Problem is this is time consuming and tree gets complex quickly as the number of variables and constraints increase

Spatial Analysis: Multivariate Clustering





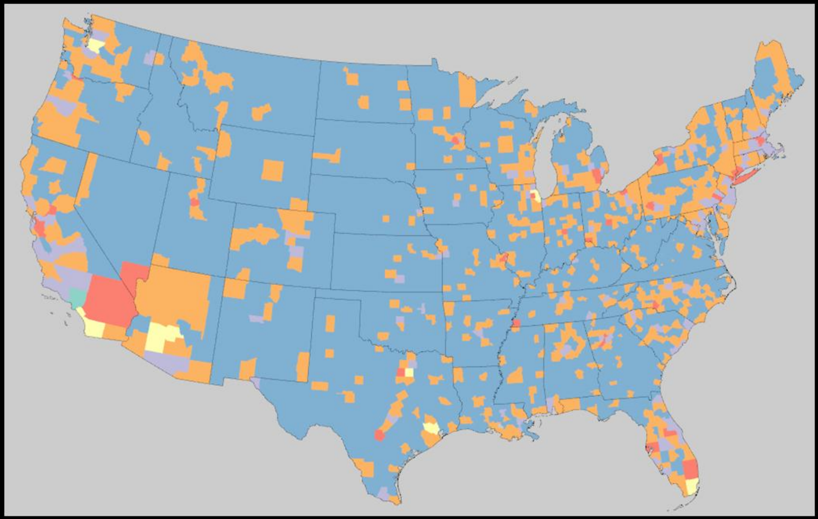
Spatial Analysis: Multivariate Clustering

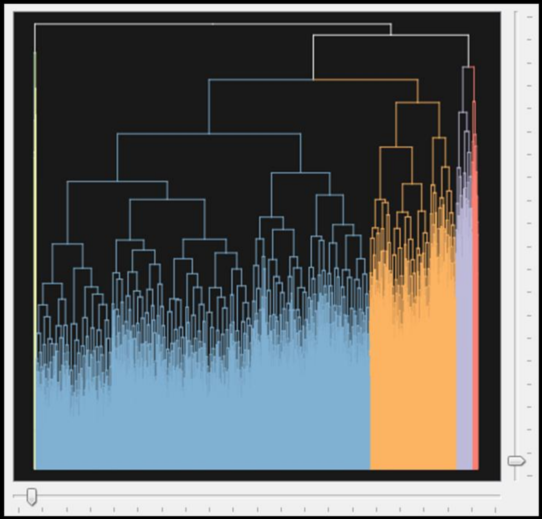
Hierarchical Clustering

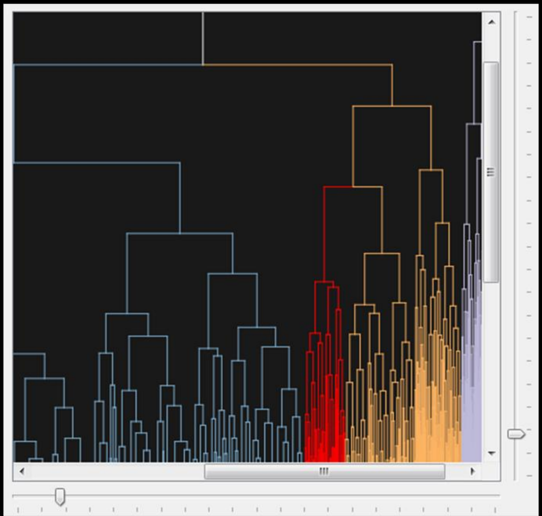
- Produces a set of nested clusters organized as a hierarchical tree
- Can be visualized as a dendrogram

No assumptions on the number of clusters

- Any desired number of clusters can be obtained by ‘cutting’ the dendogram at the proper level







Spatial Analysis: Multivariate Distance Function

In these high-dimensional clusters, we don't necessarily have a concept of how far away two points are

We want to know how spatial units are related in high-dimensional space and observe these patterns in the geographic projection

We can directly link a distance function to the brush tip

Metric Names	Formula
Euclidean	$dist = \sqrt{\sum_i (v_i^1 - v_i^2)^2}$
Manhattan	$dist = \sum_i \lvert v_i^1 - v_i^2 \rvert$
Canberra	$dist = \sum_i \frac{\lvert v_i^1 - v_i^2 \rvert}{\lvert v_i^1 \rvert + \lvert v_i^2 \rvert}$
Binary	$dist = \sum_i v_i^1 = v_i^2$

Indices of Industrial Diversity

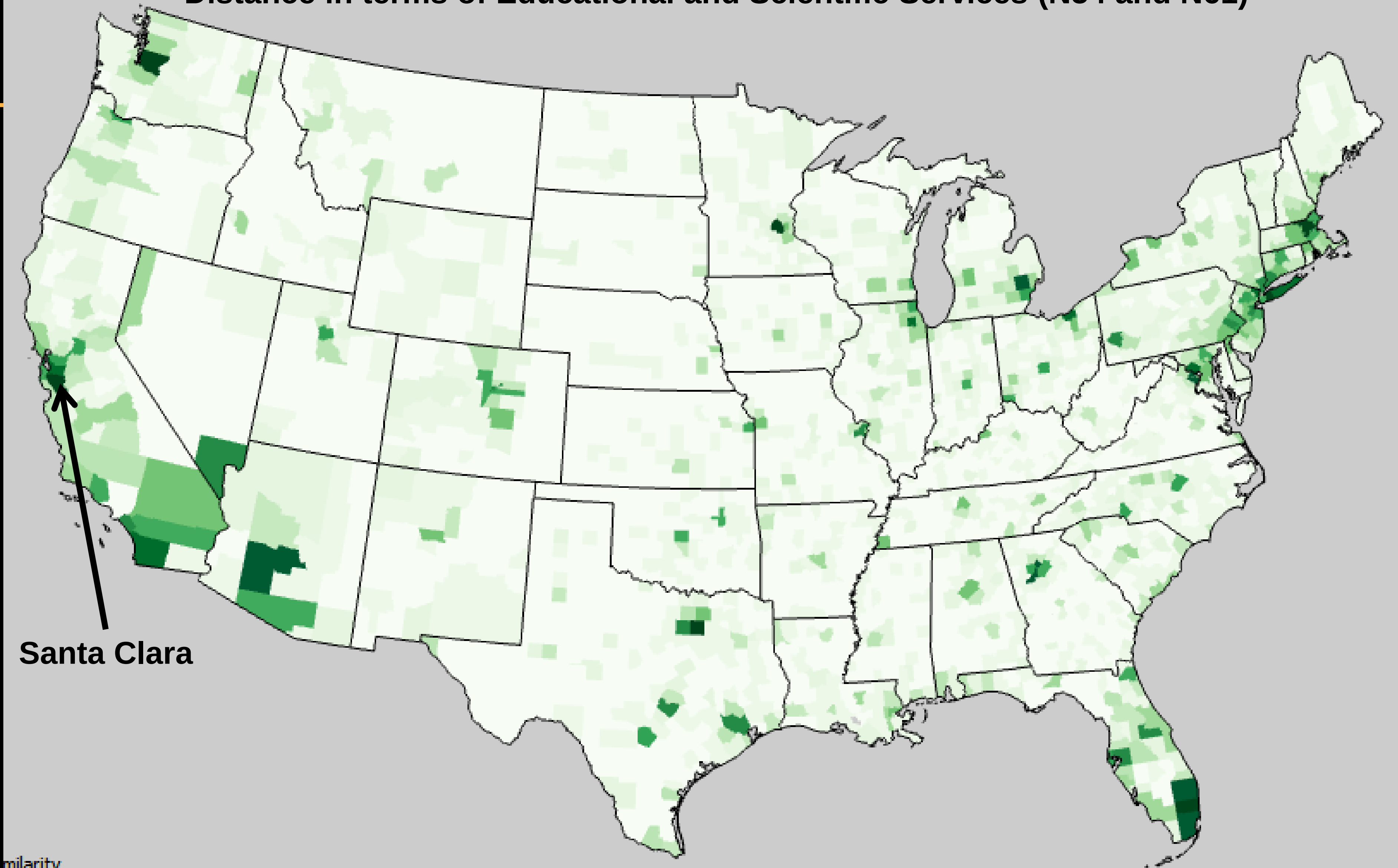
Indices of industrial diversity measure the relative concentration of a particular economic variable across a specified number of industries

Industrial diversity is measured by the number of firms as reported by the North American Industrial Classification System (NAICS)

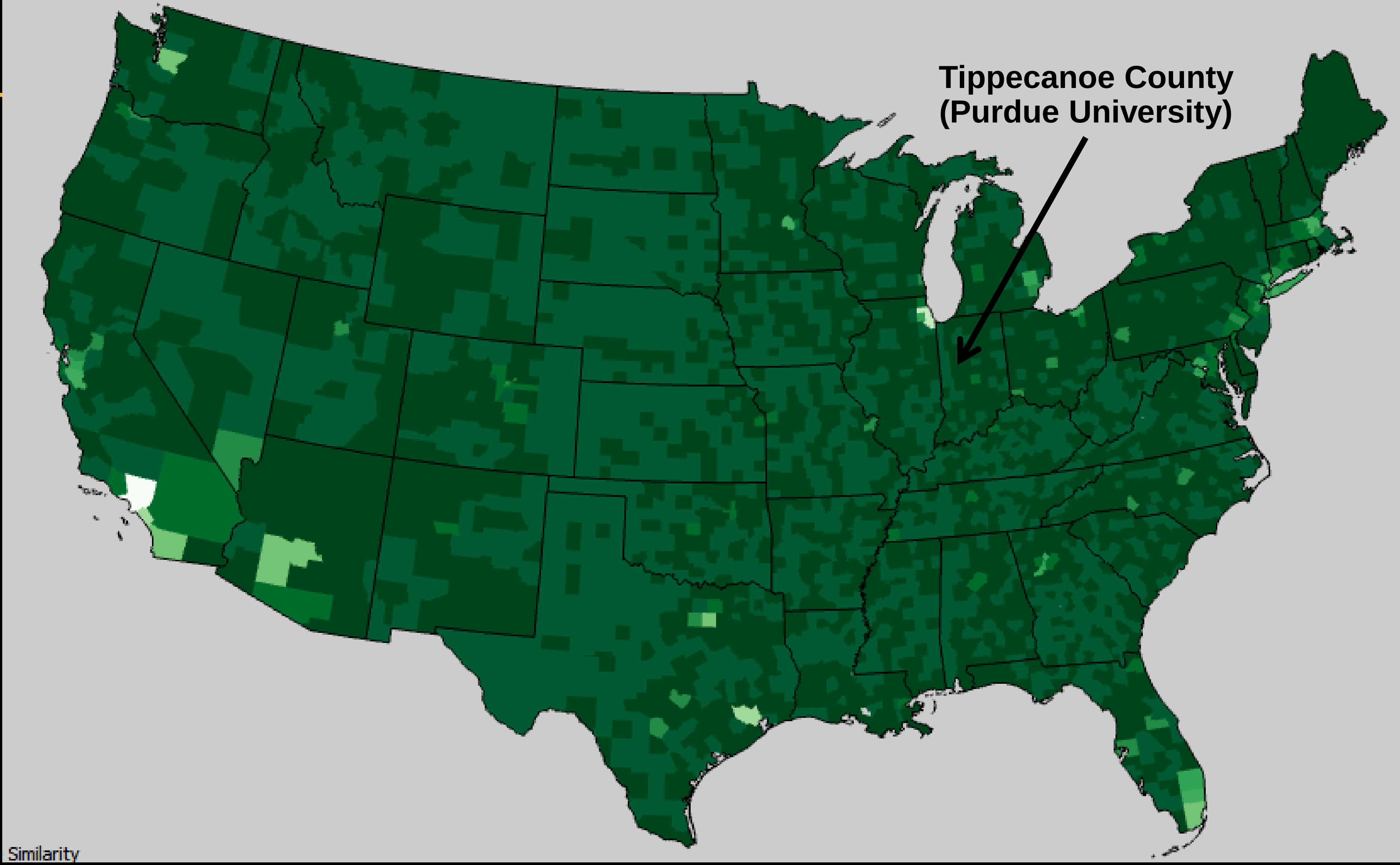
N11	Agriculture
N21	Mining
N22	Utilities
N23	Construction
N31	Manufacturing
N42	Wholesale Trade
N44	Retail Trade
N48	Transportation and Warehousing
N62	Health Care

N51	Information
N52	Finance and Insurance
N53	Real Estate
N54	Scientific and Technical Service
N55	Management of Companies
N56	Administrative Waste Services
N61	Educational Services
N71	Arts and Recreation
N72	Accommodation and Food Service

Distance in terms of Educational and Scientific Services (N54 and N61)

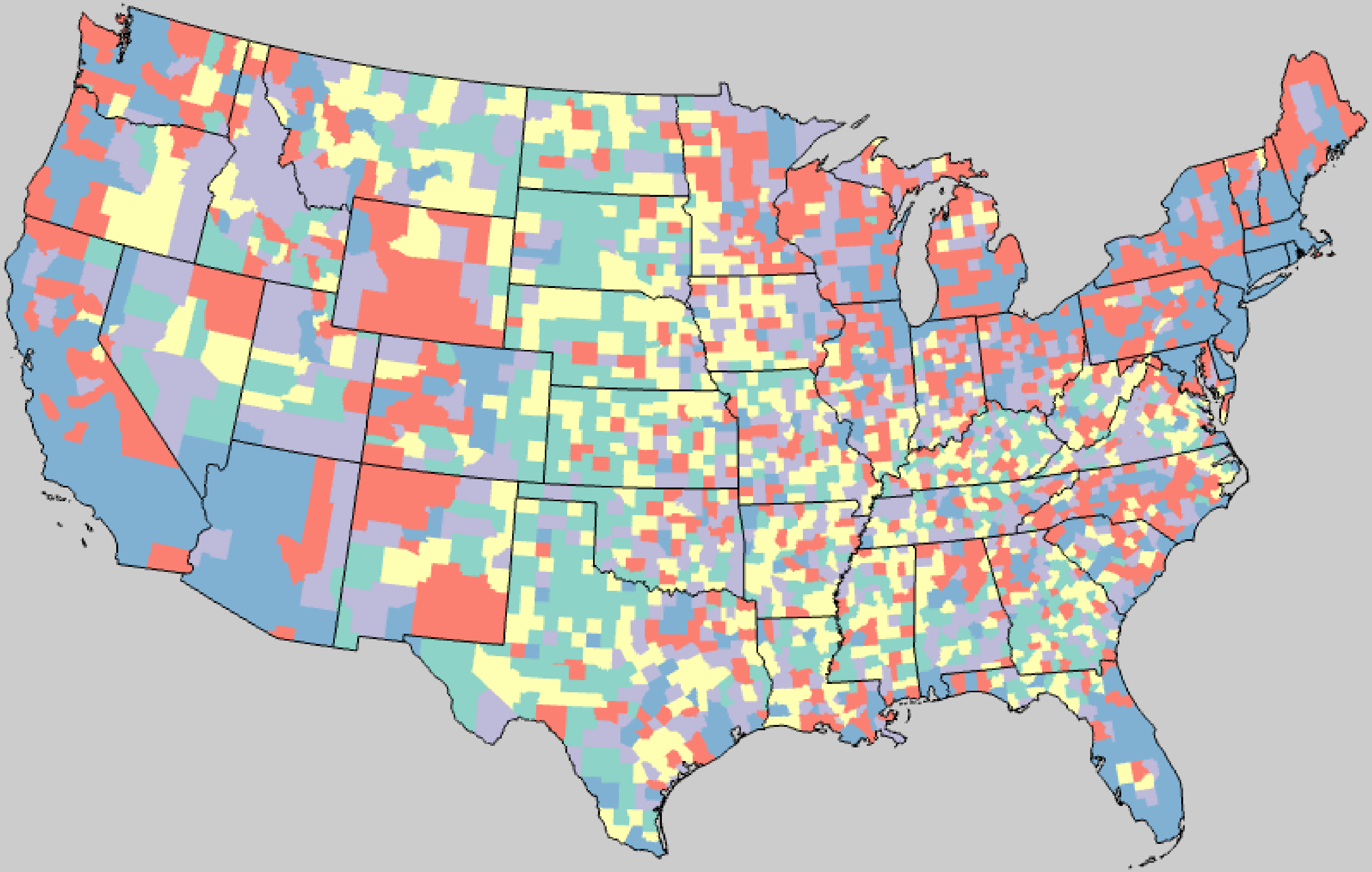


Distance in terms of Educational and Scientific Services (N54 and N61)



Distance in terms of Food and Entertainment Services (N71 and N72)

K-means clustering (k=5) in terms of Food and Entertainment Services (N71 and N72)



Spatial Analysis: Interactive Clustering

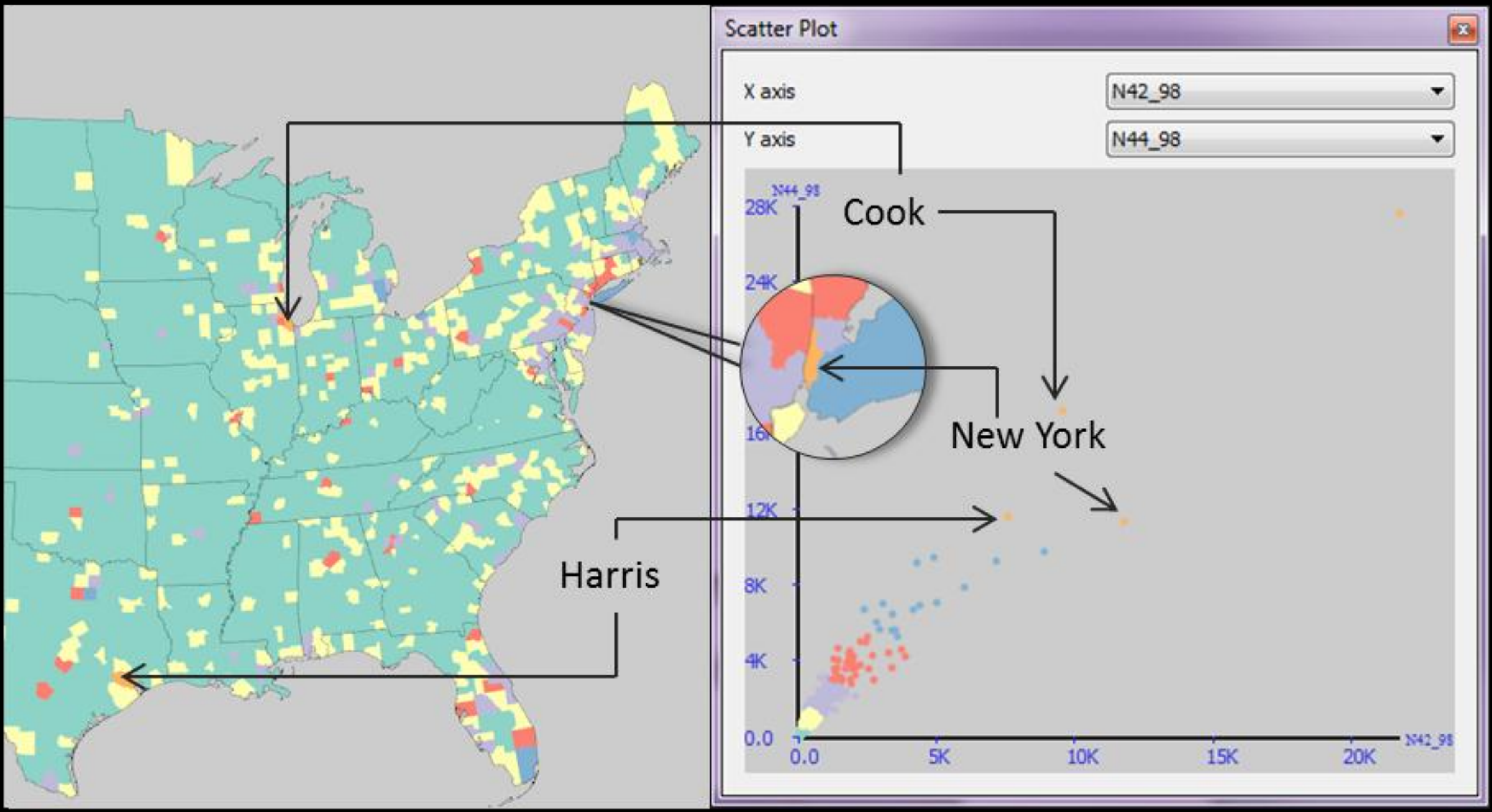
Adapt k-means so that each spatial unit has a unique weight (initially set so all weights are equal)

Analysts may now modify weights based on original k-means output, this is done by assigning spatial units to different clusters

Upon cluster reassignment, the weight of that spatial unit is updated to be $w_i = 1 + \sqrt{s_j}$ *where s_j is the size of cluster j*

Now the cluster centroids are recomputed and k-means rerun such that cluster centroids are $c_j = \frac{\sum_{k \in J} w_k u_k}{\sum_{k \in J} w_k}$

Spatial Analysis: Interactive Clustering



Spatiotemporal Analysis

Key spatial questions tend to be about causal relationships

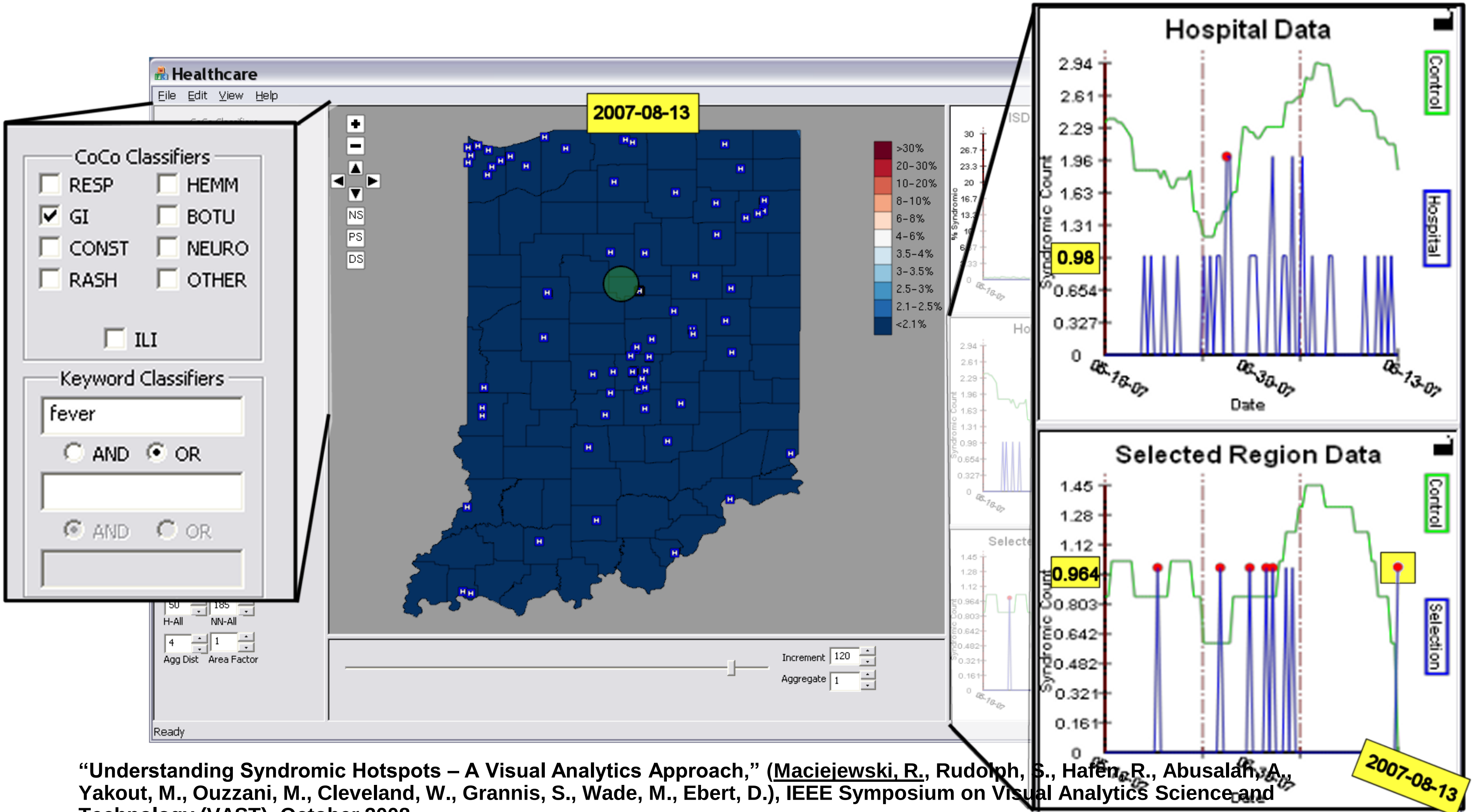
- Do neighboring spatial regions have similar values?
- Do neighboring spatial regions have similar temporal trajectories?
- Are disparate spatial regions showing similar trajectories or values?
- Are temporal trends lagging or leading in neighboring spatial regions?

Key Temporal questions tend to focus on change[1]

- What is the magnitude of the change?
- What is the shape of the changes?
- How fast is the change occurring (what is the velocity of the change)?
- What is the direction of the change?

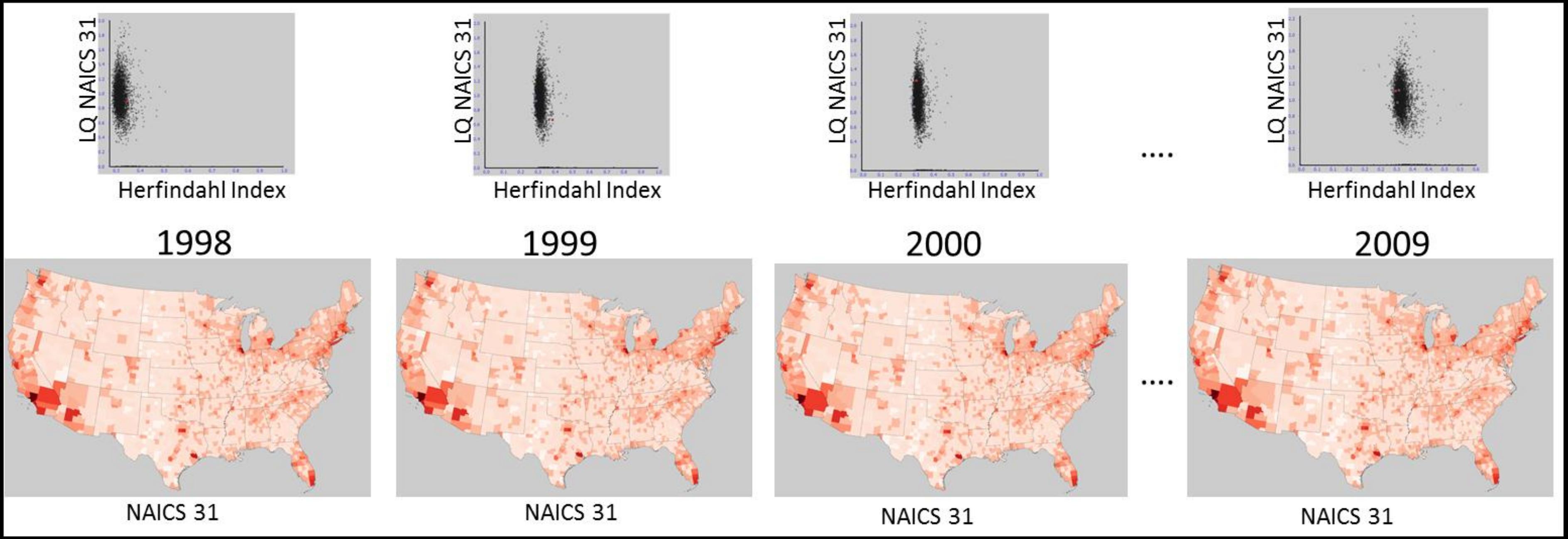
1 - S. Few. *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press, 2009.

Facilitating Spatiotemporal Analysis: Linked Time Series



"Understanding Syndromic Hotspots – A Visual Analytics Approach," (Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W., Grannis, S., Wade, M., Ebert, D.), IEEE Symposium on Visual Analytics Science and Technology (VAST), October 2008

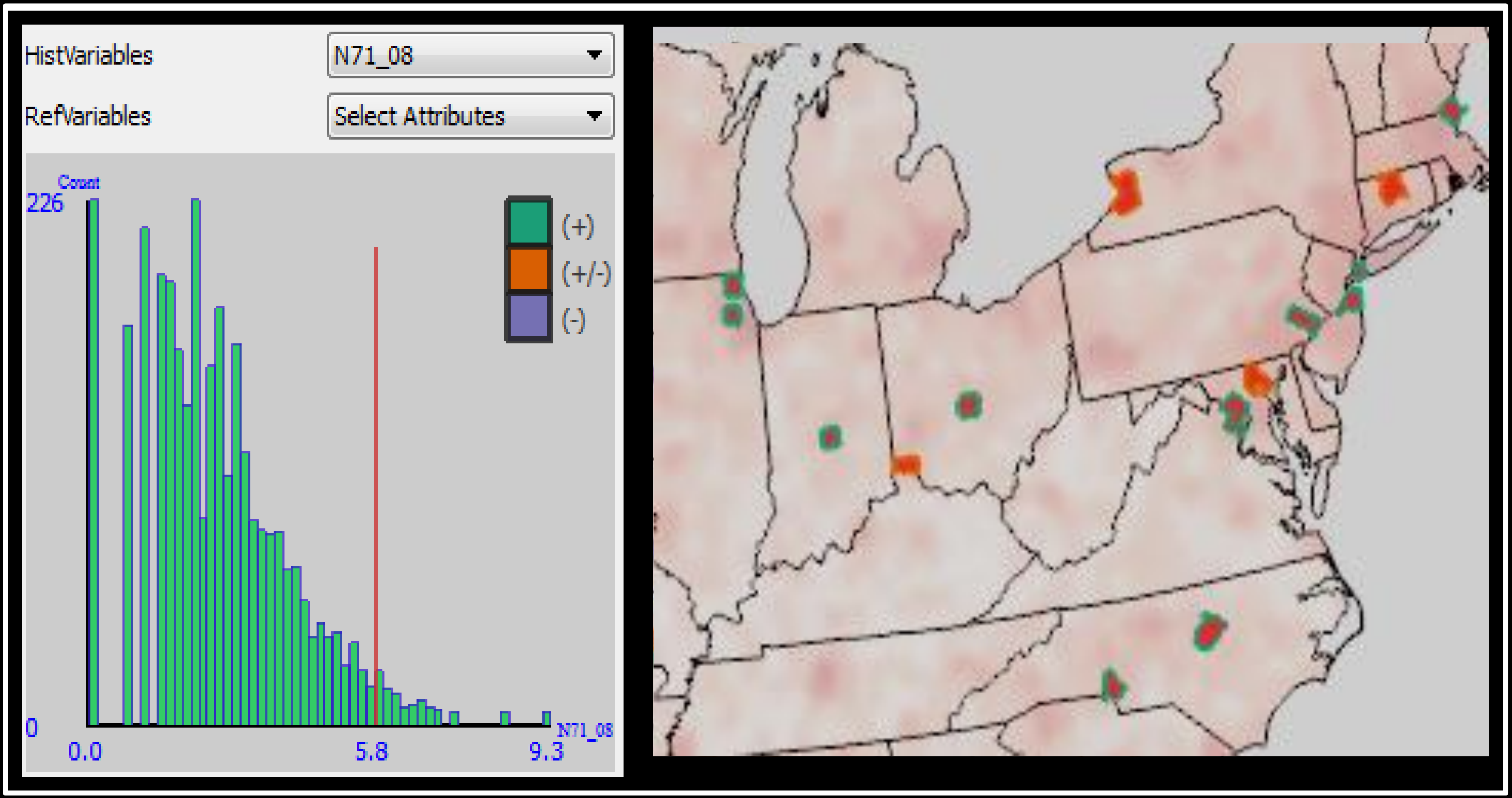
Facilitating Spatiotemporal Analysis: Animation and Small Multiples



1 – B. Tversky, J. B. Morrison, and M. Bertrancourt. Animation: Can it facilitate? *International Journal of Human-Computer Studies*, 57:247– 262, 2002.

2 – M. Harrower. The cognitive limits of animated maps. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 42(4):349–357, 2009.

Facilitating Spatiotemporal Analysis: Identifying Temporal Thresholds



Facilitating Spatiotemporal Analysis: Temporal Similarity Brushing

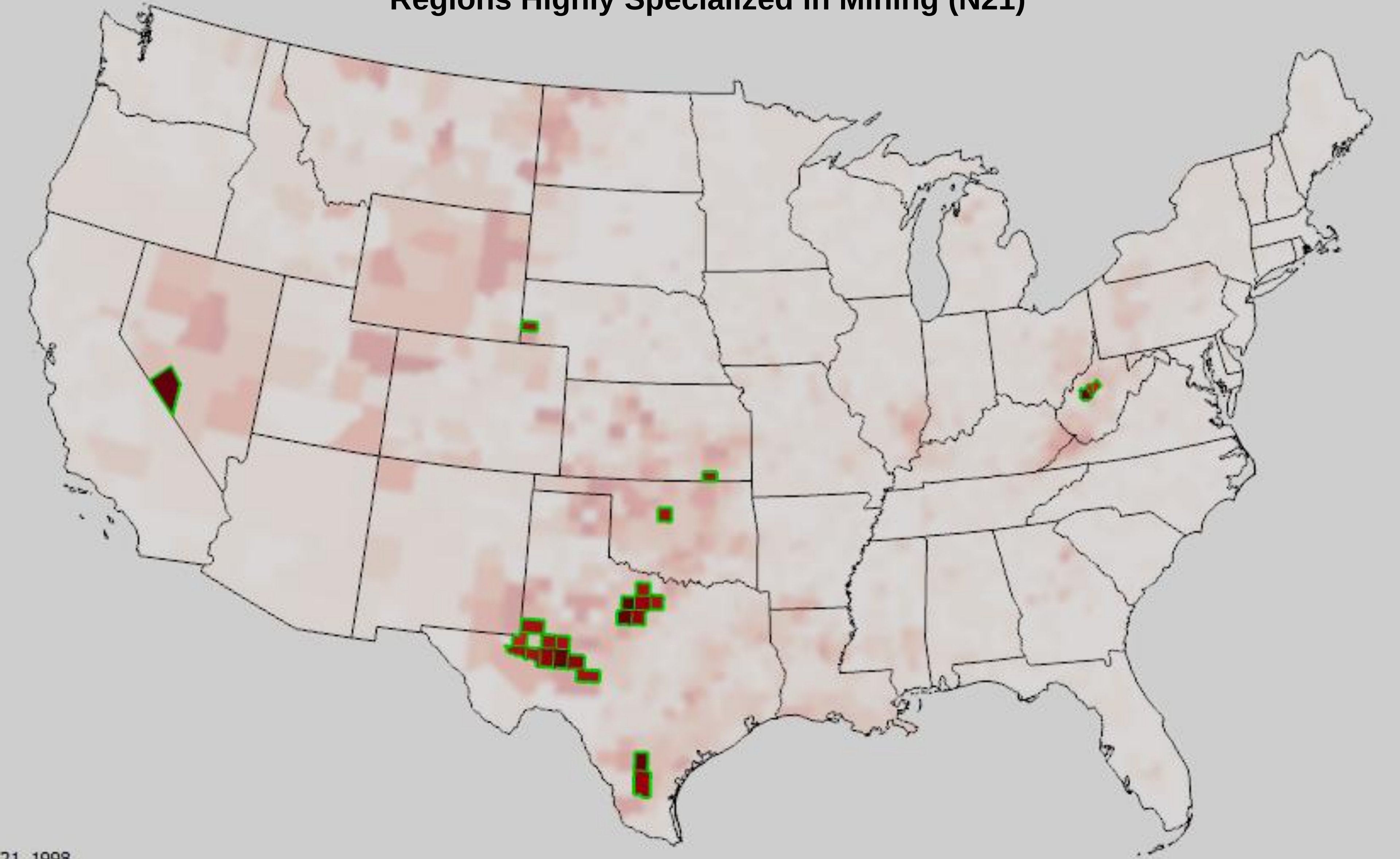
Many of these questions depend on temporal trajectories

We can apply time series similarity work as a means of determining which regions have similar trajectories

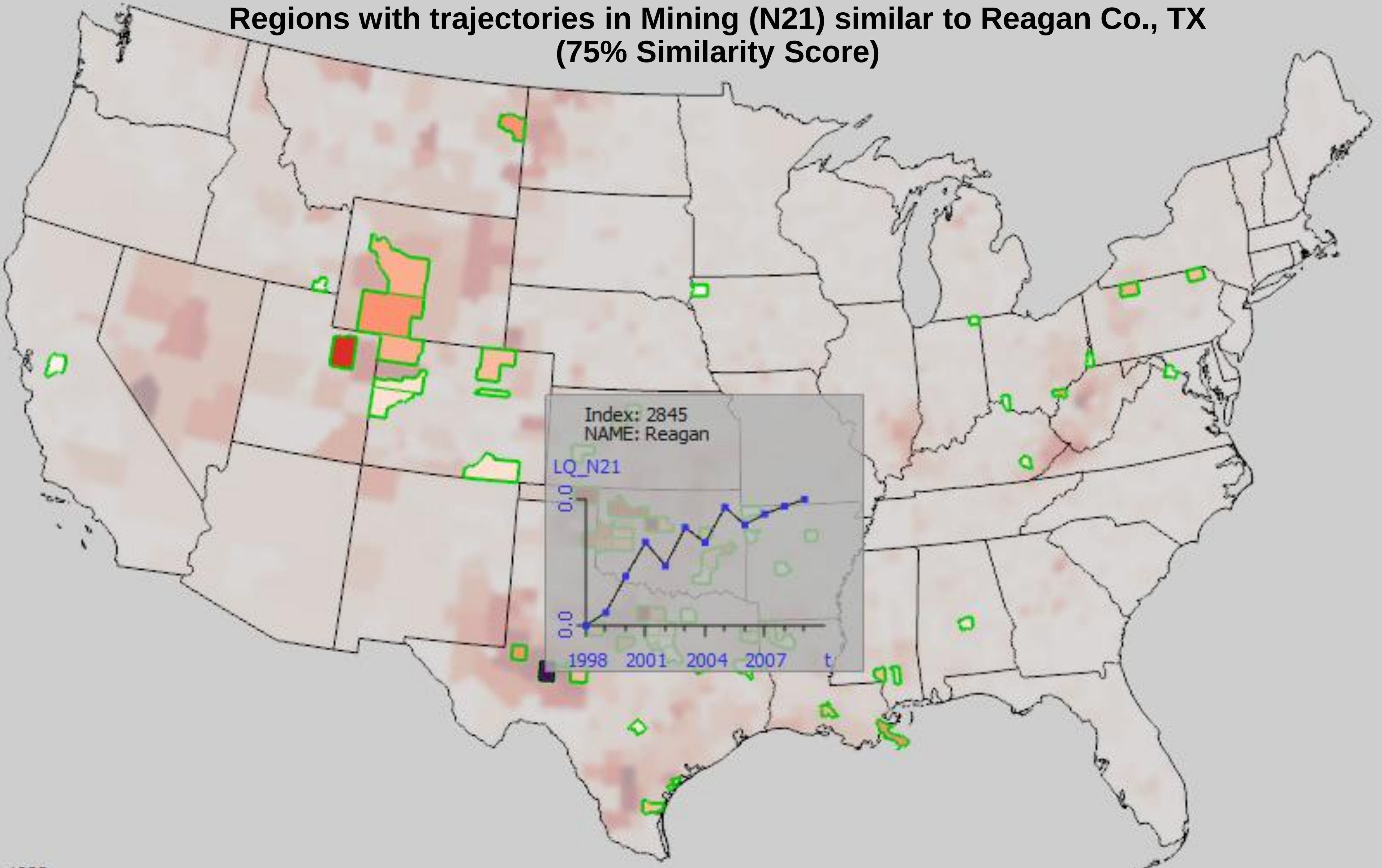
One common metric is the sequence normalized metric which will focus on matching trends as opposed to exact values

$$dist_{SNE}(\mathbf{p},\mathbf{q}) = \sqrt{\sum_{i=1}^{n}\left(\frac{q_i-\mu_q}{\sigma_q}-\frac{p_i-\mu_p}{\sigma_p}\right)^2}.$$

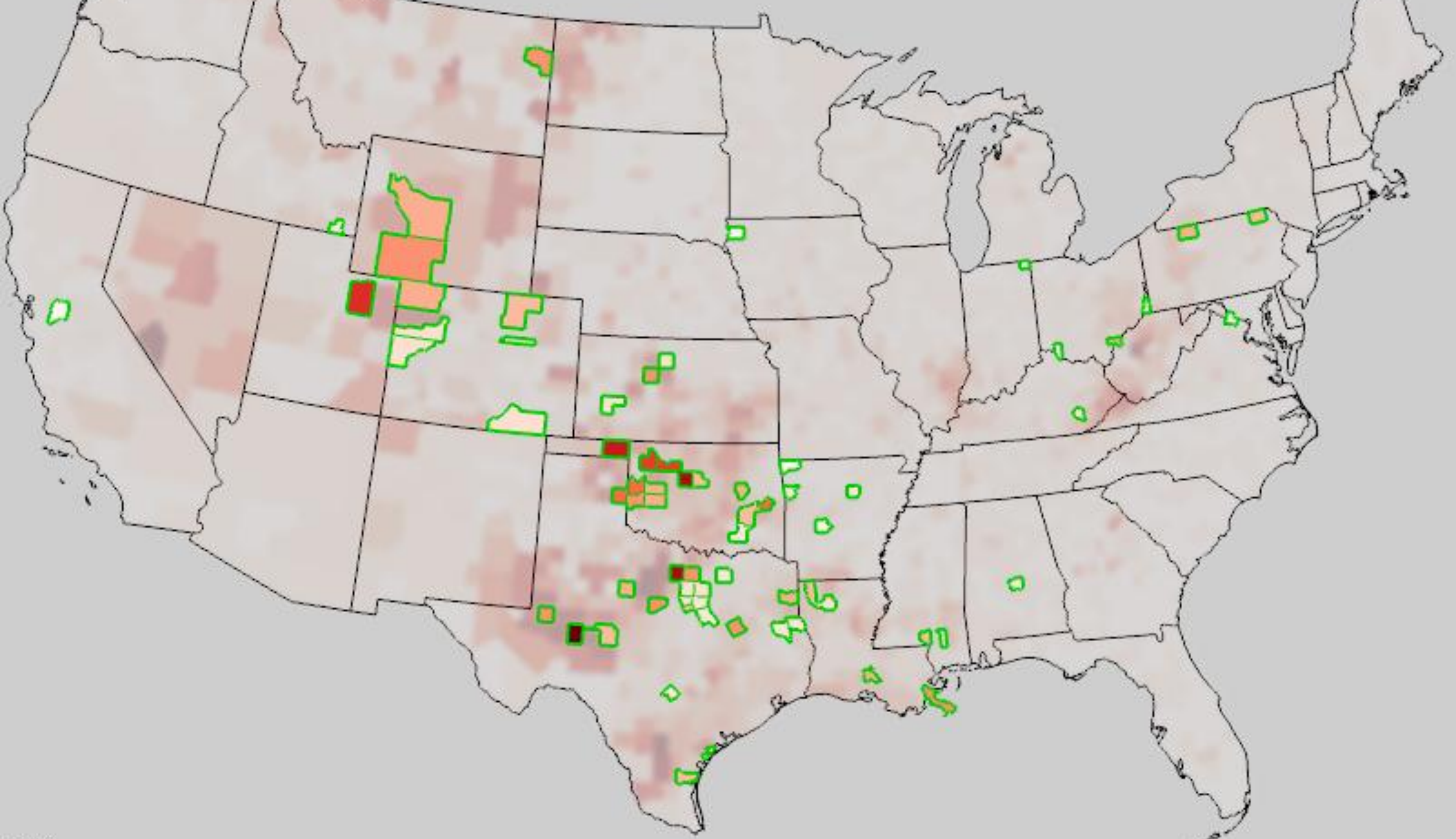
Regions Highly Specialized in Mining (N21)



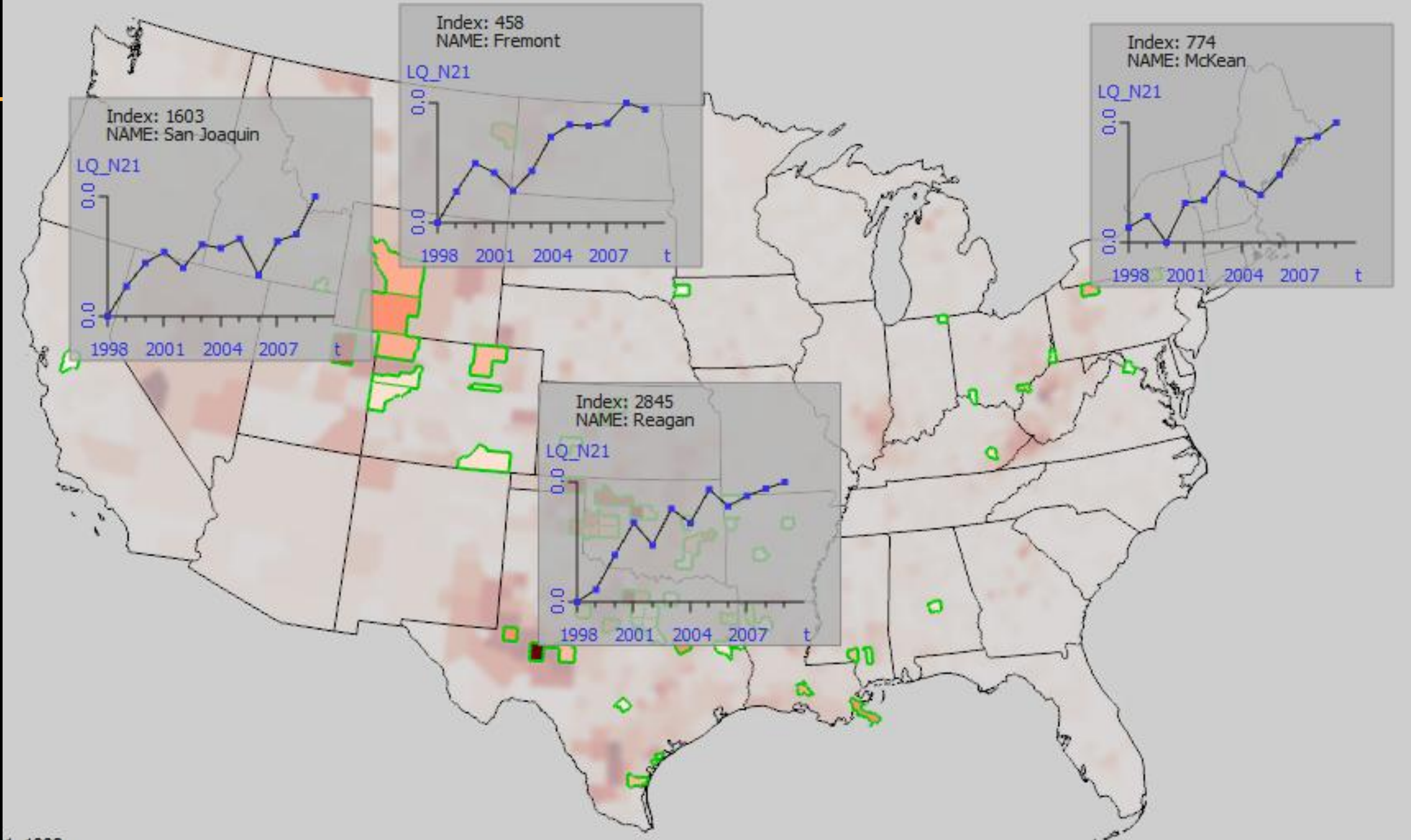
Regions with trajectories in Mining (N21) similar to Reagan Co., TX (75% Similarity Score)



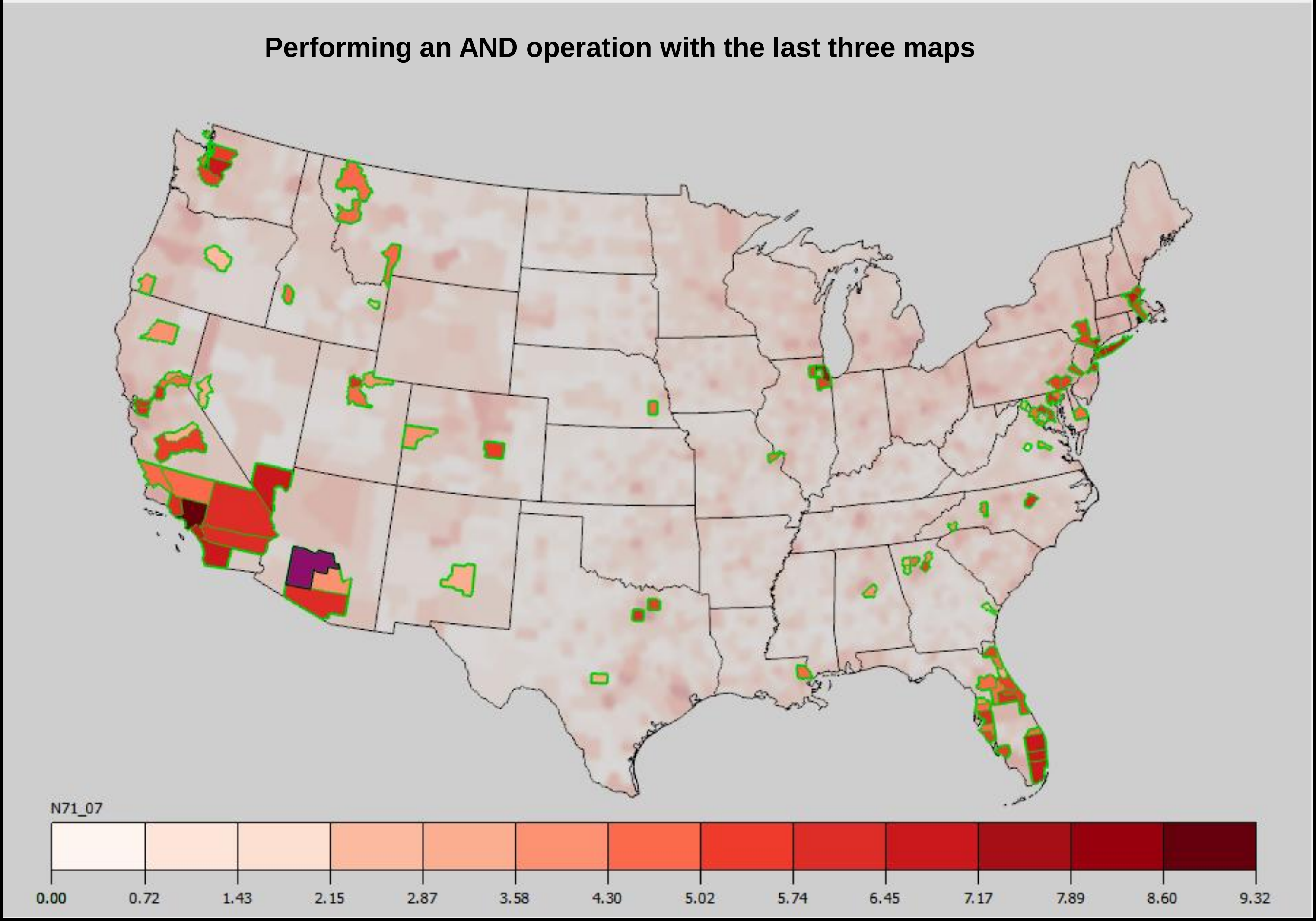
Regions with trajectories in Mining (N21) similar to Reagan Co., TX
(75% Similarity Score)



What Does 75% Similar Look Like?



Applied Analytical Brushing: A Case Study in Industrial Diversity

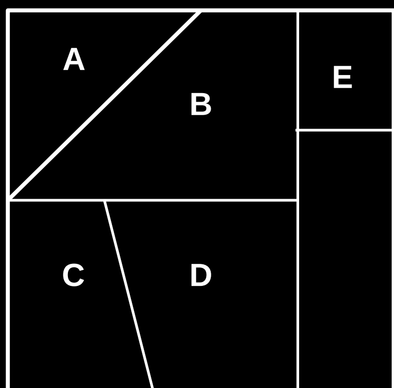


Temporally Coherent Spatial Clusters

What is the best way to cluster spatiotemporal data?

- ***Allow time to be a vector in the cluster algorithm***
  - Now spatial units can only belong to one cluster
- ***Cluster data per time step and then link clusters across time step by space***
  - Now spatial units can belong to multiple clusters, but need to maintain coherence between cluster labels at each time step

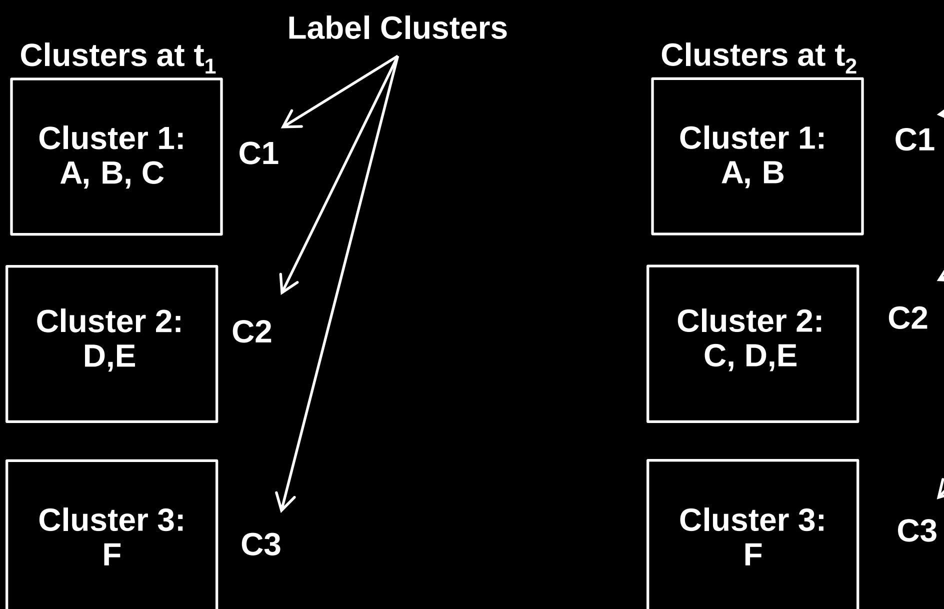
Temporally Coherent Spatial Clusters



Assume that we have some spatial arrangement of data with multivariate measures per time step

For each time step, we perform k-means clustering

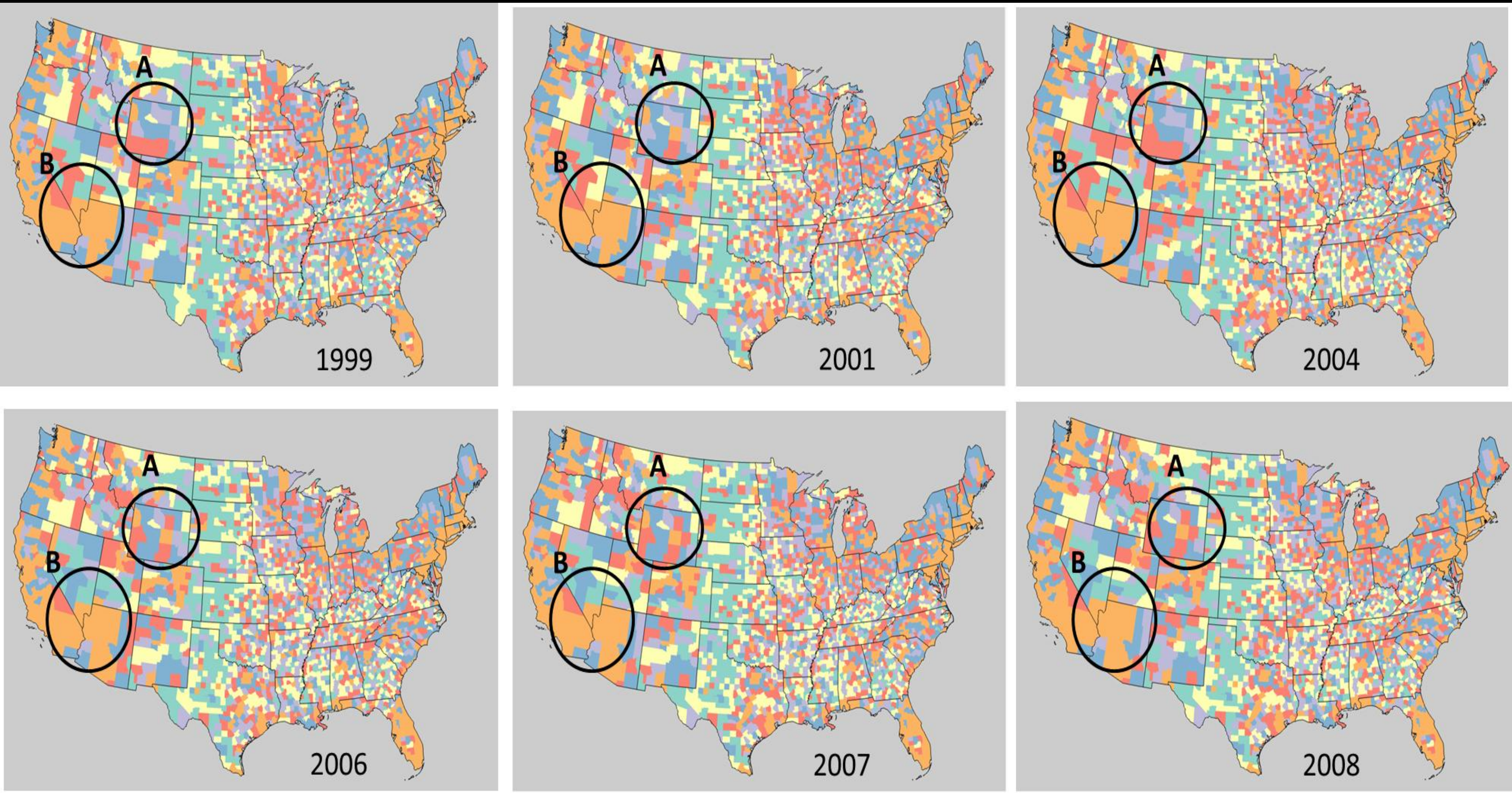
Label Clusters, but use information from previous time step, minimize the number of elements for which the cluster label will change (Kuhn-Munkres (KM) Algorithm[1])



So all spatial units maintain the same label from t_1 to t_2 except for element C

1 - J. Munkres. Algorithms for the assignment and transportation problems, Journal of the Society for Industrial & Applied Mathematics, 5(1):32–38, 1957.

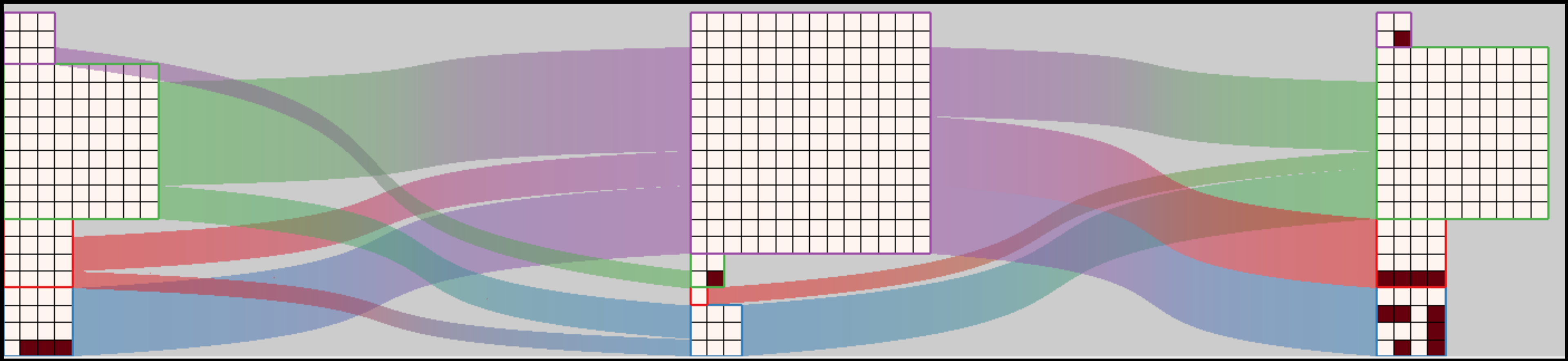
Temporally Coherent Spatial Clusters



Temporally Coherent Spatial Clusters

What if we ignore the space and instead visualize how the element cluster labels change?

- What time steps should we present?
- Can we globally minimize the label changes?
- Measures of stability based on label changes?



Challenges

Scalability

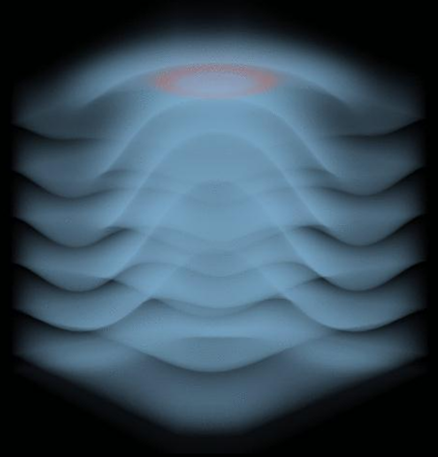
- ***Unfortunately, as dimensionality increases, performance of similarity metrics decreases[1]***
  - Is the use of fractional distance metrics intuitive for users?
  - Should dimensional reduction/clustering be done spatially, temporally or spatiotemporally?

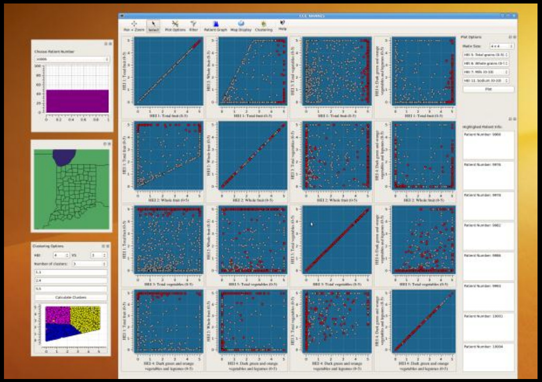
How do we move from hypothesis generation to hypothesis testing?

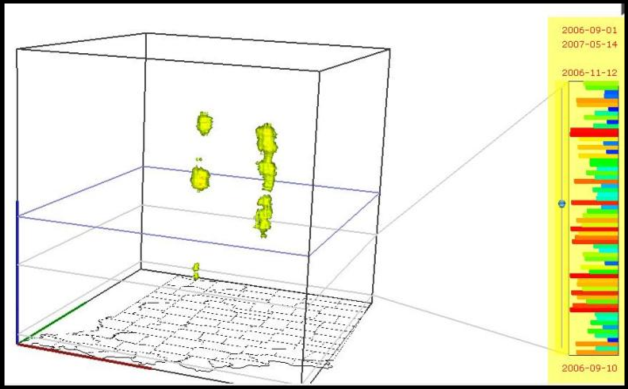
- ***Traditional techniques such as multiple regression do not scale well to high-dimensions***

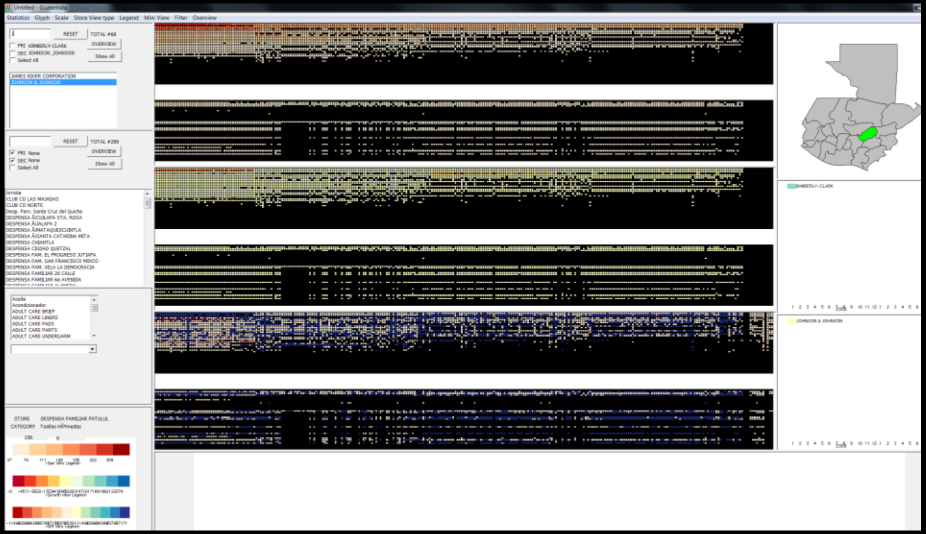
1 – Charu C. Aggarwal, Alexander Hinneburg, and Daniel A. Keim. 2001. On the Surprising Behavior of Distance Metrics in High Dimensional Spaces. In *Proceedings of the 8th International Conference on Database Theory* (ICDT '01), Jan Van den Bussche and Victor Vianu (Eds.). Springer-Verlag, London, UK, 420-434.

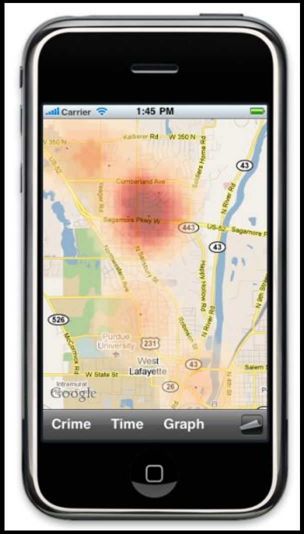
Questions?






2006-09-01
2007-05-14
2006-11-12
2006-09-10




Carrier
1:45 PM
West Lafayette
Google
Crime
Time
Graph

Facilitating Spatial Analysis: Multivariate Representations

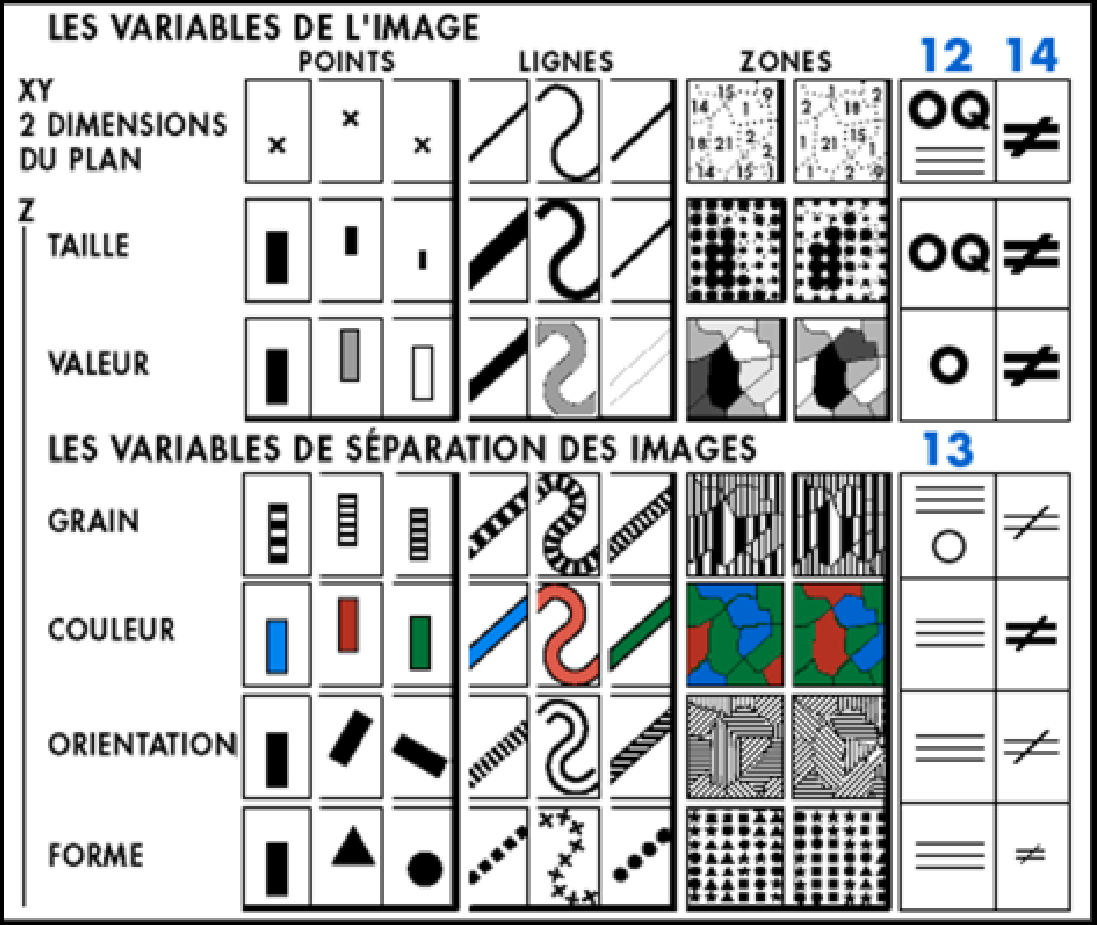
Visualization is concerned primarily with a mapping to visual form

[x,y]

- ~~Position~~

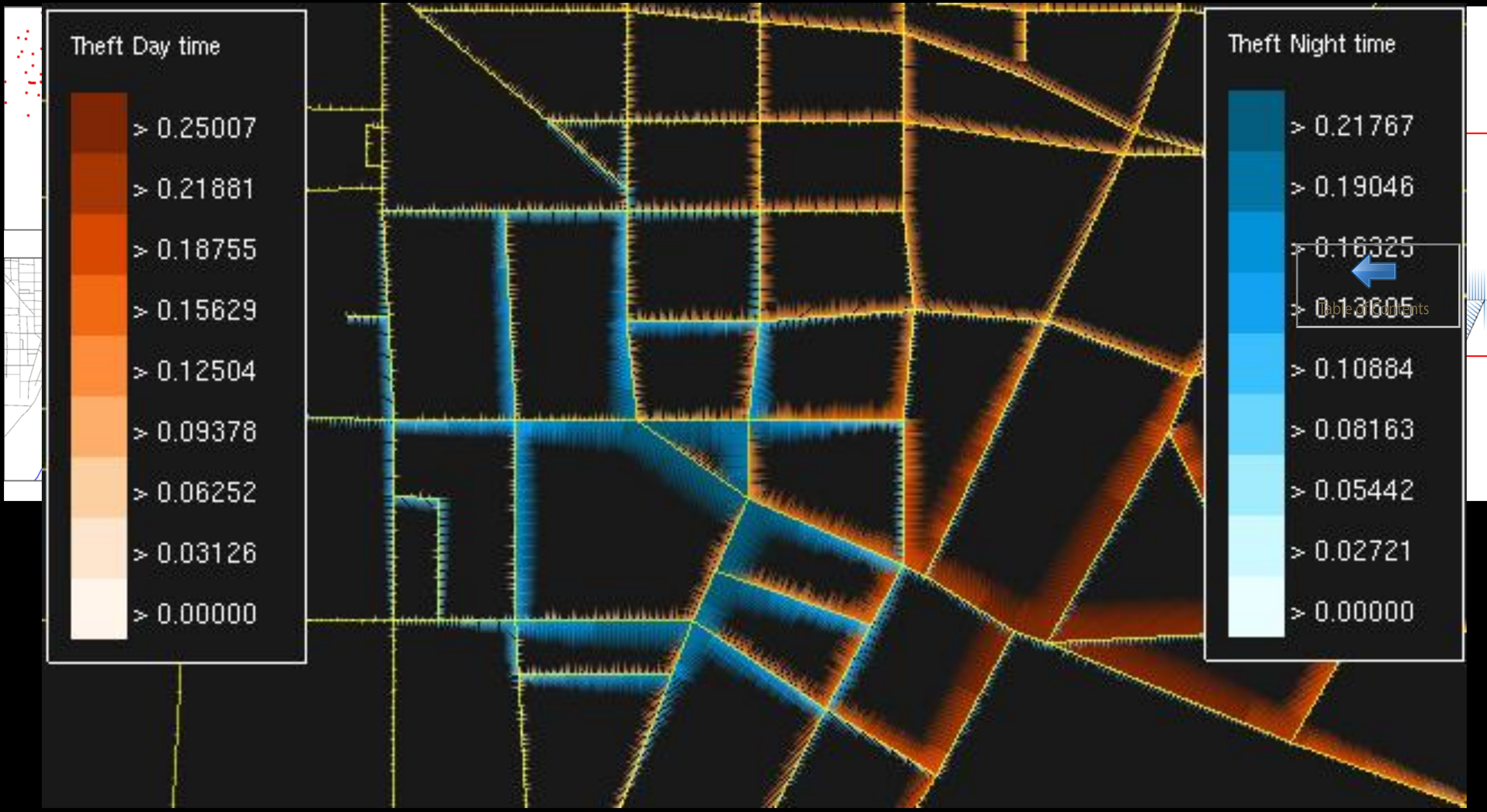
[z]

- ~~Size (Taille)~~
- Value (Valeur)
- Color (Couleur)
- Texture (Grain)
- ~~Orientation~~
- ~~Shape (Forme)~~



J Bertin (1967), *The Semiology of Graphics*

Bristle Maps: A Multivariate Abstraction Technique for Geovisualization

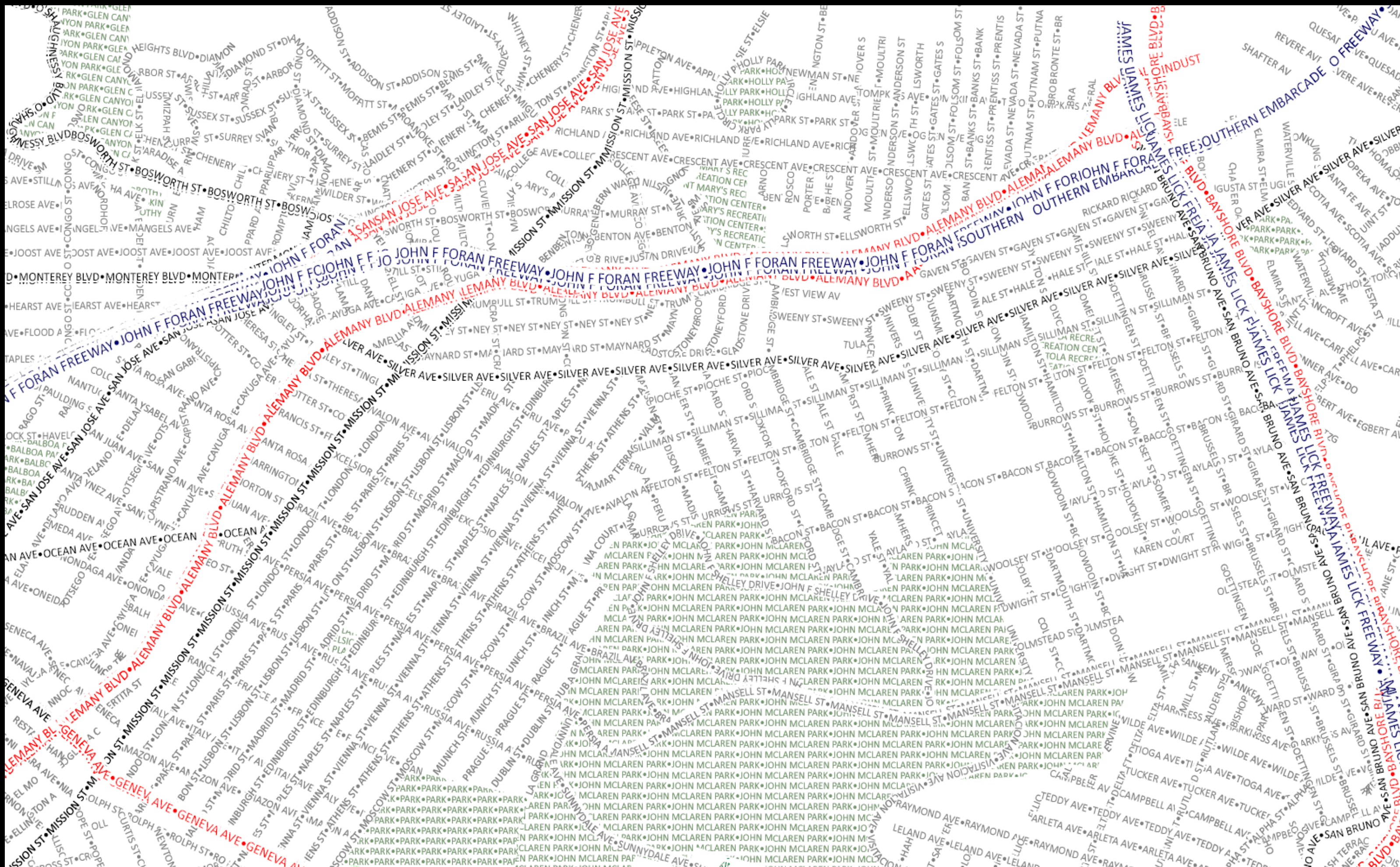


Kim, S., Maciejewski, R., Malik, A., Jang, Y., Ebert, D. S., Isenberg, T., "Bristle Maps: A Multivariate Abstraction Technique for Geovisualization," *IEEE Transactions on Visualization and Computer Graphics*, (To appear)

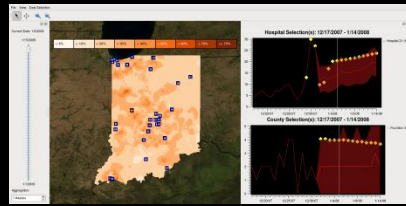
Table of Contents

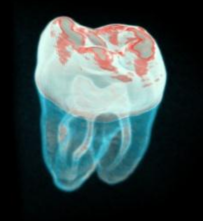
Typographic Maps: An Illustrative Abstraction Technique

Table of Contents



1 – **Afzal, S., <u>Maciejewski, R.</u>, Jang, Y., Elmqvist, N., Ebert. D. S., "Spatial Text Visualization Using Automatic Typographic Maps,"** ***IEEE Transactions on Visualization and Computer Graphics*****, 18(12): 2556-2564, 2012.**





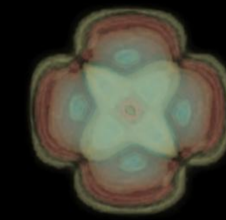
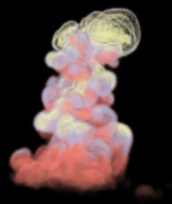


Table of Contents

Applied Visual Analytics for Exploring the National Health and Nutrition Examination Survey

The Ongoing Data Deluge



- *Since 2003, digital information has accounted for 90% of all information produced*[1]
- *In 2009, drones from Iraq and Afghanistan recorded 24 years of video footage*
- *In 2010, the amount of information added annually to the digital universe was estimated to be nearly 1 ZB*
- *Wal-mart process > 1 million transactions per hour*
- *By 2013 Cisco estimates the annual internet traffic will be 667 EBs*

1 – C. R. Johnson, R. Moorhead, T. Munzner, H. Pfister, P. Rheingans, and T. S. Yoo, Eds., *NIH-NSF Visualization Research Challenges Report*, IEEE Press, ISBN 0-7695-2733-7, 2006

Research Challenges in the Data Deluge

Long-term fundamental challenges

- Mobility
- Scale
- Adaptation
- Uncertainty
- Multisource, cross-media analysis
- Numerical to visual analysis

"All this data must be fused with relevant contextual or situational information and visualized in such a way to give our forces a clear picture of threats, options and potential consequences.

The science and technology community has been successful in giving our forces a data advantage; now it's time to give them a decision advantage."

-Zachary J. Lemnios Director, Defense Research and Engineering Department of Defense February 17, 2010

Exploiting the Data Deluge

- ***Credit card companies monitor purchases for fraudulent activity***
- ***Mobile operators analyze call patterns to see if people you frequently call use a rival network***
- ***The oil industry explores seismic data before drilling***
- ***The US government collects tons of population based data that we can exploit for analysis!***

National Health and Nutrition Examination Survey (NHANES)

Designed to assess health and nutritional status of adults and children in the US

- *Combines interviews and physical examinations*
- *Exams a nationally representative sample of about 5,000 persons per year*
- *Includes demographic, socioeconomic, dietary and health-related questions*

Healthy Eating Index (HEI)

Measure of diet quality that assesses conformance to federal dietary guidance

USDA uses HEI to monitor diet quality of the US population

HEI-2005 uses recommendations of the MyPyramid guidance as the baseline

1 - Guenther, P. M., Reedy, J., Krebs-Smith, S. M., Reeve, B. B., "Evaluation of the Healthy Eating Index-2005," *Journal of the American Dietetic Association*, 108:1854-1896, 2008.

2 – Freedman, L. S., Guenther, P. M., Krebs-Smith, S. M., Kott, P. S., "A Population's Mean Healthy Eating Index-2005 Scores are Best Estimated by the Score of the Population Ratio when One 24-Hour Recall is Available," Journal of Nutrition, 138:1725-1729, 2008.

Healthy Eating Index–2005 components and standards for scoring[1]

Component		Maximum points	Standard for maximum score	Standard for minimum score of zero
Total Fruit (includes 100% juice)	(HEI-1)	5	≥0.8 cup equiv. per 1,000 kcal	No Fruit
Whole Fruit (not juice)	(HEI-2)	5	≥0.4 cup equiv. per 1,000 kcal	No Whole Fruit
Total Vegetables	(HEI-3)	5	≥1.1 cup equiv. per 1,000 kcal	No Vegetables
Dark Green and Orange Vegetables and Legumes[2]	(HEI-4)	5	≥0.4 cup equiv. per 1,000 kcal	No Dark Green or Orange Vegetables or Legumes
Total Grains	(HEI-5)	5	≥3.0 oz equiv. per 1,000 kcal	No Grains
Whole Grains	(HEI-6)	5	≥1.5 oz equiv. per 1,000 kcal	No Whole Grains
Milk[3]	(HEI-7)	10	≥1.3 cup equiv. per 1,000 kcal	No Milk
Meat and Beans	(HEI-8)	10	≥2.5 oz equiv. per 1,000 kcal	No Meat or Beans
Oils[4]	(HEI-9)	10	≥12 grams per 1,000 kcal	No Oil
Saturated Fat	(HEI-10)	10	≤7% of energy[5]	≥15% of energy
Sodium	(HEI-11)	10	≤0.7 gram per 1,000 kcal[5]	≥2.0 grams per 1,000 kcal
Calories from Solid Fats, Alcoholic beverages, and Added Sugars (SoFAAS)	(HEI-12)	20	≤20% of energy	≥50% of energy

[1]Intakes between the minimum and maximum levels are scored proportionately, except for Saturated Fat and Sodium (see note 5).
[2]Legumes counted as vegetables only after Meat and Beans standard is met.
[3]Includes all milk products, such as fluid milk, yogurt, and cheese, and soy beverages.
[4]Includes nonhydrogenated vegetable oils and oils in fish, nuts, and seeds.
[5]Saturated Fat and Sodium get a score of 8 for the intake levels that reflect the 2005 Dietary Guidelines, <10% of calories from saturated fat and 1.1 grams of sodium/1,000 kcal, respectively.

http://www.cnpp.usda.gov/Publications/HEI/healthyeatingindex2005factsheet.pdf

Playing Data Detective

So, we've collected all of these surveys, now what?

Well, how about Exploratory Data Analysis where we visually inspect the data and ask questions about it to help form hypotheses.

What are some easy questions to ask?

- *What is the average … (total fruit consumption)?*
- *What is the average … (milk consumption)?*

1 – J. W. Tukey. Exploratory Data Analysis

NHANES

Utilizing Survey Cycle 2-3

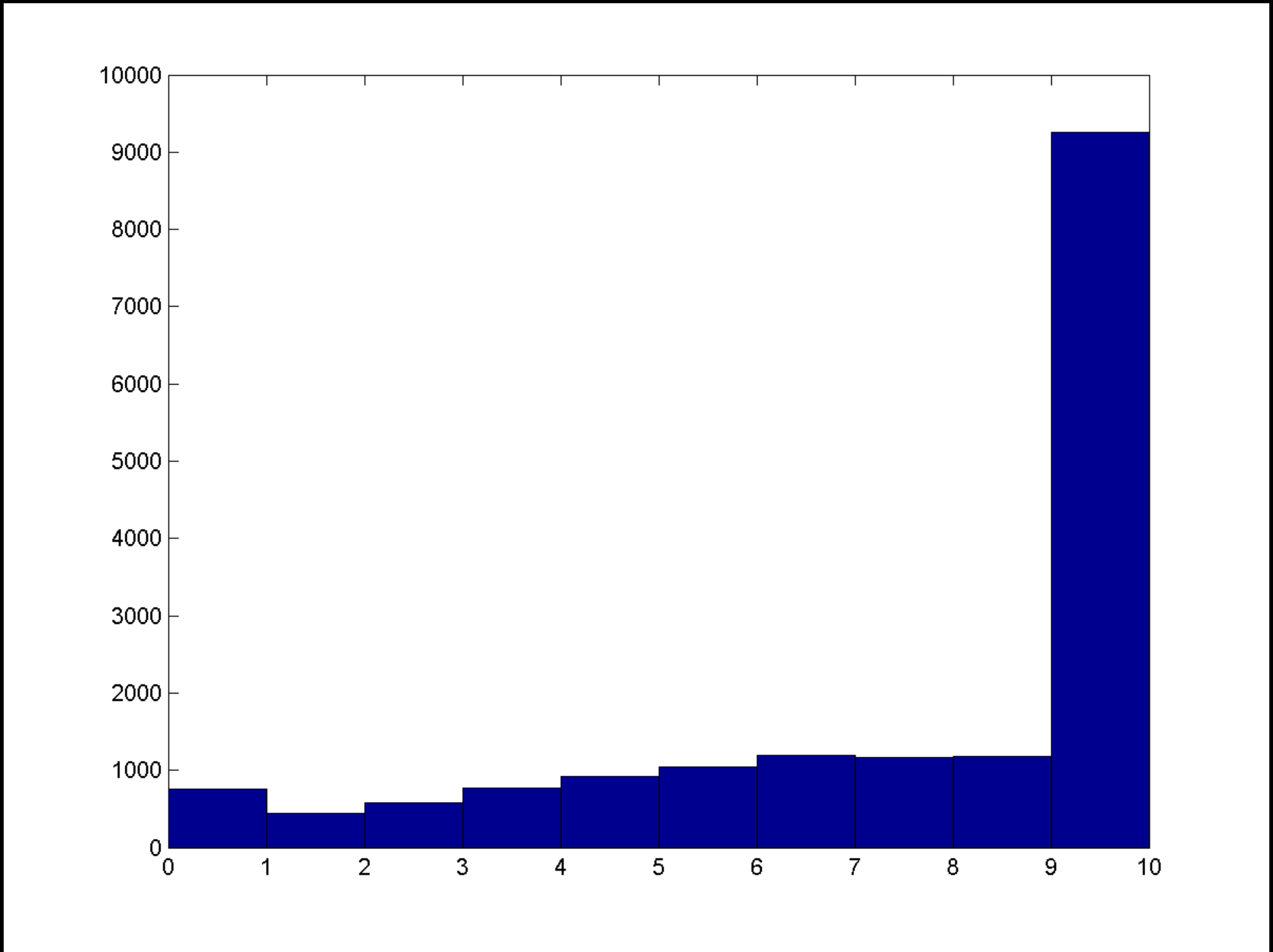
- ***Cycle 2: 2001-2002***
- ***Cycle 3: 2003-2004***
- ***Sample size of ~15000***

Other variables:

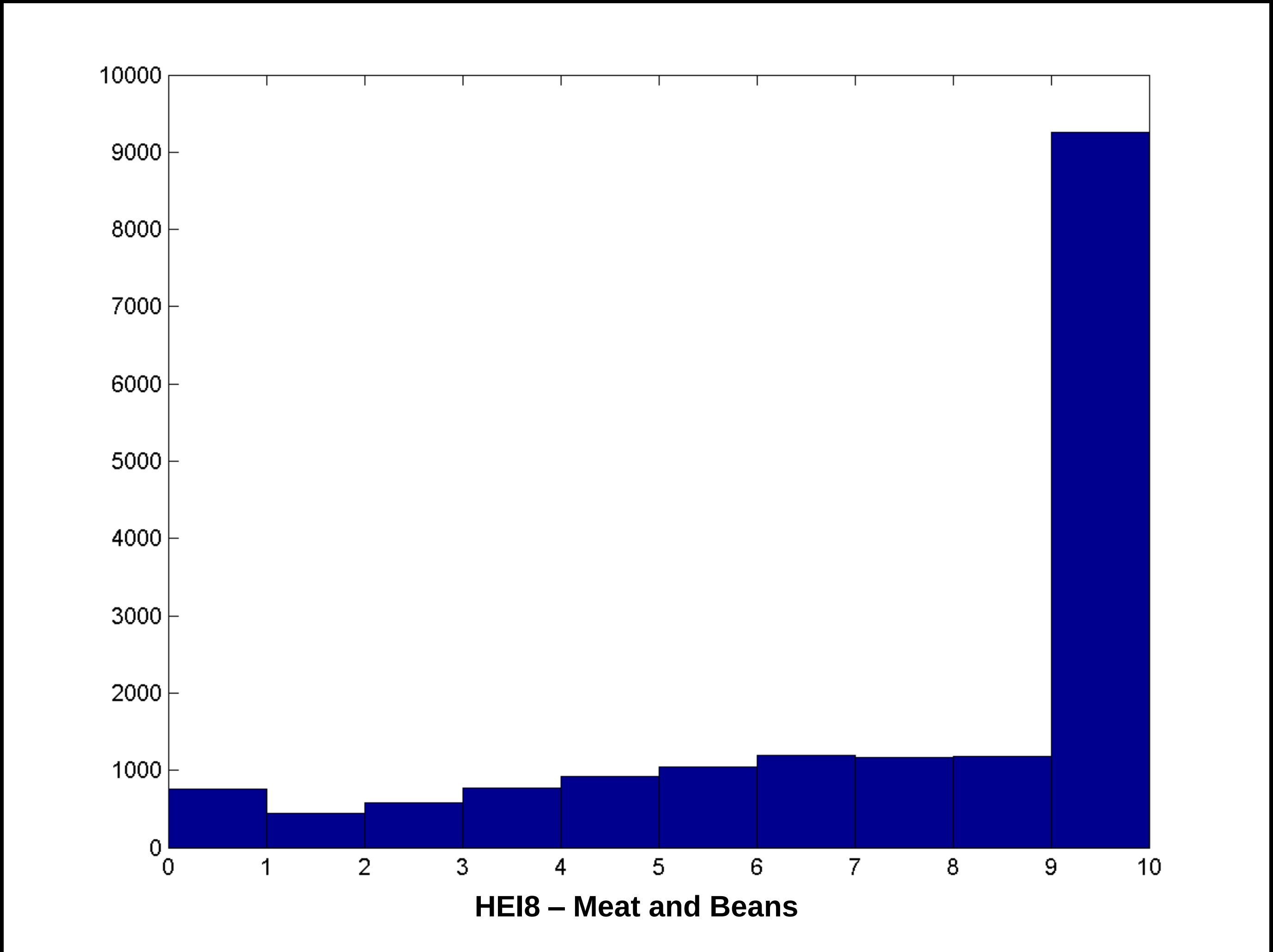
•Gender
•Age
•Race
•Total energy for the day (kcal)
•Total saturated fatty acid (gm)
•BMI

So with more data, maybe we can think of more questions!

What is the average … ?



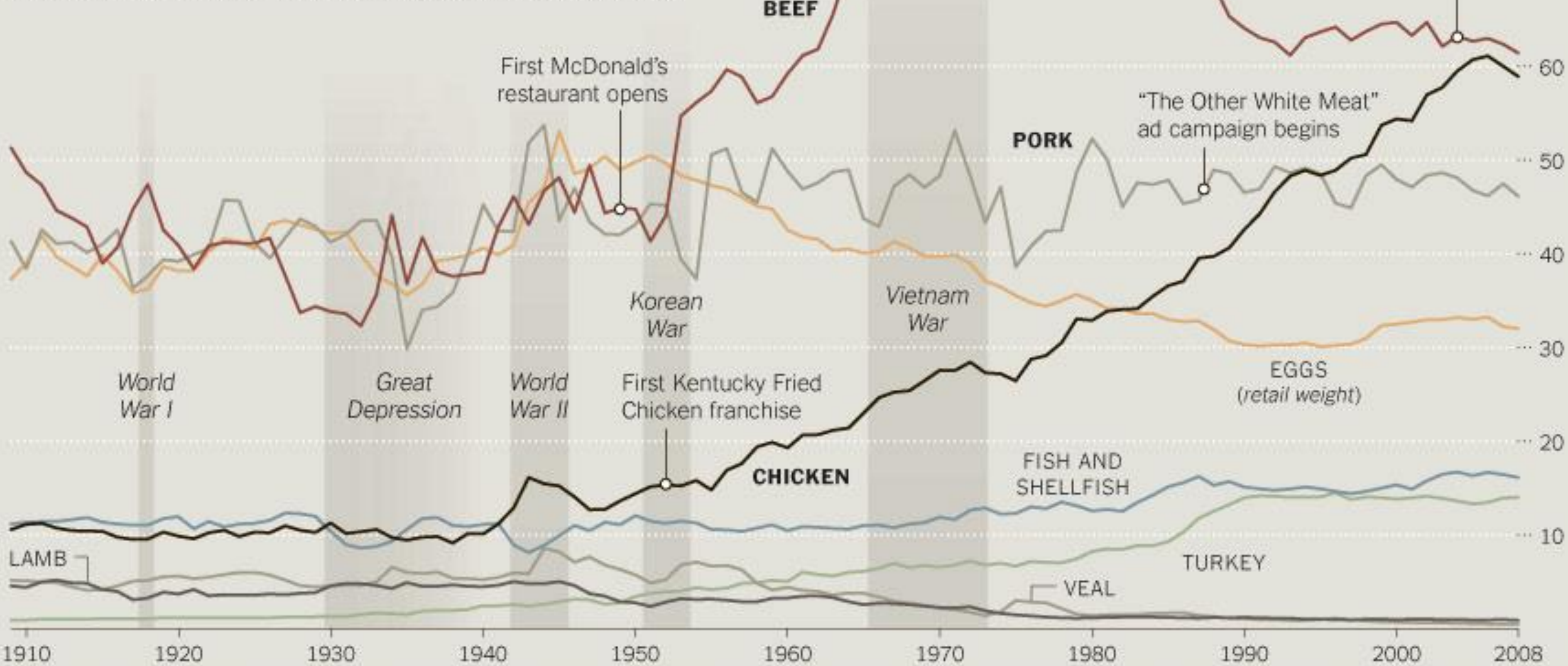
What is the average … ? Meat Consumption



A Century of Meat

American consumption of chicken and beef rose substantially after World War II, aided by the development of intensive farming methods, the proliferation of fast-food restaurant chains and supermarkets and the adoption of reliable home refrigeration.

Beef consumption peaked in 1976 but then declined, in part because of the publication of new dietary guidelines and studies that associated saturated fats and cholesterol with heart disease.



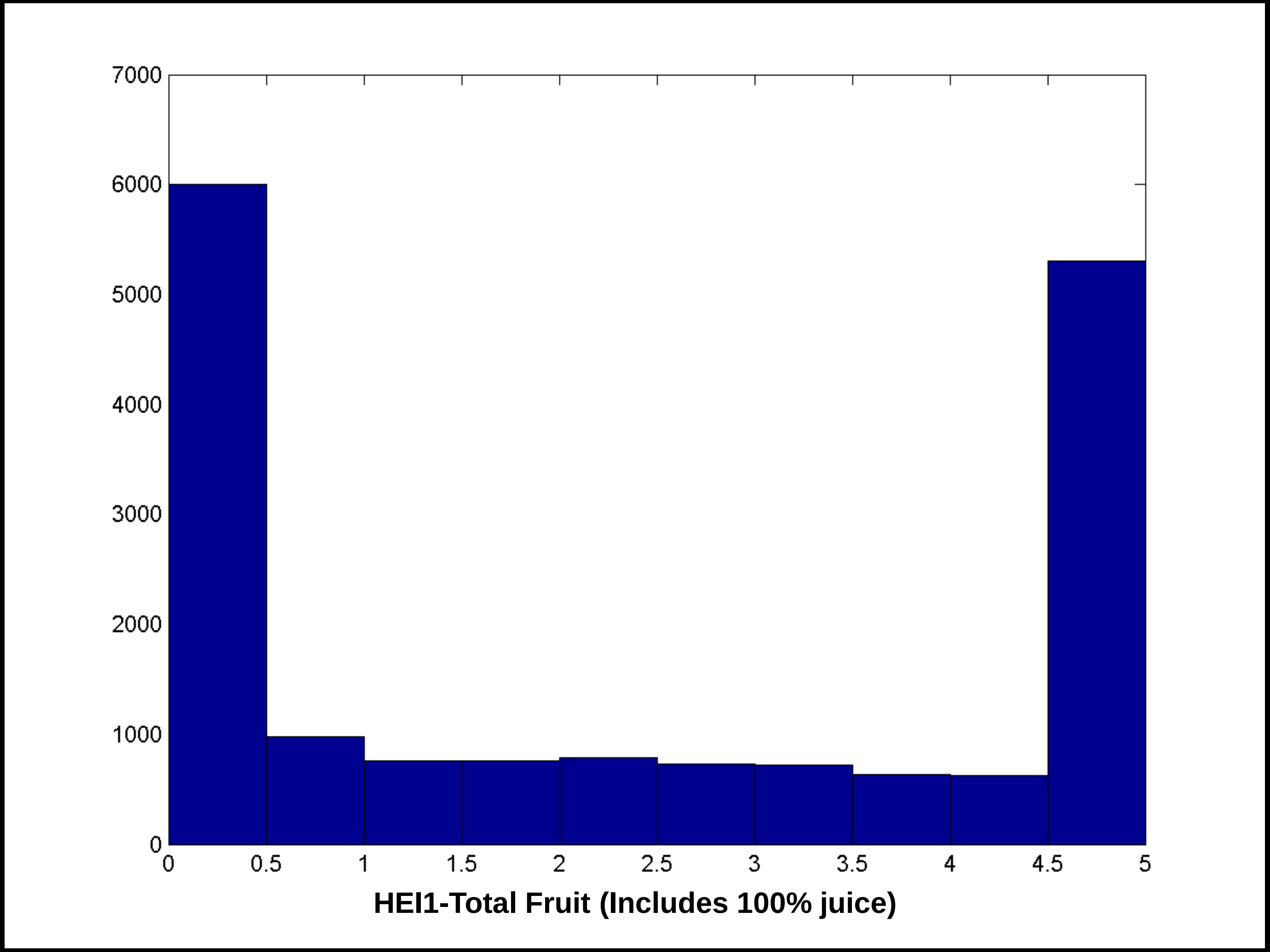
*Note: per capita availability of boneless meat is a proxy for human consumption, and is lower than retail weight or carcass weight. Bones, offal and game are excluded.

Sources: U.S. Department of Agriculture (data); news and company reports; "Putting Meat on the American Table," by Roger Horowitz

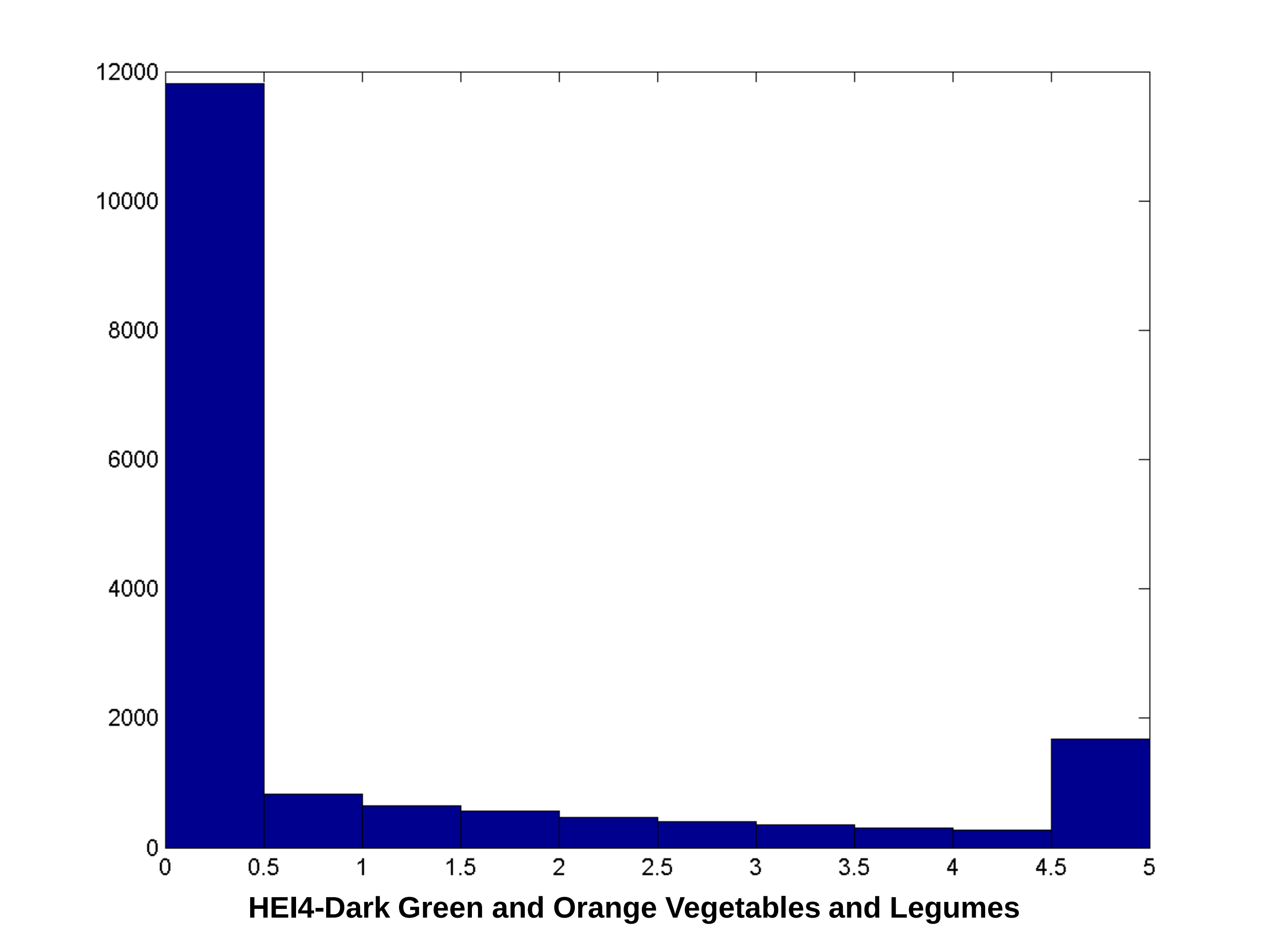
JONATHAN CORUM/THE NEW YORK TIMES

http://www.nytimes.com/imagepages/2011/03/15/science/15food_graphic.html

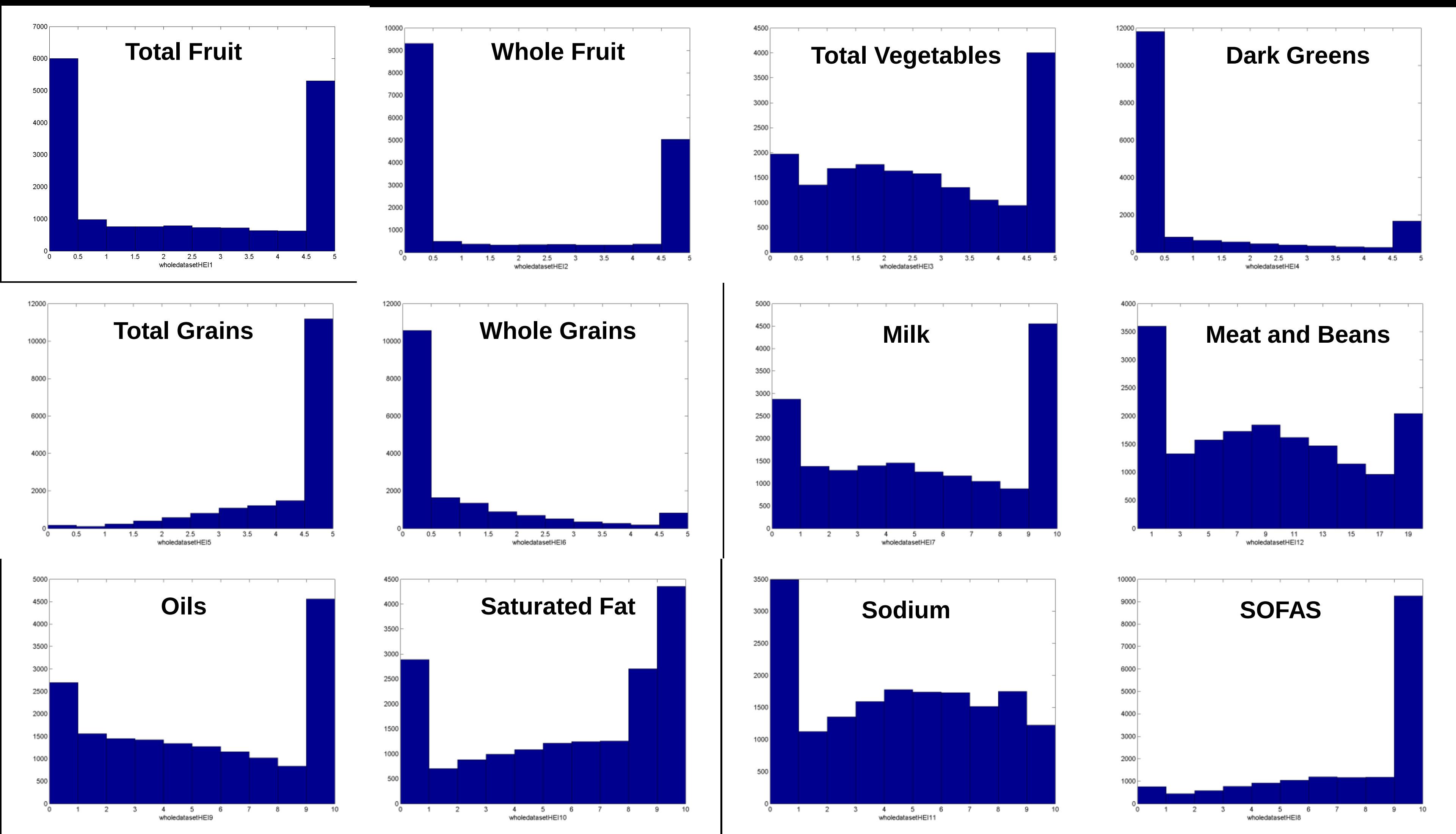
What is the average … ? Total Fruit Consumption



What is the average … ? Dark Green and Orange Vegetable Consumption



How Do These Compare?



So, What Can I say About the Population as a Whole?

- *We consume lots of SOFAs*
- *We are not good at eating our vegetables*
- *We eat lots of grains …*
- *But not lots of whole grains*

- *But what does this say about someone like you?*

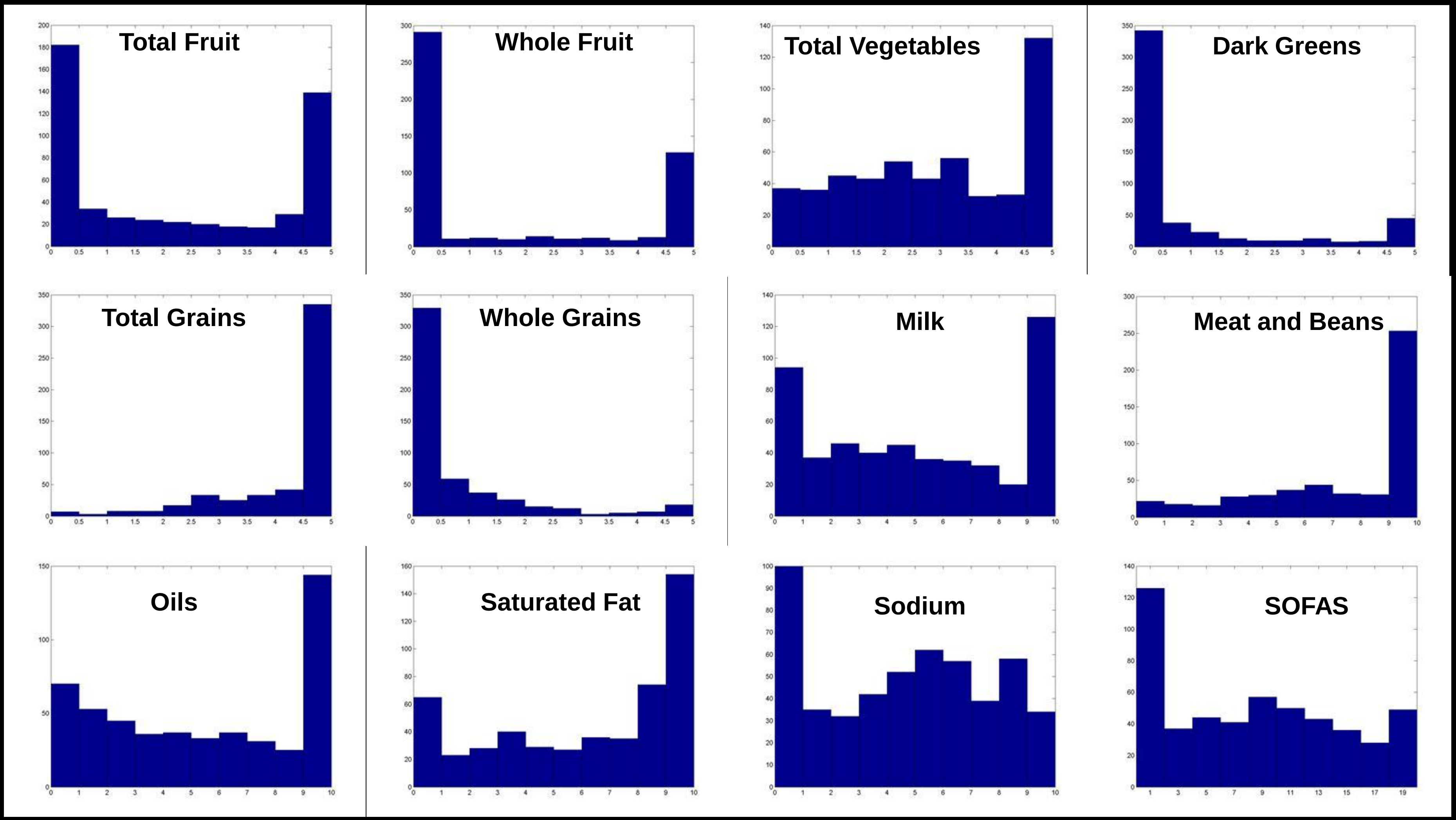
What if I further segment the data? Filter by Age, Gender, Etc.?

Let's look at Females, Age 19-30

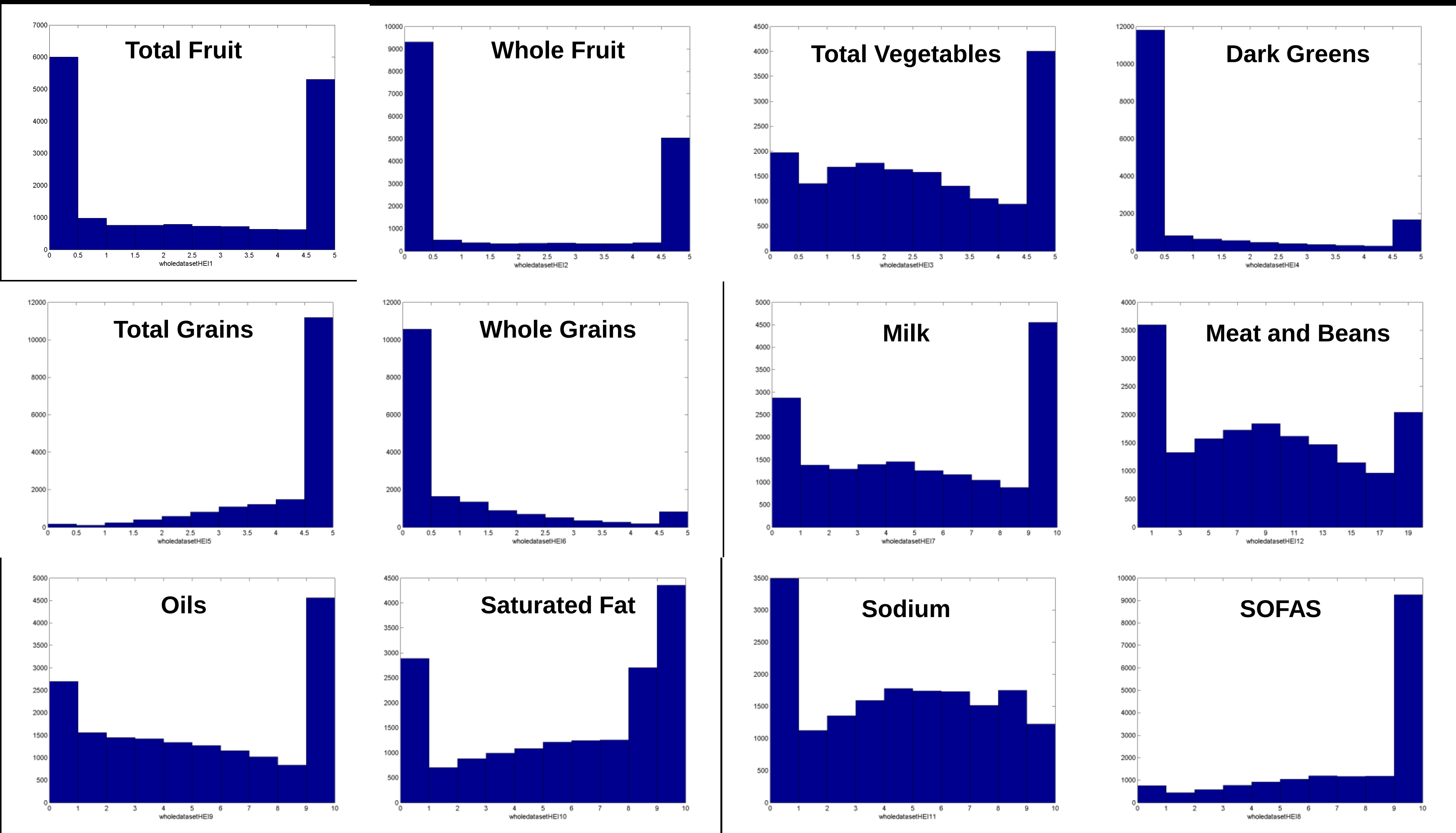
18.5 < BMI < 25.0

Total population in our sample = 511

NHANES in Females Age 19-30, Average BMI



Back to the Entire Population!



Forming Hypotheses

What do we think we see when comparing this group of females to the entire population?

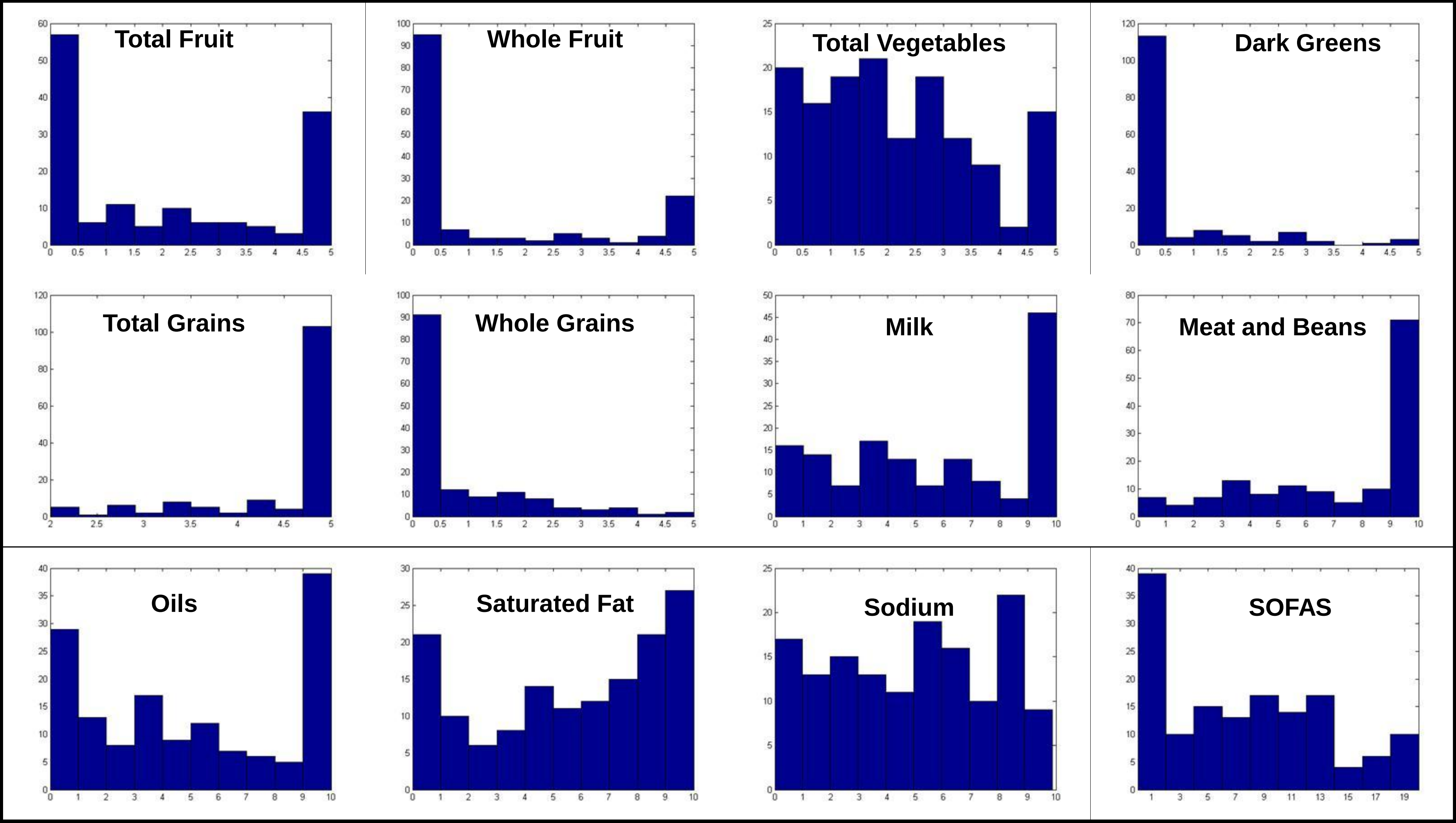
- ***Their SOFA consumption is more evenly distributed***
- ***Their meat and bean consumption is more skewed***

What if I further segment the data? Filter by Age, Gender, Etc.?

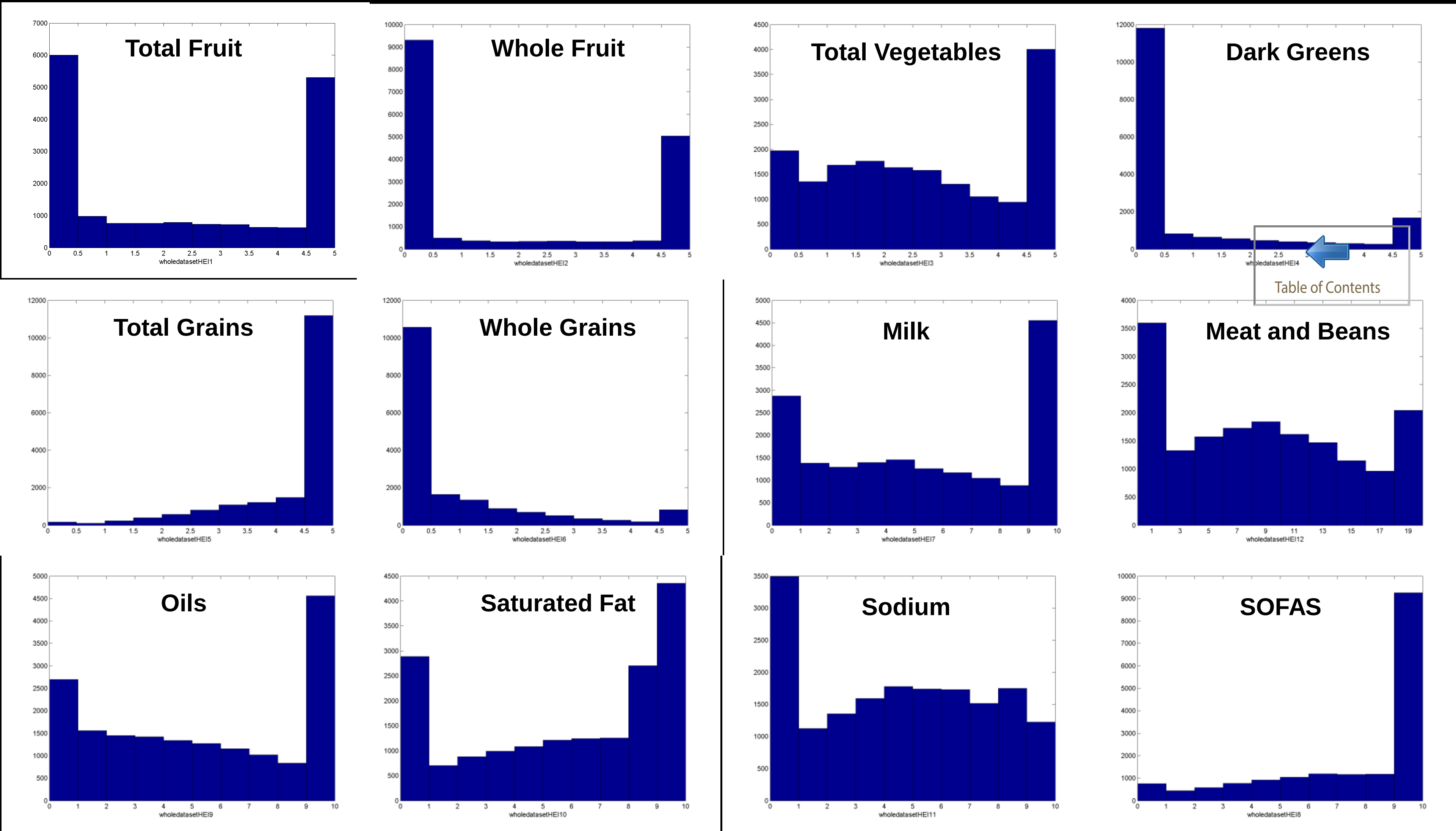
Let’s look at Males, Age 14-18, BMI < 18.5

Total population in our sample = 145

NHANES in Adolescent Underweight Males



Back to the Entire Population!



Forming Hypotheses

What do we think we see when comparing this group of females to the entire population?

- ***Their vegetable consumption is more evenly distributed***
- ***Their meat and bean consumption is more skewed***
- ***Their SOFAs are more evenly distributed***

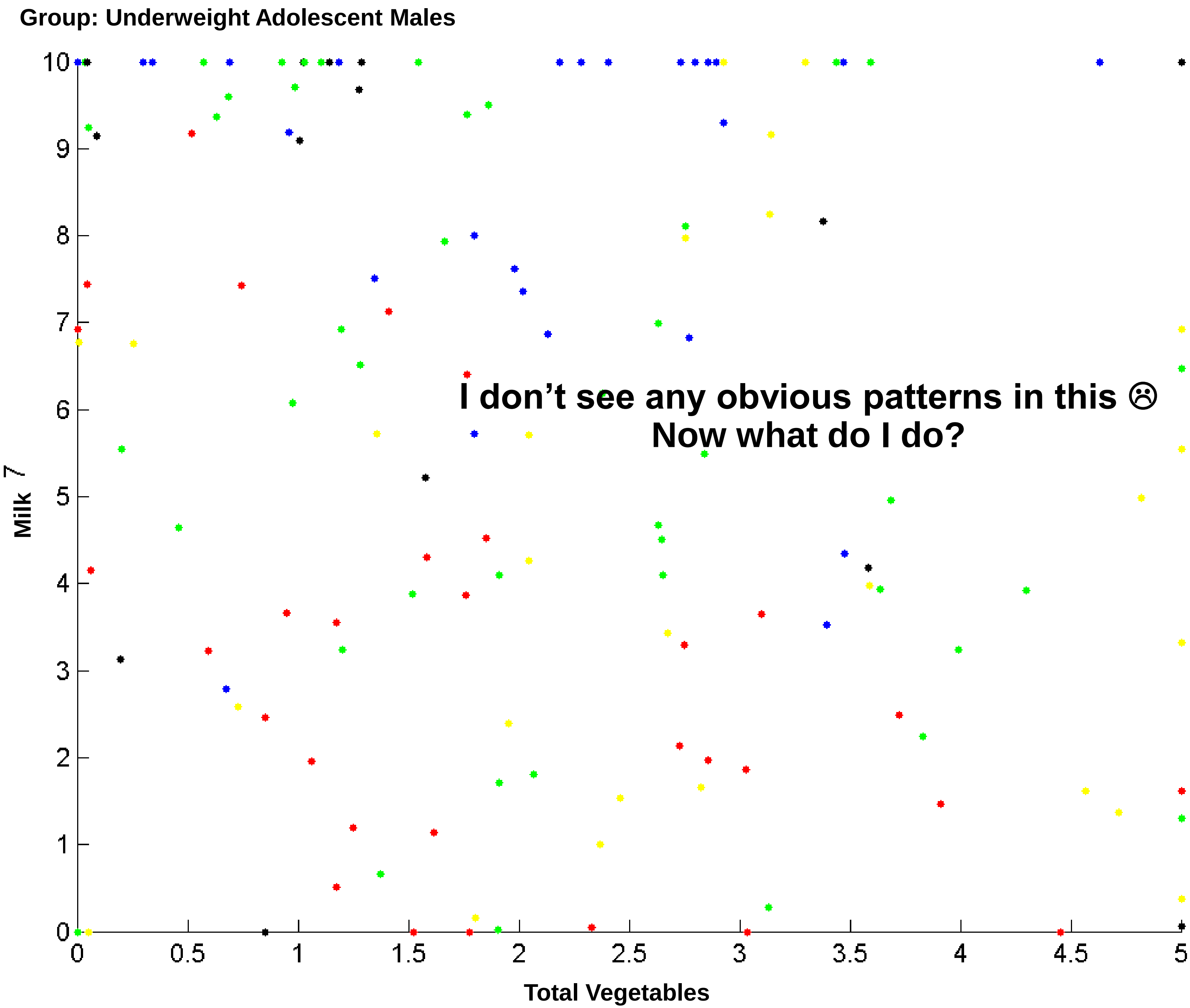
What if I want to ask:
How is … related to … ?

From our histograms, we came up with interesting insights into our data

But what if I want to know about relationships between the data?

More like:
Do people that consume more milk consume more vegetables?

We can try a scatterplot!


Group: Underweight Adolescent Males
I don’t see any obvious patterns in this ☹
Now what do I do?
Milk
Total Vegetables
0
1
2
3
4
5
6
7
8
9
10
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

What if I further question the data? All relationships by Age, Gender?

How many of these did you really expect me to do?

These histograms were made in Excel

One can also use SAS, JMP, MatLab

In all of those programs we have to write scripts, modify variables, etc.

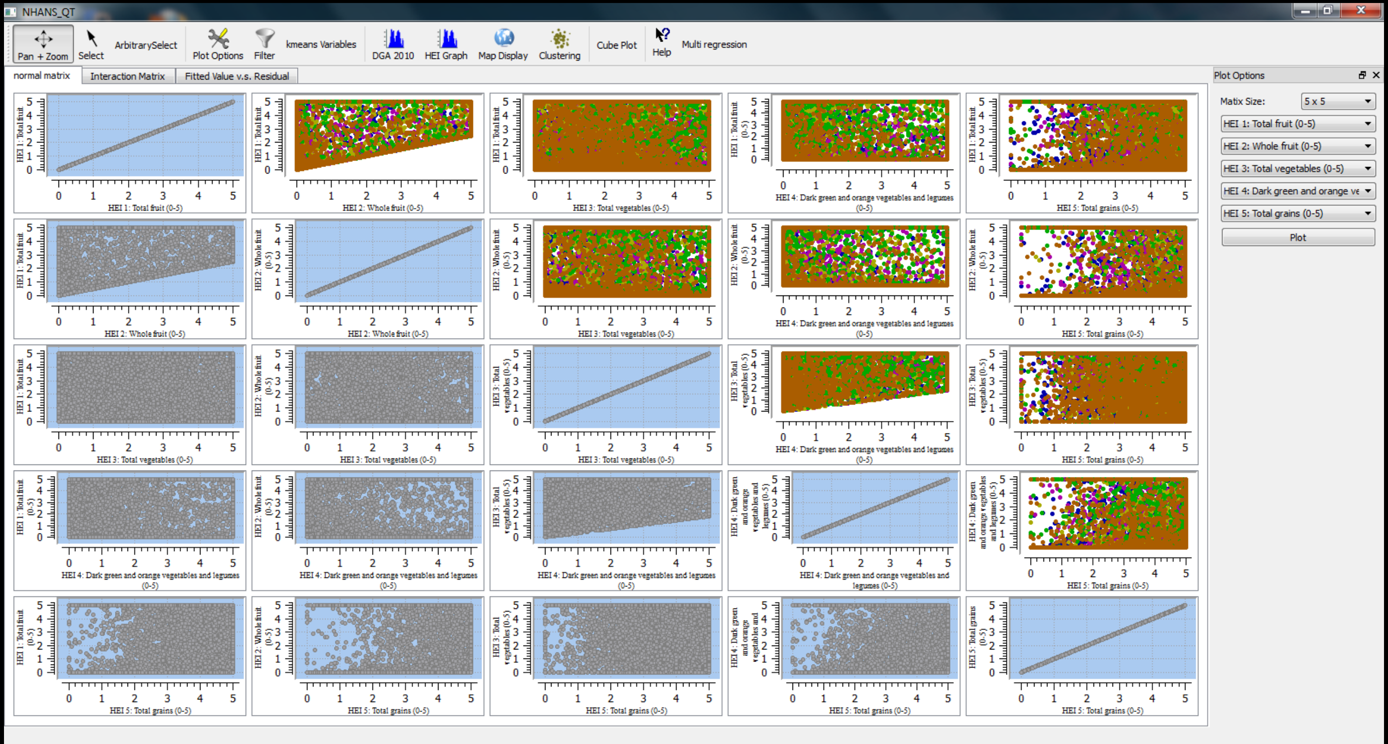
We can’t directly manipulate our variables

So, being a data detective can be very tedious

And these were just simple comparisons!

So what do we do now?

We can utilize interactive graphics combined with underlying analytical processes to help us be better detectives!



The Curse of Dimensionality

- *A term coined by Bellman in 1961*
- *Refers to the problems associated with multivariate data analysis as the dimensionality increases the available data becomes sparse*
- *Sparsity is a problem for any method that requires statistical significance*
- *Sometimes data dimensions are redundant and can be reduced with minimal information loss*
- *In visualization we are also limited with screen space and the number of available visual variables, so choosing the most appropriate dimensions is key*

Richard Ernest Bellman; Rand Corporation (1957). *Dynamic programming*. Princeton University Press.

The Curse of Dimensionality

So, what can we do?

- We can incorporate prior knowledge of the data
- We can smooth the target function
- We can reduce the dimensionality

The Practicality of the Curse of Dimensionality

For a given sample size, there is a maximum number of features above which the performance of classifying samples will degrade rather than improve

In most cases, the additional information that is lost when discarding some features is compensated by a more accurate mapping in the lower-dimensional space

So, how do we know what features we can throw away?

For visualization this implies that there are some features of a dataset that will be better to visualize (contain more information) than others!

Curses! I’m Still Overwhelmed!

Probably so, we are creating a lot of data in the world

Overall, NHANES is actually a relatively modest set of data

As our datasets get more variables we need ways to question the data

What I really want to know is ***are there relationships/groups/clusters in the data****, how can I see these when we have so many dimensions?*

One way is to just cluster the data and draw these clusters in an arbitrary space (we saw this in the demo)

Another is to try dimensional reduction!

Dimensional Reduction

Two approaches are available to reduce dimensionality

- **Feature extraction:** creating a subset of new features by combinations of the existing features
- **Feature Selection:** choosing a subset of all the features
- Given a feature space $x_i \in R^N$ find a mapping $y = f(x): R^N \to R^M$ with M<N such that the transformed feature vector $y \in R^M$ preserves (most of) the information or structure in R^N
- An optimal mapping is one that does not increase error

What Does Dimensional Reduction Mean For You?

Feature Selection: choosing a subset of all the features

- Well, that's what we were already doing
- We use our own expert knowledge to select features of interest
- Note we could also use machine learning and statistics here too if we were so inclined

What Does Dimensional Reduction Mean For You?

Feature extraction: creating a subset of new features by combinations of the existing features

- You can think of this in terms of the NHANES data
- Let's say I do an analysis and I find a correlation between SOFA consumption and sodium consumption (the more SOFAs, the more sodium)
- I could combine SOFAs and sodium into a new variable – SSOFAS
- I make up some formula to do this:
  - $SSOFAS = \alpha * SOFAs + \beta * sodium$
- So basically we just combine variables that are correlated

Principle Components Analysis

One of the most commonly applied dimension reduction techniques

PCA is a deterministic analytical procedure that utilizes an orthogonal transformation to reduce a set of sample observations with potentially correlated variables into a set of uncorrelated variables called principal components

The number of principal components will always be less than or equal to the original number of variables in the sample set

. T. Jolliffe, “Principal Component Analysis," *Springer Series in Statistics*, Springer, NY, 2002.

Principle Component Analysis

The main limitation of PCA is that it does not consider class separability since it does not take into account the class label of the feature vector

- PCA simple performs a coordinate rotation that aligns the transformed axes with the directions of maximum variance
- There is no guarantee that the directions of maximum variance will contain the most interesting/important features

That sounded really hard!

- *Yes, but we have software to help us!*
- *In JMP Principal components can be accessed through the Principal Components command on the Analyze > Multivariate Methods menu.*

- *But, how does principal component analysis help me?*
- *What can principal component analysis tell me?*

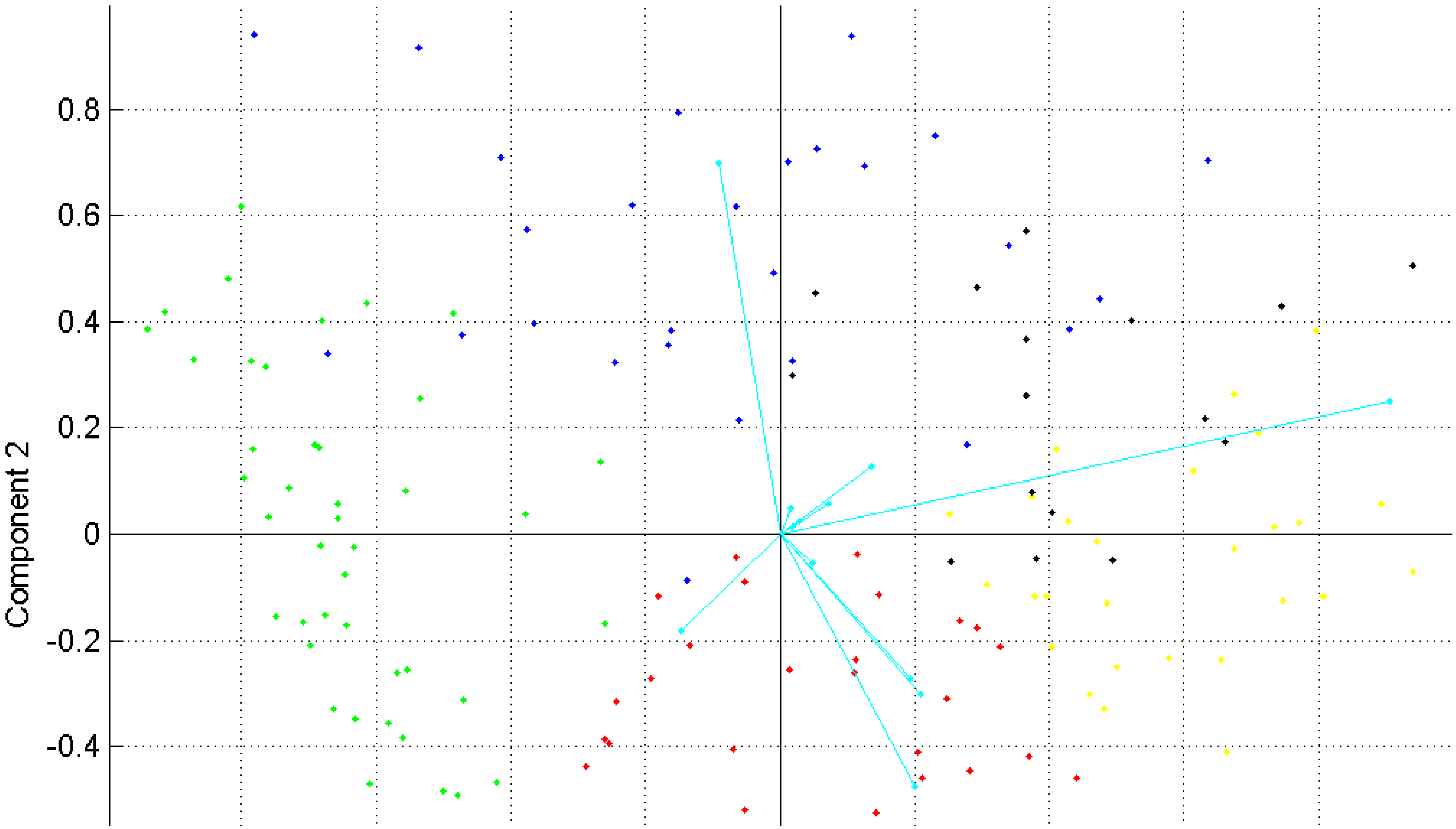
Principal Component Analysis

- *Let's go back to NHANES for an example and focus on the HEI values.*
- *Each person in my data set has 12 HEI values, so you can think of each person as a point in 12-dimensional space*
- *It is hard to understand all the relationships in this space, PCA tries to figure out what dimensions to combine for you automatically*
- *After you run PCA you will get a set of what you can think of as new variables*
- *Since there were 12 variables in our NHANES data, we will get at most 12 principal components (but hopefully less!)*

Principal Component Analysis

- *Ok, you've convinced me to run PCA on NHANES and I get 8 components, but now I'm confused!*
- *The output is such that your first component accounts for as much of the variability in the data set as possible*
- *The second component accounts for the second most variability, etc.*
- *Typically we will plot the data using the first two principal components as axes*
- *We can also write out each component to analyze it*

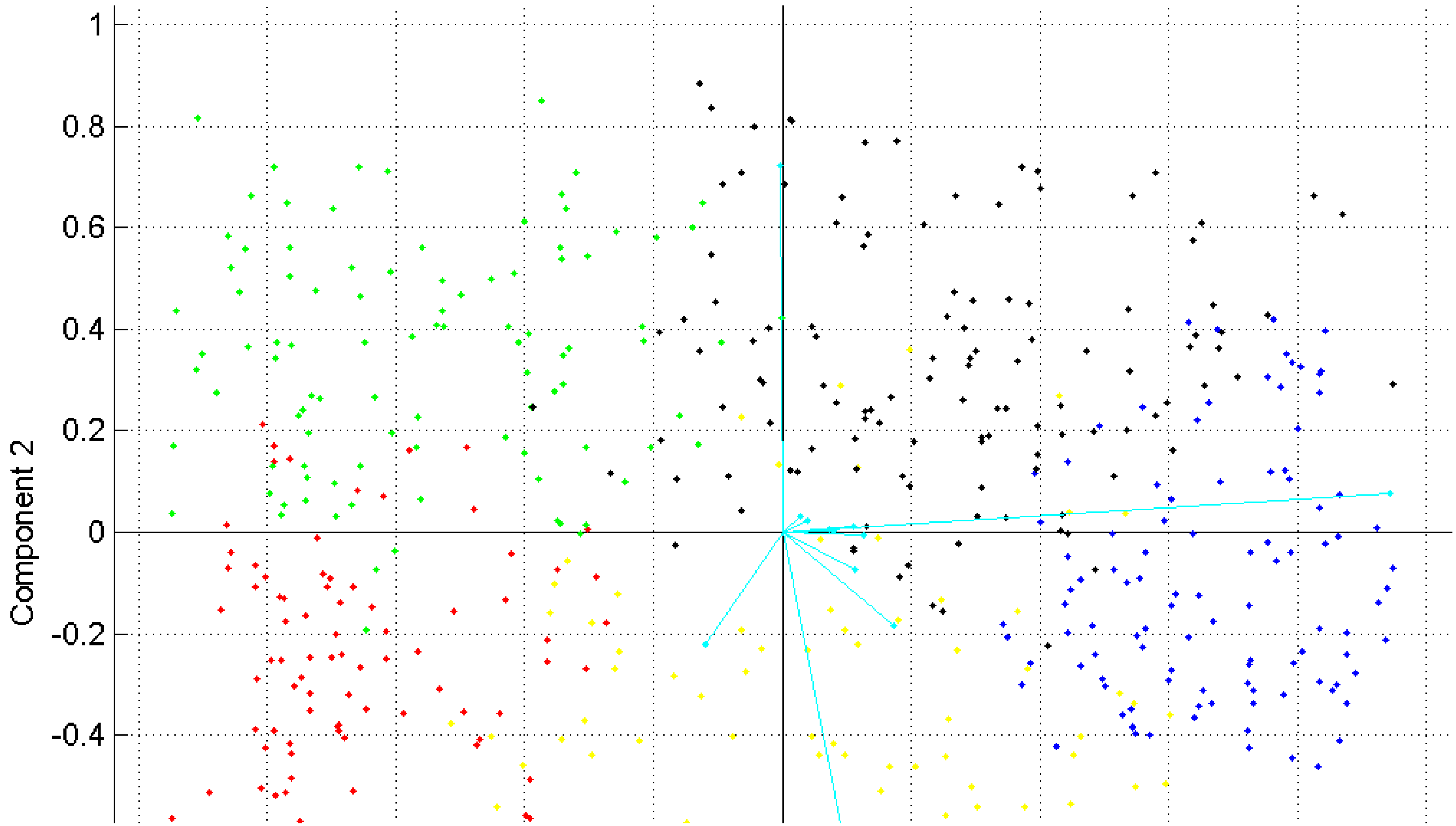
Underweight Adolescent Males



FV1 ≈ 0.14 HEI1 + 0.07 HEI2 +0.02 HEI5 + 0.02 HEI6 + -0.09 HEI7 + 0.19 HEI8 +0.05 HEI3 + 0.03 HEI4 + 0.20 HEI9 + 0.20 HEI10 + -0.15 HEI11 + 0.91 HEI12

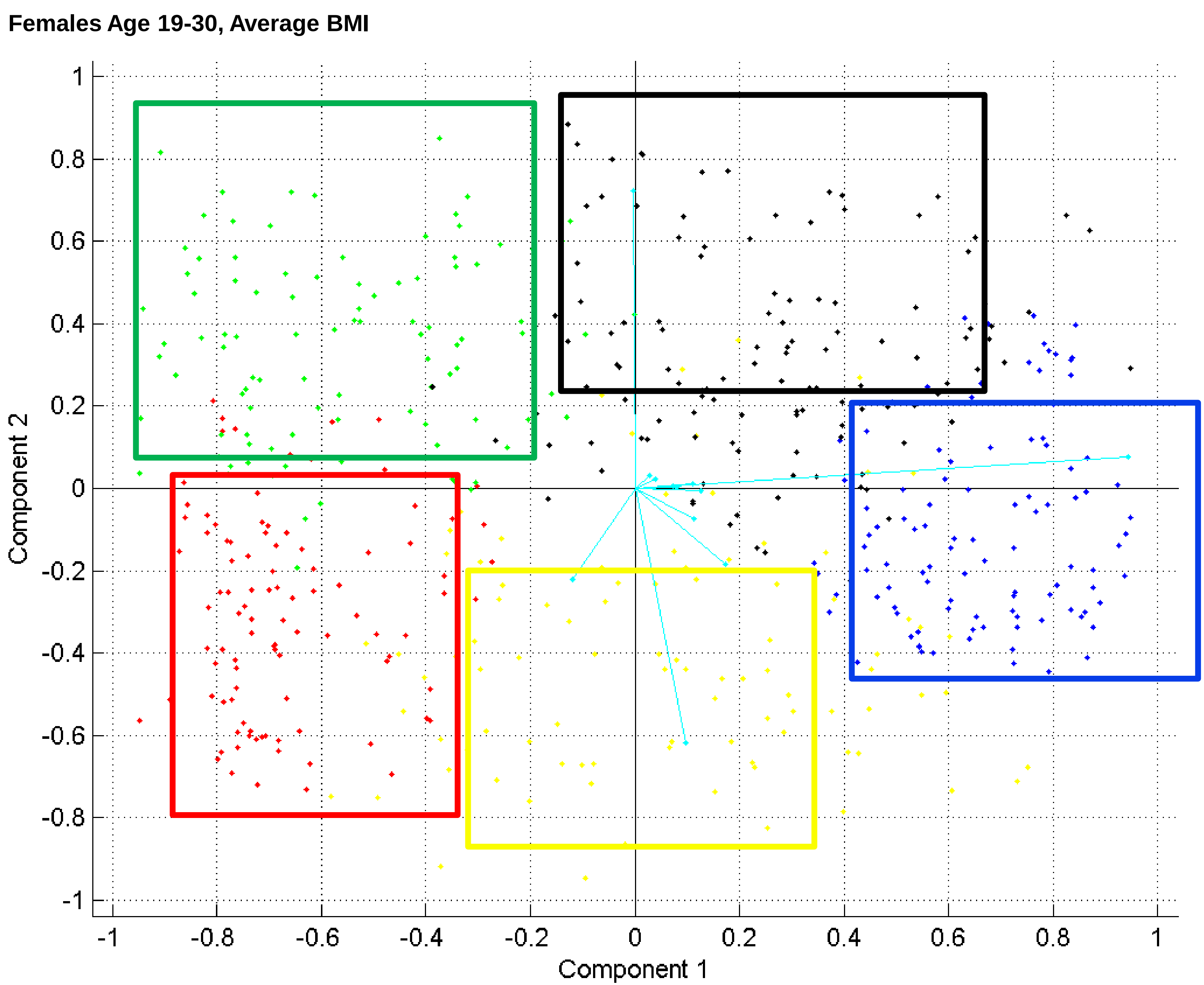
FV2 ≈ 0.13 HEI1 + 0.06 HEI2 + 0.05 HEI5 + 0.01 HEI6 + 0.70 HEI7 + -0.27 HEI8 + -0.05 HEI3 + 0.02 HEI4 + -0.48 HEI9 + -0.30 HEI10 + -0.18 HEI11 + 0.25 HEI12

Females Age 19-30, Average BMI



FV1 ≈ 0.13 HEI1 +0.11 HEI2 +0.04 HEI5 + 0.03 HEI6 + -0.00 HEI7 + 0.11 HEI8 +0.08 HEI3 + 0.07 HEI4 + 0.17 HEI9 + 0.10 HEI10 +-0.12 HEI11 + 0.94 HEI12

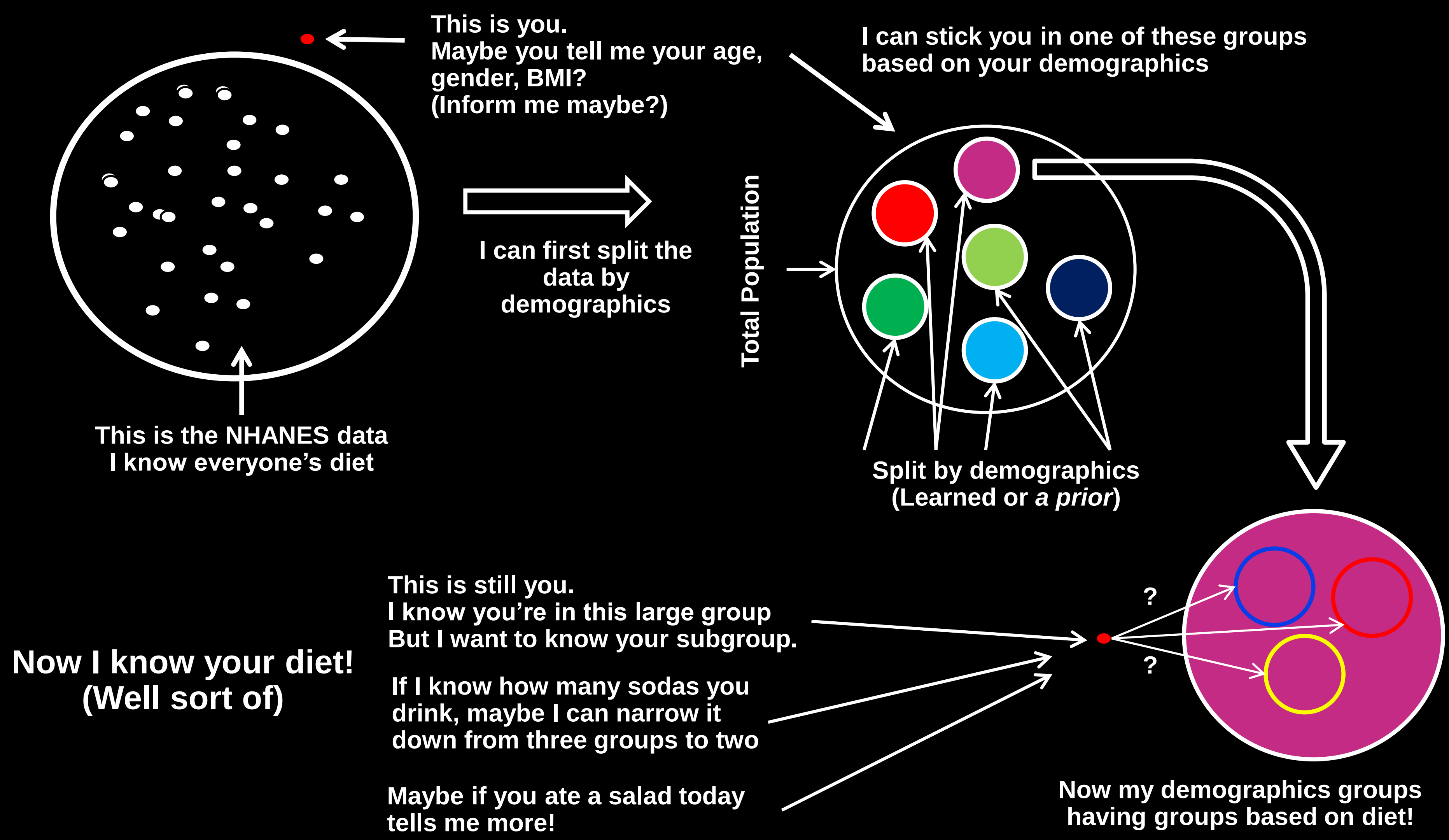
FV2 ≈-0.01 HEI1 + 0.01 HEI2 +0.02 HEI5 + 0.03 HEI6 + 0.72 HEI7 + -0.07 HEI8 + 0.00 HEI3 + 0.00 HEI4 + -0.18 HEI9 +-0.62 HEI10 +-0.22 HEI11 +0.08 HEI12


Females Age 19-30, Average BMI
Component 2
Component 1
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Putting it all together

- ***If I know things about you …***
  - Age, gender, BMI
- ***Can I guess your typical daily diet?***
  - Probably not very accurately
- ***But, if I ask you a few questions, maybe I can get closer!***

Putting it all together

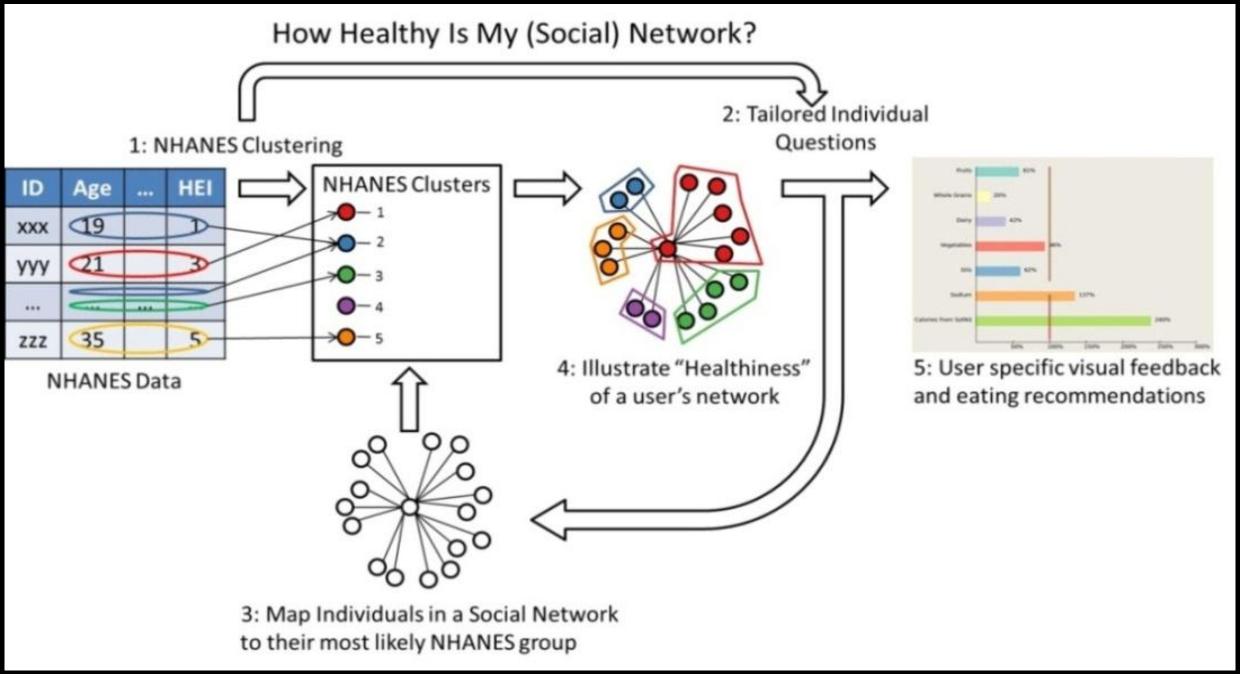


Summary

- *Data analysis is hard!*
- *We need better tools to be data detectives*
- *By being data detectives, we can ask questions and maybe create predictions about unknown data based on available data (we make a model of the data)*
- *Visual analytics helps us in the data detective work!*
- *But we still need to understand what the analyses do for us so that we can put them into practice.*

The Future!

- *Right now, I have a Facebook app that can guess your income based on your location and your friends (still beta testing so it is not publically available ☹)*
- *Soon, we'll have a Facebook app that will start narrowing down your diet*



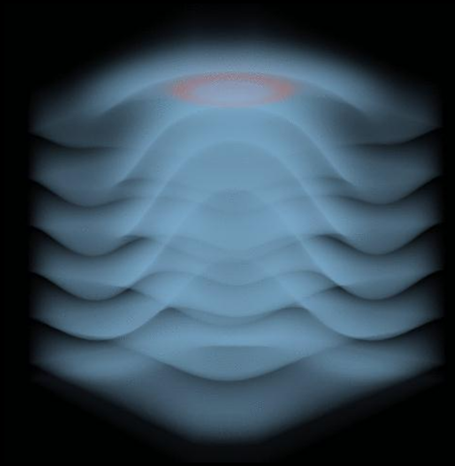
Dr. Ross's Shameless Plug

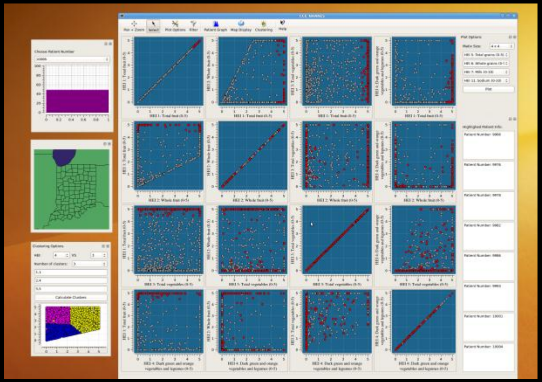
- ***You can play with my other app though***
- Search for hedcutr on Facebook and you can make this:

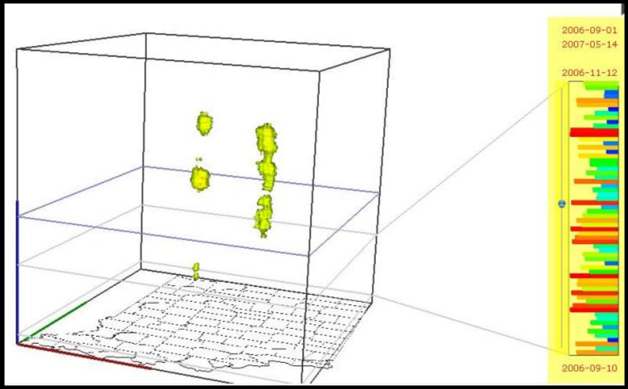
- I'm collecting data on the rendering parameters to see what people like!

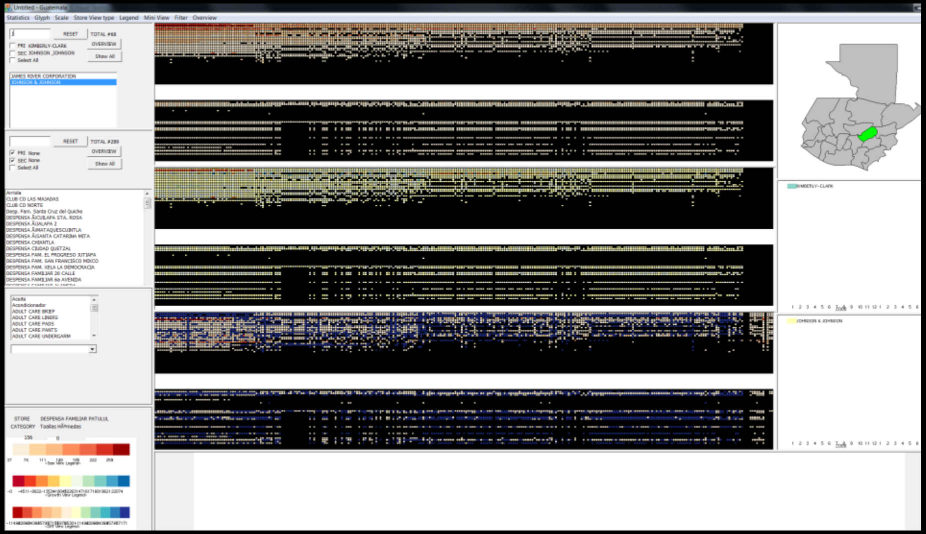
Table of Contents

Questions? (http://rmaciejewski.faculty.asu.edu)









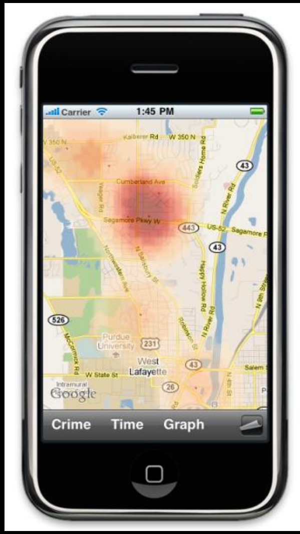


Table of Contents

VisWeek 2012
SCIVIS • INFOVIS • VAST
BIOVIS • LDAV

Examining the Use of a Visual Analytics System for Sensemaking Tasks: Case Studies with Domain Experts

Youn-ah Kang and John Stasko
Georgia Institute of Technology

Evaluation in Visual Analytics

We evaluate systems because..

- We hope our technologies are making an impact
- We hope they are helping people gain value from their information

Evaluation is challenging and not very common

- Usability testing and controlled experiments remain crucial
- Actual case studies of prolonged system use by analysts working with their own data are still rare

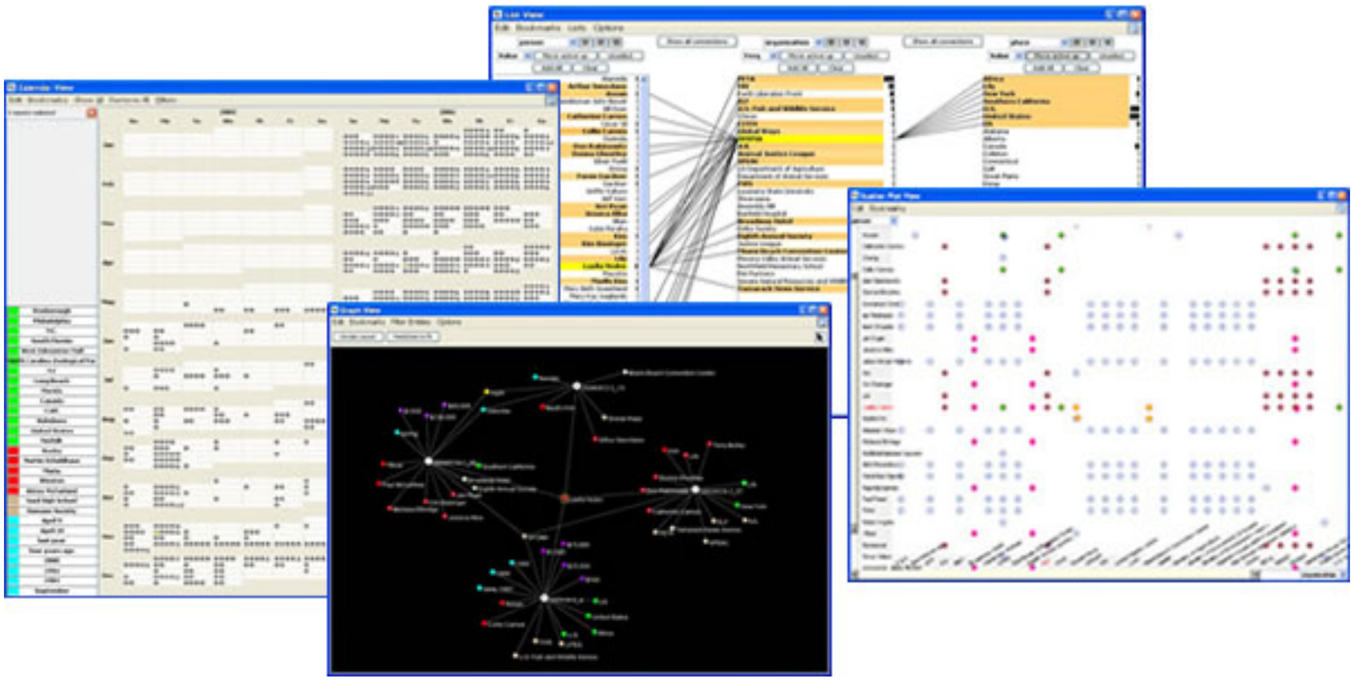
Case Studies

- Case studies can provide valuable findings and insights
  - Yield a description of how a tool was used and where users had problems
  - Difficult to achieve through controlled lab studies

- Multi-dimensional in-depth long-term case studies (MILC)
  - B. Shneiderman and C. Plaisant (2006). Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies.

System of Study

- Jigsaw (http://www.cc.gatech.edu/gvu/ii/jigsaw/)



Motivation and Goals

- Is Jigsaw helping analysts with their tasks and problems?
- For what types of documents and analyses does Jigsaw help?
- What are particularly useful features/capabilities as well as missing or problematic ones?

Recruitment and Study Protocol

- 6 working/practicing investigators who were using the system
  - 3 intelligence analysts, 2 academic researchers, and 1 business analyst
- Used Jigsaw for a range of 2-14 months
- Semi-structured interviews
- Follow-up email conversations

Case Studies

P1: Aerospace Engineering Researcher

- Task: Compare two major air traffic control programs and examine their compatibility

OI-0320 Initial Surface Traffic Management

Description: Departures are sequenced and staged to maintain throughput. Air Navigation Service Provider (ANSP) automation uses departure-scheduling tools to flow surface traffic at high-density airports. Automation provides surface sequencing and staging lists for departures and average departure delay (current and predicted). ANSP automated decision support tools integrate surveillance data. This includes weather data, departure queues, aircraft flight plan information, runway configuration, expected departure times, and gate assignments. Automation provides surface sequencing and staging lists for departures and average departure delay (current and predicted). Local collaboration between ANSP and airport stakeholders improves information flow to decision support as well as the ability for aircraft operators to meet their operational and business objectives.

Functional Drivers: The use of improved departure scheduling and surface management will reduce delays and environmental impacts resulting in more efficient operations.

SOPR: FAA ***SOPR Unique Reference:*** 104209

SOCR: ***Primary Supported OIs:*** OI-0322, OI-0321

OI Group: Trajectory Management - Surface Operational Improvments

	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
OI-0320: Initial Surface Traffic Management					O	2012												
EN-6020: Environmentally Improved Surface Operations - Level 1 - Initial			E	2010														
EN-0007: High-Density Arrival/Departure Detail Operational Concept				E	2011													
PI-0024: Secure Information Exchange	P	2008																

P1: Aerospace Engineering Researcher

- Goal: Identify similarities/differences and create a mapping between the two programs
  - Does a concept or capability suggested in one program also appear in the other program?

Program A	Program B
OI-0320: Initial Surface Traffic	**L07-02 TS-0201:** Basic Departure Management (DMAN)
	L07-02 TS-0202: Departure Management Synchronized with Pre-Departure Sequencing
	L07-02 TS-0203: Integration of Surface Management Constraint into Departure Management
	L07-02 TS-0306: Optimized Departure Management in the Queue Management Process
	L10-02 AO-0205: Automated Assistance to Controller for Surface Movement Planning and Routing
	L10-03 AO-0501: Improved Operations in Adverse Conditions through Airport Collaborative Decision Making

P1: Aerospace Engineering Researcher

- Originally done manually using MS word and search
  - Search for descriptions of program A -> identify keywords -> review descriptions of program B containing matching keywords one by one

- Jigsaw helped: Review and compare the huge document collection and complete the mapping between the two
  - Published the work at *Aviation Technology Integration and Operations ATIO Conference*

P2: Business Analyst at an Accounting Firm

- Task: Analyze unstructured data and identify any linkages between people/companies relevant to a financial fraud
- Goal: Find evidence for a financial fraud
- Before Jigsaw: Put all documents into an Excel spreadsheet, search for keywords, and read all returned documents
- Jigsaw helped: Reveal connections between people & companies that were not easily identifiable
  - Found evidence of a financial fraud after analyzing 100,000 emails

P3: Industrial & Systems Engineering Researcher

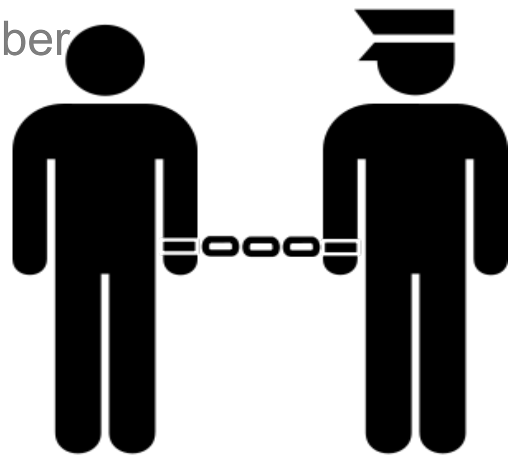
- Task: Validate her model about company transformation by combining historical company data
  (5,000+ announcements and news articles of 9 IT firms for 10 years)

- Goal: Make sense of the documents and extract keywords for the next step – data mining

- Jigsaw helped: Attain a clear understanding of the documents in a short amount of time

P4: Intelligence Analyst at a Police Department

- Task: Make sense of daily incident reports and identify patterns, trends, and any top issues in the city
- Goal: Find connections between individuals, places, and other incidents within accumulated crime reports
- Originally read all the reports individually and tried to remember different connections using printed copies of the documents
- Jigsaw helped: Develop a repository of important connections
  - Helped the police arrest a criminal by identifying where he might be



P5: Intelligence Analyst at a National Lab

- Task: Review resumes and find a good candidate with a certain specialty

- Goal: Examine connections in candidate info and find an expert in a specialized area
  - Skills, publications, co-authors, education, employment history.

- Performed using Analyst's Notebook but felt limited

- Jigsaw helped: Identify possible connections between people and technology

P6: Intelligence Analyst with the US Air Force

- Task: Examine budget summaries of R&D programs in the Department of Defense and identify common themes (10,000+ documents from 20+ agencies such as Air Force, Navy, and DARPA)

- Goal: Highlight what programs/topics are similar, what makes them similar, and who are working on similar topics

- Jigsaw helped: Effectively search for similar tools and technologies that required further investigation

Types of Tasks

- Relationship / connection between entities
  - Targeted investigation rather than seeing the big picture
- Search / comparison
  - If the documents contain specific keywords
- Understanding
  - Getting an overview of documents
- As a communication aid / shared understanding of data
  - Persuasive power in communication

Learning the System

- Learning curve existed
  - "How to better analyze my data using this tool"

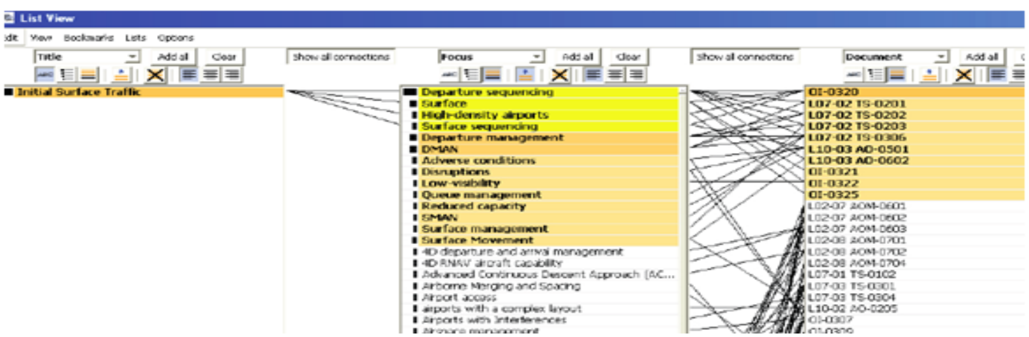
- Constructing a frame
  - Which views are most appropriate for my data and task?
  - What entity types do I want to put in this column?

➔ Finding the optimal approach
in their own way

Unexpected Use of the System

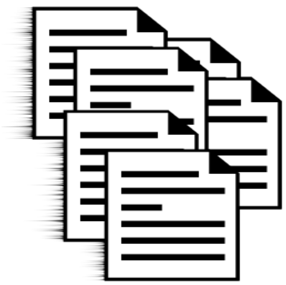
- Views for evidence/output generation

NextGen Operational Improvements	SESAR Operational Improvements and Steps
OI-0320: Initial Surface Traffic	**L07-02 TS-0201:** Basic Departure Management (DMAN)
	L07-02 TS-0202: Departure Management Synchronized with Pre-Departure Sequencing
	L07-02 TS-0203: Integration of Surface Management Constraint into Departure Management
	L07-02 TS-0306: Optimized Departure Management in the Queue Management Process
	L10-02 AO-0205: Automated Assistance to Controller for Surface Movement Planning and Routing
	L10-03 AO-0501: Improved Operations in Adverse Conditions through Airport Collaborative Decision Making
	L10-03 AO-0602: Collaborative Pre Departure Sequencing



A mapping created manually (top) and by Jigsaw (bottom)

Unexpected Use of the System



Information-dense documents

Separate docs into several projects

Merge new incoming documents with an existing Jigsaw project

Build a historical dataset

Issues and problems

- Technical issues in the preparation stage
  - Importing data into Jigsaw
  - Identifying entities
- Limited filtering options
  - Not being able to easily select a subset of data in the views

Design Implications

- Supplement automatic entity identification
- Allow flexible data (document) management
  - Provide an ability to easily select a subset of documents
- Empower with numbers
  - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
  - Limit user interaction vs. give more power
- Invest in tutorial
  - Break down into subtopics with use-cases and examples

Design Implications

- Supplement automatic entity identification
- Allow flexible data (document) management
  - Provide an ability to easily select a subset of documents
- Empower with numbers
  - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
  - Limit user interaction vs. give more power
- Invest in tutorial
  - Break down into subtopics with use-cases and examples

Design Implications

- Supplement automatic entity identification
- Allow flexible data (document) management
  - Provide an ability to easily select a subset of documents
- Empower with numbers
  - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
  - Limit user interaction vs. give more power
- Invest in tutorial
  - Break down into subtopics with use-cases and examples

Design Implications

- Supplement automatic entity identification
- Allow flexible data (document) management
  - Provide an ability to easily select a subset of documents
- Empower with numbers
  - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
  - Limit user interaction vs. give more power
- Invest in tutorial
  - Break down into subtopics with use-cases and examples

Design Implications

- Supplement automatic entity identification
- Allow flexible data (document) management
  - Provide an ability to easily select a subset of documents
- Empower with numbers
  - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
  - Limit user interaction vs. give more power
- Invest in tutorial
  - Break down into subtopics with use-cases and examples

Contributions

- Identified real-world cases of how an interactive visual system for investigative analysis assisted document sensemaking in various domains and tasks
- Discussed issues and findings that emerged upon the use of the visual analytic system
- Provided design recommendations for the system and future visual analytics tools.

Table of Contents

Acknowledgements

- We thank our six professionals for sharing their experience with Jigsaw.

- This work was supported by the National Science Foundation under awards IIS-0915788, CCF-0808863, and the VACCINE Center, a Department of Homeland Security's Center of Excellence in Command, Control and Interoperability.

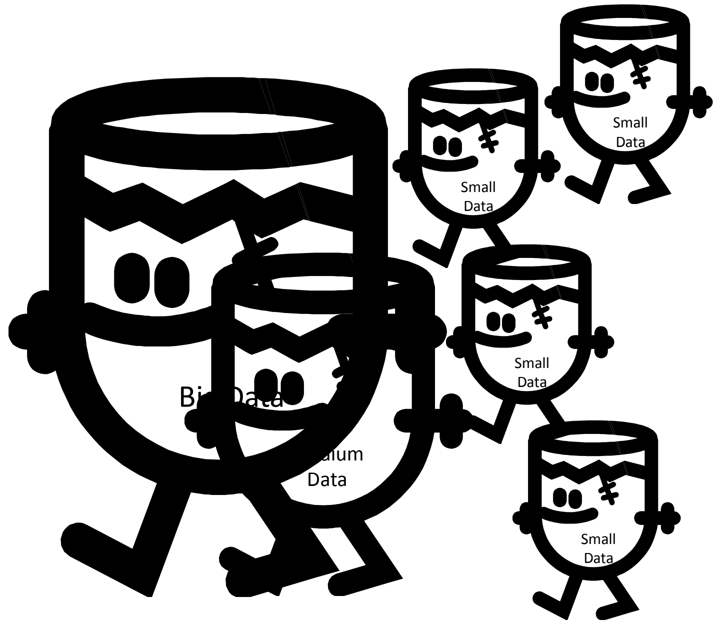
VisWeek 2012 SCIVIS • INFOVIS • VAST BIOVIS • LDAV

Table of Contents

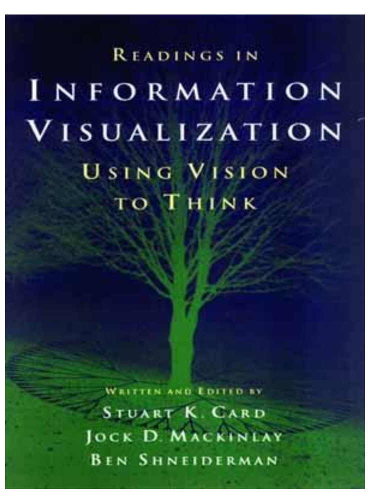
February 19, 2013

The Value of Visualization for Exploring and Understanding Data

John Stasko
School of Interactive Computing
Georgia Institute of Technology
stasko@cc.gatech.edu



3



Visualization

"The use of computer-supported, interactive visual representations of data to amplify cognition"

4

Visualization

~~Making pretty pictures~~

A cognitive process

  Internalize an understanding

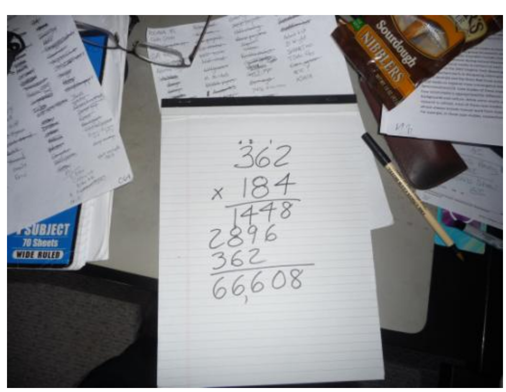
5

Visuals help us think

Provides a frame of reference, temporary storage area

Cognition → Perception

Pattern matching



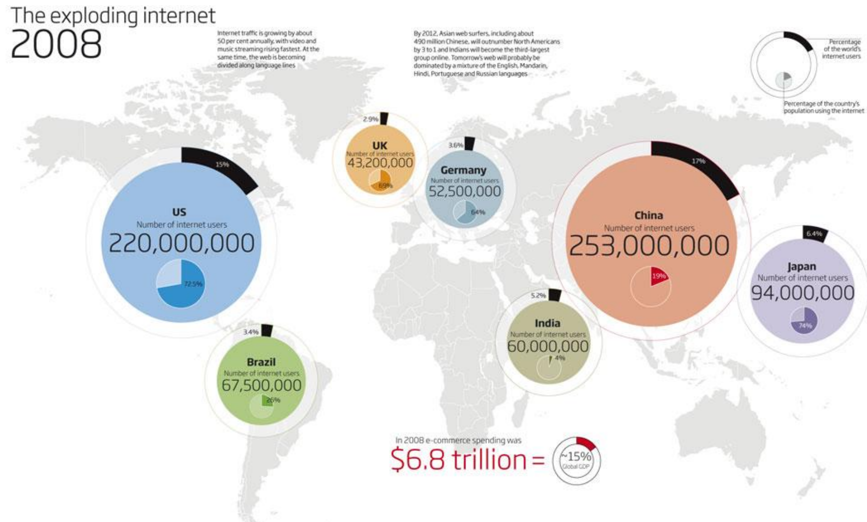
6

Applications of Visualization

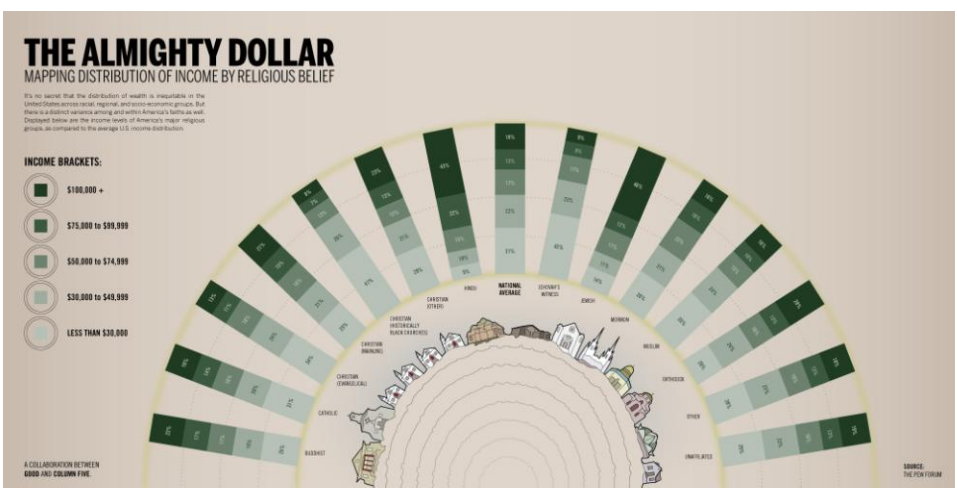
Presentation
Analysis

1. Presentation

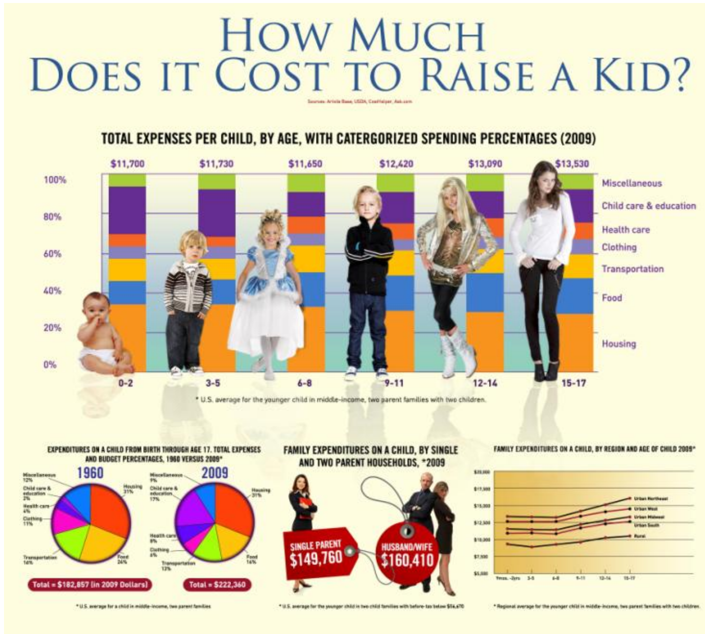
Communicate data and ideas
Explain and inform
Influence and persuade
Provide evidence and support



http://www.newscientist.com/data/images/ns/cms/mg20227062.200/mg20227062.200-6_1000.jpg



http://awesome.good.is.s3.amazonaws.com/transparency/web/1002/almighty-dollar/transparency.jpg

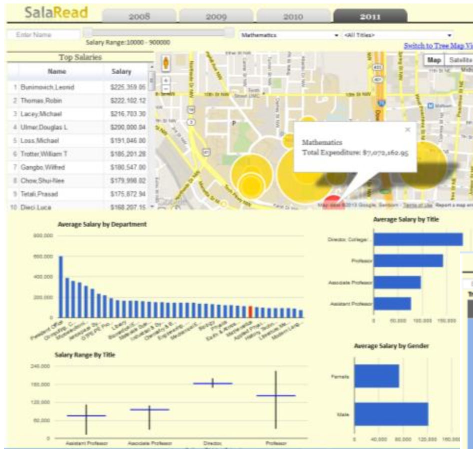


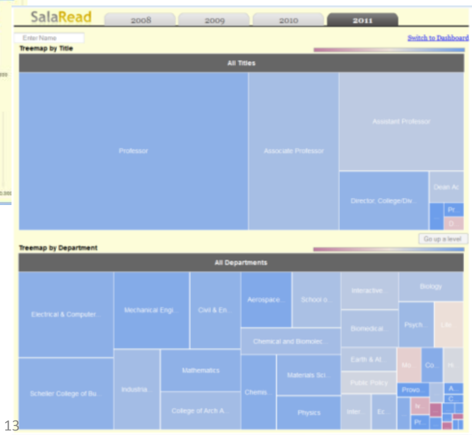
http://visualeconomics.creditloan.com/wp-content/uploads/2010/11/kidcosts-FINAL.jpg

11

Simply presenting data ***visually*** can have a profound impact

12

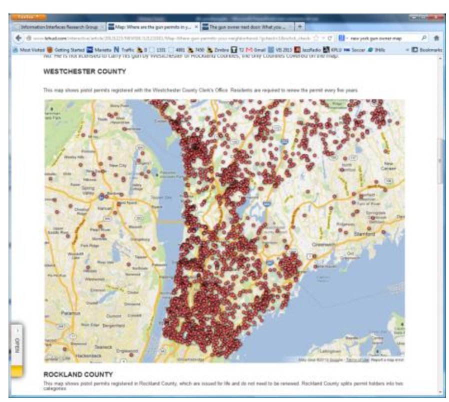




CS 7450
Fall '12
Design project

Nate Osborne
Nitya Noronha
Ameya Zambre
Pratik Zaveri

Gun ownership in New York counties



http://www.lohud.com/apps/pbcs.dll/article?AID=2012312230056&nclick_check=1

http://www.lohud.com/interactive/article/20121223/NEWS01/121221011/Map-Where-gun-permits-your-neighborhood-?gcheck=1&nclick_check=1

Frequent presentation goals

Clarify
Focus
Highlight
Simplify

May just show a few variables
and/or a subset of the data cases

2. Analysis

understand, compare, decide, judge, evaluate,
assess, determine, …

Many Data Analysis Approaches

Statistics
Database & information retrieval
Data mining
Machine learning

"Contained within the data of any investigation is information that can yield conclusions to questions not even originally asked. That is, there can be surprises in the data…To regularly miss surprises by failing to probe thoroughly with visualization tools is terribly inefficient because the cost of intensive data analysis is typically very small compared with the cost of data collection."

W. Cleveland
The Elements of Graphing Data

Visualization Techniques for Analysis

When the standard charts work, use them

http://infosthetics.com/archives/2008/09/funniest_pie_chart_ever.html


Drunkness
jigsaw
GINGERBREAD
ALE
Middle Sister
Limoncello
Capri
Yuengling
PORTER
Leinenkugel's
22

Frequent analysis goals

Show many variables
Illustrate overview and detail
Facilitate comparison

Display may not be easy to interpret at first

Preconceptions about Visualization Utility

Answering specific questions and accomplishing specific analytic tasks
Generating unexpected, serendipitous discoveries and insights
"Finding a needle in a haystack"

Yes, but not what it's best for

1. Visualization is more than just answering a specific question (as is often the case for automated analysis methods)
It also is about the investigative analysis process, which helps us to learn about, develop awareness of, and generate trust in the data, its domain, and its context.

Learning, awareness, trust, context

25

Many Data Analysis Approaches
Statistics
Database & information retrieval
Data mining
Machine learning

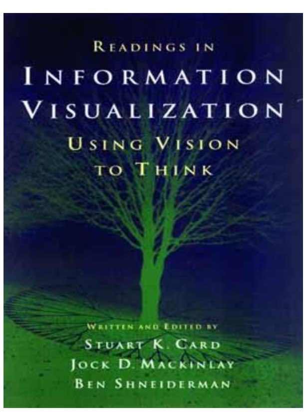
Use them when they get the job done!

26

2. Visualization, primarily through its interactive capabilities, promotes a dialog between analysts and their data by allowing a diverse and flexible set of questions to be asked and answered about a data collection and by spurring the generation of new questions.

Q & A dialog through interaction

27



Visualization
"The use of computer-supported, **interactive** visual representations of data to amplify cognition"

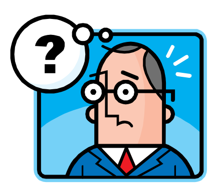
28

3. Visualization rapidly and efficiently facilitates flexible exploration to foster both broad and deep understanding of the information in a data collection.

Broad and deep understanding quickly

Visualization most useful in **exploratory data analysis**

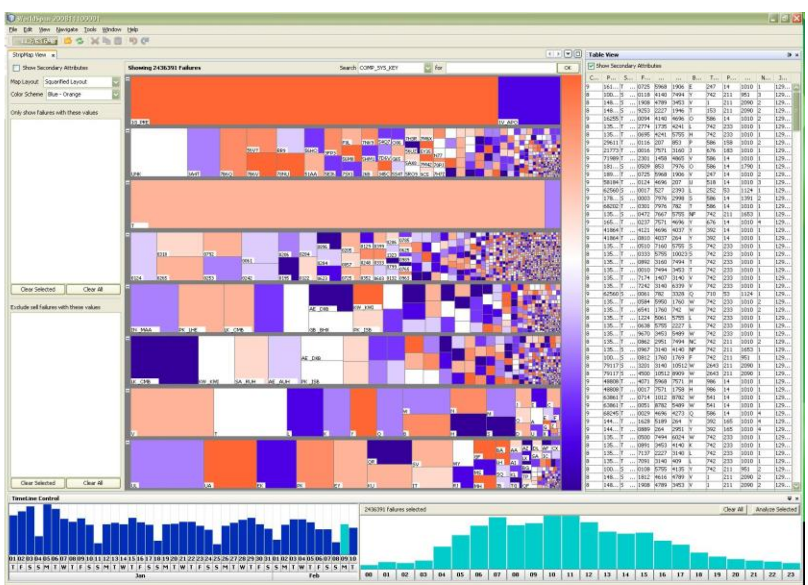
- Don't know what you're looking for
- Don't have a priori questions
- Want to know what questions to ask



insights

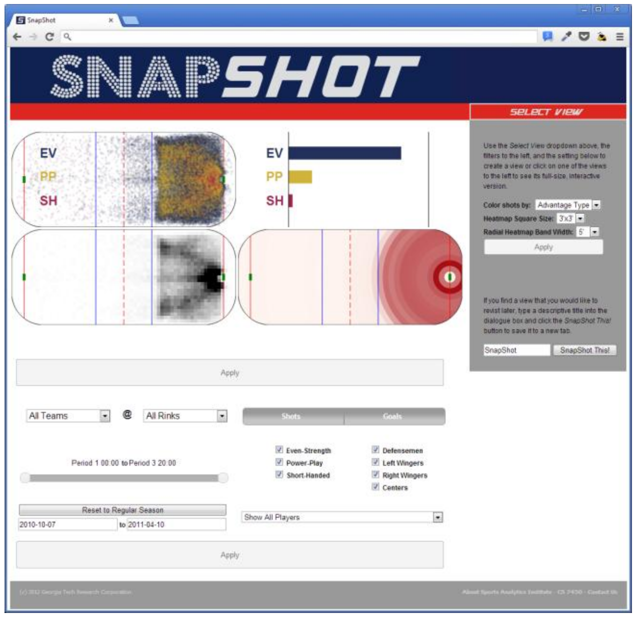
spontaneous aha! moments
vs.
knowledge-building & model confirmation

Chang et al
IEEE CG&A '09



SellTrend

Liu, Stasko, Sullivan
TVCG (InfoVis) '09



SnapShot

Pileggi, Stolper, Boyle, Stasko
TVCG (InfoVis) '12

33

"Putting the pieces together"

Jigsaw

Computational analysis of documents' text

- Entity identification, document similarity, clustering, summarization, sentiment

Multiple visualizations of documents, analysis results, entities, and their connections

- Views are highly coordinated

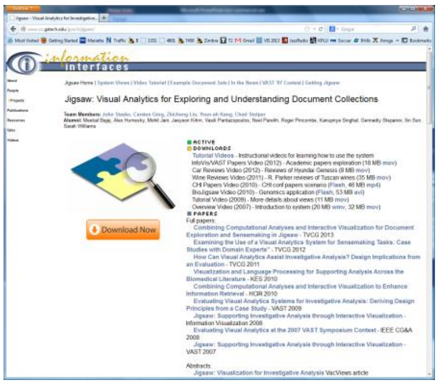
34

Example

Reviews of wines from Tuscany from ‘02, ‘03, ‘07-’10
Text: review narrative
Entities: variety, producer, rating, vintage, color, location, producer, “descriptor”, …
Descriptor terms (~9000), eg: abrasive, oaky, cherry, mocha, textured
1657 reviews
from database of 150,000 reviews

Download for free

http://www.cc.gatech.edu/gvu/ii/jigsaw



37

1. **Visualization** is more than just answering a specific question (as is often the case for automated analysis methods)
It also is about the investigative analysis process, which helps us to learn about, develop awareness of, and generate trust in the data, its domain, and its context.

Learning, awareness, trust, context

Data cleaning example

38

2. Visualization, primarily through its interactive capabilities, promotes a dialog between analysts and their data by allowing a diverse and flexible set of questions to be asked and answered about a data collection and by spurring the generation of new questions.

Q & A dialog through interaction

Repeated interactive queries

39

3. Visualization rapidly and efficiently facilitates flexible exploration to foster both broad and deep understanding of the information in a data collection.

Broad and deep understanding quickly

Varieties, colors, reviewers, producers, concepts, …

40

Insights

No great aha! – More learning & knowledge building
Domain knowledge (wines) matters

41

More visualization benefits

Ease of specifying queries
Opportunistic discovery of relevant data
Spurs the generation of new questions

42

Take Aways

Presentation & analysis

Interaction provides the power

Exploring & developing questions

43

Acknowledgments

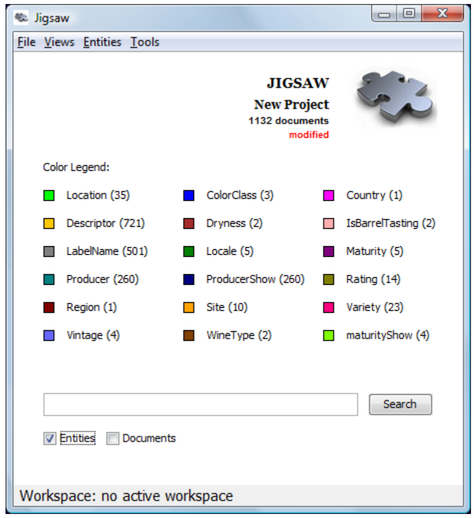
- Work conducted as part of the Southeastern Regional Visualization and Analytics Center, supported by DHS and NVAC and the DHS Center of Excellence in Command, Control & Interoperability (VACCINE Center)

- Supported by NSF IIS-0414667, CCF-0808863 (FODAVA lead), NSF IIS-0915788

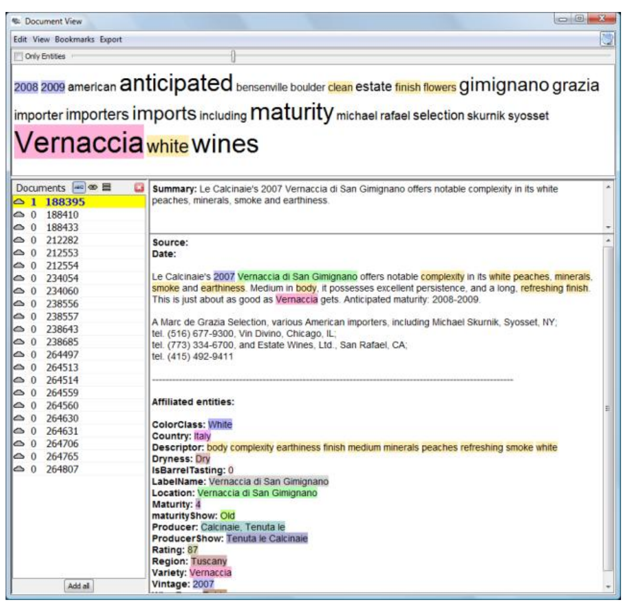
This material is supported in part by: National Science Foundation Grant #0414667

44

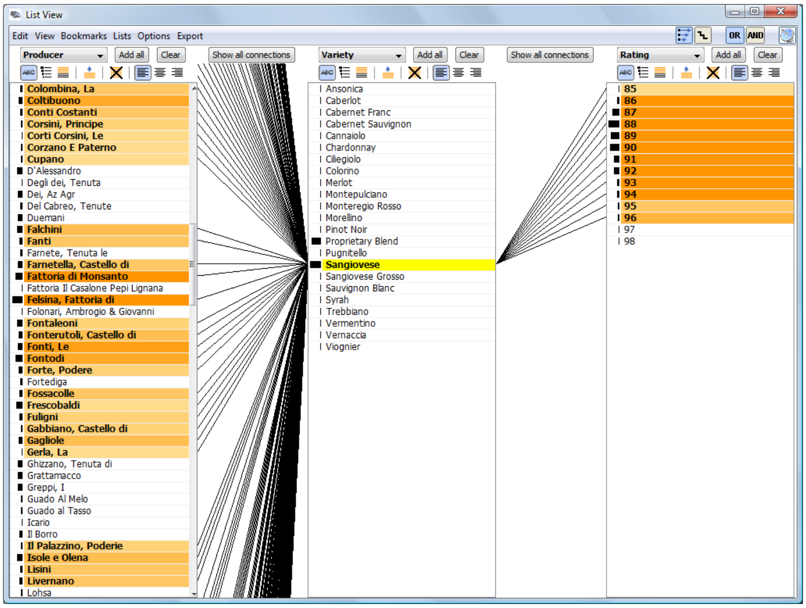
Console



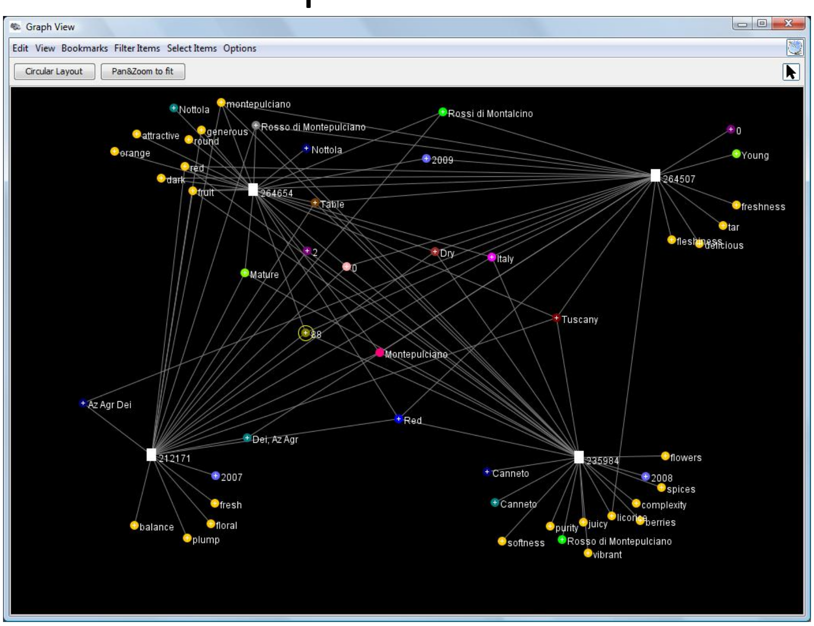
Document View



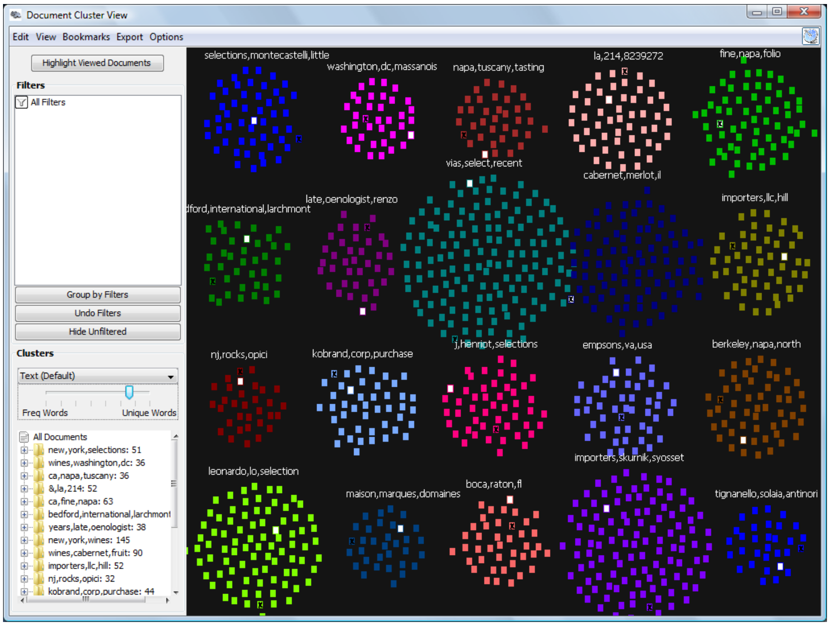
List View



Graph View

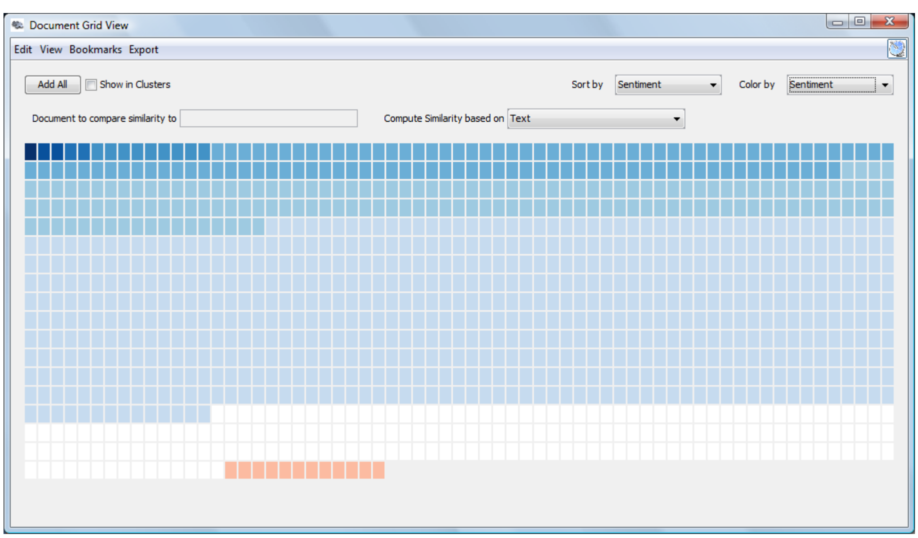


Document Cluster View



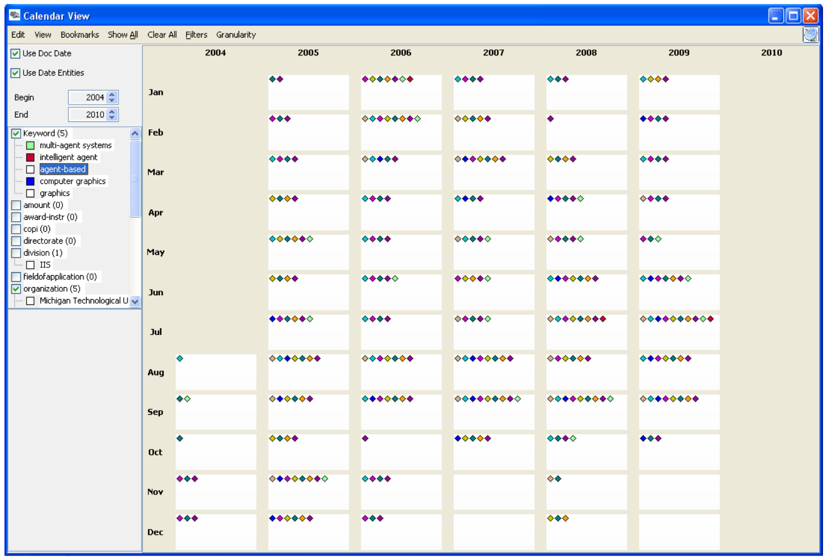
49

Document Grid View



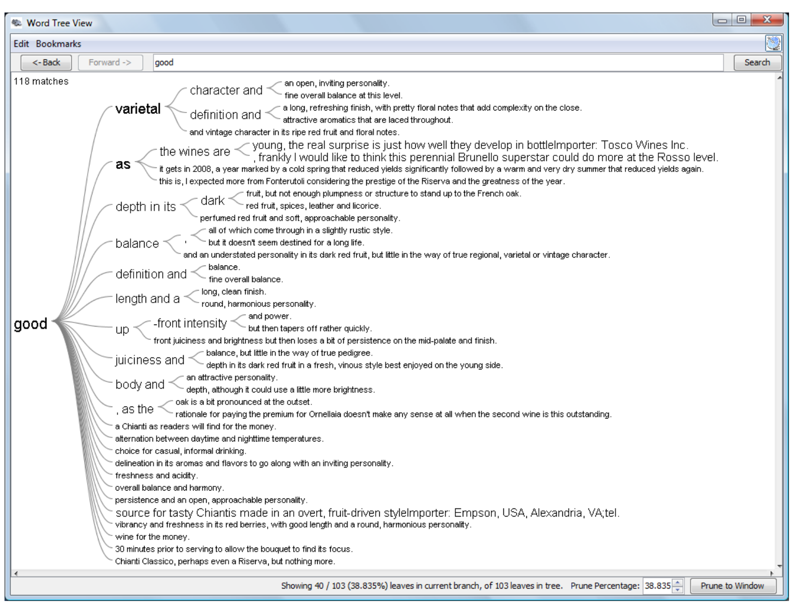
50

Calendar View



51

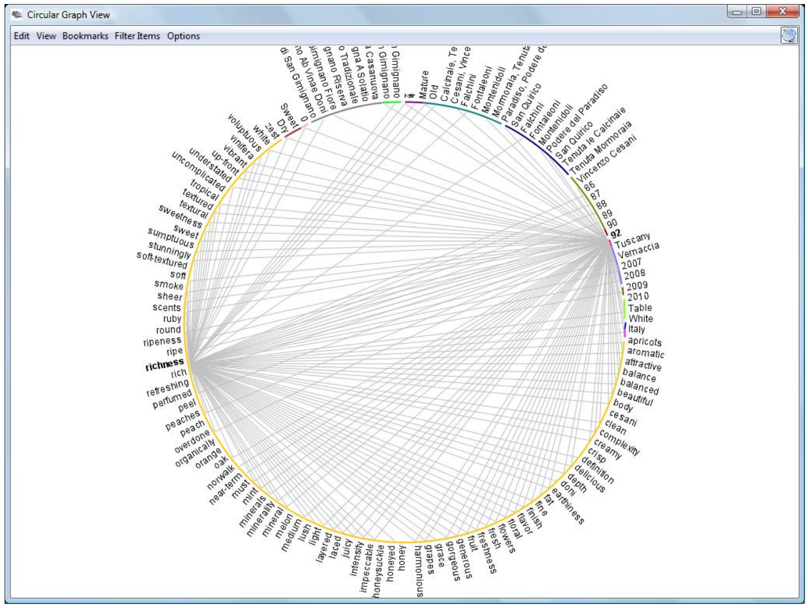
WordTree View



52

Table of Contents

Circular Graph View



Scatterplot View

Table of Contents

May 29, 2013

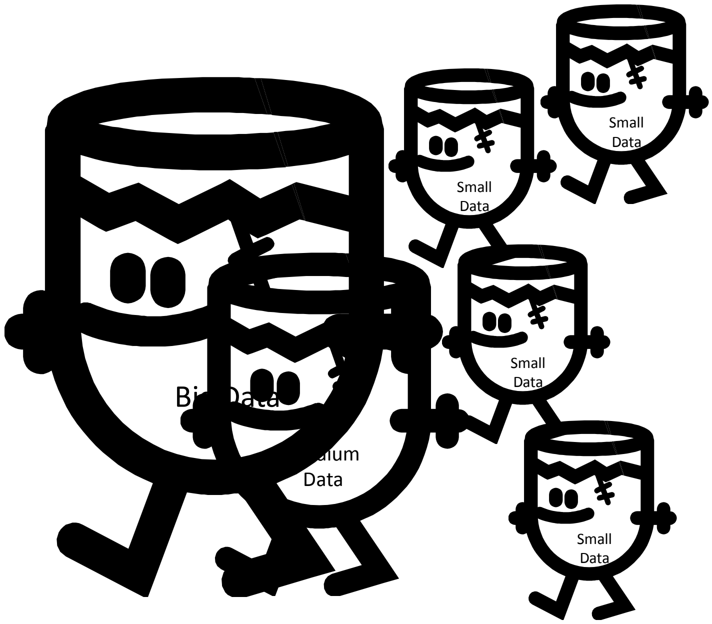
The Value of Visualization for Exploring and Understanding Data

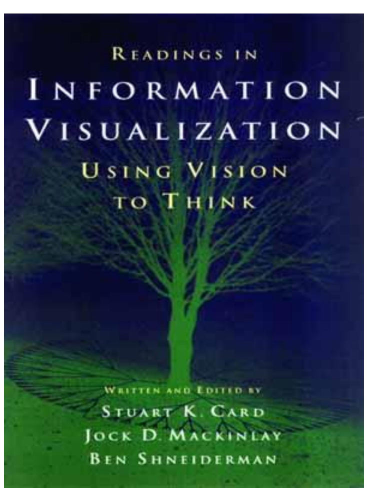
John Stasko
School of Interactive Computing
Georgia Institute of Technology
stasko@cc.gatech.edu

Graphics Interface 2013

information interfaces

Georgia Tech College of Computing School of Interactive Computing





Visualization

“The use of computer-supported, interactive visual representations of data to amplify cognition”

Visualization

~~Making pretty pictures~~

A cognitive process

- Internalize an understanding

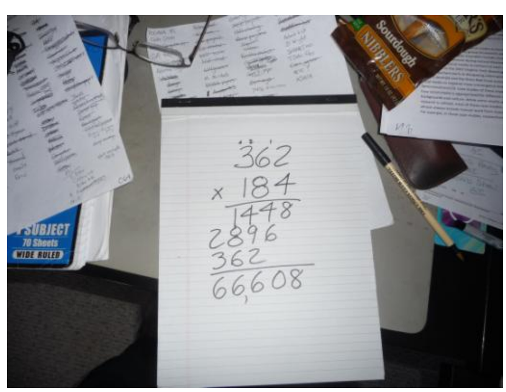
5

Visuals help us think

Provides a frame of reference, temporary storage area

Cognition → Perception

Pattern matching



6

Applications of Visualization

Presentation

Analysis

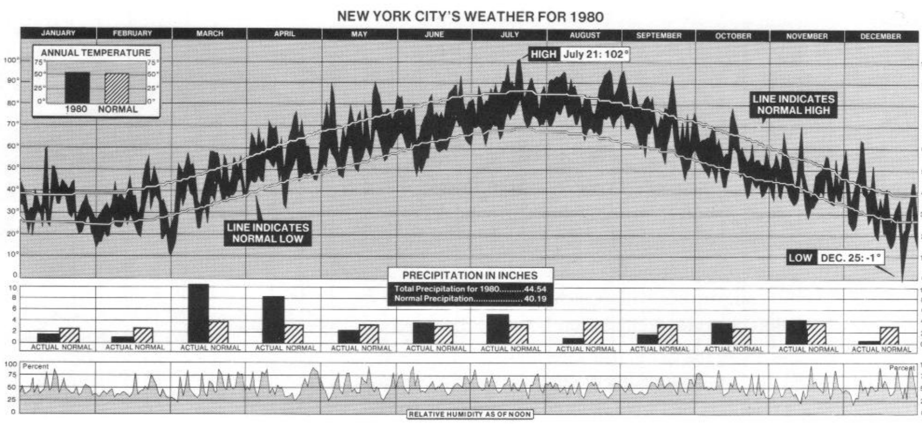
1. Presentation

Communicate data and ideas

Explain and inform

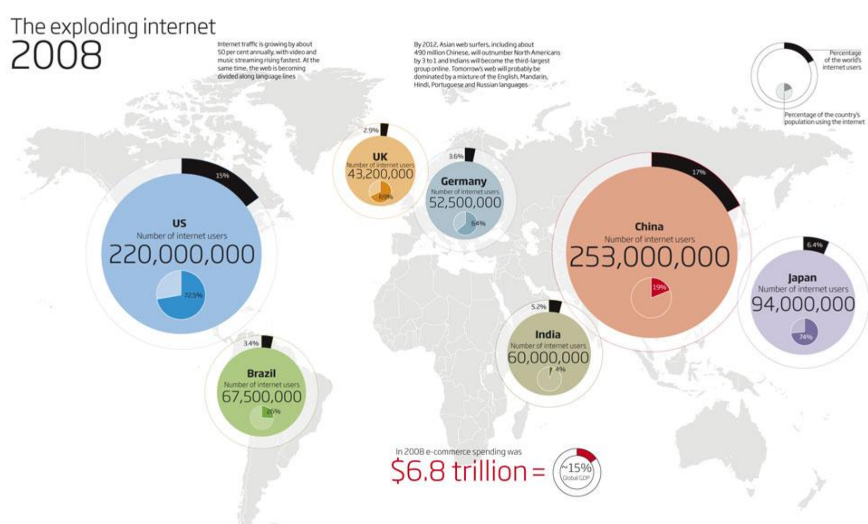
Influence and persuade

Provide evidence and support

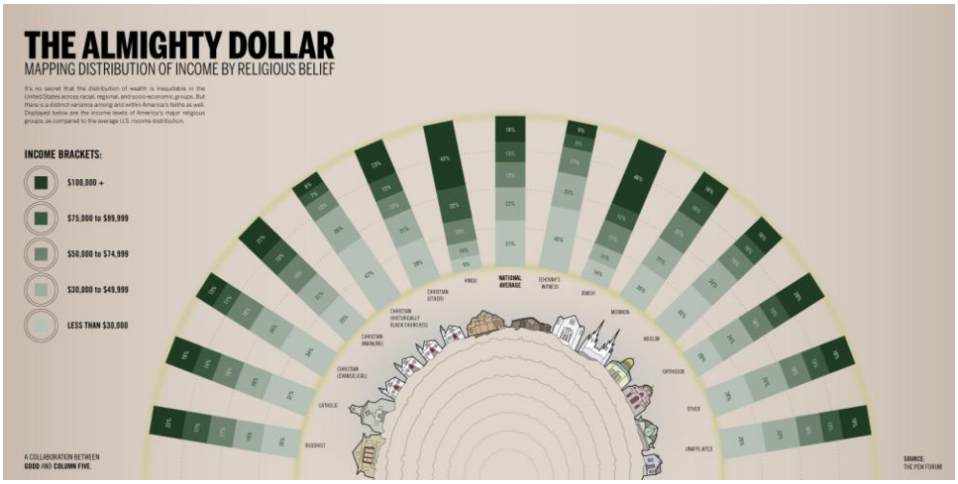


New York Times, January 11, 1981, p. 32.

Tufte, Vol. 1



http://www.newscientist.com/data/images/ns/cms/mg20227062.200/mg20227062.200-6_1000.jpg



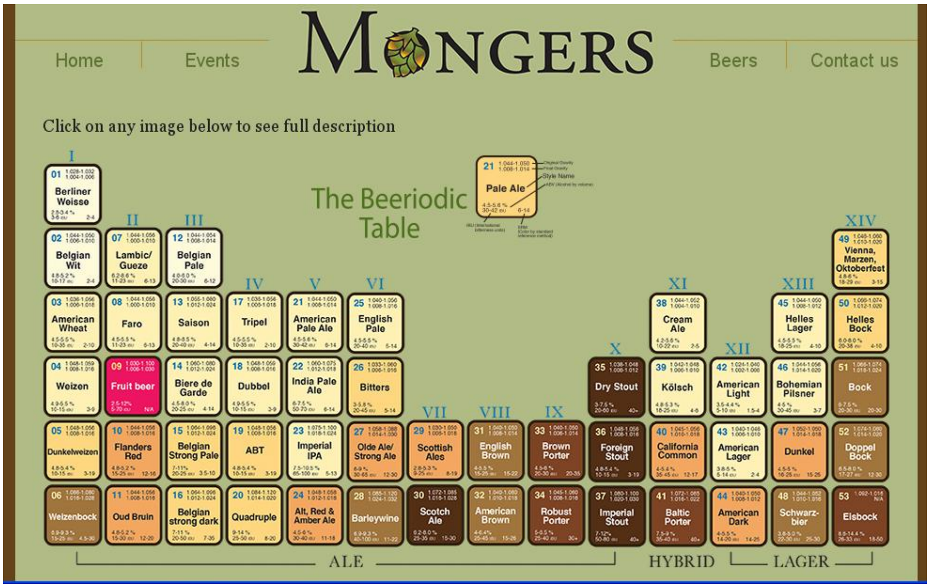
http://awesome.good.is.s3.amazonaws.com/transparency/web/1002/almighty-dollar/transparency.jpg

11

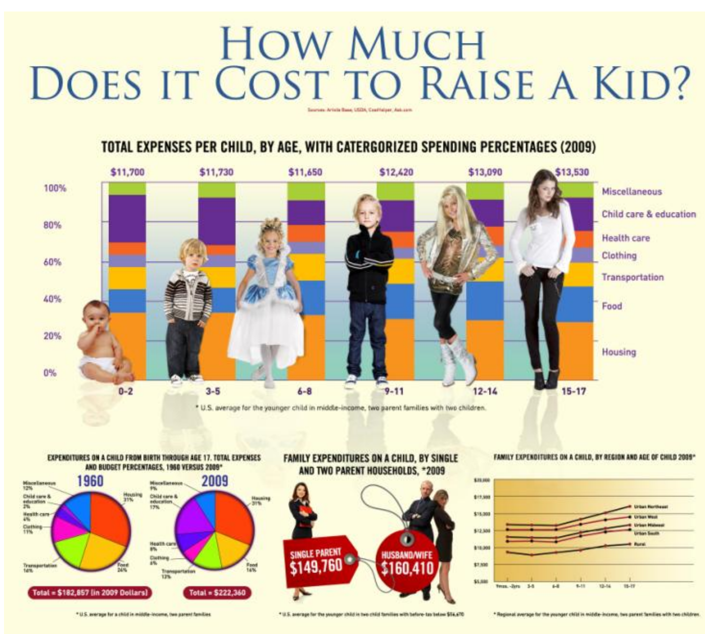
Simon Scarr
South China Morning Post

http://www.visualisingdata.com/blog/wp-content/uploads/2013/04/GOALS.jpg

12



http://thebeermongers.com/beers/



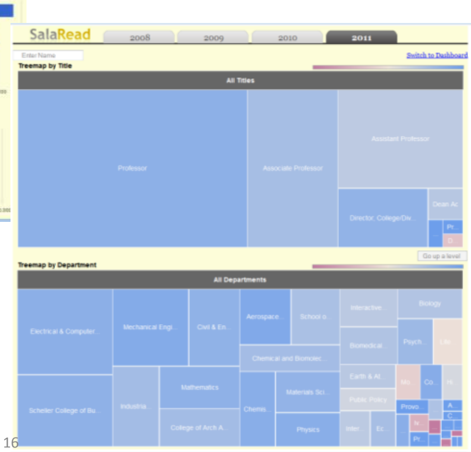
http://visualeconomics.creditloan.com/wp-content/uploads/2010/11/kidcosts-FINAL.jpg

Simply presenting data ***visually*** can have a profound impact

15

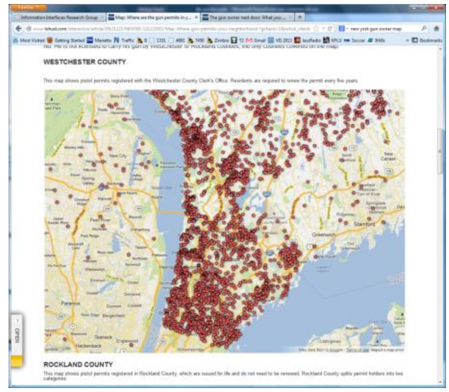
CS 7450
Fall '12
Design project

Nate Osborne
Nitya Noronha
Ameya Zambre
Pratik Zaveri



16

Gun ownership in New York counties



http://www.lohud.com/apps/pbcs.dll/article?AID=2012312230056&nclick_check=1

http://www.lohud.com/interactive/article/20121223/NEWS01/121221011/
Map-Where-gun-permits-your-neighborhood-?gcheck=1&nclick_check=1

17

Frequent presentation goals

Clarify
Focus
Highlight
Simplify

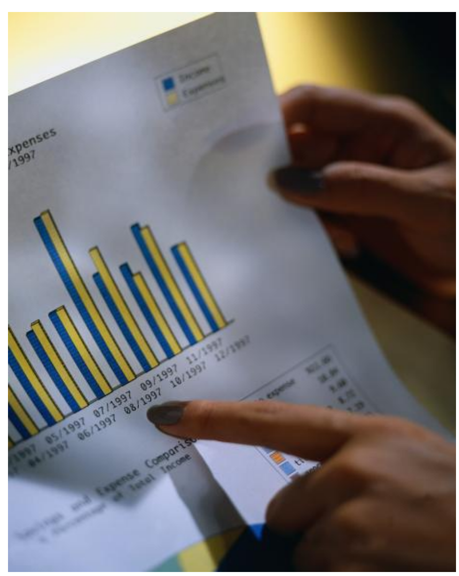
May just show a few variables
and/or a subset of the data cases

18

2. Analysis
understand, compare, decide, judge, evaluate, assess, determine, …

19

Many Data Analysis Approaches
Statistics
Database & information retrieval
Data mining
Machine learning

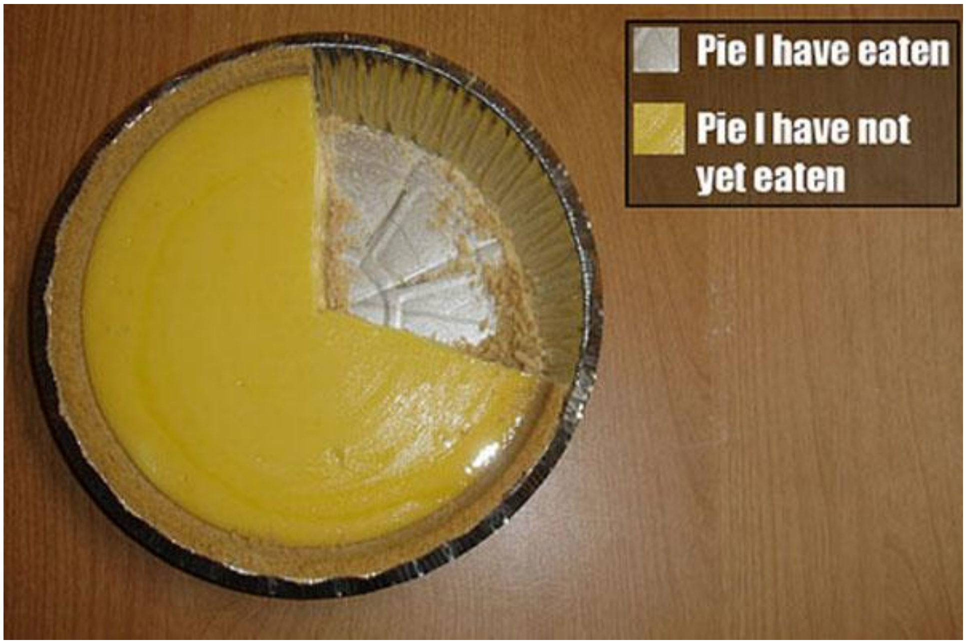


20

"Contained within the data of any investigation is information that can yield conclusions to questions not even originally asked. That is, there can be surprises in the data…To regularly miss surprises by failing to probe thoroughly with visualization tools is terribly inefficient because the cost of intensive data analysis is typically very small compared with the cost of data collection."

W. Cleveland
The Elements of Graphing Data

Visualization Techniques for Analysis
When the standard charts work, use them



http://infosthetics.com/archives/2008/09/funniest_pie_chart_ever.html

23

25

Frequent analysis objectives

Show many variables
Illustrate overview and detail
Facilitate comparison

Display may not be easy to interpret at first

26

Preconceptions about Visualization Utility

Answering a specific question and accomplishing a specific analytic task
Generating unexpected, serendipitous discoveries and insights
"Finding a needle in a haystack"

Yes, but not what it's best for

27

So what is visualization most useful for?

28

1. Visualization is more than just answering a specific question (as is often the case for automated analysis methods)
It also is about the investigative analysis process, which helps us to learn about, develop awareness of, and generate trust in the data, its domain, and its context.

Learning, awareness, trust, context

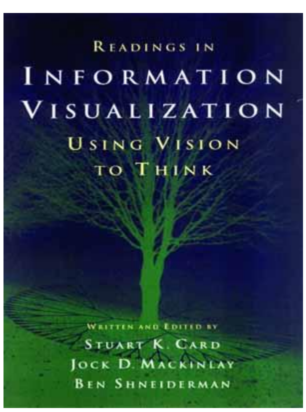
Many Data Analysis Approaches
Statistics
Database & information retrieval
Data mining
Machine learning

Use them when they get the job done!

2. Visualization, primarily through its interactive capabilities, promotes a dialog between analysts and their data by allowing a diverse and flexible set of questions to be asked and answered about a data collection and by spurring the generation of new questions.

Q & A dialog through interaction

31



Visualization
"The use of computer-supported, **interactive** visual representations of data to amplify cognition"

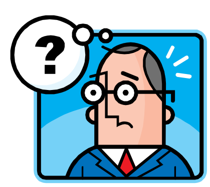
32

3. Visualization rapidly and efficiently facilitates flexible exploration to foster both broad and deep understanding of the information in a data collection.

Broad and deep understanding quickly

Visualization most useful in **exploratory data analysis**

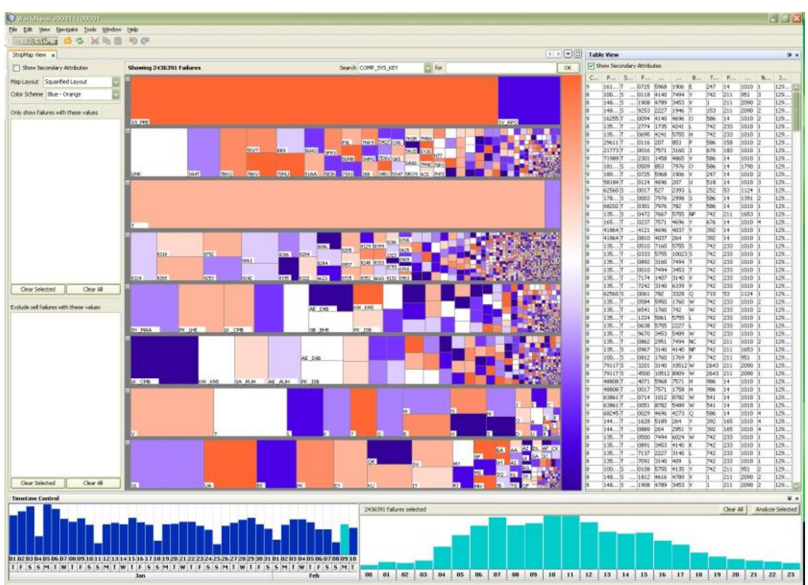
- Don't know what you're looking for
- Don't have a priori questions
- Want to know what questions to ask



insights

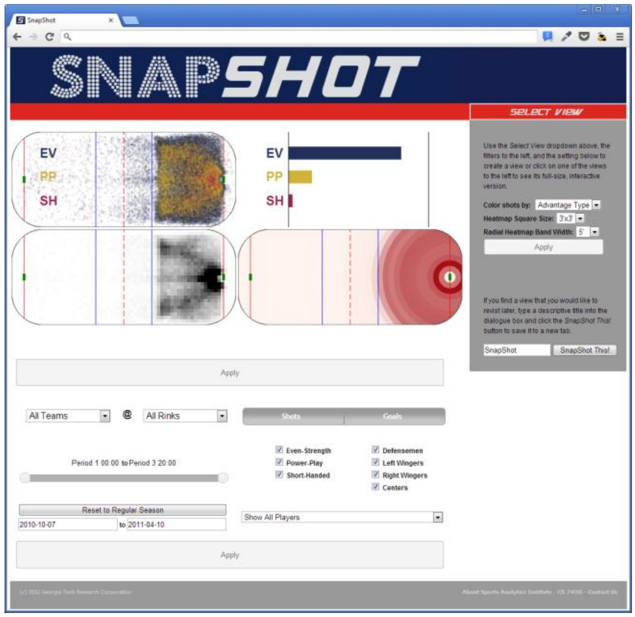
spontaneous aha! moments
vs.
knowledge-building & model confirmation

Chang et al
IEEE CG&A '09



SellTrend

Liu, Stasko, Sullivan
TVCG (InfoVis) '09



SnapShot

Pileggi, Stolper, Boyle, Stasko
TVCG (InfoVis) '12

"Putting the pieces together"

Jigsaw

Computational analysis of documents' text
Entity identification, document similarity, clustering, summarization, sentiment
Multiple visualizations of documents, analysis results, entities, and their connections
Views are highly coordinated

39

Example

Reviews of wines from Tuscany from ‘02, ‘03, ‘07-’10

Text: review narrative

Entities: variety, producer, rating, vintage, color, location, producer, “descriptor”, …

Descriptor terms (~9000), eg: abrasive, oaky, cherry, mocha, textured

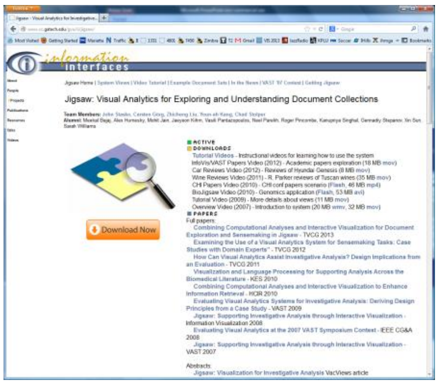
1657 reviews

from database of 150,000 reviews

40

Download for free

http://www.cc.gatech.edu/gvu/ii/jigsaw



41

1. Visualization is more than just answering a specific question (as is often the case for automated analysis methods)
It also is about the investigative analysis process, which helps us to learn about, develop awareness of, and generate trust in the data, its domain, and its context.

Learning, awareness, trust, context

Data cleaning example

42

2. Visualization, primarily through its interactive capabilities, promotes a dialog between analysts and their data by allowing a diverse and flexible set of questions to be asked and answered about a data collection and by spurring the generation of new questions.

Q & A dialog through interaction

Repeated interactive queries

3. Visualization rapidly and efficiently facilitates flexible exploration to foster both broad and deep understanding of the information in a data collection.

Broad and deep understanding quickly

Varieties, colors, reviewers, producers, concepts, …

Insights

No great aha! – More learning & knowledge building
Domain knowledge (wines) matters

Evaluation

Perhaps key open problem/challenge in visualization

By definition, extremely difficult
Proposed approaches:
- Insight-based (Saraiya et al, *TVCG* ‘05)
- Case study (Shneiderman & Plaisant, BELIV ‘06)

More visualization benefits

Ease of specifying queries

Opportunistic discovery of relevant data

Spurs the generation of new questions

Take Aways

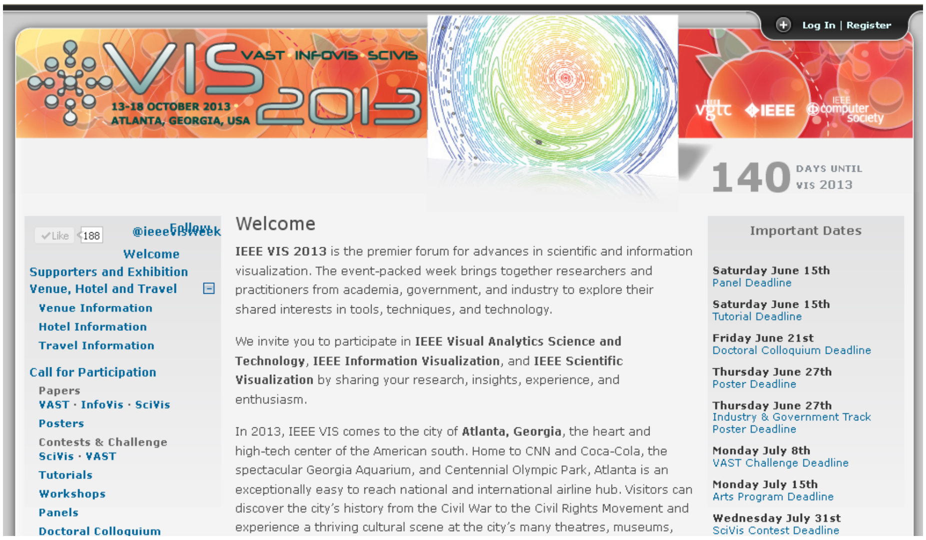
Presentation & analysis

Interaction provides the power

Exploring & developing questions

Table of Contents

VIS 2013 October 13-18 Atlanta, GA



http://ieeevis.org

49

Acknowledgments

- Work conducted as part of the Southeastern Regional Visualization and Analytics Center, supported by DHS and NVAC and the DHS Center of Excellence in Command, Control & Interoperability (VACCINE Center)

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

- Supported by NSF IIS-0414667, CCF-0808863 (FODAVA lead), NSF IIS-0915788

This material is supported in part by: National Science Foundation *Grant #0414667*

50

Table of Contents

Morgan State University iLEAPS Radio Interview Presentation

http://pixel.ecn.purdue.edu:8080/~purpl/VACCINE_Resources/iLEAPS_WEAA_%20Interview.zip

Table of Contents

DEFINING THE TYPICAL DHS MAP

Raechel A. Bianchetti
Robert E. Roth
Justine I. Blanford
Anthony C. Robinson
Alan M. MacEachren

GeoVISTA Center

OVERVIEW

- Mission
- Background
- Method
- Results
- Conclusions

GeoVISTA Center

MISSION

- To determine whether a common map design existed across a range of DHS maps.
  - Evaluate the use of symbol standards across maps
  - Evaluate the types of symbols used on those maps

GeoVISTA Center

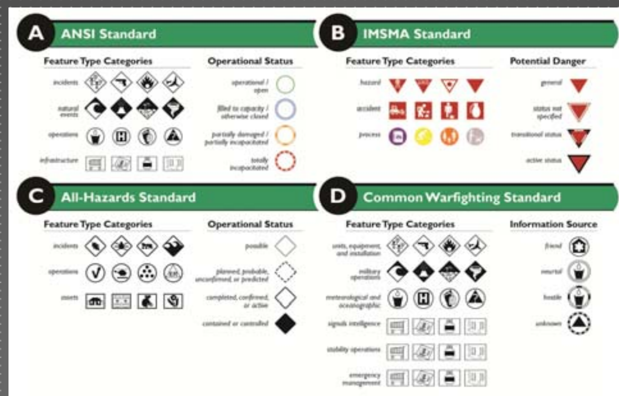
BACKGROUND

interviews
discussions
communication

card sorting
content analysis
action

GeoVISTA Center

EXISTING STANDARDS



GeoVISTA Center

METHODS

- Content Analysis (Berelson 1952)
  - A method used to uncover themes within a collection of non-numerical documents
    - Literature
    - Art
    - Maps
  - Quantitative and Qualitative methods for evaluating the code

GeoVISTA Center

CONTENT ANALYSIS & MAPS

- *National Geographic Magazine* (Lutz and Collins 1993)
- provincial travel maps in Canada (Grant and Keller 1999)
- state-level public health maps depicting cancer rates (Parrott et al. 2007)
- maps generated in support of political discourse (Edsall 2007, Muehlenhaus 2010)
- national topographic maps (Kent 2009, Kent and Vujakovic 2009, Kent and Vujakovic 2011)
- reference maps of European urban centers (Ciolkosz-Styk 2011a, Ciolkosz-Styk 2011b)
- *Goode's World Atlas maps* (Muehlenhaus 2011)

GeoVISTA Center

METHOD OF CONTENT ANALYSIS

1. Gather documents
   1. 76 static maps provided by DHS at interviews or through HSIN
   2. Maps are from 5 agencies- OIP, FEMA, IICD,
2. Develop a set of codes relating to the key research theme
   1. Chose cartographic design choices
3. Apply the codes to the maps
   1. Codes were recorded using excel spreadsheet

GeoVISTA Center

LIMITATIONS

- The map must
  - include point symbols and the array of point symbols must be visually discriminable in order to allow for comparison to the ANSI standard and the other emergency and hazards symbol standards reviewed in Section 2.
  - cannot be a screenshot from a web portal in order to ensure that all map content was visible in the document (e.g., all included feature types/map symbols, map elements, etc.).
- The contents of inset maps included as part of a larger map product were not considered in order to eliminate double counting of map features and to eliminate redundancy of thematic information; the presence/absence of inset maps instead was included as a code in the coding scheme for subsequent description of inset map contents.

GeoVISTA Center

CODING SCHEME

THEMES

- Map Purpose
- Map Content
- Map Design

Categories

- Map objective (o)
- Map source (s)
- Feature type (f)
- Feature condition (c)
- Iconicity (i)
- Symbol Visual Variable (v)
- Basemap (b)
- Map Extent (x)
- Map Elements (e)

GeoVISTA Center

CATEGORIES

Map Objective

- Preparedness
- Response
- Recovery
- mitgatio

Map Source

- Customs and Border Control
- Federal Emergency Management Agency
- National Operations Center
- Office of Infrastructure Protection
- US Coast Guard

GeoVISTA Center

CATEGORIES

Feature Type

- ANSI
- IAL
- MISC

Feature Condition

- Operational Status
- Level of Danger
- Information Trustworthiness

GeoVISTA Center

CATEGORIES

Symbol Visual Variables

- Color Hue
- Pattern
- Shape

Iconicity

- Pictorial
- Associative
- Geometric

GeoVISTA Center

CATEGORIES

Basemap

- Vector
- Remotely Sensed Imagery
- Shaded Relief

Map Extent

- City
- City-to-State
- Regional
- National

GeoVISTA Center

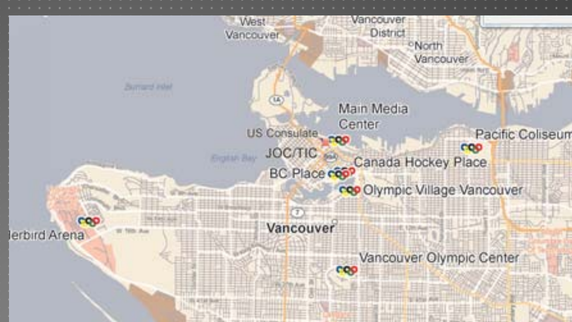
CATEGORIES

Map Elements

- Indication of Orientation
- Indication of Scale
- Inset Maps
- Map Legend

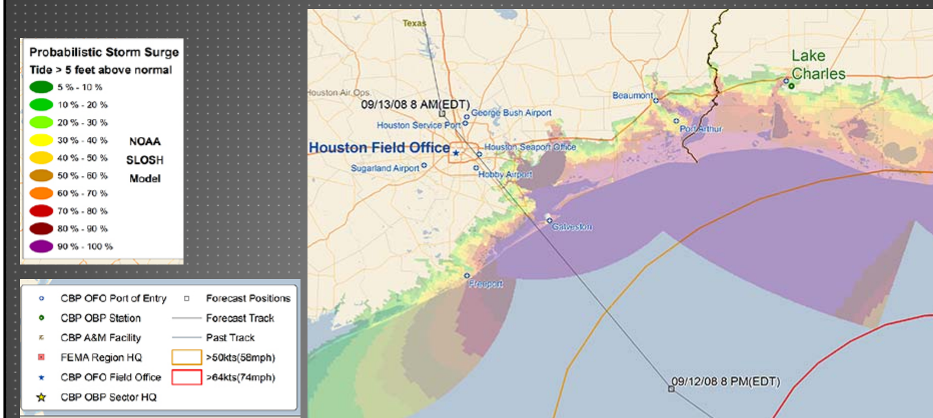
GeoVISTA Center

THE MAPS: SPORTING EVENTS



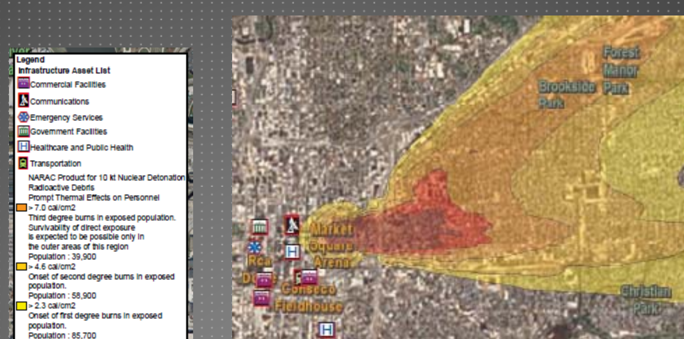
GeoVISTA Center

THE MAPS: HURRICANE



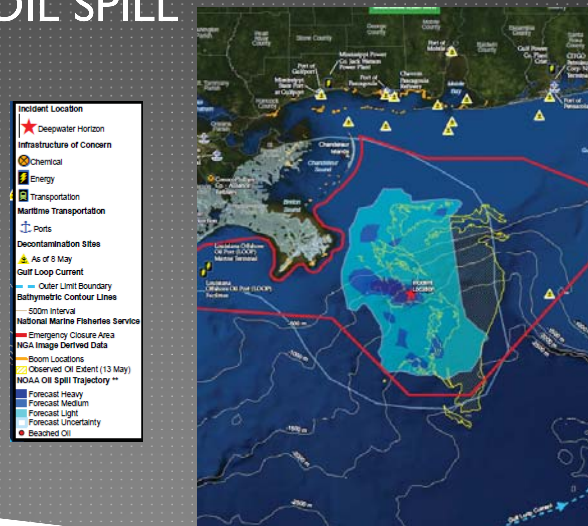
GeoVISTA Center

THE MAPS: NUCLEAR DETONATION

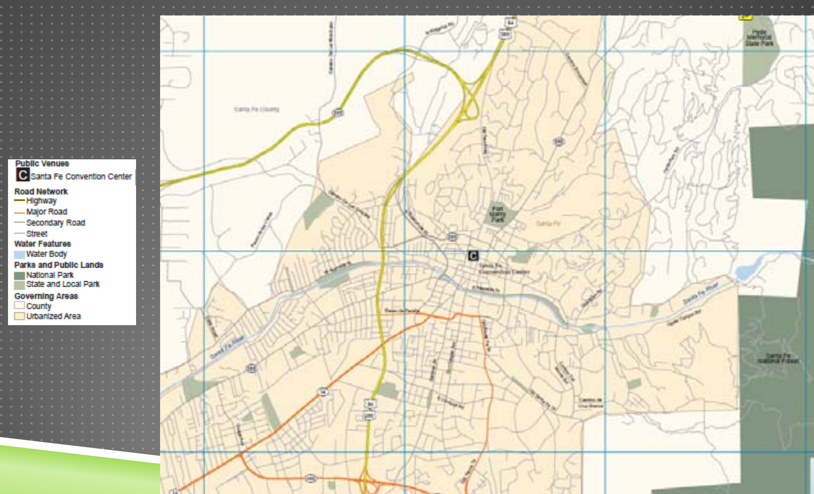


GeoVISTA Center

THE MAPS: OIL SPILL



THE MAPS: ELECTIONS



THE MAPS: SPORTING EVENT

RESULTS: MAP OBJECTIVES

Map objectives

- Not all of the maps aligned with the emergency. management cycle
  - Thematic maps vs. reference maps

Map source

- Most maps were OIP (n=42)
- Map context indicated the agency

RESULTS

Feature Type

- Starting with 198 ANSI symbols we ended with only 26
- Additional 84 feature types added

Feature Type Conditions

- Only 7 examples of levels being used
- 2-5 levels per map

GeoVISTA Center

RESULTS

Iconicity

- Geometric was most common
- Star Symbols for "everything"

Visual Variable

- Only 8 levels instances
- Most common variable used is size

GeoVISTA Center

RESULTS

Base Map

- Vector Maps
- Only FEMA and OIP used shaded relief maps
- MODIS Case

Map Extent

- City (n=40) and City-State(n=27) were most prominent
- City had much fewer points per map than City-State

GeoVISTA Center

RESULTS

MAP ELEMENTS

- 73/76 used map elements
- Orientation n=71
- Legend n=73
- Scale bar n=69
- Inset Maps n=58**

GeoVISTA Center

RESULTS: MAP ELEMENTS

	SCALE	ORIENTATION	LEGENDS	INSET MAPS
FEMA	87%	93%	93%	93%
CBP	100%	83%	100%	33%
USCG	91%	64%	91%	18%

GeoVISTA Center

CONCLUSIONS

- This content analysis is a method for critically analyzing the maps created at DHS to support a bottom-up process of symbol standard development
- There is no typical DHS map (surprised??)
- The ANSI standard
  - Not typically used
  - Not used as they were meant to be

GeoVISTA Center

Table of Contents

THANK YOU.

- U.S. Department of Homeland Security

GeoVISTA Center

Table of Contents

Design & Use Guidelines

For Interactive Maps: A Case Study

http://www.slideshare.net/reroth

@RobertERoth
Robert E. Roth | reroth@wisc.edu

@alanGeoVISTA
Alan M. MacEachren | maceachren@psu.edu

#aag2013
Los Angeles, CA | April 12th, 2013

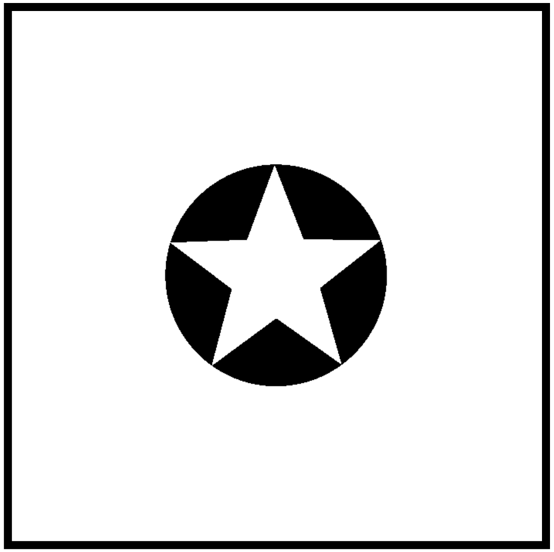
THE UNIVERSITY of WISCONSIN MADISON

GeoVISTA Center

REPRESENTATION *CREATES* UNCERTAINTY

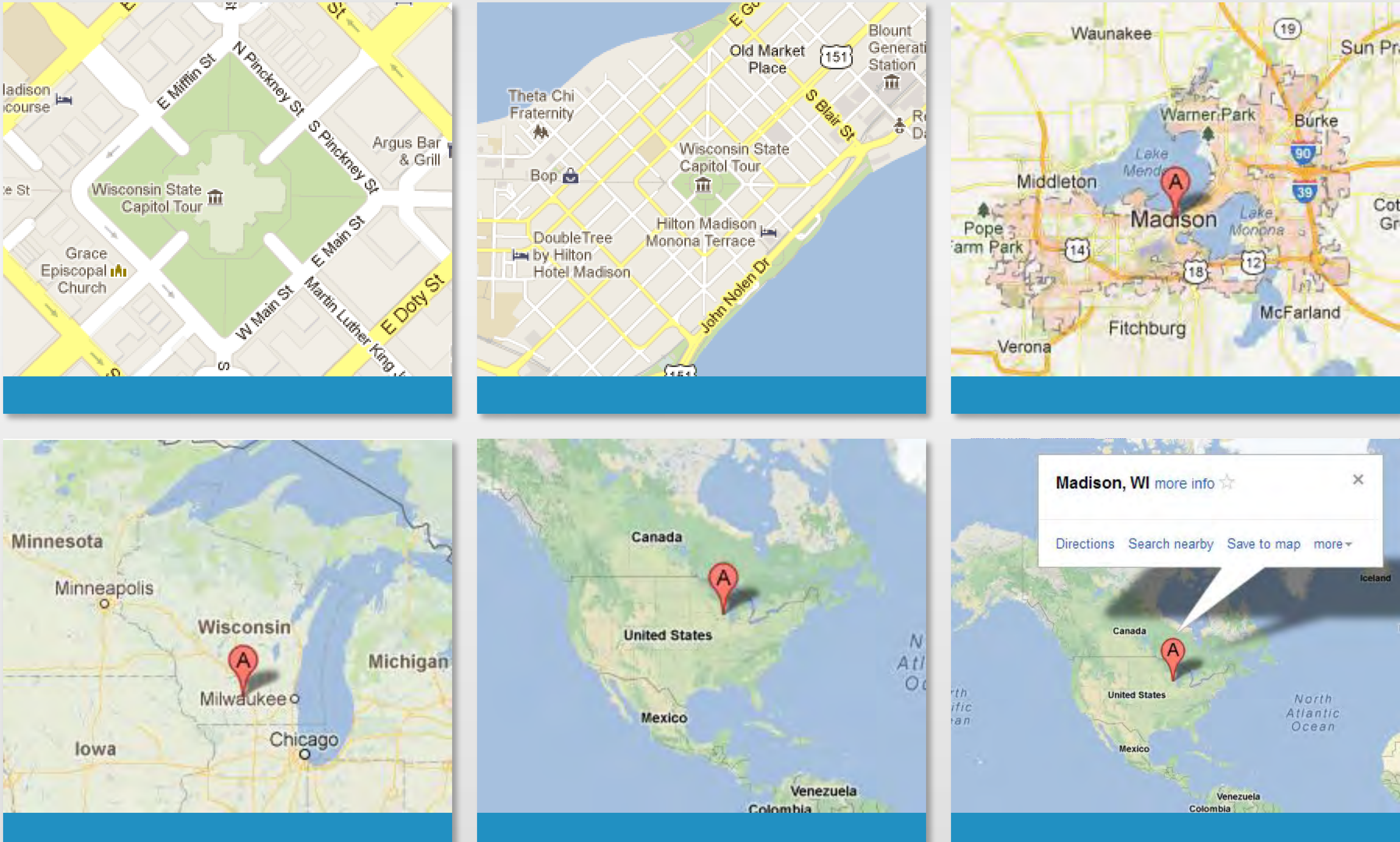
power (useful and usable)

generalization



certainty

INTERACTION *OVERCOMES* UNCERTAINTY(?)



@RobertERoth @alanGeoVISTA
Robert E. Roth Alan M. MacEachren

BACKGROUND: INTERACTION SCIENCE

Roth (2013)
Cartographic Interaction: What We Know and What We Need to Know

Question	*Topics*
What?	*interaction vs. interface • digital vs. analog • stages of interaction • system response time • interactive maps vs. mapping systems vs. map mashups*
Why?	*visual thinking • insight • exploration/confirmation/synthesis/presentation • discussion / debate • decision making*
When?	*interface complexity • interface freedom & flexibility • convention & standardization • work productivity • efficiency and effectiveness*
Who?	*user-centered design • user ability • user expertise • user motivation • adaptive cartography • role-based interfaces • geocollaboration*
Where?	*input capabilities • bandwidth size • processing power • display capabilities • interoperability • mobile mapping • open source development*
How?	*interaction primitives • stages of interaction • interaction strategies • objectives, operators, operands • interface styles • interface design*

Background: How?

Extant Taxonomies of Interaction Primitives

Blok et al. (1999)

Yi et al. (2007)

Buja et al. (1996)

Haber & McNabb (1990)

Crampton (2002)

MacEachren et al. (1999)

Andrienko et al. (2003)

Chi & Riedl (1998)

Zhou & Feiner (1998)

Amar et al. (2005)

Shepherd (1995)

Wehrend & Lewis (1990)

Peuquet (1994)

Shneiderman (1996)

Keim (2002)

Becker & Cleveland (1987)

Wehrend (1993)

Dykes (1997)

Masters & Edsall (2000)

Persson et al. (2006)

Dix & Ellis (1998)

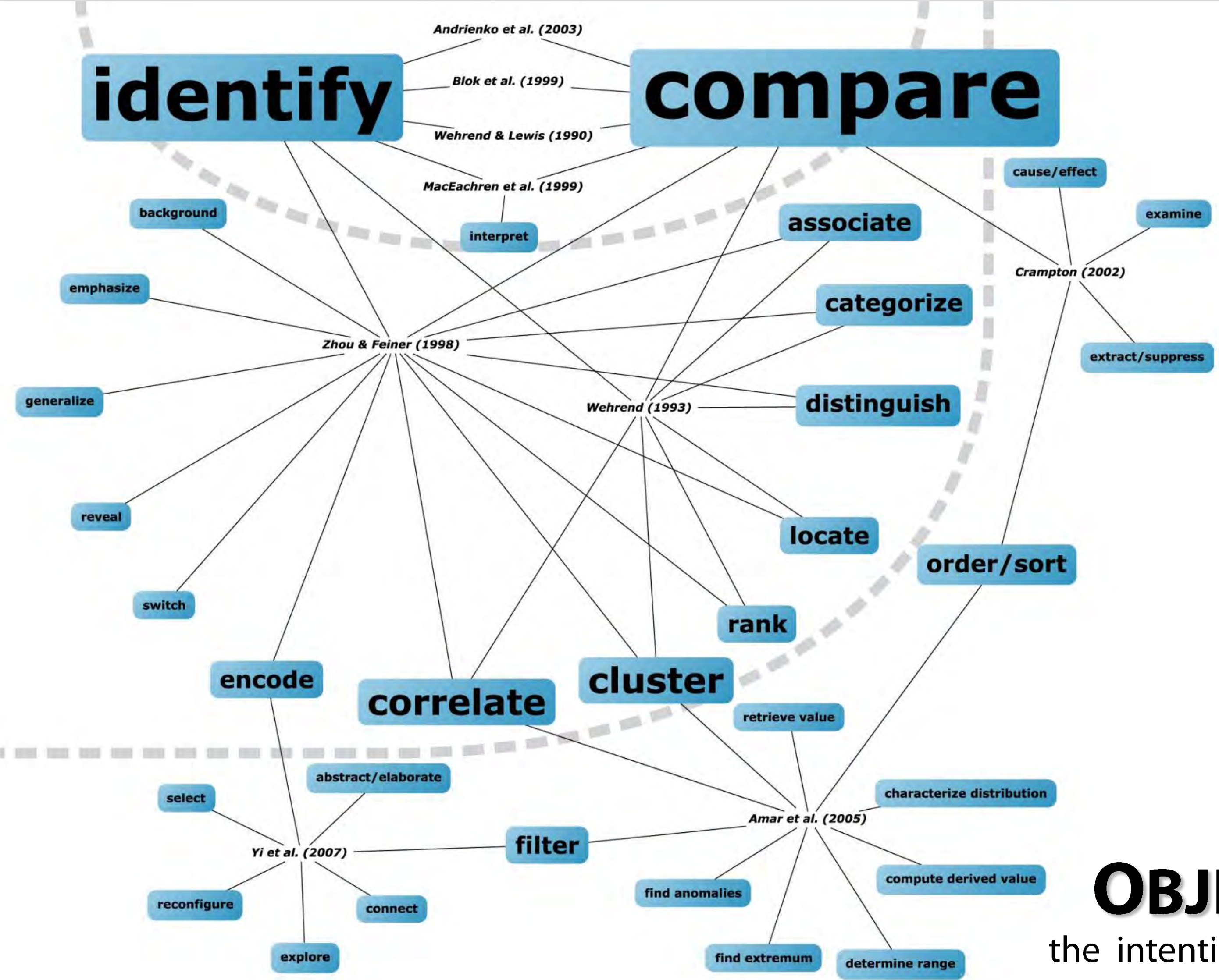
Ward & Yang (2003)

Chuah & Roth (1996)

Roth (2012)

Cartographic Interaction Primitives: Framework & Synthesis

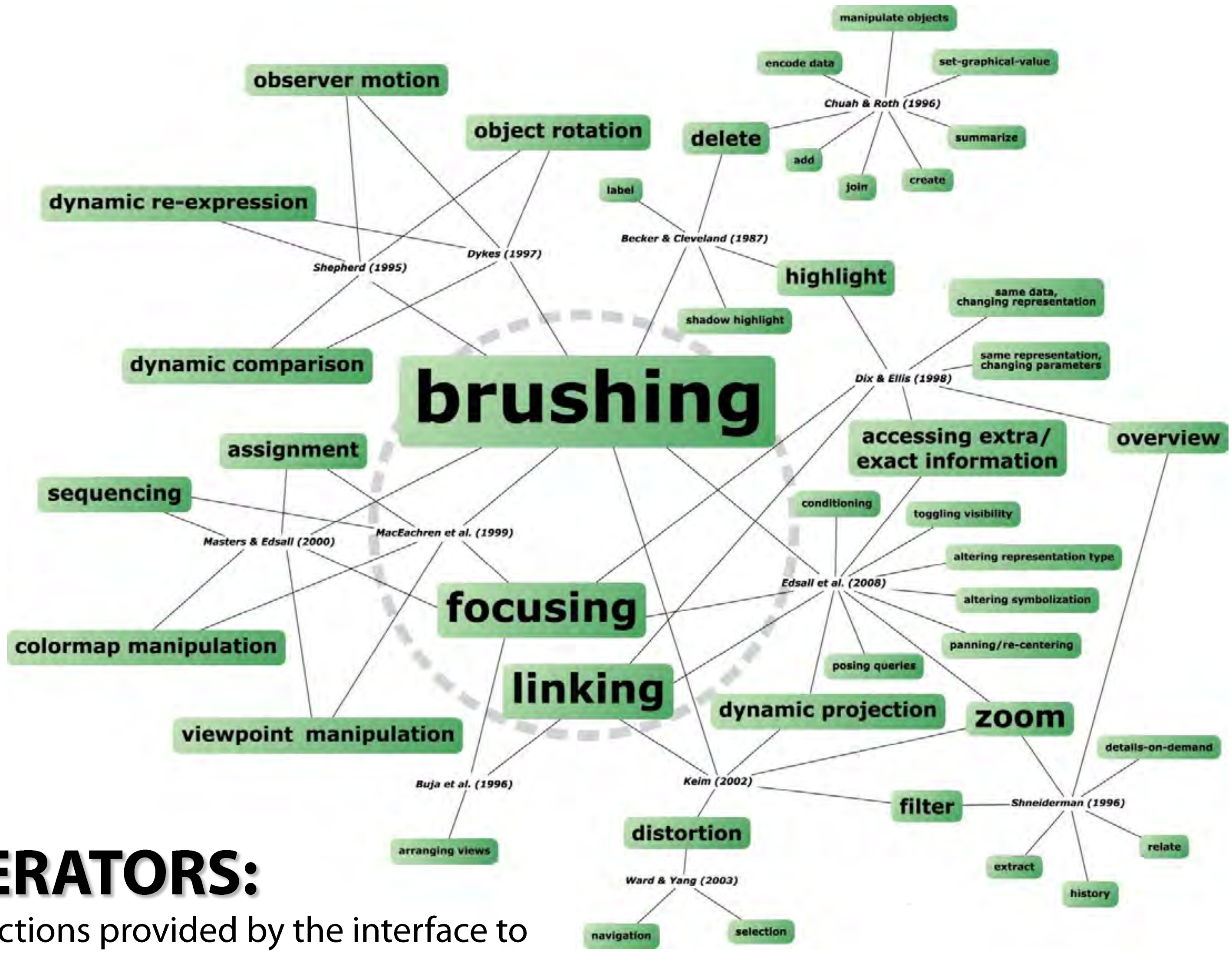
Edsall et al. (2008)



OBJECTIVES:

the intention of the user in completing the interaction

@RobertERoth @alanGeoVISTA
Robert E. Roth Alan M. MacEachren



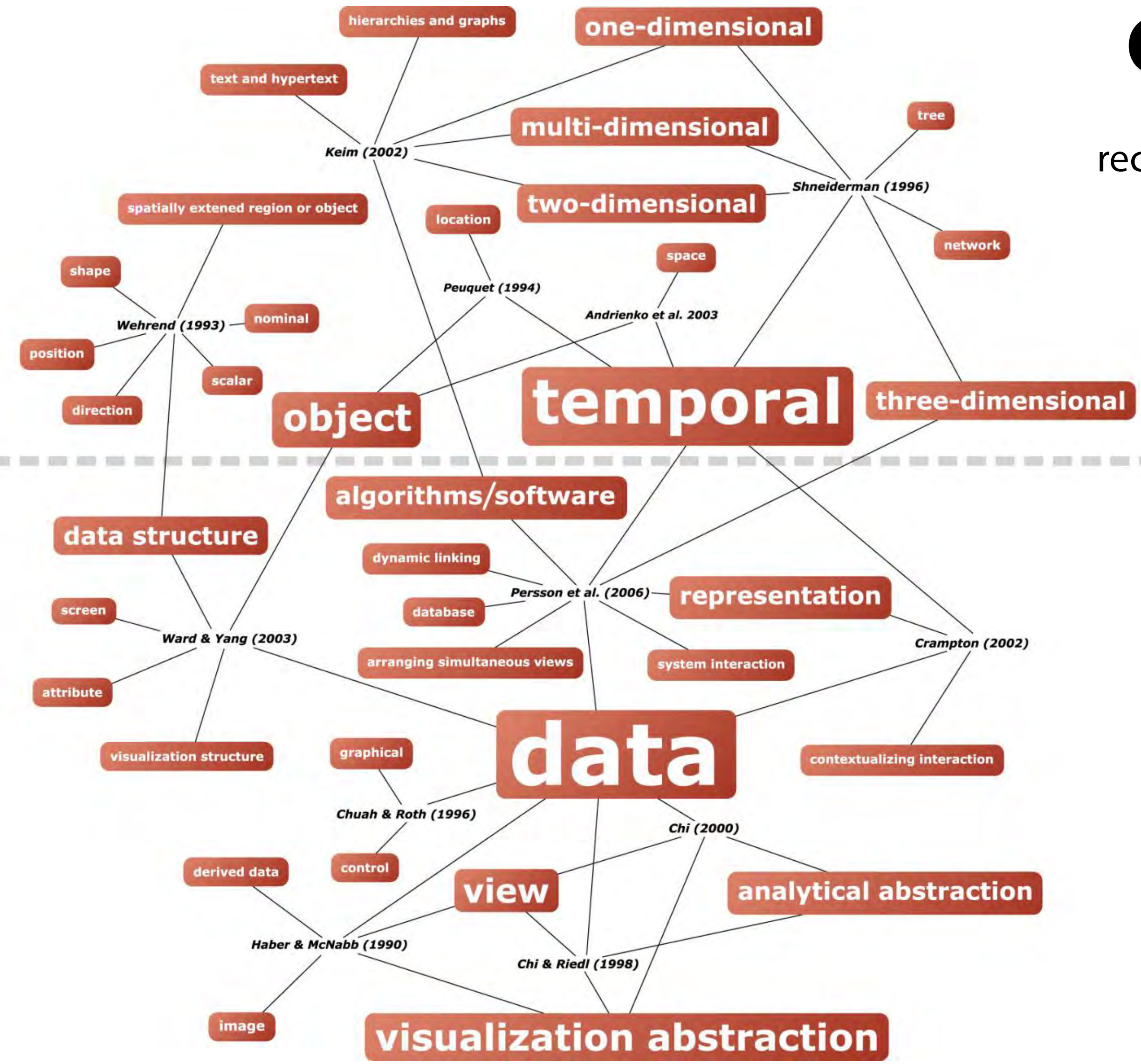
OPERATORS:

the functions provided by the interface to support the user's objectives

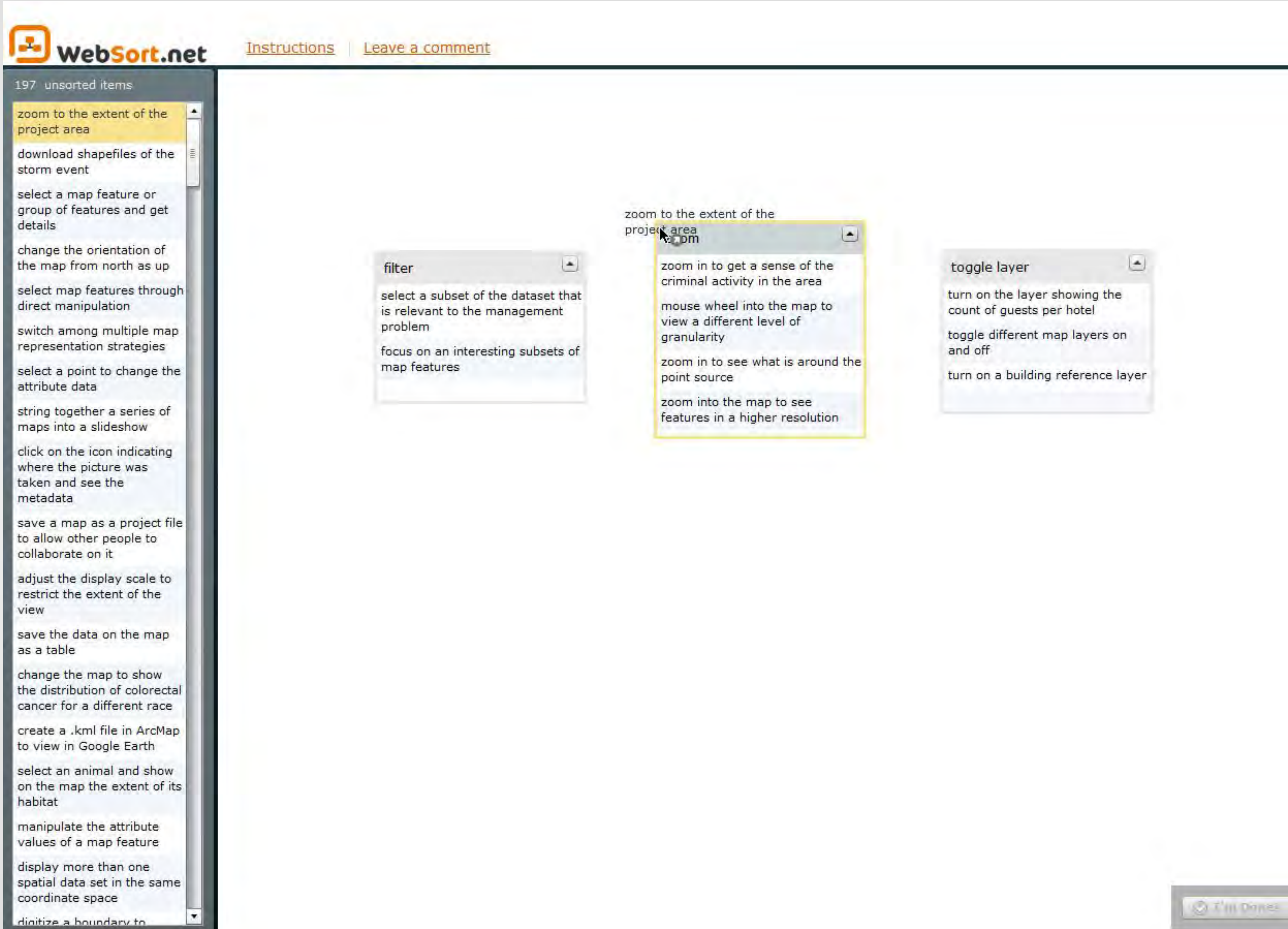
@RobertERoth @alanGeoVISTA
Robert E. Roth Alan M. MacEachren

OPERANDS:

characteristics of the recipient of the operator



9


WebSort.net
Instructions
Leave a comment
197 unsorted items
zoom to the extent of the project area
download shapefiles of the storm event
select a map feature or group of features and get details
change the orientation of the map from north as up
select map features through direct manipulation
switch among multiple map representation strategies
select a point to change the attribute data
string together a series of maps into a slideshow
click on the icon indicating where the picture was taken and see the metadata
save a map as a project file to allow other people to collaborate on it
adjust the display scale to restrict the extent of the view
save the data on the map as a table
change the map to show the distribution of colorectal cancer for a different race
create a .kml file in ArcMap to view in Google Earth
select an animal and show on the map the extent of its habitat
manipulate the attribute values of a map feature
display more than one spatial data set in the same coordinate space
filter
select a subset of the dataset that is relevant to the management problem
focus on an interesting subsets of map features
zoom to the extent of the project area
zoom in to get a sense of the criminal activity in the area
mouse wheel into the map to view a different level of granularity
zoom in to see what is around the point source
zoom into the map to see features in a higher resolution
toggle layer
turn on the layer showing the count of guests per hotel
toggle different map layers on and off
turn on a building reference layer

BACKGROUND: PRIMITIVE TAXONOMY

objectives	operators		operands
identify	*reexpress*	*pan*	*space-alone*
compare	*arrange*	*zoom*	*attributes-in-space*
rank	*sequence*	*filter*	*space-in-time*
associate	*resymbolize*	*search*	
delineate	*overlay*	*retrieve*	
	reproject	*calculate*	

increase in sophistication

Roth (2012)
An Empirically-Derived Taxonomy of Cartographic Interaction Primitives

CASE STUDY: GEOVISTA *CRIMEVIZ*

http://www.geovista.psu.edu/CrimeViz

Project Personnel:

Kevin S. Ross | ***kevin@kross.com***

Benjamin G. Finch | ***bgf111@psu.edu***

Wei Luo | ***wul132@psu.edu***

Craig A. McCabe | ***CMcCabe@esri.com***

Ryan Mullins | ***ryanmullins@psu.edu***

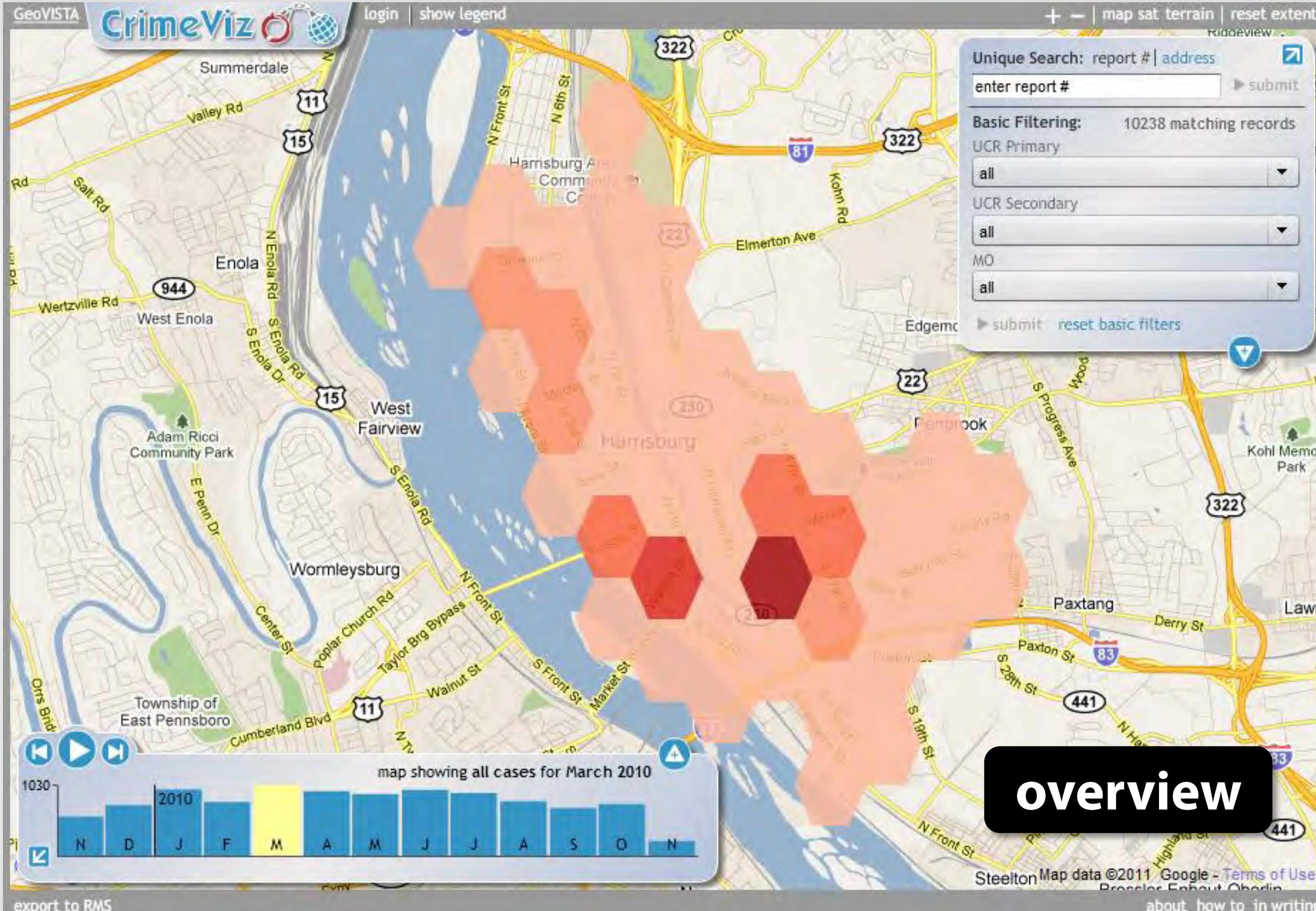
Scott Pezanowski | ***spezanowski@psu.edu***

Camilla Robinson | ***clr281@psu.edu***

Contact:

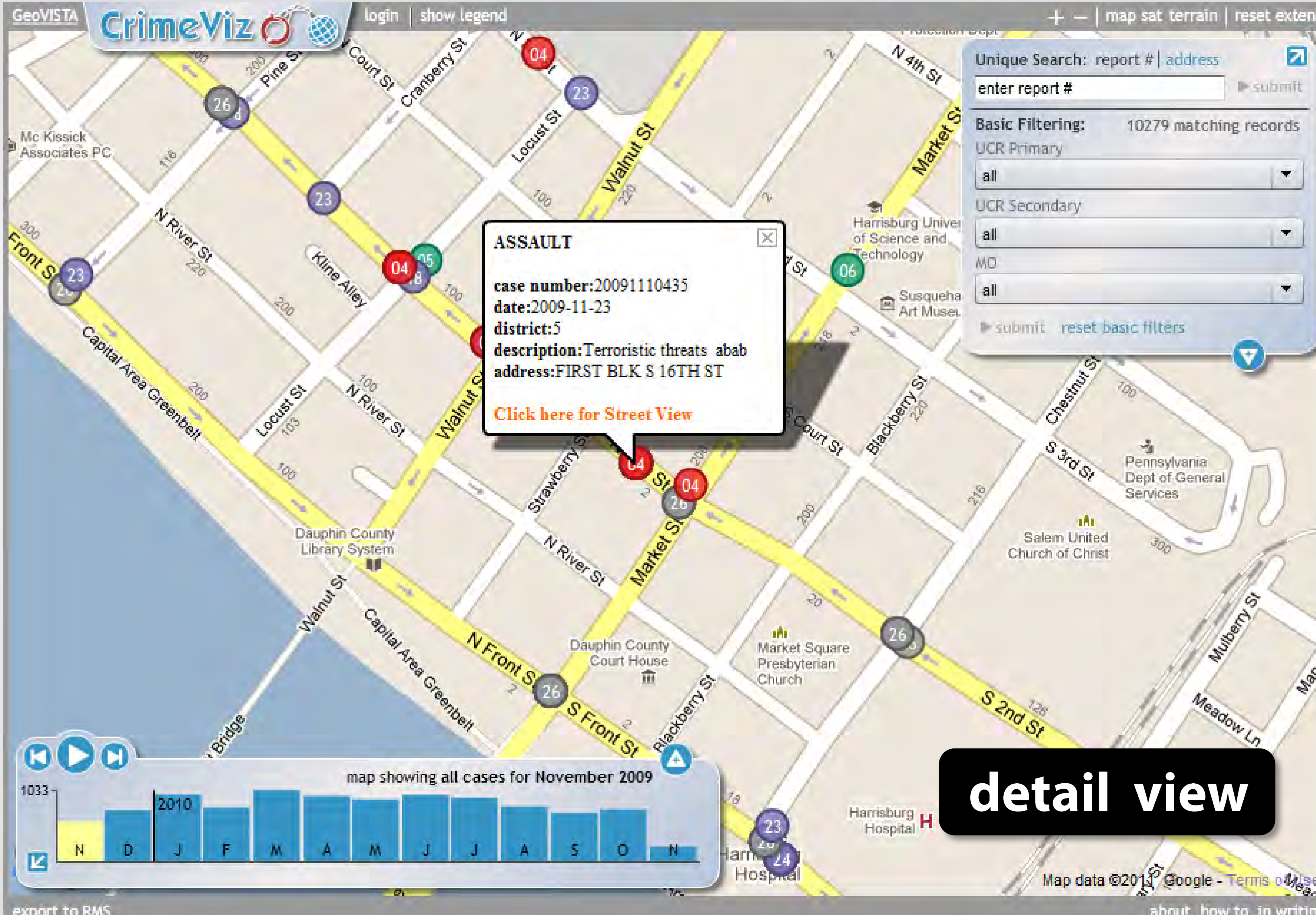
Alan M. MacEachren | ***maceachren@psu.edu***

Robert E. Roth | ***reroth@wisc.edu***



#aag2013
Los Angeles, CA | April 12th, 2013

@RobertERoth **@alanGeoVISTA**
Robert E. Roth *Alan M. MacEachren*

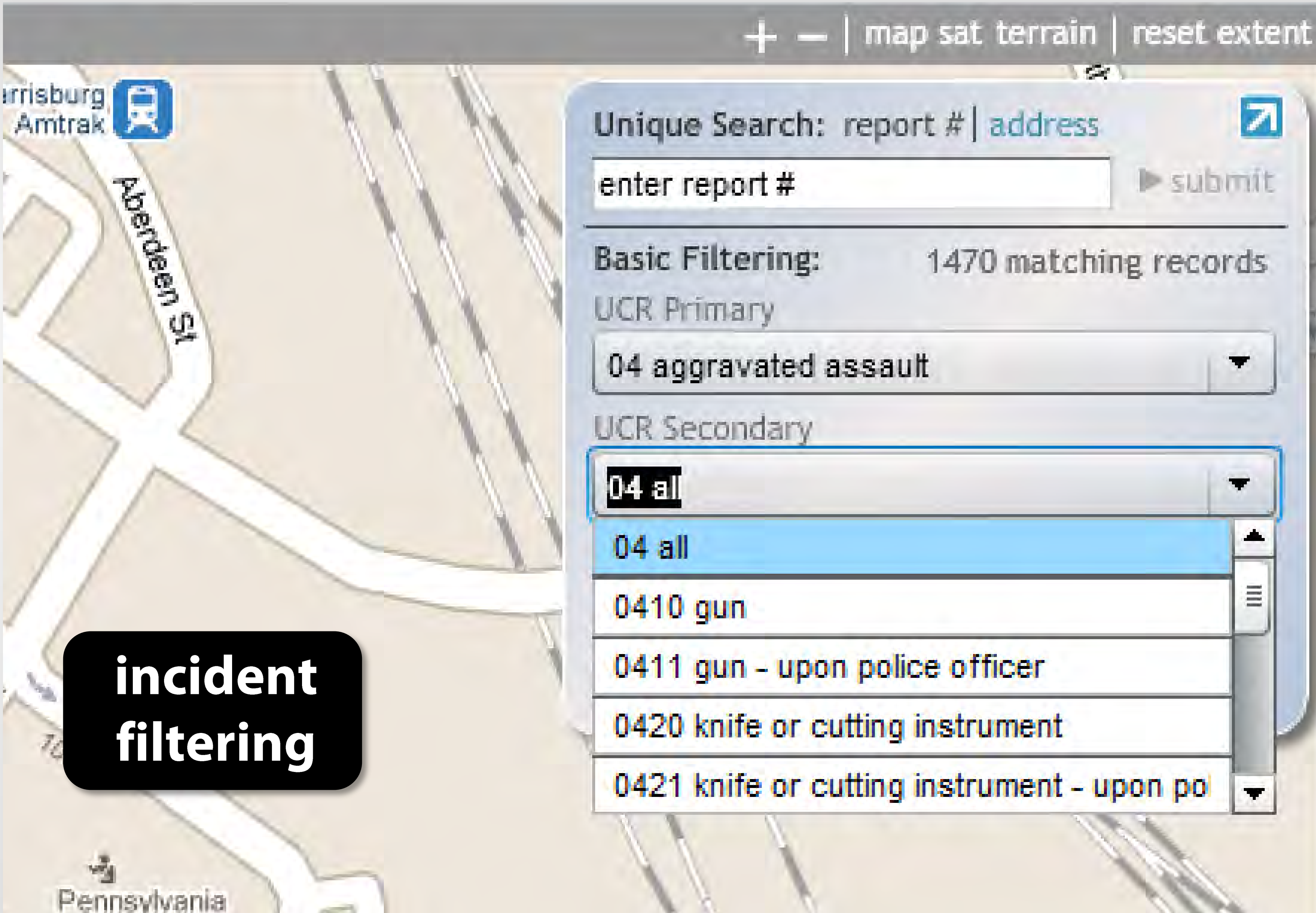


#aag2013

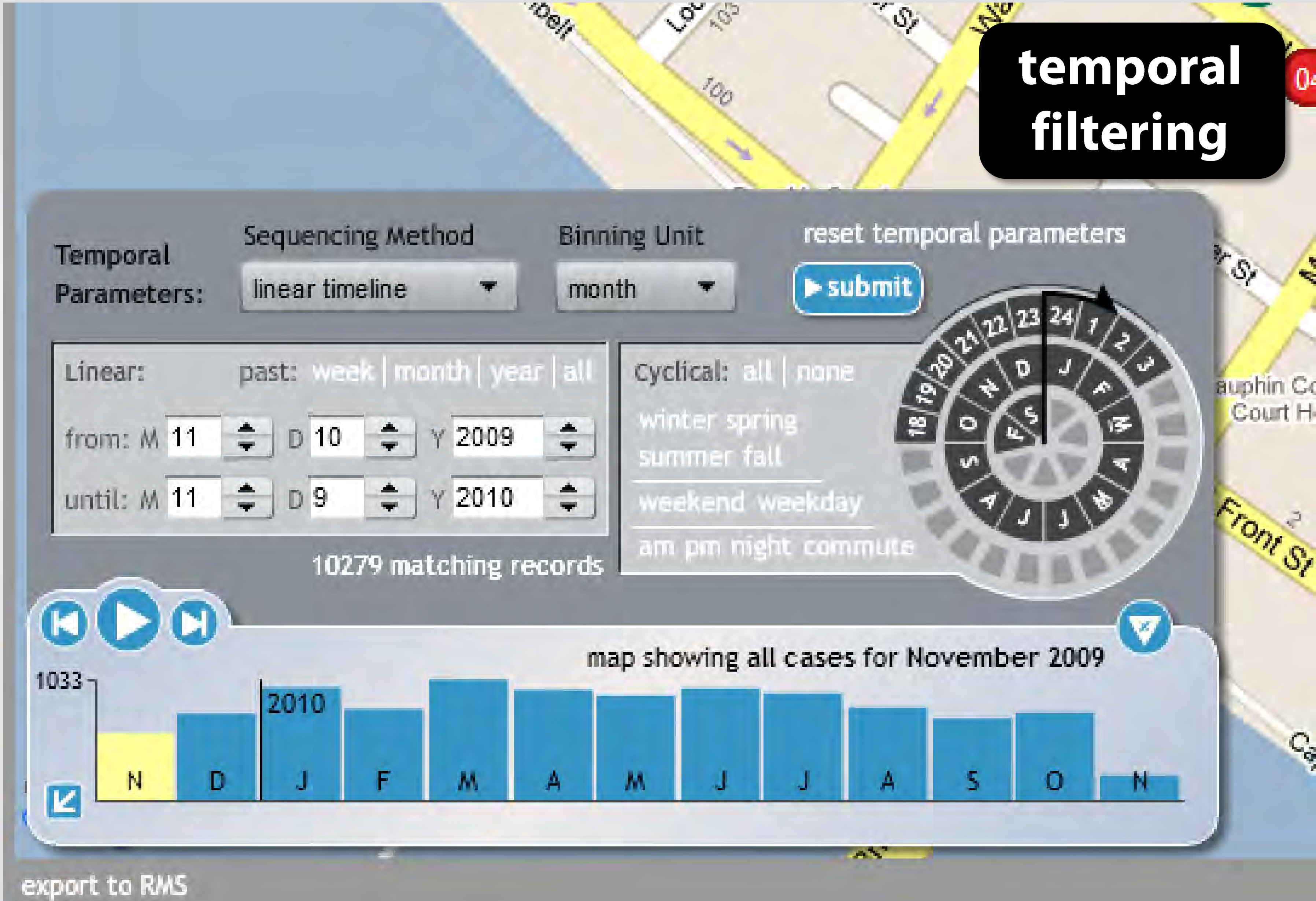
Los Angeles, CA | April 12th, 2013

@RobertERoth **@alanGeoVISTA**

Robert E. Roth *Alan M. MacEachren*



+ – | map sat terrain | reset extent
rrisburg
Amtrak
Aberdeen St
Unique Search: report # | address
enter report #
submit
Basic Filtering: 1470 matching records
UCR Primary
04 aggravated assault
UCR Secondary
04 all
04 all
0410 gun
0411 gun - upon police officer
0420 knife or cutting instrument
0421 knife or cutting instrument - upon po
incident filtering
Pennsylvania


temporal filtering
Temporal Parameters:
Sequencing Method
linear timeline
Binning Unit
month
reset temporal parameters
submit
Linear:
past: week | month | year | all
from: M 11 D 10 Y 2009
until: M 11 D 9 Y 2010
10279 matching records
Cyclical: all | none
winter spring summer fall
weekend weekday
am pm night commute
map showing all cases for November 2009
1033
2010
N D J F M A M J J A S O N
export to RMS

METHOD: INTERACTION STUDY

identifying prototypically successful interaction strategies

Participants:

- n=**10**
- law enforcement personnel at Harrisburg Bureau of Police

Interaction Study Protocol:

- **15** tasks: **5** objectives x **3** operands
- closed-ended: questions had **"correct"** answers
- **3** minute time limit per question

Interaction Analysis

- interaction logging by operator

Operator Logging

Filter (F)

Space	*Menu Selection Numerical Stepper by 'District'*
	Form Fill-in by 'District'
	Menu Selection for 'Reset Advanced Features'
Attribute	*Menu Selection by 'UCR Primary,' 'UCR Secondary' and 'MO'*
	Form Fill-in by 'UCR Primary,' 'UCR Secondary' and 'MO'
	Menu Selection for 'Reset Basic Filters'
Time	*Menu Selection Numerical Stepper for 'From' and 'To' Linear Filtering*
	Form Fill-in 'From' and 'To' Linear Filtering
	Menu Selection Shortcuts for Linear Filtering ('Week', 'Month', 'Year', 'All')
	Direct Manipulation of 'Hours', 'Months', and 'Days' Widgets for Cyclical
	Menu Selection Shortcuts for Cyclical Filtering ('All', 'None', 'Winter', etc.)
	Menu Selection for 'Reset Temporal Parameters'

Search (S)

Space	*Form Fill-in Search by 'Address'*
Attribute	*Form Fill-in Search by 'Report #'*

Retrieve (R)

Attribute	*Direct Manipulation Mouse-Over of Hexagon Bin*
	Direct Manipulation Mouse-Over of Crime Incident
	Direct Manipulation Click of Crime Incident
	Menu Selection to Activate Street View
	Direct Manipulation Mouse-Over of Data Layer Element
Time	*Direct Manipulation Mouse-Over of Histogram Bin*

Operator Logging

Reexpress (X)	
Time	*Menu Selection for Linear versus Composite Time*
Sequence (Q)	
Time	*Direct Manipulation Click of the 'Play' (Loop) and 'Pause' VCR Controls*
Overlay (O)	
Space	*Menu Selection for Basemap Type ('Map', 'Sat', 'Terrain')*
Attribute	*Menu Selection Checkboxes for Point/Line Data Layers ('Schools', etc.)*
	Menu Selection Radio Buttons for Polygonal Data Layers ('Districts', etc.)
	Menu Selection 'Reset' Additional Data Layers
Pan (P)	
Space	*Direct Manipulation Click+Drag on Map*
	Direct Manipulation 'Resent Extent' Control
Time	*Direct Manipulation Click on Histogram Bin*
	Direct Manipulation Click on 'Back' and 'Step' VCR Controls
	Direct Manipulation of Histogram Slider Bar (When Entirety is Not Displayed)
Zoom (Z)	
Space	*Direct Manipulation Double-Click on Map*
	Direct Manipulation Click on Hexagon Grid
	Direct Manipulation Click on Data Layer Element
	Direct Manipulation '+' and '-' Controls
	Direct Manipulation 'Resent Extent' Control
Time	*Menu Selection for Binning Unit*

Task	Complete	Correct	Avg Time	Freq.	Diversity
By Objective					
Identify	30 (100%)	26 (87%)	0:31	3.1	2.1
Compare	30 (100%)	25 (83%)	0:58	7.4	2.7
Rank	29 (96.7%)	28 (93%)	0:56	9.1	4.3
Associate	30 (100%)	24 (80%)	1:06	7.6	3.2
Delineate	27 (90%)	20 (67%)	1:27	9.8	4.6
By Operand					
Space	50 (100%)	47 (94%)	0:34	5.9	2.9
Attribute	46 (92%)	41 (82%)	1:16	8.1	3.5
Time	50 (100%)	35 (70%)	1:09	8.2	3.7
Total	**146 (97%)**	**123 (82%)**	**1:00**	**7.4**	**3.4**

@RobertERoth *Robert E. Roth*
@alanGeoVISTA *Alan M. MacEachren*

Task	Complete	Correct	Avg Time	Freq.	Diversity
By Objective					
Identify	30 (100%)	26 (87%)	**0:31**	3.1	2.1
Compare	30 (100%)	25 (83%)	**0:58**	7.4	2.7
Rank	29 (96.7%)	28 (93%)	**0:56**	9.1	4.3
Associate	30 (100%)	24 (80%)	**1:06**	7.6	3.2
Delineate	27 (90%)	20 (67%)	**1:27**	9.8	4.6
By Operand					
Space	50 (100%)	47 (94%)	0:34	5.9	2.9
Attribute	46 (92%)	41 (82%)	1:16	8.1	3.5
Time	50 (100%)	35 (70%)	1:09	8.2	3.7
Total	**146 (97%)**	**123 (82%)**	**1:00**	**7.4**	**3.4**

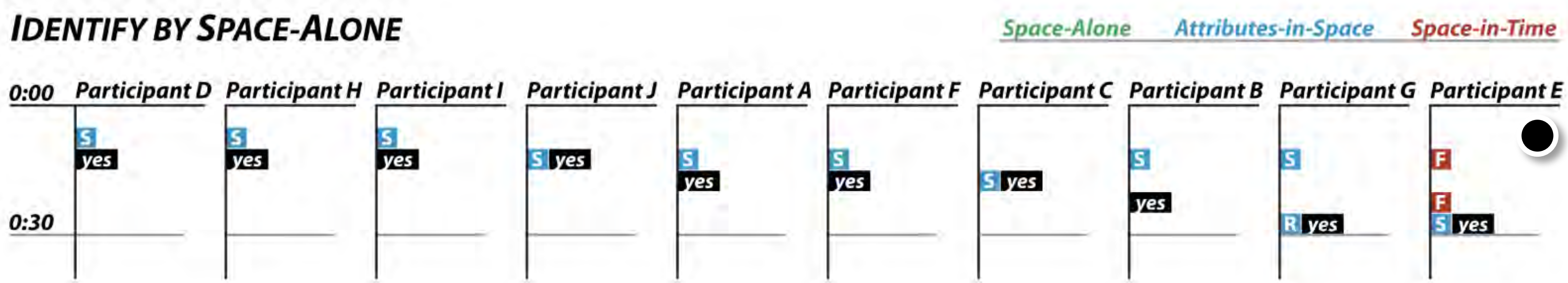
Task	Complete	Correct	Avg Time	Freq.	Diversity
By Objective					
Identify	30 (100%)	26 (87%)	0:31	3.1	2.1
Compare	30 (100%)	25 (83%)	0:58	7.4	2.7
Rank	29 (96.7%)	28 (93%)	0:56	9.1	4.3
Associate	30 (100%)	24 (80%)	1:06	7.6	3.2
Delineate	27 (90%)	20 (67%)	1:27	9.8	4.6
By Operand					
Space	50 (100%)	**47 (94%)**	0:34	5.9	2.9
Attribute	46 (92%)	**41 (82%)**	1:16	8.1	3.5
Time	50 (100%)	**35 (70%)**	1:09	8.2	3.7
Total	**146 (97%)**	**123 (82%)**	**1:00**	**7.4**	**3.4**

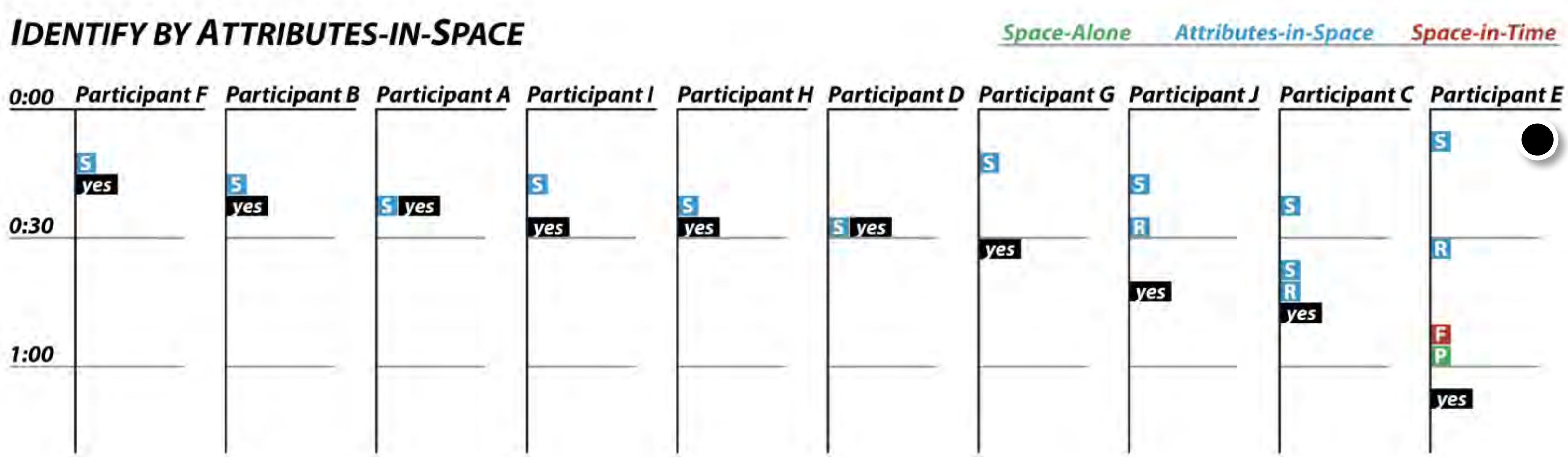
Task		Frequency	Extensiveness
By Operator			
	Reexpress (X)	78	50 (33%)
	Sequence (Q)	32	14 (9%)
	Overlay (O)	57	46 (31%)
	Pan (P)	121	41 (27%)
	Zoom (Z)	**127**	**66 (44%)**
	Filter (F)	**240**	**90 (60%)**
	Search (S)	42	29 (19%)
	Retrieve (R)	**395**	**106 (71%)**
By Operand			
	Space	175	49 (33%)
	Attribute	248	117 (78%)
	Time	543	94 (63%)
Total		**1092**	**150 (100%)**

Task	Frequency	Extensiveness
By Operator		
Reexpress (X)	78	50 (33%)
Sequence (Q)	32	14 (9%)
Overlay (O)	57	46 (31%)
Pan (P)	121	41 (27%)
Zoom (Z)	127	66 (44%)
Filter (F)	240	90 (60%)
Search (S)	42	29 (19%)
Retrieve (R)	395	106 (71%)
By Operand		
Space	**175**	49 (33%)
Attribute	**248**	117 (78%)
Time	**543**	94 (63%)
Total	**1092**	**150 (100%)**

RESULTS: SUBOPTIMAL USE PROFILES

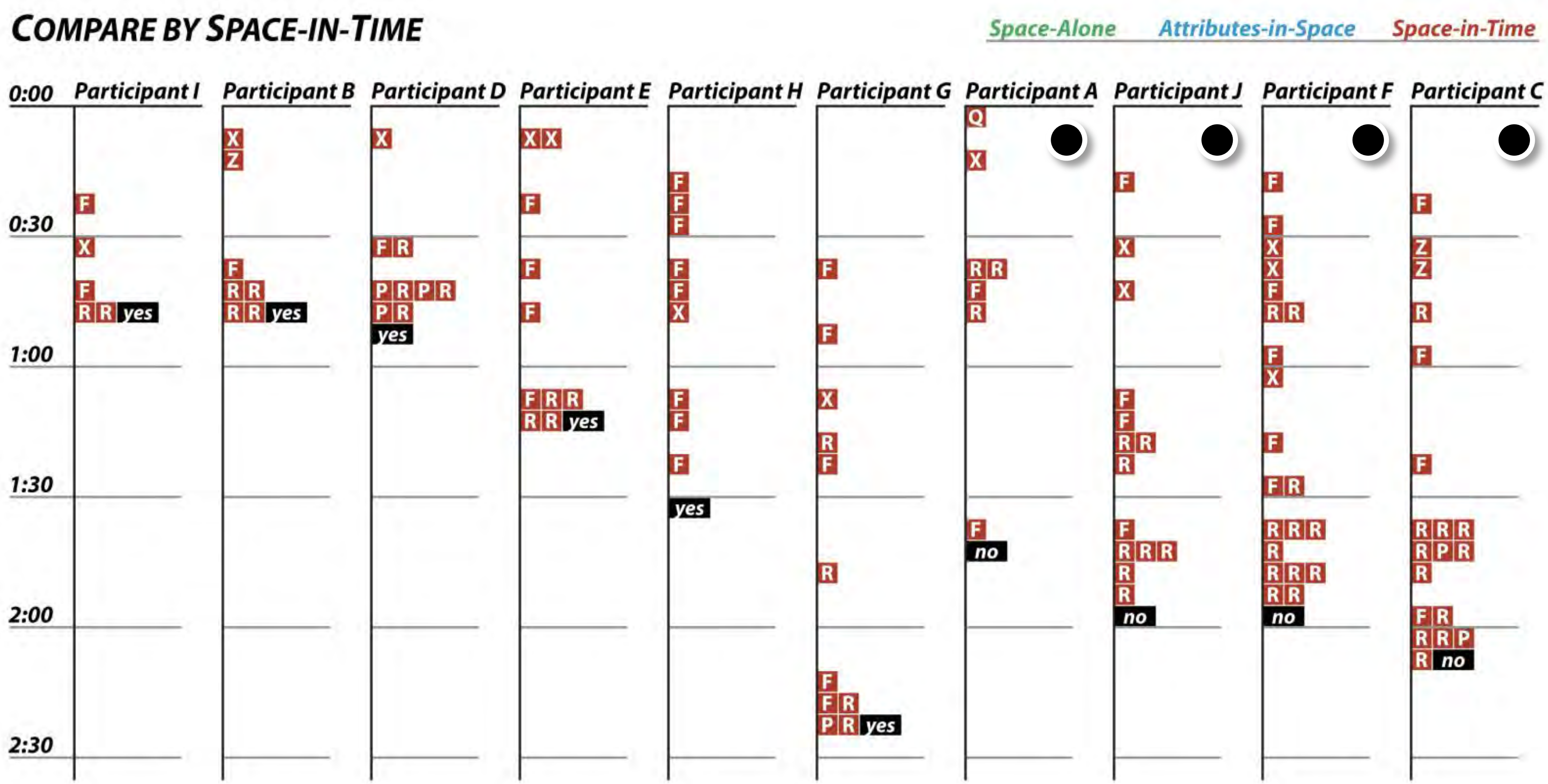
Persona: Excessive-Filterer; *search* vs. *filter* by sophistication





Results: Suboptimal Use Profiles

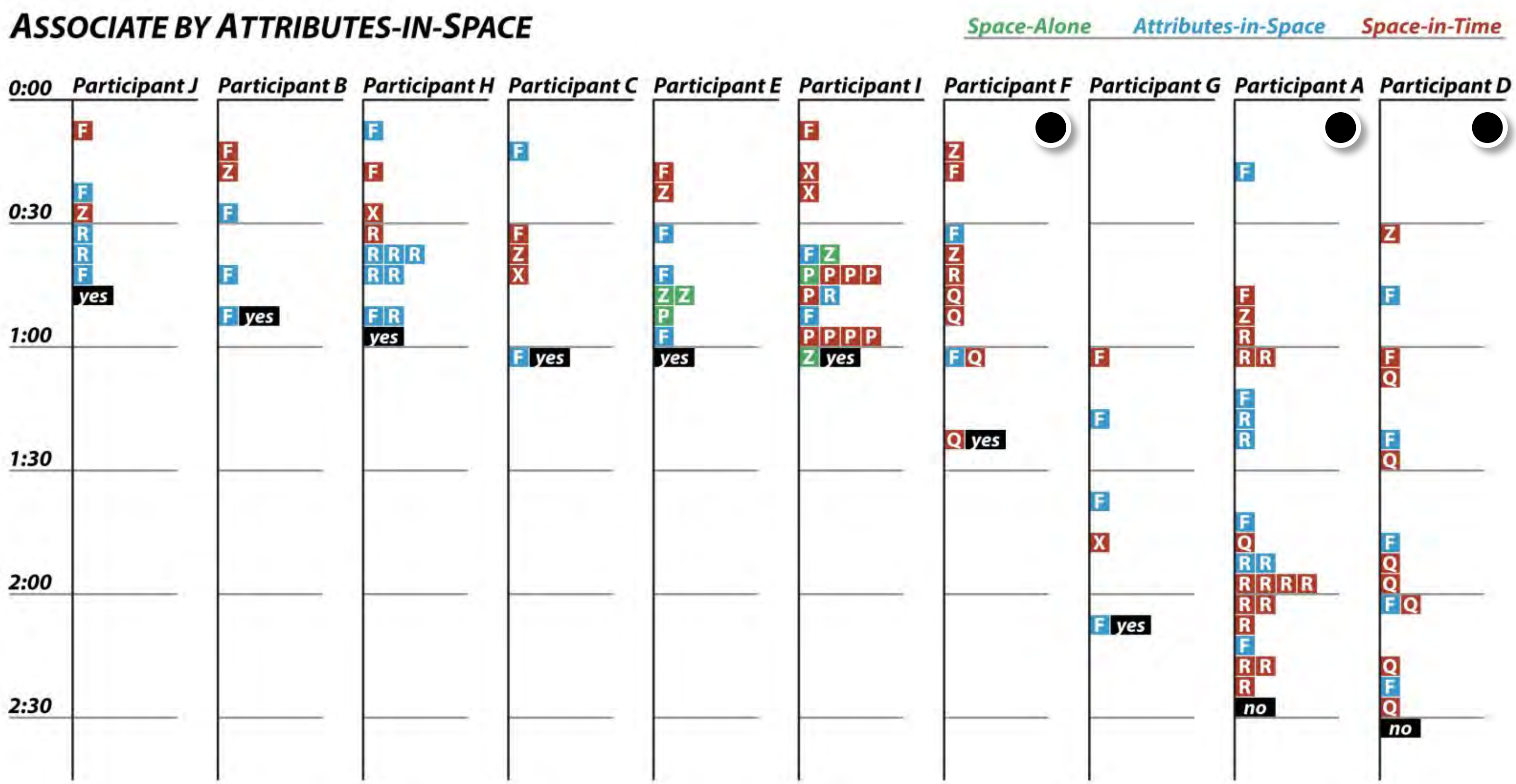
Persona: Mistaken-Reexpresser, Unsure-Retriever



@RobertERoth @alanGeoVISTA
Robert E. Roth Alan M. MacEachren

RESULTS: SUBOPTIMAL USE PROFILES

Persona: Blind-Sequencer



QUESTIONS?

http://www.slideshare.net/reroth

Design and Use Guidelines for Interactive Maps: A Case Study

@RobertERoth

Robert E. Roth | reroth@wisc.edu

@alanGeoVISTA

Alan M. MacEachren | maceachren@psu.edu

#aag2013

Los Angeles, CA | April 12th, 2013

THE UNIVERSITY of WISCONSIN MADISON

GeoVISTA Center

Table of Contents

Task (Objective+Operand) Protocol

Identify

Space	*On what street did incident #20101100945 occur?*
Attribute	*What type of crime is incident #20101100894, which occurred at 200 Herr Street?*
Time	*How many total crime incidents occurred in District #5 on September 1st, 2010?*

Compare

Space	*Are Fire Station #2 and Fire Station #8 in the same police district?*
Attribute	*Is incident #20101100945, which occurred on Market Street, the same type of crime as incident #20101100608, which occurred on 3rd Street?*
Time	*In October 2010, how many more crime incidents occurred within Harrisburg on Sundays compared to Mondays?*

Rank

Space	*What school in Harrisburg is closest to Interstate-83?*
Attribute	*Which crime type was the most common in District #1 on November 5th, 2010?*
Time	*From 2006 through 2010, which month exhibited the highest frequency of crime incidents across Harrisburg?*

Associate

Space	*Which route should Harrisburg citizens take to get to the west bank of the Susquehanna River during an evacuation related to Three Mile Island?*
Attribute	*From 2006 through 2010, is the geographic pattern of prostitution (16) related to the geographic pattern of sex offenses (17)?*
Time	*From 2006 through 2010, does the trend in crime increase or decrease across Harrisburg from noon (12:00) to midnight (24:00)?*

Delineate

Space	*Which police districts exhibit clusters of increased criminal activity from 2006 through 2010?*
Attribute	*From 2006 through 2010 , how many different ways was fraud (11) committed across Harrisburg?*
Time	*The 2008 spike in robbery (03) incidents in Harrisburg spanned across which months?*

Table of Contents

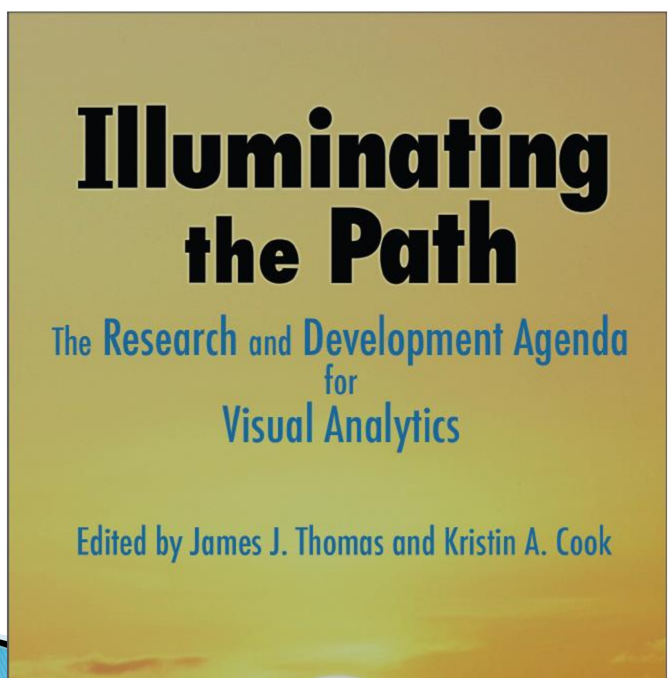
GeoVISTA Center

Geo-Social Visual Analytics with Applications to Catastrophic Risk Management

Wei Luo
Ph.D Candidate
GeoVISTA Center, Department of Geography
Penn State University

Global Catastrophic Risk INSTITUTE

GeoVISTA Center



GeoVISTA Center

Visual Analytics

- Visual analytics is defined as the science of analytical reasoning facilitated by interactive visual interfaces (Thomas and Cook 2005).
- Objective: Reducing America's vulnerability to terrorism and natural disasters.

GeoVISTA Center

Future Visual Analytics

- http://www.youtube.com/watch?v=5tfnmhl-A54&feature=player_embedded

GeoVISTA Center

Introduction

- Geographical context and social network context demonstrate strong conceptual and observed overlaps, but the integration of geography and social network contexts has not received sufficient attention to develop a comprehensive understanding of their interaction or their impact on outcomes of interest.

GeoVISTA Center

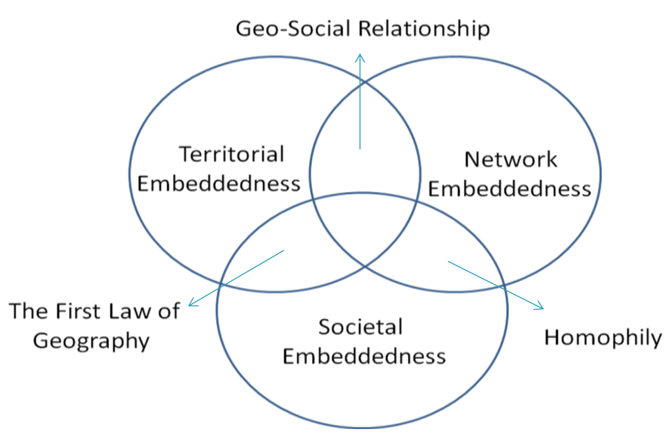
Goal

- This research develops theoretical and practical models of geosocial context through geovisual analytics approaches that account for geographical, social network, and conceptual dimensions of that context.

Objectives

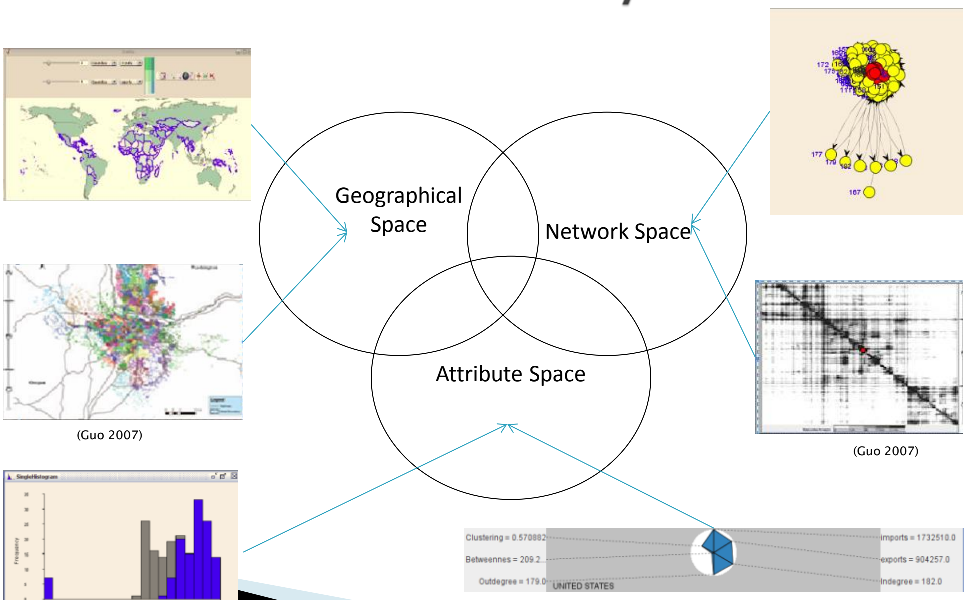
- Develop a conceptual model to understand the interaction of geography and social network at the theoretical level.
- Develop a geosocial visual analytics tool at a geographical region level (e.g., county, state, and country) to understand the role of spatial proximity in shaping the international trade network (ITN) across different geographical regions.
- Develop a geosocial visual analytics tool to support decision-making processes in airborne disease control (e.g., influenza) with agent-based epidemic models.

GeoSocial Conceptual Framework



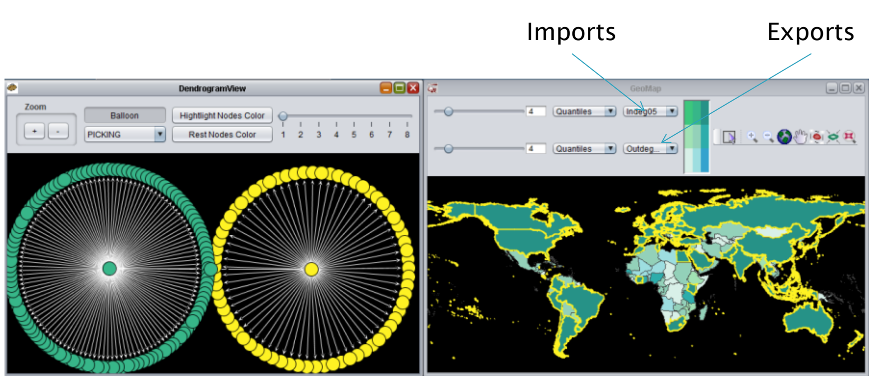
Based on fundamental categories of embededness (Hess 2004)

GeoSocial Visual Analytics



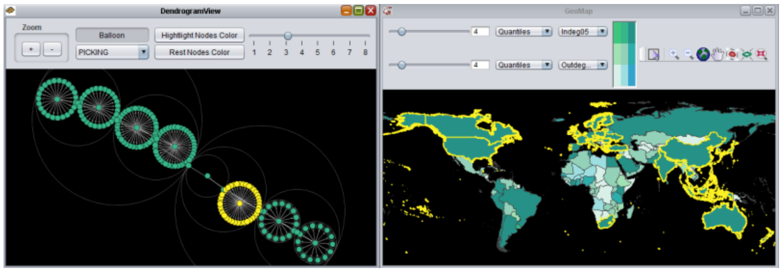
http://www.geovista.psu.edu/GeoSocialApp/

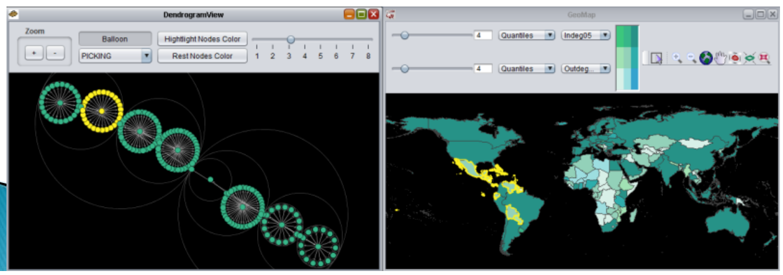
GeoSocialApp with International Trade Network in 2005



The convergence of the iterated correlations (CONCOR) algorithm (Breiger et al. 1975).

Selected Results

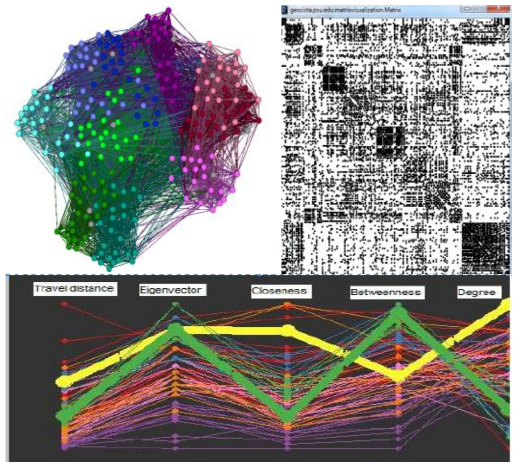




Potential Applications

- Financial contagion on the international trade network (Kali and Reyes 2010).
- The impact of climate change on human migrants within the U.S.

GeoSocial Agent-Based Epidemic Visual Analytics



Expected Insights

- Does the proposed multi-scale geovisual analytics approach enable new control scenarios compared to agent-based epidemic models alone?
- Does the distribution of high-risk individuals enable new control scenarios?
- Does the visualization of disease transmission simulation process enable new control scenarios ?
- To achieve three goals, I will do a task analysis with and without the tools and follow up with post-task interviews of the expert participants.

GeoVISTA Center

Contributions

- The proposed conceptually geo-social model provides a theoretical framework to study geo-social relationships.
- The second case study transforms identified patterns into useful knowledge to directly support decision making in epidemics. This contribution fills one major gap in visual analytics: how to transform knowledge developed through visual exploration into decision-making support directly.
- Geosocial visual analytics presented here are applicable to other geographical and network contexts among geographical regions (e.g., human migrants among different states in the U.S.), as well as at the individual level (e.g., facebook and twitter network over space).

Questions ?

Table of Contents

References

- Adams, j., K. Faust & G. S. Lovasi (2012) Capturing context: Integrating spatial and social network analyses. *Social networks, 34, 1-5.*
- Guo, D. (2007) Visual analytics of spatial interaction patterns for pandemic decision support. *International Journal of Geographical Information Science, 21, 859-877.*
- Hess, M. (2004) "Spatial" relationships? Towards a reconceptualization of embeddedness. *Progress in human geography, 28, 165-186.*
- Kali, R. & J. Reyes (2010) Financial contagion on the international trade network. *Economic Inquiry,* 48**,** 1072-1101.
- Luo, W., A. M. MacEachren, P. Yin & F. Hardisty. 2011. Spatial-Social Network Visualization for Exploratory Data Analysis. In *SIGSPATIAL International Workshop on Location-Based Social Networks (LBSN 2011) in conjunction with the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS) 2011. 1-4 November, 2011, Chicago, Illinois.: ACM.*
- McPherson, M., L. Smith-Lovin and J. M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology*: 415-444.
- Thomas, J. & K. Cook. 2005. *Illuminating the path: The research and development agenda for visual analytics*. IEEE Computer Society.
- Tobler, W. (1970) A computer movie simulating urban growth in the Detroit region. *Economic geography, 234-240.*

Table of Contents

Geovisualization to Geovisual Analytics: Visual Reasoning with Big & Messy Data

Alan M. MacEachren

GeoVISTA Center
Department of Geography
Penn State University
maceachren@psu.edu

---

from cartography, through geovisualization, to geovisual analytics:
data → info → knowledge → knowledge application

- (late 1980s) **Visualization in Scientific Computing (VisSC)**: prompting a shift from **cartography → geovisualization** –

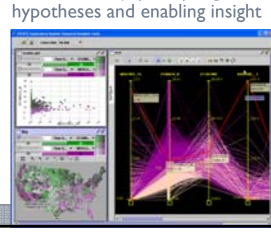
Outline

- Geovisualization and the roots of Geovisual Analytics:
  - *cartography → geovisualization → geovisualization + KDD*
  - sample applications
- → Geovisual analytics:
  - *Definition, Drivers, Conceptual frameworks*
  - *Challenges*
    - scaling humans to cope with big data
    - supporting analytical reasoning to generate and apply the information
    - extracting meaningful information from messy, heterogeneous data
  - *2 Case study projects: info foraging and contextualization, sensemaking and situational awareness*
- *Summary & GIScience challenges*

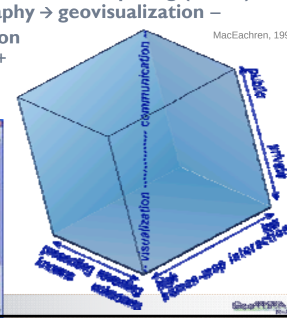
---

from cartography, through geovisualization, to geovisual analytics:
data → info → knowledge → knowledge application

- (late 1980s) **Visualization in Scientific Computing (VisSC)**: prompting a shift from **cartography → geovisualization** –
- (1990s) dynamic **geovisualization** – (Cartography + VisSC + EDA + HCI + InfoVis): prompting hypotheses and enabling insight

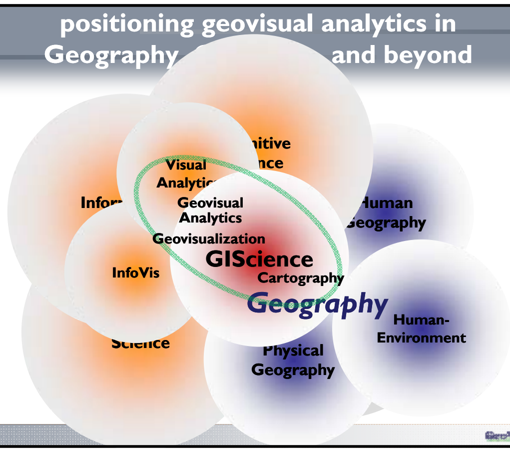


MacEachren, 1994



---

positioning geovisual analytics in Geography and beyond

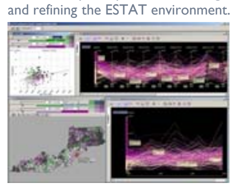


Exploratory Spatial-Temporal Analysis Toolkit (ESTAT): user-centered design / usability engineering

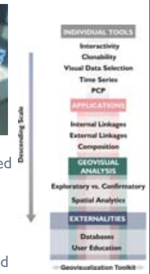
Robinson, A. C. 2007. 'A design framework for exploratory geovisualization in epidemiology', Information Visualization, *6, pp. 197-214.*

ESTAT, (exploratory spatio-temporal analysis toolkit) is designed to facilitate visual exploration for cancer research.

18 epidemiologists at NCI and Penn State HMC participated in assessing and refining the ESTAT environment.



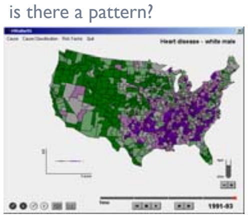
Evaluating ESTAT has involved the following knowledge elicitation techniques: focus groups; card-sorting; verbal protocol analysis; case study collaboration … and resulted in a hierarchical approach to user-centered design for exploratory geovis in epi



Lengerich, E., Meyer, H., & MacEachren, A. 2005, Combining Usability Techniques to Design Geovisualization Tools for Epidemiology. Cartography and GIScience, 32(4), 243-255.

Dynamic representation = user and data-driven change

is there a pattern?



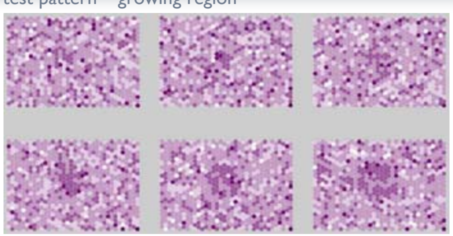
is your answer different?

MacEachren, A., Boscoe, F., Haug, D., & Pickle, L. 1998. Geographic Visualization: Designing Manipulable Maps for Exploring Temporally Varying Georeferenced Statistics. Proc., InfoVis '98, Raleigh-Durham, NC, Oct. 19-20, 1998, pp. 87-94.

GeoViz Toolkit: Flu Data Analysis

Animation: Fundamental issues in visual cognition

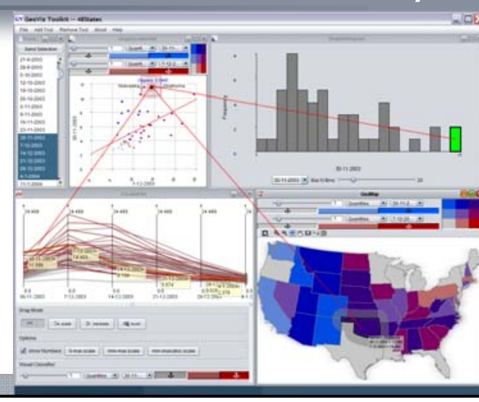
pace: test pattern – growing region



- map readers answer more quickly and identify more patterns correctly when using animated maps than when using static small-multiple maps
- pace and cluster coherence interact so that different paces are more effective for identifying certain types of clusters (none vs. subtle vs. strong)
- there are some gender differences in the animated condition (males faster)

Griffin, A.L., MacEachren, A.M., Hardisty, F., Steiner, E. and Li, B., 2006, A comparison of animated with static small multiple maps for visually identifying space-time clusters. *Annals of the Association of American Geographers, 96, pp. 740-753.*

GeoViz Toolkit: Flu Data Analysis



GeoViz Toolkit: Flu Data Analysis

Frank Hardisty
hardisty@psu.edu

http://www.geovista.psu.edu/grants/cdcesda/software/
http://code.google.com/p/geoviz/

http://www.google.org/about/flutrends/download.html

Google Flu Trends for the U.S.: Historic Seasons Copyright 2008 Google Inc.

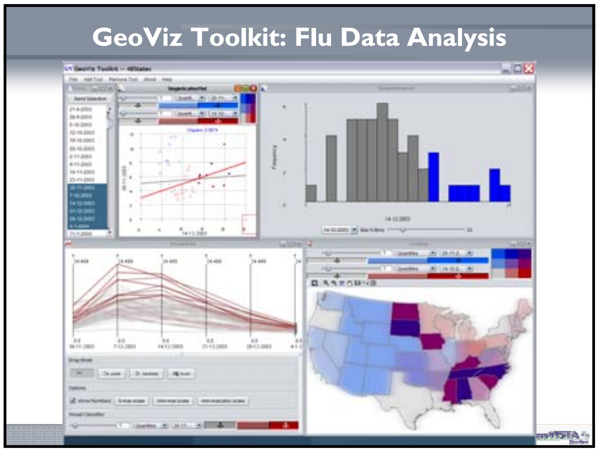
Use subject to Terms of Service (http://www.google.com/accounts/TOS)

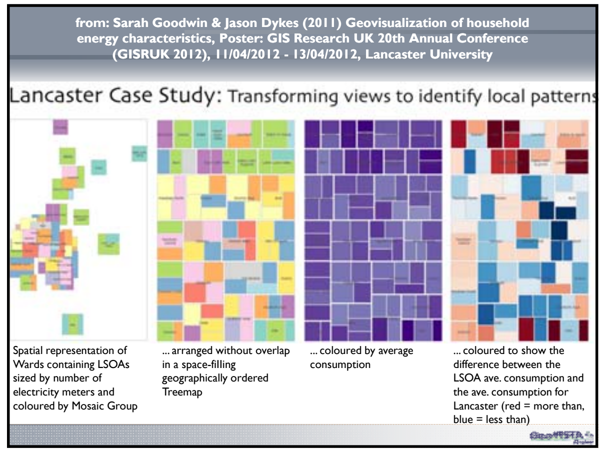
United States Weekly ILI Percentages, 2003-2004 through 2007-2008

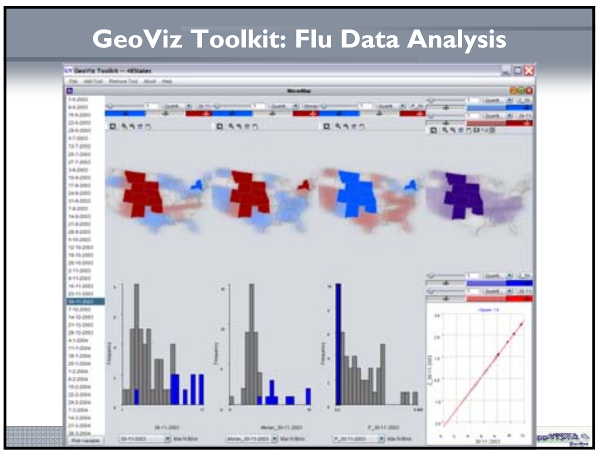
Each week begins on the Sunday indicated for the row

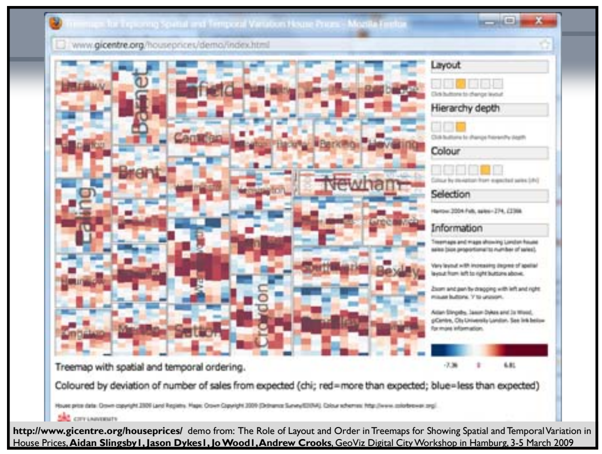
For more information, please visit http://google.org/flutrends

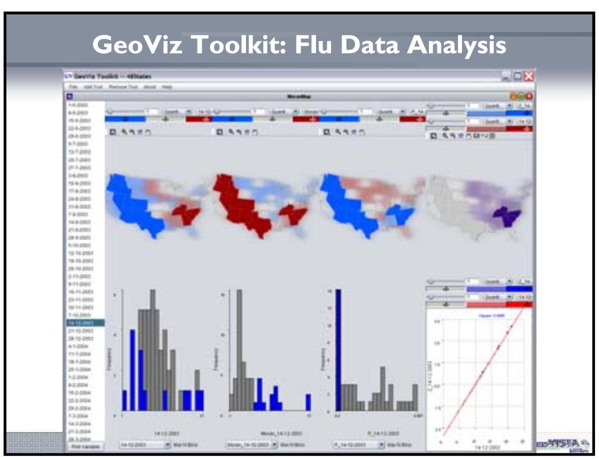
GeoViz Toolkit: Flu Data Analysis

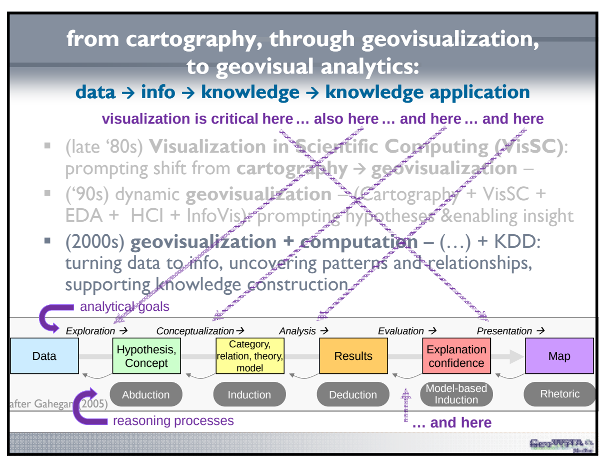








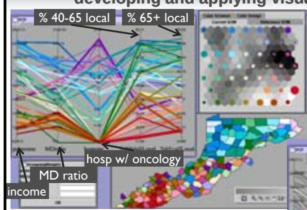




Author: Alan M. MacEachren – Please do not post online

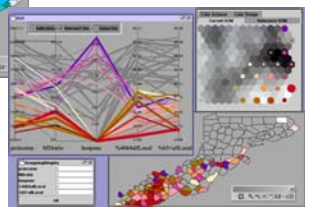
Multivariate patterns related to undesirable cancer control situations

developing and applying visual analytic methods and tools



Self-organized map based computational categorization of cancer + covariates

- more similar colors represent more similar places for cancer and covariates
- "local" for cancer diagnosis is better = cancer identified early (localized in body)



Counties grouped due to (a) well below average % of cancer cases diagnosed at "local" stage for two age groups: 40-64 and 65+ and (b) low income, low rate of MDs, and low # of hospitals with oncology service

Guo, D., Gahegan, M., MacEachren, A., Zhou, B. (2005) Multivariate Analysis and Geovisualization with an Integrated Geographic Knowledge Discovery Approach. Cartography & GIS, 32(2): 113-132

Geovisual analytics: 3 challenges

- *scaling humans to cope with big data*
- *extracting meaningful information from messy, heterogeneous data*
- *understanding and supporting the analytical reasoning processes needed to utilize the data*

Georg Gartner (Tues. keynote): maps "enable spatial problem solving (reasoning, planning)"

from cartography, through geovisualization, to geovisual analytics:

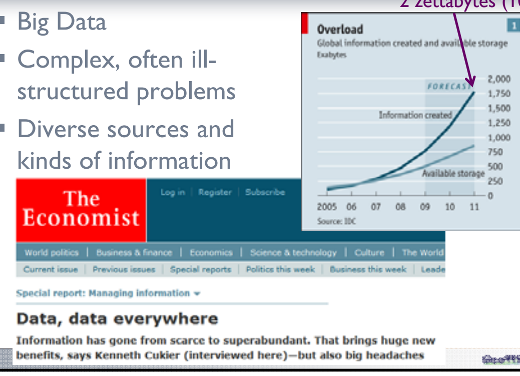
data → info → knowledge → knowledge application

- (late '80s) **Visualization in Scientific Computing (VisSC)**: prompting shift from **cartography → geovisualization** –
- ('90s) dynamic **geovisualization** – (Cartography + VisSC + EDA + HCI + InfoVis): prompting hypotheses &enabling insight
- (2000s) **geovisualization + computation** – (…) + KDD: turning data to info, uncovering patterns and relationships, supporting knowledge construction
- **(geo)visual analytics →**

Some scientific and social challenges

- Big Data
- Complex, often ill-structured problems
- Diverse sources and kinds of information

2 zettabytes (10^{21})

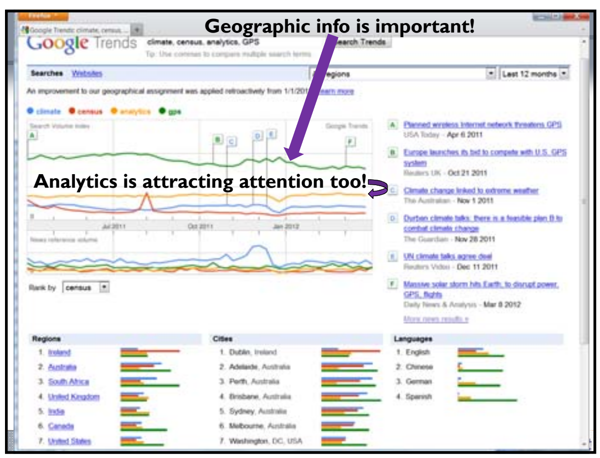


Geovisual Analytics

- *Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces*

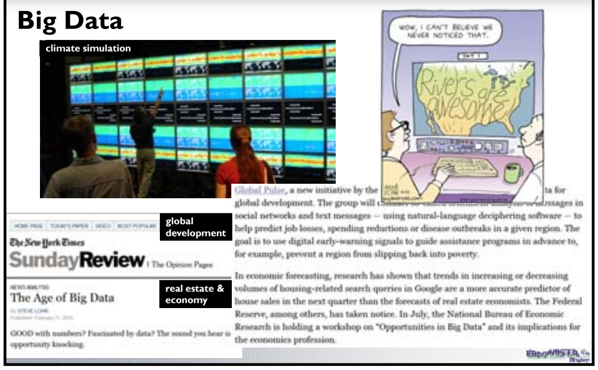
  … thus, the focus is on assembling evidence, generating inferences and explanations from evidence, comparing / assessing those inferences and explanations, and reporting results

- ***Geo**visual analytics focuses on visual interfaces to analytical methods that support *reasoning* with and about geo-information – *to enable insights about something for which place matters**



Geovisual analytics: scaling analysts *to cope with information deluge*

Big Data



NSF IGERT: Big Data Social Science

- PI: Burt Monroe (Political Science)
- CoPIs:
  - Alan MacEachren (Geography)
  - Lee Giles (IST & CS&E)
  - Aleksandra B. Slavkovic (Stats & Public Health Sci)
  - Melissa Hardy (Sociology)

from: Gennady Andrienko & Natalia Andrienko, Visual Analytics of Movement, tutorials at MODAP workshop (Istanbul, June 2011)

Spatio-Temporal Big Data Analysis

Example of visual analytics: analysis of city traffic

E.g. Milan car movement data: 2,075,216 position records of 17,241 cars during 1 week

This is far too much for processing in RAM!

Approach:

1) aggregate the data in the database using the standard DBMS operations;
2) load the aggregates in RAM and visualise for enabling analysis

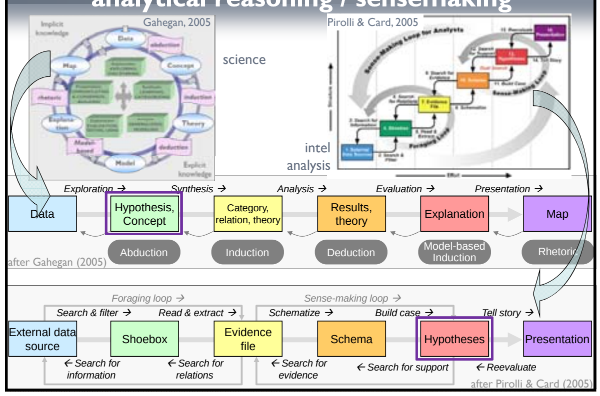
Spatial aggregation: by compartments (cells) of a territory division

Temporal aggregation: by time intervals
- linear time
- cyclic time

see also: Andrienko, G., Andrienko, N., Hurter, C., Rinzivillo, S. and Wrobel, S. 2011: From movement tracks through events to places: Extracting and characterizing significant places from mobility data. *IEEE VAST, 2011, 161-170.*

Fraunhofer

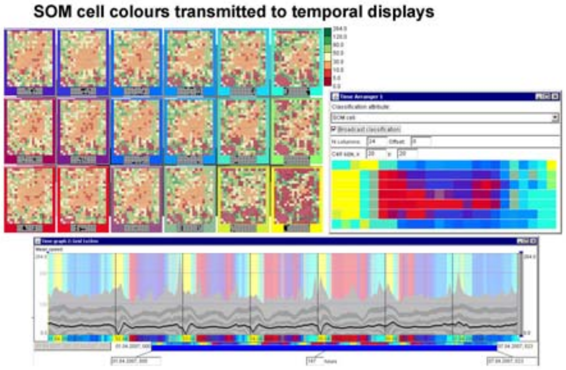
Geovisual analytics: supporting analytical reasoning / sensemaking



from: Gennady Andrienko & Natalia Andrienko, Visual Analytics of Movement, tutorials at MODAP workshop (Istanbul, June 2011)

Spatio-Temporal Big Data Analysis

SOM cell colours transmitted to temporal displays



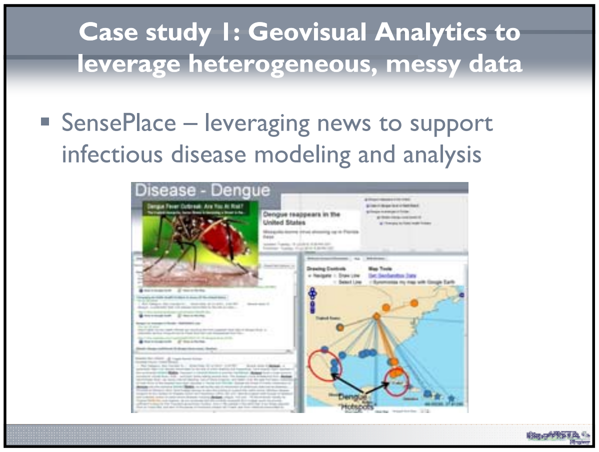
Information Processing Challenges for Sensemaking (reasoning) & Decision-making

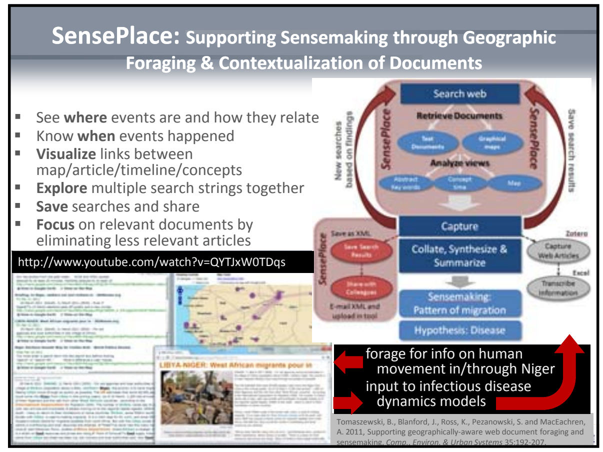
– from Zack, 2007 as adapted by Muhren & van de Walle, 2010 –

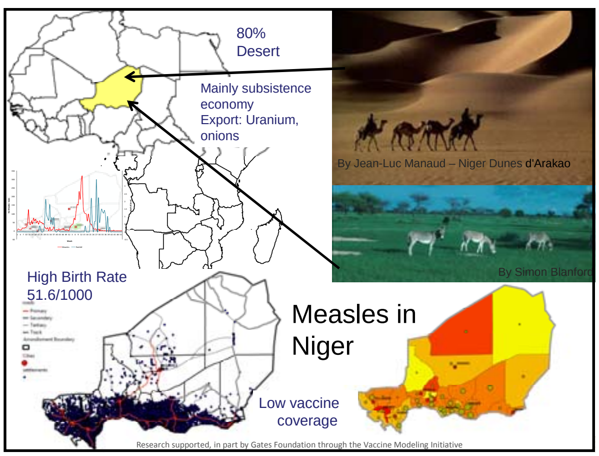
what is processed / processing problem	Information	Frame(s) of Reference
Lack of variety	uncertainty	ambiguity
Diversity of …	complexity	equivocation

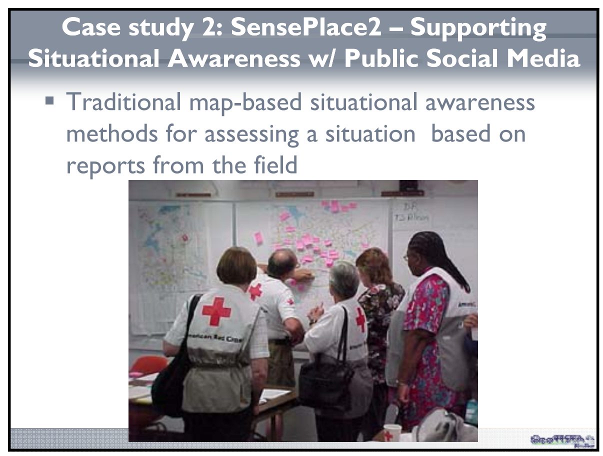
- **uncertainty**: not enough info to perform task – requires tools for *information foraging*
- **ambiguity**: lack of framework to interpret – requires tools to *build context*
- **complexity**: more info than can be processed – requires tools for *information filtering, fusion, synthesis*
- **equivocation**: competing/contradictory frameworks – requires tools for *re-framing, analysis of competing hypotheses, deliberation, negotiation, etc.*

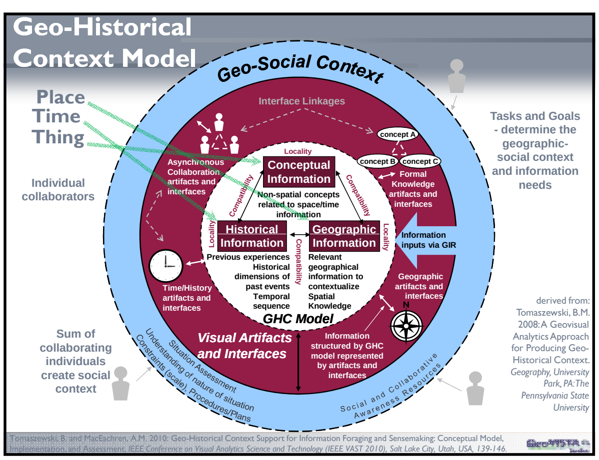
Muhren, W. and Van de Walle, B. 2010: A Call for Sensemaking Support Systems in Crisis Management. *Interactive Collaborative Information Systems, Berlin: Springer, 425-452.*
Zack, M.H. 2007: The role of decision support systems in an indeterminate world. Decision Support Systems 43, 1664-1674.

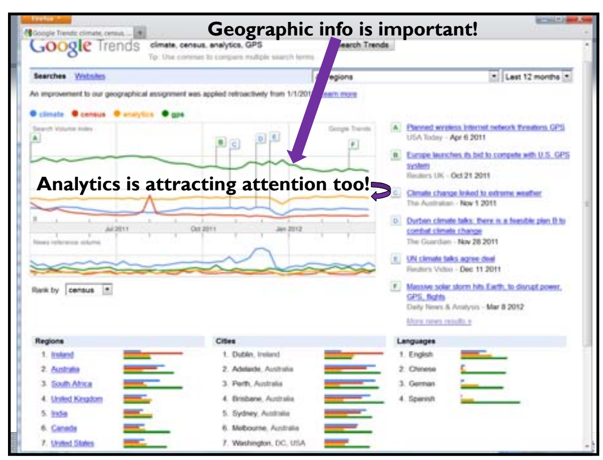

Case study 1: Geovisual Analytics to leverage heterogeneous, messy data
SensePlace – leveraging news to support infectious disease modeling and analysis
Disease - Dengue
Dengue Hotspots


SensePlace: Supporting Sensemaking through Geographic Foraging & Contextualization of Documents
See where events are and how they relate
Know when events happened
Visualize links between map/article/timeline/concepts
Explore multiple search strings together
Save searches and share
Focus on relevant documents by eliminating less relevant articles
http://www.youtube.com/watch?v=QYTJxW0TDqs
Search web
Retrieve Documents
Analyze views
Capture
Collate, Synthesize & Summarize
Sensemaking: Pattern of migration
Hypothesis: Disease
Save search results
New searches based on findings
forage for info on human movement in/through Niger
input to infectious disease dynamics models
Tomaszewski, B., Blanford, J., Ross, K., Pezanowski, S. and MacEachren, A. 2011, Supporting geographically-aware web document foraging and sensemaking. Comp., Environ. & Urban Systems 35:192-207


80% Desert
Mainly subsistence economy
Export: Uranium, onions
By Jean-Luc Manaud – Niger Dunes d'Arakao
By Simon Blanford
High Birth Rate 51.6/1000
Measles in Niger
Low vaccine coverage
Research supported, in part by Gates Foundation through the Vaccine Modeling Initiative


Case study 2: SensePlace2 – Supporting Situational Awareness w/ Public Social Media
Traditional map-based situational awareness methods for assessing a situation based on reports from the field


Geo-Historical Context Model
Geo-Social Context
Place
Time
Thing
Individual collaborators
Tasks and Goals - determine the geographic-social context and information needs
Conceptual Information
Historical Information
Geographic Information
GHC Model
Visual Artifacts and Interfaces
Sum of collaborating individuals create social context
derived from: Tomaszewski, B.M. 2008: A Geovisual Analytics Approach for Producing Geo-Historical Context. Geography, University Park, PA: The Pennsylvania State University
Tomaszewski, B. and MacEachren, A.M. 2010: Geo-Historical Context Support for Information Foraging and Sensemaking: Conceptual Model, Implementation, and Assessment. IEEE Conference on Visual Analytics Science and Technology (IEEE VAST 2010), Salt Lake City, Utah, USA, 139-146.

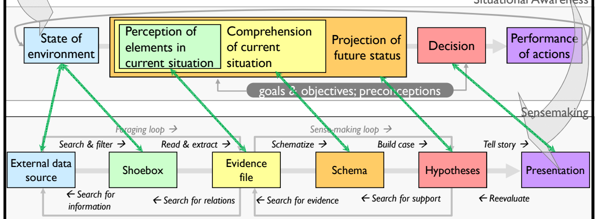

Geographic info is important!
Google Trends
Analytics is attracting attention too!

But, social media is more important!

Situational Awareness ←→ Sensemaking

situational awareness as a sensemaking process

- Situational awareness model (after Endsley, 1995; top) compared to model of sense making (after Pirolli & Card, 2005; bottom)



Potential of social media for SA

Heading off disaster, one tweet at a time

By Jim Spellman, CNN
September 22, 2010 1:13 p.m. EDT | Filed under: Social Media

The Boulder, Colorado, wildfire triggered a strong response from iReport, often CNN's first source for disaster images.

Denver, Colorado (CNN) -- When word came of a huge explosion in San Bruno, California, FEMA chief Craig Fugate grabbed his cell phone and logged on to Twitter.

"I got out my little Android phone and went on Twitter and pulled up the grid to search for 'explosion' and got tweets coming out of the area," he said.

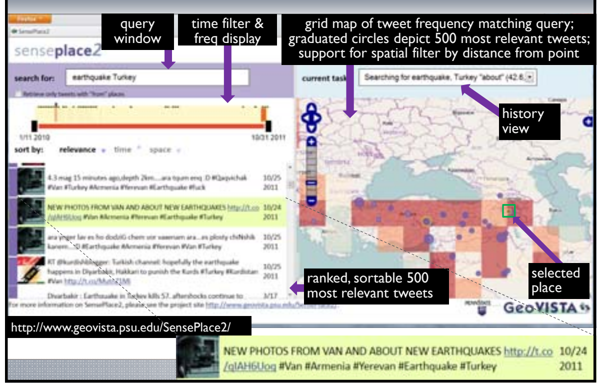
After a few minutes he says he determined the gas explosion and ensuing fire, though horrible, was a localized event and wouldn't spread to other communities.

"I got better situational awareness [from Twitter] before we got official word," he said. "Four or five years ago I wouldn't have gotten that quality of information."

STORY HIGHLIGHTS

- Government officials are usually wary of information on Twitter and other social media
- FEMA chief Craig Fugate uses the data for "situational awareness"

SensePlace2: Place-Time-Concept Analytics



Turkey turns to Twitter after quake

CNN Video: http://www.cnn.com/video/#/video/world/2011/10/25/ctw-anderson-bayulgen-turkey-twitter.cnn
Connect the world|Added on October 25, 2011

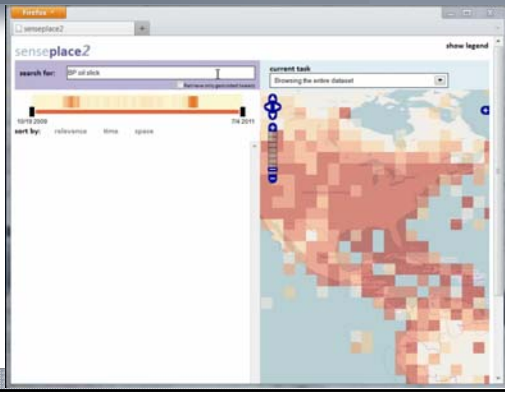
Okan Bayülgen: "Twitter has worked well. It has been the most effective social web site in terms of organizing help. Several victims were found and rescued thanks to the posts of Twitter."

SensePlace2: Computational process

- Collecting and processing Tweets:
  - crawl Tweets & load in database
  - analyze tweets for where they are from
    - "location" setting – often imprecise and inaccurate
    - time zones – accurate but imprecise
    - GPS coordinates – accurate & precise but uncommon
  - extract named-entities; e.g., locations, organizations, persons, hashtags, URLs, etc (what and where tweets are about) - uses customized version of GATE's ANNIE named-entity extraction
  - store extracted entities: PostGRES DB w/ GIS extension
  - index tweets: Lucene index provides fast query and relevance ranks based on "boosting" (e.g., for tweets w/ places, hashtags, organizations, time, money, etc.)

SensePlace 2: Implimentation

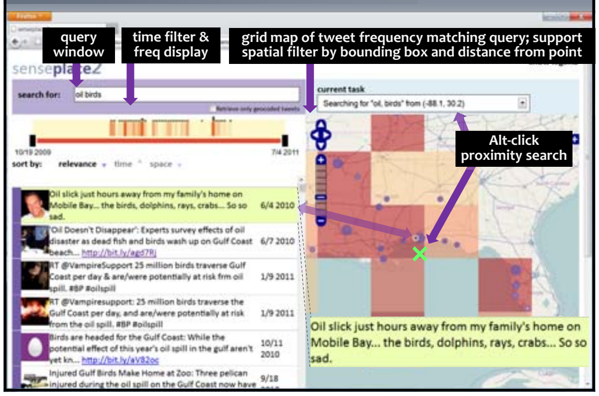
SensePlace2: Place-Time Web Analytics



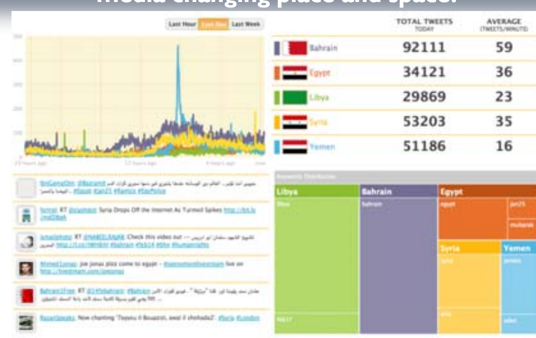
Summary: Insights thus far

- *Situational Awareness:*
  - can be effectively conceptualized as a process to which the sensemaking model of visual analytics is applicable
- *Social media & crisis management:*
  - practitioners consider social media a potential SA source
  - social media have uses for both info push and info foraging
  - for info foraging, practitioners want info on: incidents, times, places, people, organizations
- *System design:*
  - combination of spatial relational database and Lucene indexing enables flexible query on 10s of millions of info fragments

SensePlace2: Place-Time Web Analytics

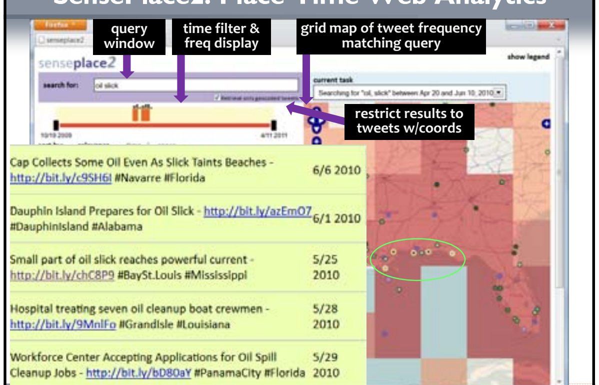


Beyond situational awareness: How are social media changing place and space?



Al Jazeera's 'Arab Spring' Twitter dashboard, Jun 4, 2011 by Dirk Singer, http://liesdamnedliesstatistics.com/2011/06/al-jazeeras-arab-spring-twitter-dashboard.html

SensePlace2: Place-Time Web Analytics



Some specific GIScience challenges for Geovisual Analytics

- develop scientific understanding and related methods/tool that specifically:
  - support analytical reasoning with and about place-based information – this will require deeper understanding of what "reasoning" is in relation to crisis management tasks such as info (geo) foraging, sensemaking, situation assessment, decision-making, etc. and methods / tools for representing and utilizing context
  - support an extended *process* of work – this will require work domain analysis coupled with structures and tools to support processes rather than tasks (particularly ones dealing with uncertain, complex, ambiguous, and equivocal information).
  - recognize and capture *geographic insight* and its *provenance* – this will require attention to knowledge management


That's all: Thanks
maceachren@psu.edu
www.geovista.psu.edu
Acknowledgements: Parts of the work represented here were supported by funds from: National Science Foundation: EIA-0306845 & 9983451; Gates Foundation funding of the Vaccine Modeling Initiative (https://www.vaccinemodeling.org/); National Visualization and Analytics Center, a U.S. Dept. of Homeland Security program operated by the Pacific Northwest National Laboratory (PNNL), and the Dept of Homeland Security Visual Analytics for Command, Control, and Interoperability Environments Center of Excellence. A large number of people contributed to work represented. These include Justine Blanford, Jin Chen, Ping Du, Mark Gahegan, Bryan Grenfell, Diansheng Guo, Frank Hardisty, Anuj Jaiswal, Gene Lengerich, Presenjit Mitra, Donna Peuquet, Scott Pezanowski, ChiChun Pan, Anthony Robinson, Robert Roth, Alexander Savelyev, Michael Stryker, Brian Tomaszewski, Ian Turton, Chaoqing, Yu, Ying Wang, Chris Weaver, Xiao Zhang

Table of Contents

Mobile Symbol Design and Event Reporting for Mobile Devices:

A Preview of a Mobile Application and Developer API Based on an Online Symbol Repository

Joshua Stevens
Scott Pezanowski
Elaine Guidero
Anthony C. Robinson
Alan M. MacEachren

Overview

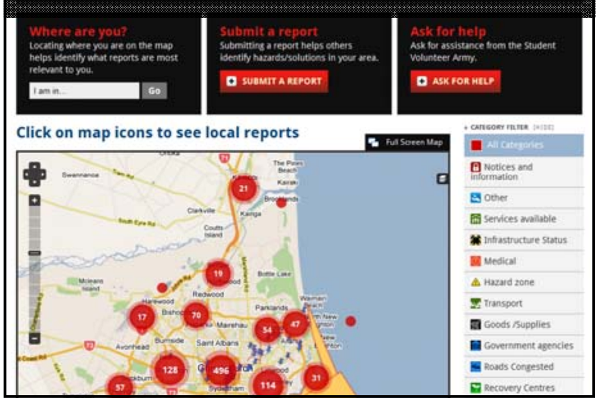
1. Event reporting + current strategies
2. Preview of our mobile app
3. What makes our app possible
4. Symbol design for mobile devices

Event Reporting

Event Reporting

- Interest of emergency management, crisis mapping, and law enforcement
- ...but also the general public
- Played a major role in...
  - Haiti earthquake, 2010
  - Arab Spring, 2010
  - Christchurch earthquake, 2011
  - Public demonstrations since (e.g., #OWS)

Christchurch Recovery Map, Built w/ Ushahidi



The Good:

Event maps are popular + utilized

The Bad:

Event maps are largely desktop-oriented

The Ugly:

Event maps typically use the same abstract symbology

The Uglier:

Event mapping platforms center on a single geographic area and/or event type

A Preview of Our Mobile App

MapAlert

Streamlined Event Symbolization

Streamlined Event Reporting

How MapAlert is Possible

What Makes MapAlert Possible:

Symbol Store: An online repository of 2,400+ map symbols

Symbol Store Web Services

Symbol Store Web Services

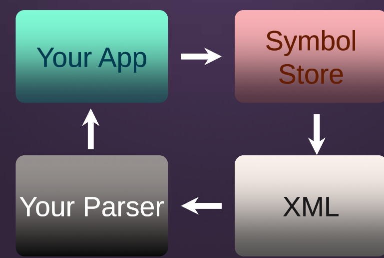
- Developers can create their own apps to access the Symbol Store
- Main feature: Search
  - Symbol name
  - Tags
  - Category
  - Agency
  - Org:Set

Extended Features

- Edit
- Upload
- Parse and generate style files, SVG, PNG

charles river analytics

How Symbol Store Web Services Work



Search Example using HTML GET

Request

```
GET
/OurServiceURL/OurService.asmx/SearchLucene?value=string&timestamp=string HTTP/1.1
Host: www.geovista.psu.edu
```

Response

```
HTTP/1.1 200 OK
Content-Type: text/xml; charset=utf-8
Content-Length: length

<?xml version="1.0" encoding="utf-8"?>
<ArrayOfAnyType xmlns="http://www.geovista.psu.edu/">
  <anyType />
  <anyType />
</ArrayOfAnyType>
```


These services were made for you.

Mobile Symbol Design

Mobile Symbol Design

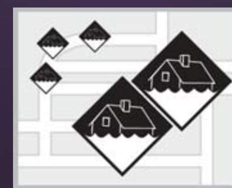
- We wanted to learn if mobile symbols were unique (spoiler alert: they are)
- Reviewed close to 80 separate studies and manufacturer guidelines (Apple + Google)
- Boiled the findings down to some “simple” recommendations...brief overview here
- Designed mobile symbology to test against existing DHS HSWG set

Mobile Symbol Recommendation 1:

Use dark symbols on light backgrounds.

Mobile Symbol Recommendation 2:

Size matters. Bigger is (mostly) better.



Mobile Symbol Recommendation 3:

Symbols should be simple.

(Can you guess what this is? You shouldn't have to.)

Mobile Symbol Recommendation 4:

Visually distinguish interactive and non-interactive symbols.

Our Mobile Symbol Design vs HSWG

Evaluation now Planned for November

- Will compare our symbols to DHS:HSWG on a tablet device
- Looking at:
  - accuracy
  - time
  - associative relationships
  - influence of base maps
  - interactive affordance

Table of Contents

Table of Contents

Sharing Map Symbology for Emergency Management

Anthony C. Robinson, Ph.D

Lead Faculty for Online Geospatial Education
John A. Dutton e-Education Institute
Assistant Director, GeoVISTA Center
Department of Geography
The Pennsylvania State University

Co-Authors:

Scott Pezanowski
Raechel Bianchetti
Elaine Guidero
Justine Blanford
Alan M. MacEachren

PENNSTATE 1855
GeoVISTA

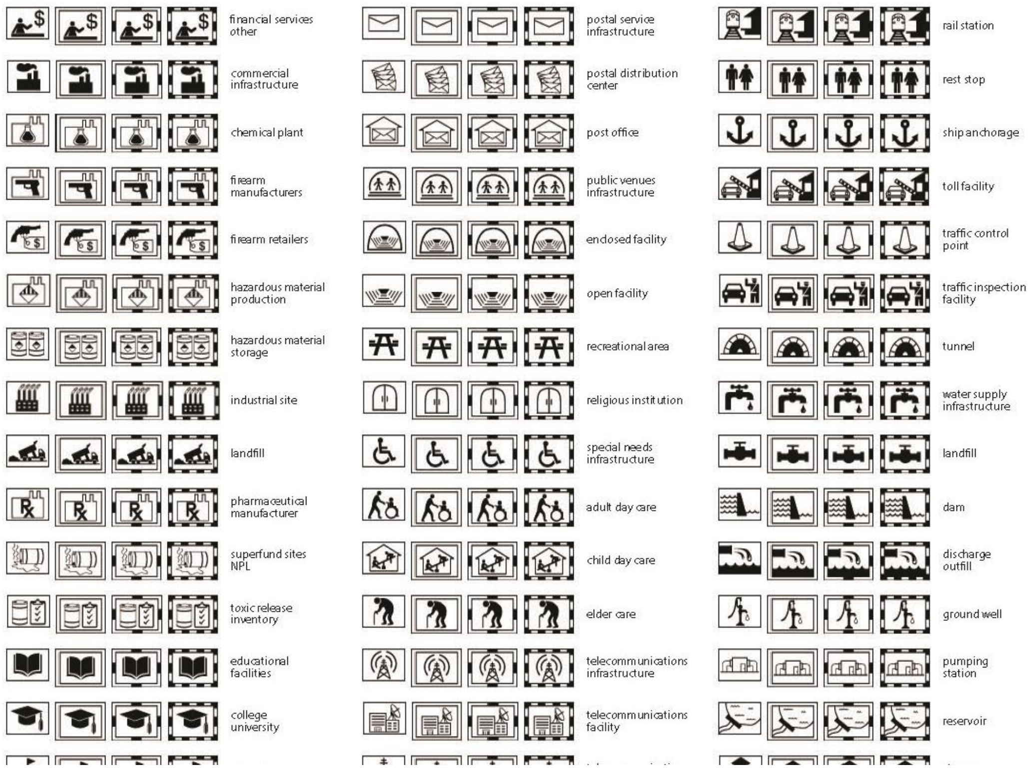
Agenda

symbology @ DHS

symbol store design

current progress

future work



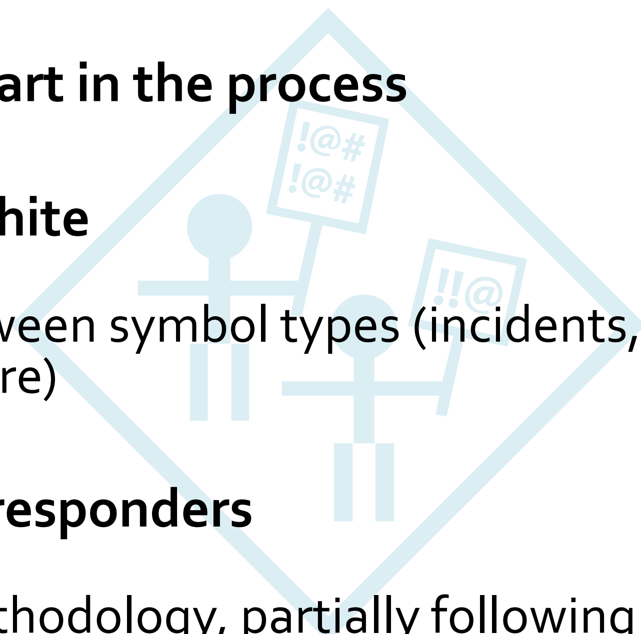
symbology @ DHS

background

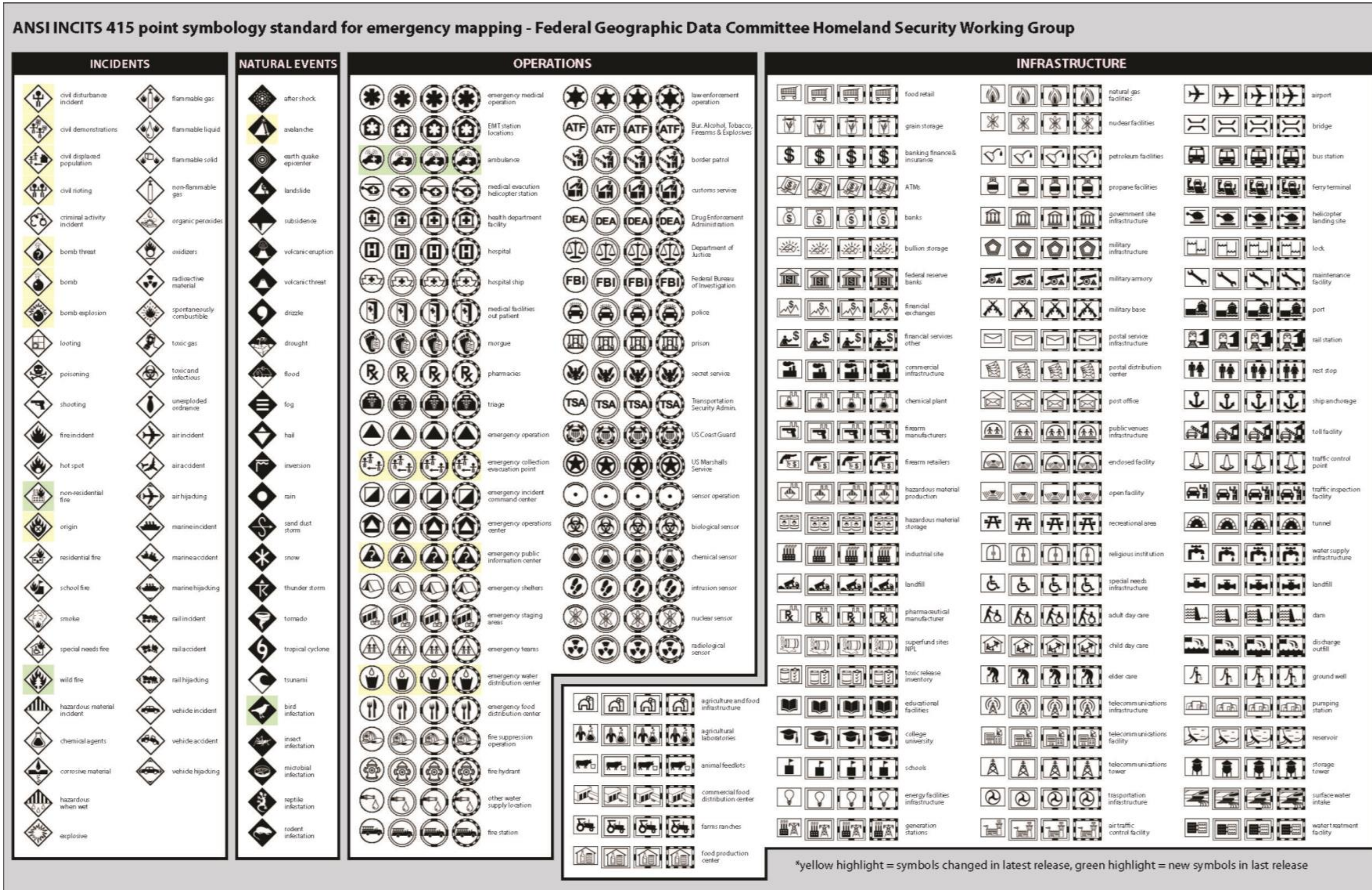
- **Diverse DHS organizations produce and use maps**
  - Audiences range from geospatial analysts to general public
- **No consistent set of map symbols used across DHS**
  - Even if we just look at point symbols
- **ANSI INCITS 415-2006 intended for emergency management mapping**
  - Poorly adopted by practitioners
- **Initial Objective: Develop process for symbol standardization**
- **Secondary Objective: Develop mechanism for symbol interoperability**

ANSI Standard

- **Point symbol set designed for emergency response**
  - Goal was to facilitate common situational awareness, support point symbol interoperability
- **Federal/state/local stakeholders took part in the process**
- **Symbols designed to work in black & white**
  - Outline shapes used to distinguish between symbol types (incidents, natural events, operations, infrastructure)
- **Evaluation conducted online with first responders**
  - Made use of an “accept” or “reject” methodology, partially following the ANSI guidelines



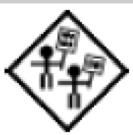
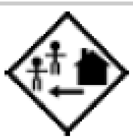
ANSI Standard



ANSI Standard

	U	*Bird Infestation (Infestation)* - A harassing or troublesome invasion of birds. (Source: derived from the definition of “infestation” found in the FactMonster.com dictionary)
	V	*Insect Infestation (Infestation)* - A harassing or troublesome invasion of insects. (Source: derived from the definition of “infestation” found in the FactMonster.com dictionary)
	W	*Microbial Infestation (Infestation)* - A harassing or troublesome invasion of a microbe. (Source: derived from the definition of “infestation” found in the FactMonster.com dictionary)
	X	*Reptile Infestation (Infestation)* - A harassing or troublesome invasion of reptiles. (Source: derived from the definition of “infestation” found in the FactMonster.com dictionary)
	Y	*Rodent Infestation (Infestation)* - A harassing or troublesome invasion of rodents. (Source: derived from the definition of “infestation” found in the FactMonster.com dictionary)

ANSI Standard

	A	*Civil Disturbance Incident (Theme)* - Human activities resulting in the disrupting of services or requiring varying levels of support, law enforcement or attention.
	B	*Civil Demonstrations (Civil Disturbance)* - A public display of group feelings toward a person or cause. (Source: Merriam-Webster Online Dictionary definition)
	C	*Civil Displaced Population (Civil Disturbance)* - Persons or groups of person who have been forced or obliged to flee or to leave their homes or places of habitual residence, in particular as a result of or in order to avoid the effects of armed conflict, violations of human rights, or natural or human-made disasters. (Source: United Nations Guiding Principles on Internally Displaced Persons, 1998)
	D	*Civil Rioting (Civil Disturbance)* - A public disturbance involving (1) an act or acts of violence by one or more persons part of an assemblage of three or more persons, which act or acts shall constitute a clear and present danger of, or shall result in, damage

Understanding the Problem

- We discovered that most mission areas had their own, 'in house' standards

- These were developed on an ad hoc basis, usually by one cartographer

- Collections of Esri markers and whatever else they could scrape together

- Such 'standards' are passed around to new employees and are promoted as default option

Robinson, A.C., Roth, R.E., and A.M. MacEachren (2011) Understanding User Needs for Map Symbol Standards in Emergency Management. *Journal of Homeland Security and Emergency Management.* 8(1): Article 33.

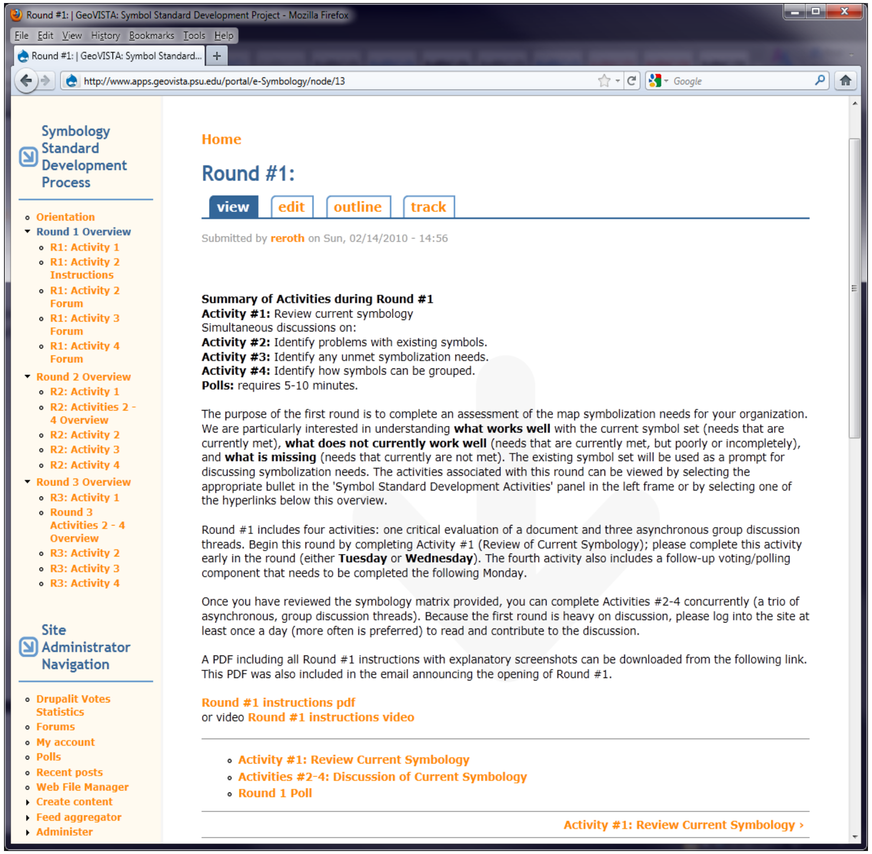
New ‘Standardization’ Process

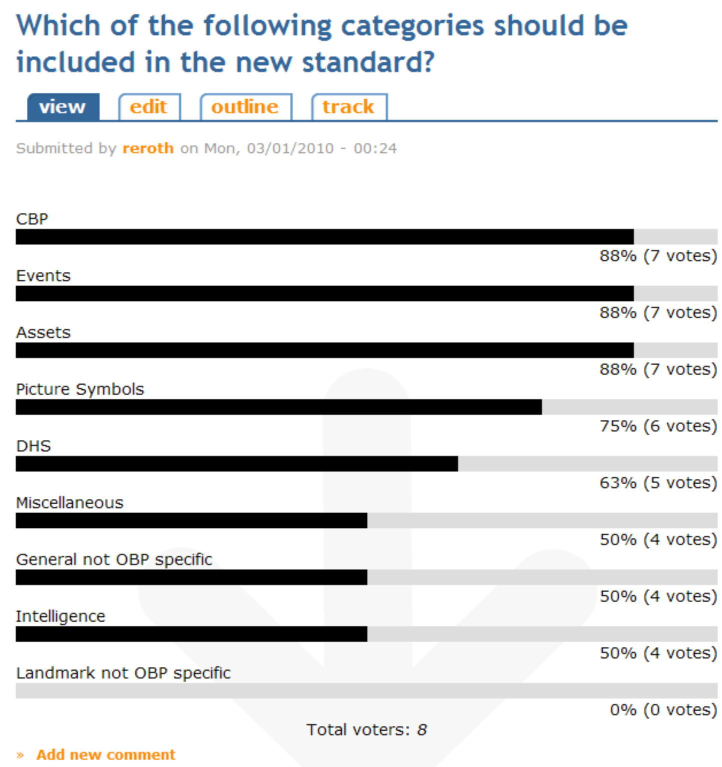
- **Distributed, web-based activities through a customized Drupal site**

- **Phase 1: Needs Assessment**
  - Review current symbology, identify new symbol needs, problems with current symbols
- **Phase 2: Initial Standard Development**
  - Develop symbol categories, vote on changes to current symbology
- **Phase 3: Standard Refinement**
  - Discuss, refine & vote on final categories
- **Phase 4: Implementation & Quality Control**
  - Test new symbology in exercise, submit standard for graphical refinement by cartographers

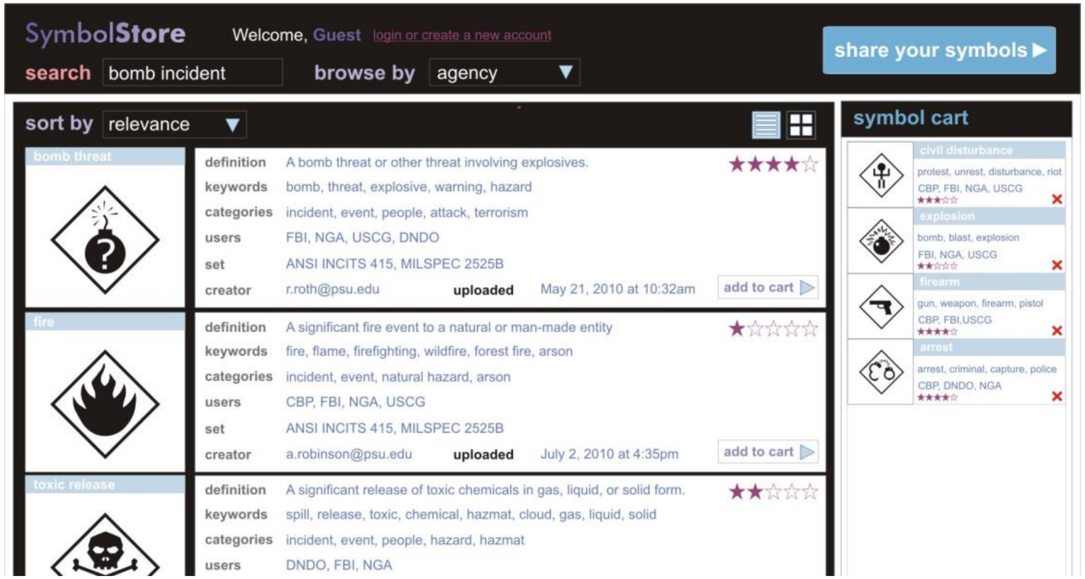
- **Methods feature**
  - Round-based discussion & voting (modified Delphi)
  - Card-sorting activities (using websort.com)
  - Anonymized participation

e-Symbology Portal





Robinson, A.C., Roth, R.E., Blanford, J., Pezanowski, S. and MacEachren, A.M. In Press: Developing map symbol standards using a distributed, asynchronous process. *Environment and Planning B: Planning and Design*.



symbol store design

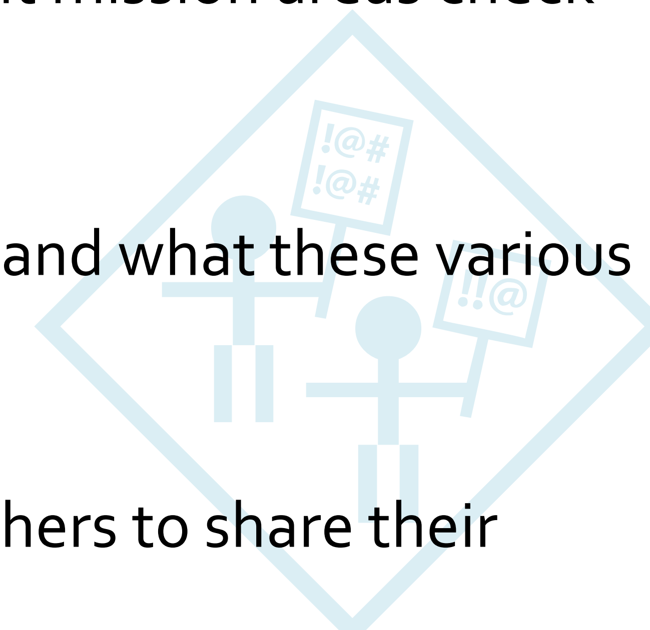
Sharing Symbols

- What happens once a group has completed the process?

- How can cartographers in different mission areas check out these ‘in house’ standards?

- What can we do to better understand what these various standards have in common?

- How can we encourage cartographers to share their symbols more widely?



Symbol Store Approach

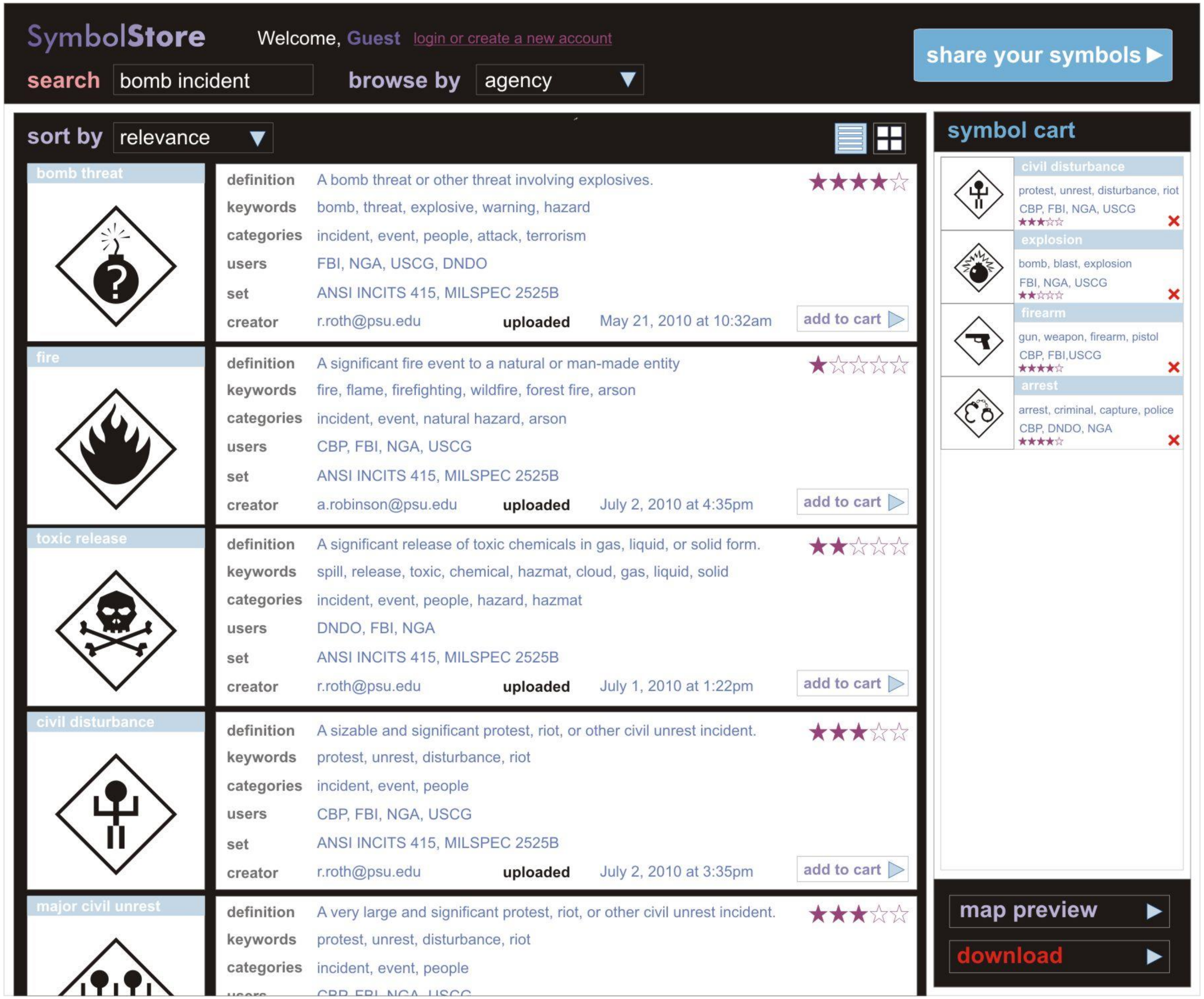
- Develop web-based solution for sharing / browsing point symbols
- Simple design – aim to support one key task really well
- Symbols that have gone through our process can receive special tags
- Iterative development process -> static mockup, dynamic prototype, refined prototype, final version

Design Goals

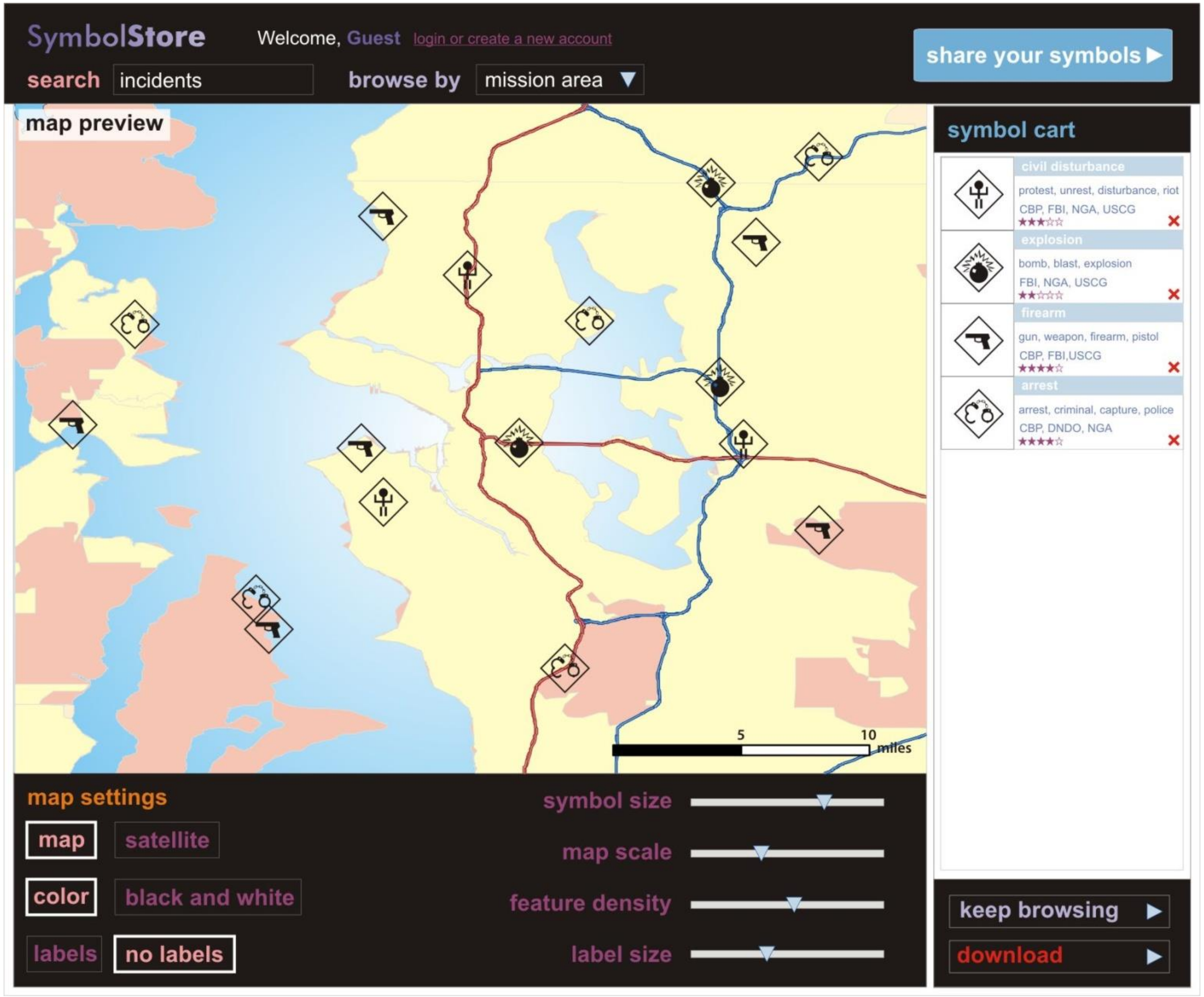
Four simple design objectives for Symbol Store:

1. **Search for and retrieve symbols**
   – by keyword
2. **Preview symbols on realistic maps**
   – like ColorBrewer, of course
3. **Browse for symbols**
   – discover new symbols by category, agency, etc…
4. **Share symbols**
   – Contribute your own symbols and add appropriate metadata

Initial Design



Initial Design

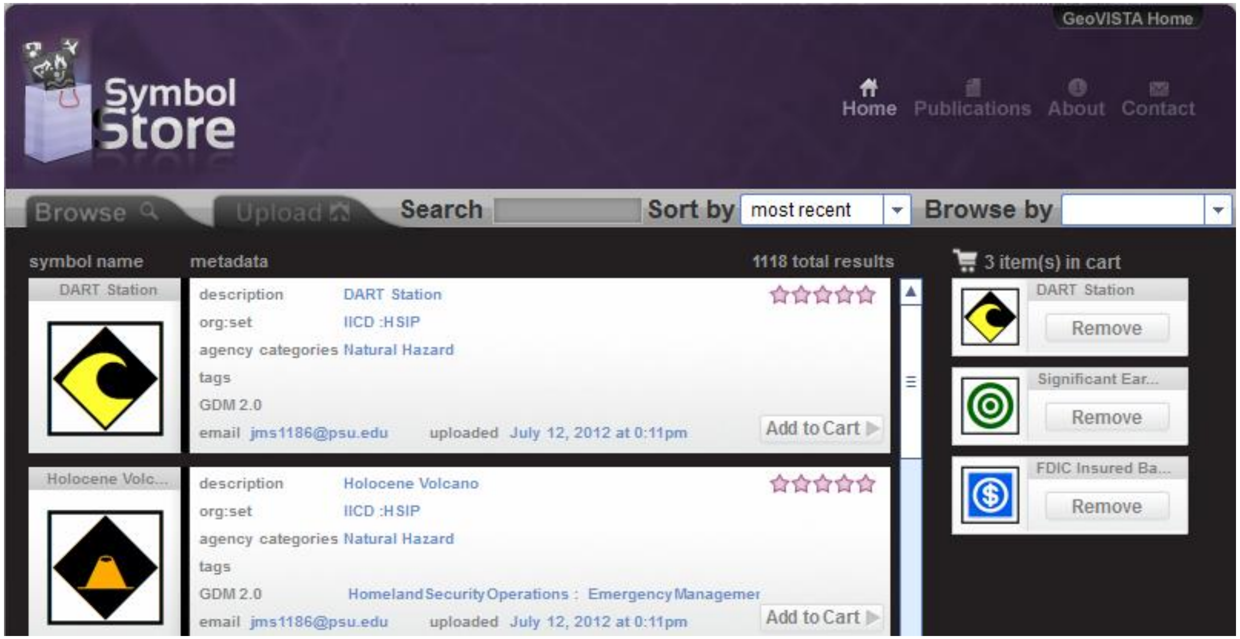


Back-end Design

- ArcObjects is used to parse Esri style files and generate preview images of the symbols
- Lucene index for text content
  - Quick text searching and retrieval of large amounts of content
  - Ranks hits based upon user searches
  - Weights search criteria to improve search results
- .NET web service supports read/write to the Lucene index and read/write for .style files
- All content stored in the Lucene index or files on the server (style files and image preview files)

User Interface

- Interface built using Flash Catalyst (converts AI & other art into UI objects)
- Functionality connected to interface using Flash Builder 4 (formerly Flex)
- Flash Builder used to get results from the web service
- Flash plugin required for web browsers, but otherwise works across platforms

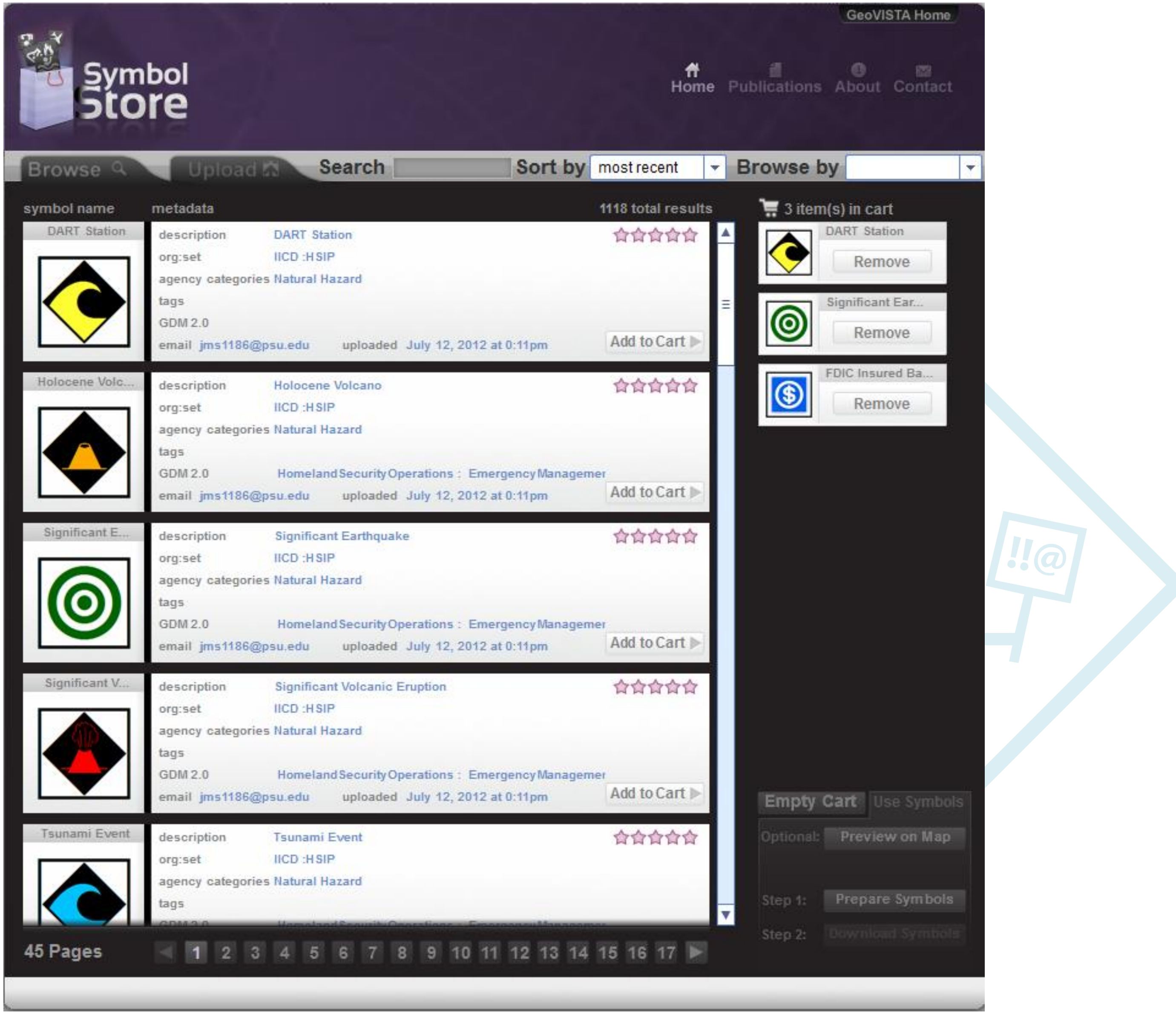


current progress

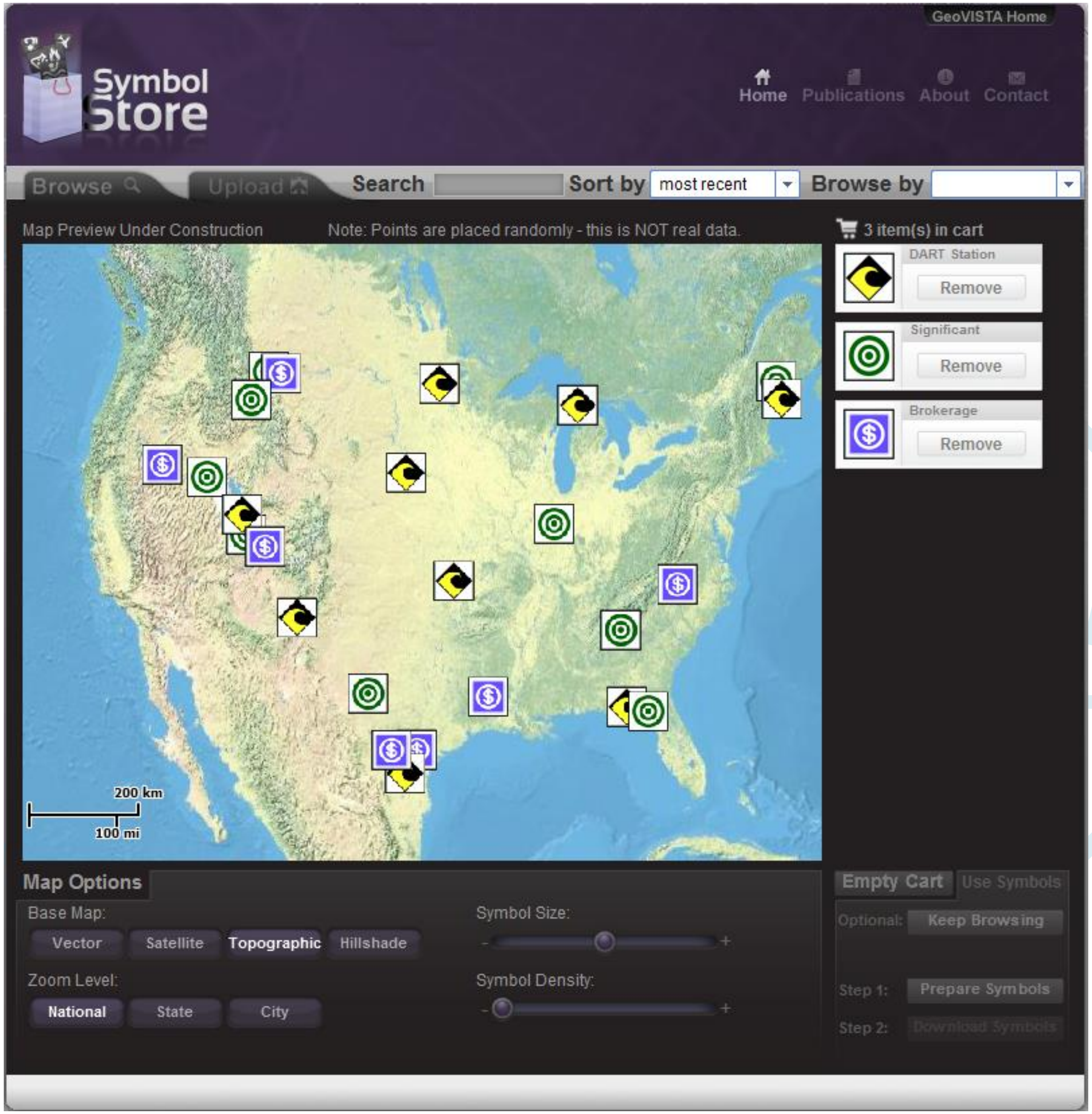
Progress

- Developed interactive prototype based on static mockup + sponsor feedback
- Prototype supports all four of our initial task goals:
  - Search
  - Preview
  - Browse
  - Share
- Public version now available at **symbolstore.org**
  - Uploading not enabled for public version yet ☹

Main UI

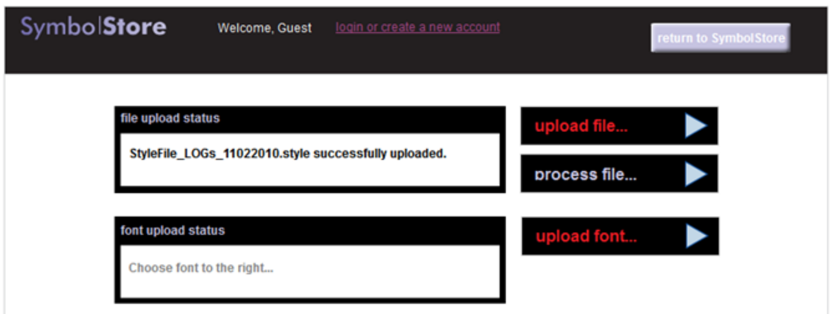

GeoVISTA Home
Symbol Store
Home
Publications
About
Contact
Browse
Upload
Search
Sort by
most recent
Browse by
symbol name
metadata
1118 total results
3 item(s) in cart
DART Station
description DART Station
org:set IICD :HSIP
agency categories Natural Hazard
tags
GDM 2.0
email jms1186@psu.edu uploaded July 12, 2012 at 0:11pm
Add to Cart
Holocene Volc...
description Holocene Volcano
org:set IICD :HSIP
agency categories Natural Hazard
tags
GDM 2.0 Homeland Security Operations : Emergency Managemen
email jms1186@psu.edu uploaded July 12, 2012 at 0:11pm
Add to Cart
Significant E...
description Significant Earthquake
org:set IICD :HSIP
agency categories Natural Hazard
tags
GDM 2.0 Homeland Security Operations : Emergency Managemen
email jms1186@psu.edu uploaded July 12, 2012 at 0:11pm
Add to Cart
Significant V...
description Significant Volcanic Eruption
org:set IICD :HSIP
agency categories Natural Hazard
tags
GDM 2.0 Homeland Security Operations : Emergency Managemen
email jms1186@psu.edu uploaded July 12, 2012 at 0:11pm
Add to Cart
Tsunami Event
description Tsunami Event
org:set IICD :HSIP
agency categories Natural Hazard
tags
DART Station
Remove
Significant Ear...
Remove
FDIC Insured Ba...
Remove
Empty Cart
Use Symbols
Optional: Preview on Map
Step 1: Prepare Symbols
Step 2: Download Symbols
45 Pages
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

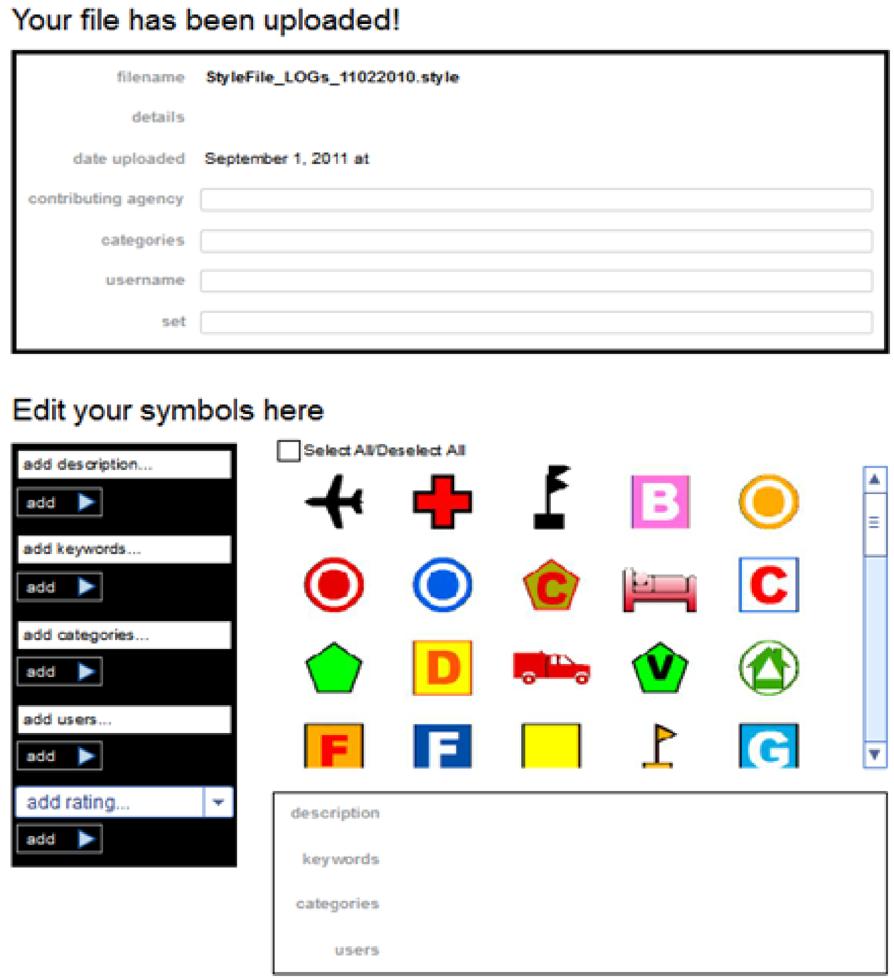
Map Preview



GeoVISTA Home
Symbol Store
Home
Publications
About
Contact
Browse
Upload
Search
Sort by
most recent
Browse by
Map Preview Under Construction
Note: Points are placed randomly - this is NOT real data.
3 item(s) in cart
DART Station
Remove
Significant
Remove
Brokerage
Remove
200 km
100 mi
Map Options
Base Map:
Vector
Satellite
Topographic
Hillshade
Zoom Level:
National
State
City
Symbol Size:
Symbol Density:
Empty Cart
Use Symbols
Optional:
Keep Browsing
Step 1:
Prepare Symbols
Step 2:
Download Symbols

Uploading & Metadata





Initial Evaluation

- 6 Cartographers from California's Dept. of Water Resources
  - Primary mission involves flood mapping
- Basic task analysis + follow-up focus group using early Symbol Store prototype
- Tasks focused on our four key design goals, overall satisfaction with the prototype was high for each feature
- Lots of suggested improvements, including:
  - Grid view to show more symbols at once
  - Use style file details to pre-populate metadata
  - Show USGS Topo design as option in Map Preview

future work

Next Steps

- Integrate portions of our symbol standardization process into the Symbol Store
- Develop support for other technical standards (SLD, for example)
- Extend tools to handle dynamic / multi-scale point symbols
- Explore Symbol Store usage patterns to identify frequently used symbols, cross-organizational commonalities, etc…
- Vector export (instead of just Esri .style files)

Table of Contents

for more information:

www.geovista.psu.edu/symbology/

try it - www.symbolstore.org

email - arobinson@psu.edu

This work is supported by a contract from the U.S. Department of Homeland Security Science and Technology Directorate, Command, Control and Interoperability Division. The views and opinions expressed here are of the authors, and do not reflect the official positions of the Department of Homeland Security or the Federal Government.

GeoVISTA

PENNSTATE

1855

Table of Contents

SymbolStore.org: An Open Resource for Map Symbols

Anthony C. Robinson, PhD

Lead Faculty for Online Geospatial Education
John A. Dutton e-Education Institute
Assistant Director, GeoVISTA Center
Department of Geography
The Pennsylvania State University

Project Team:

Scott Pezanowski
Joshua Stevens
Justine Blanford
Alan M. MacEachren
Raechel Bianchetti
Elaine Guidero
Ryan Mullins
Eun-Kyeong Kim

GeoVISTA

PENNSTATE 1855

DEPARTMENT OF GEOGRAPHY
COLLEGE OF EARTH AND MINERAL SCIENCES

Agenda

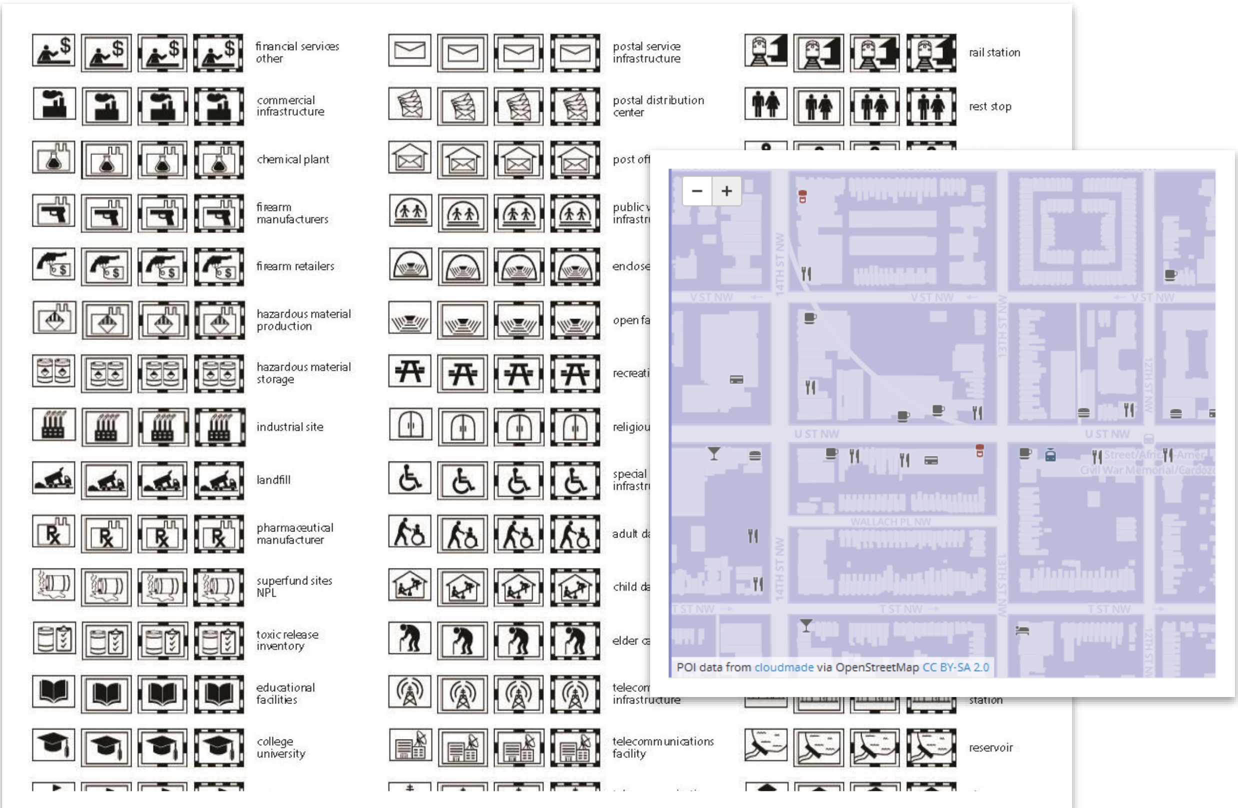
Sharing symbols

SymbolStore features

SymbolStore.org progress & demo

Next steps

GeoVISTA Center



Sharing symbols

Background

- Our work began with DHS to explore how they produce and use maps
  - Focus on symbol interoperability
- No consistent set of map symbols used across DHS
  - No mechanisms to share them, either
- Standards for various sub-domains exist, but are weakly adopted
  - Symbol discovery happens in ArcGIS and through simple web sharing
- Our goal: Develop mechanism to support symbol interoperability

Understanding the Problem

- Working with DHS, we discovered that most mission areas had their own ‘in house’ standards
- These were developed on an ad hoc basis, usually by one cartographer
- Collections of Esri markers and whatever else they could scrape together
- Such ‘standards’ are passed around to new employees and are promoted as default option

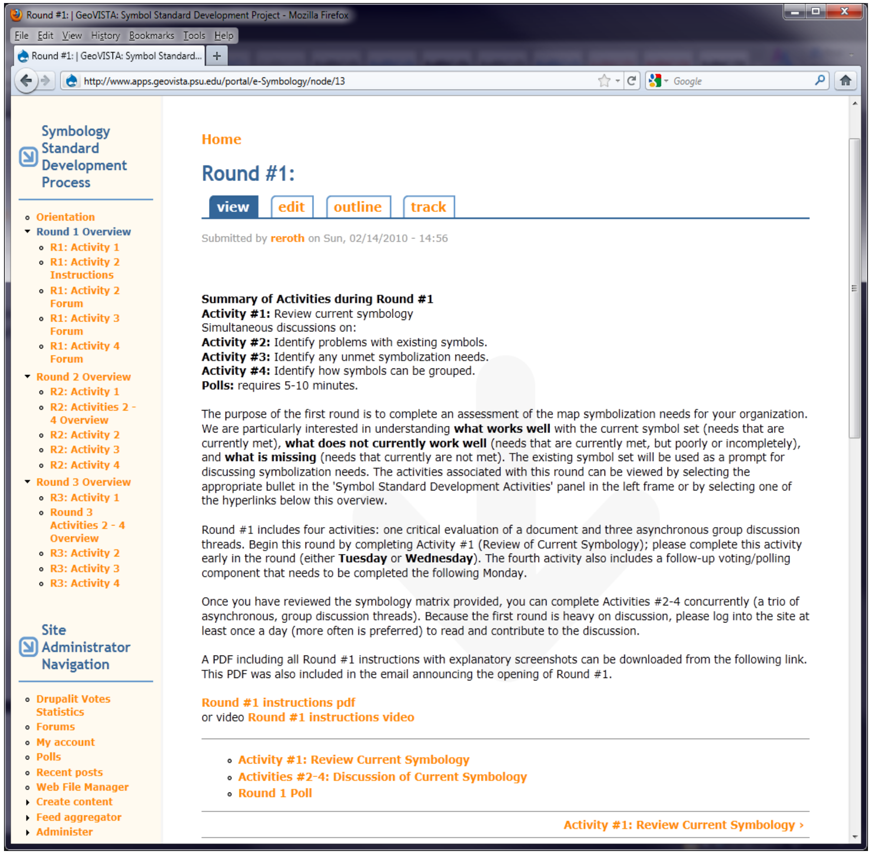
Robinson, A.C., Roth, R.E., and A.M. MacEachren (2011) Understanding User Needs for Map Symbol Standards in Emergency Management. *Journal of Homeland Security and Emergency Management.* 8(1): Article 33.

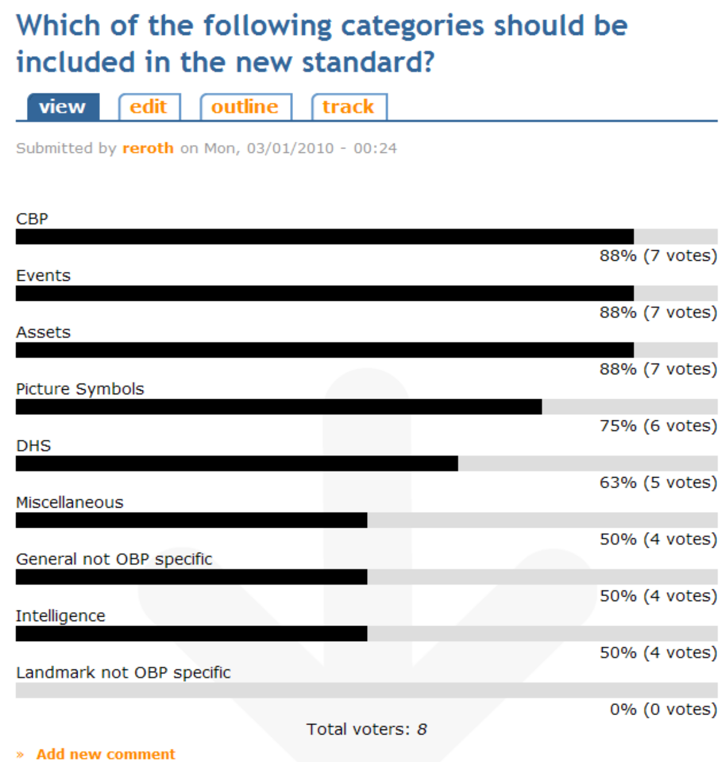
GeoVISTA Center

New ‘Standardization’ Process

- Distributed, web-based activities through a customized Drupal site
- Phase 1: Needs Assessment
  - Review current symbology, identify new symbol needs, problems with current symbols
- Phase 2: Initial Standard Development
  - Develop symbol categories, vote on changes to current symbology
- Phase 3: Standard Refinement
  - Discuss, refine & vote on final categories
- Phase 4: Implementation & Quality Control
  - Test new symbology in exercise, submit standard for graphical refinement by cartographers
- Methods feature
  - Round-based discussion & voting (modified Delphi)
  - Card-sorting activities (using websort.com)
  - Anonymized participation

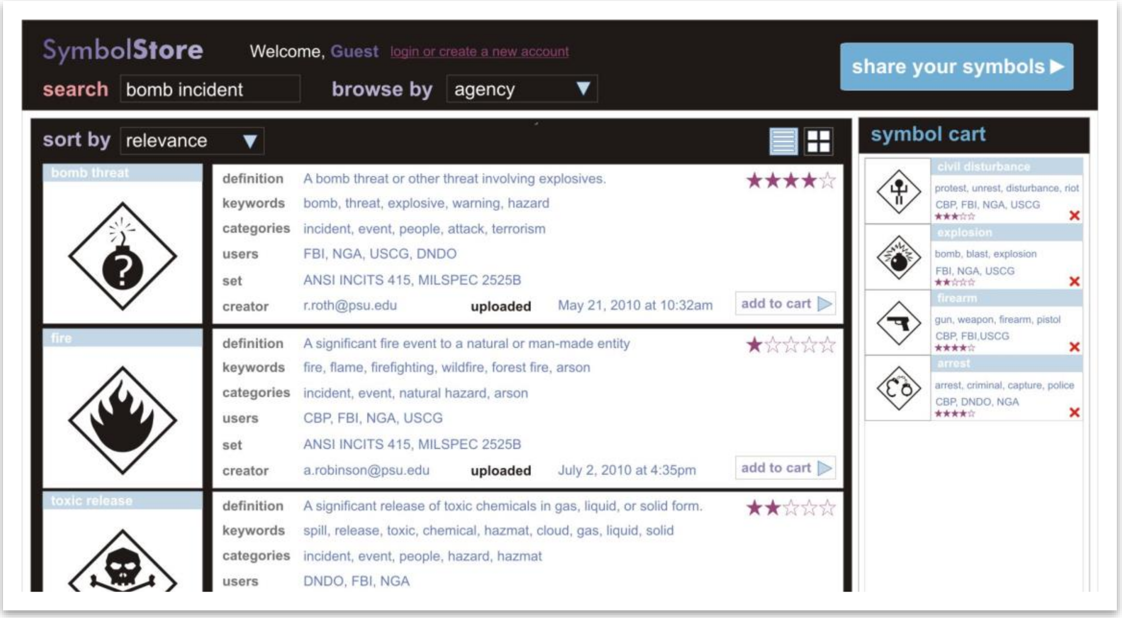
e-Symbology Portal





Robinson, A.C., Roth, R.E., Blanford, J., Pezanowski, S. and MacEachren, A.M. In Press: Developing map symbol standards using a distributed, asynchronous process. *Environment and Planning B: Planning and Design*.

GeoVISTA Center



SymbolStore features

Key design goals

- What happens once a group has completed the standardization process?
- How can other cartographers check out others standards?
- What can we do to better understand what these various standards have in common?
- How can we encourage cartographers to share their symbols more widely?

Symbol Store Approach

- Develop web-based solution for sharing / browsing point symbols
- Simple design – aim to support one key task really well
- Symbols that have gone through our process can receive special tags
- Initially just for DHS – today we're releasing the first public version

GeoVISTA Center

Design Goals

Four simple design objectives for Symbol Store:

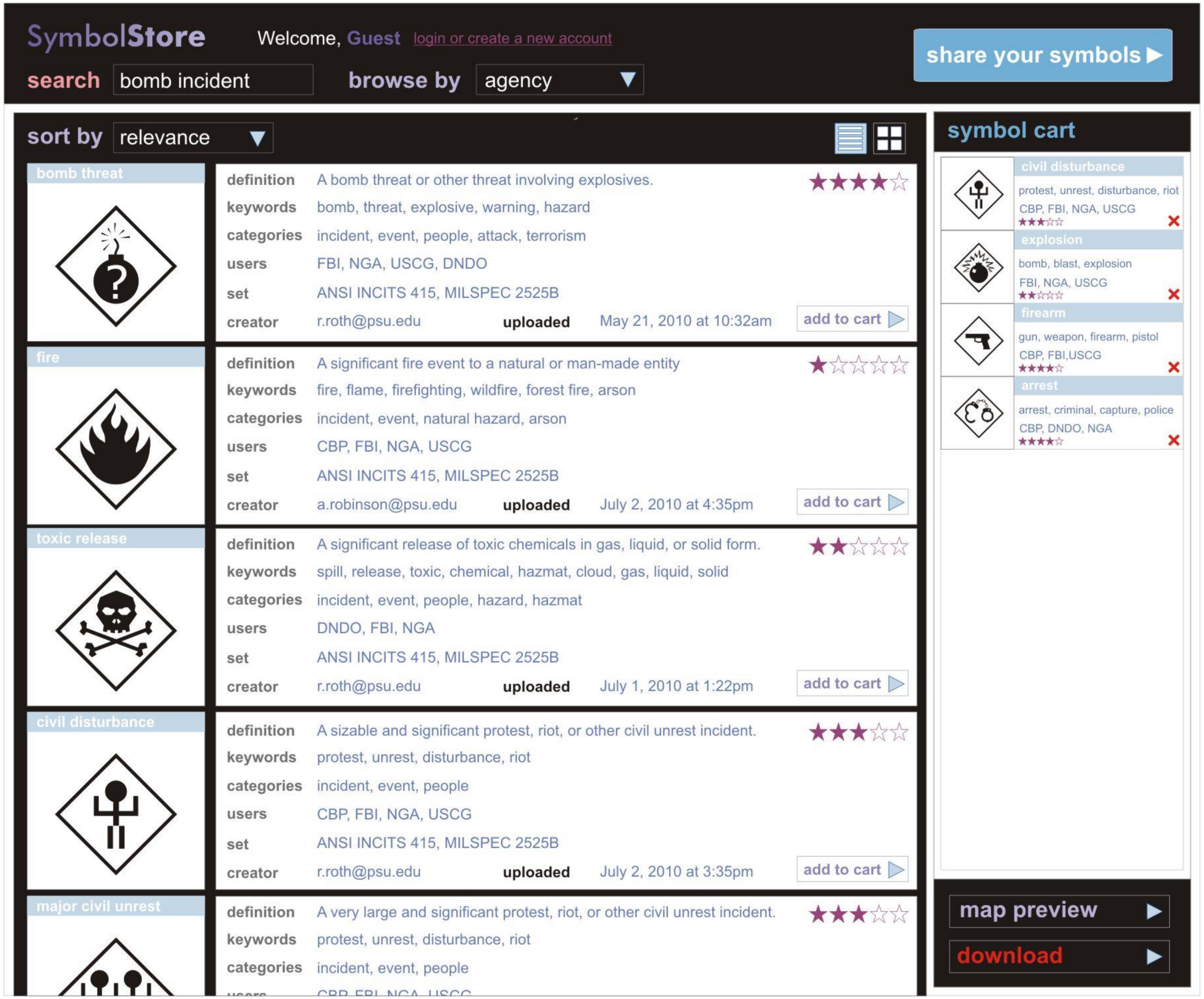
1. Search for and retrieve symbols
   – by keyword

2. Preview symbols on realistic maps
   – like ColorBrewer and other map design apps

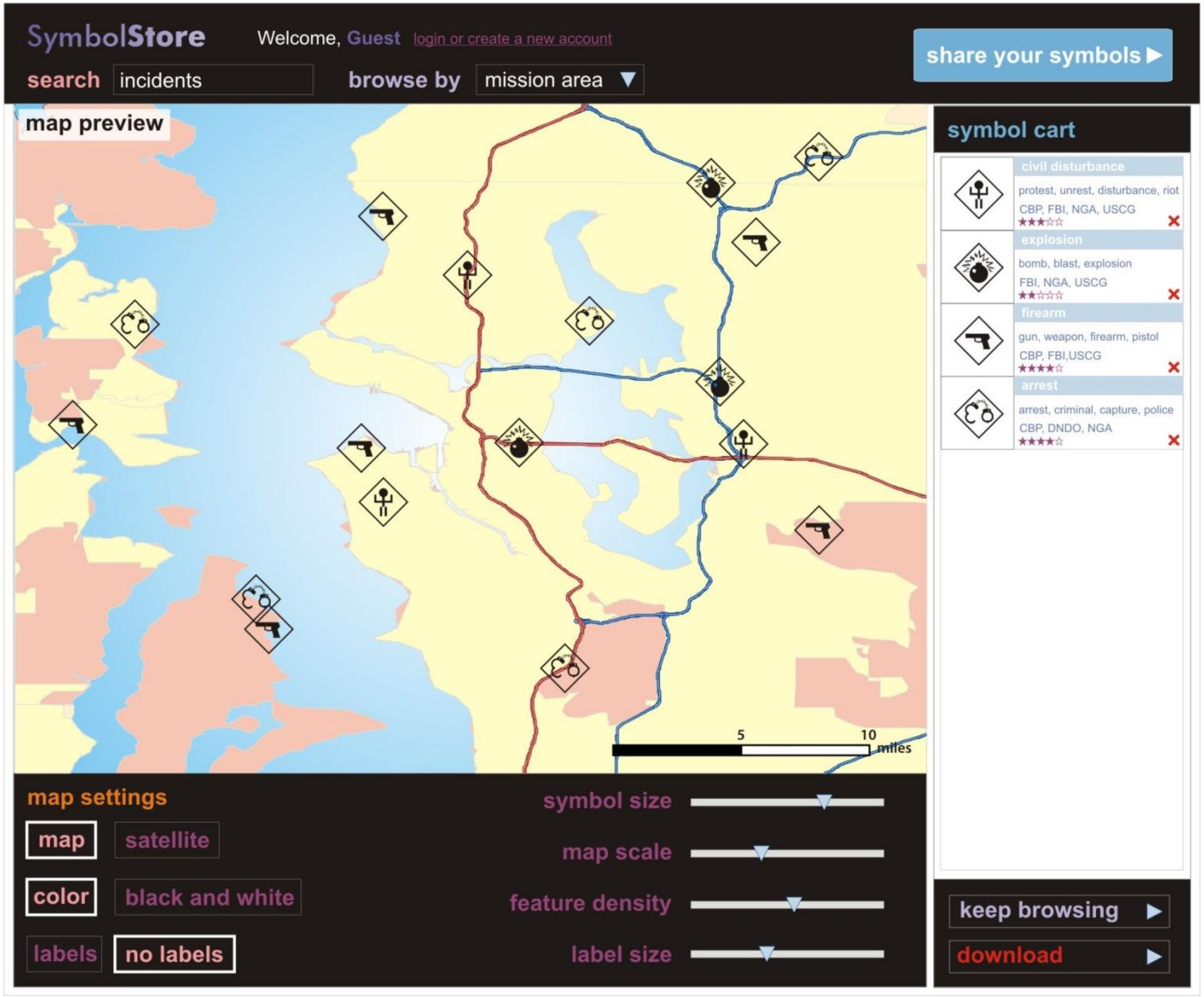
3. Browse for symbols
   – discover new symbols by category, agency, etc…

4. Share symbols
   – Contribute your own symbols and add appropriate metadata

Initial Design



Initial Design

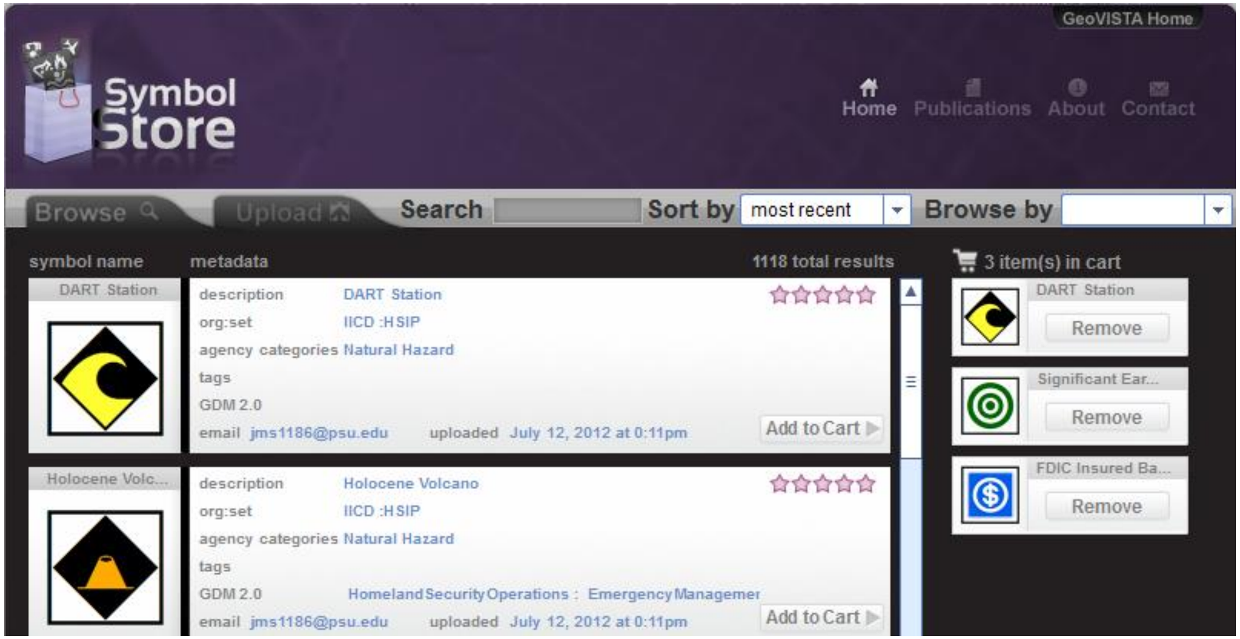


User Interface

- Interface built using Flash Catalyst
- Functionality connected to interface using Flash Builder 4 (formerly Flex)
- Flash Builder used to get results from our web service
- Yes, we plan on migrating to a Javascript-based interface in the near future

Back-end Design

- ArcObjects is used to parse Esri .style files and generate preview images of the symbols
- Lucene index for text content
  - Quick text searching and retrieval of large amounts of content
  - Ranks hits based upon user searches
  - Weights search criteria to improve search results
- .NET web service supports read/write to the Lucene index and read/write for .style files
- All content stored in the Lucene index or files on the server (style files and image files)

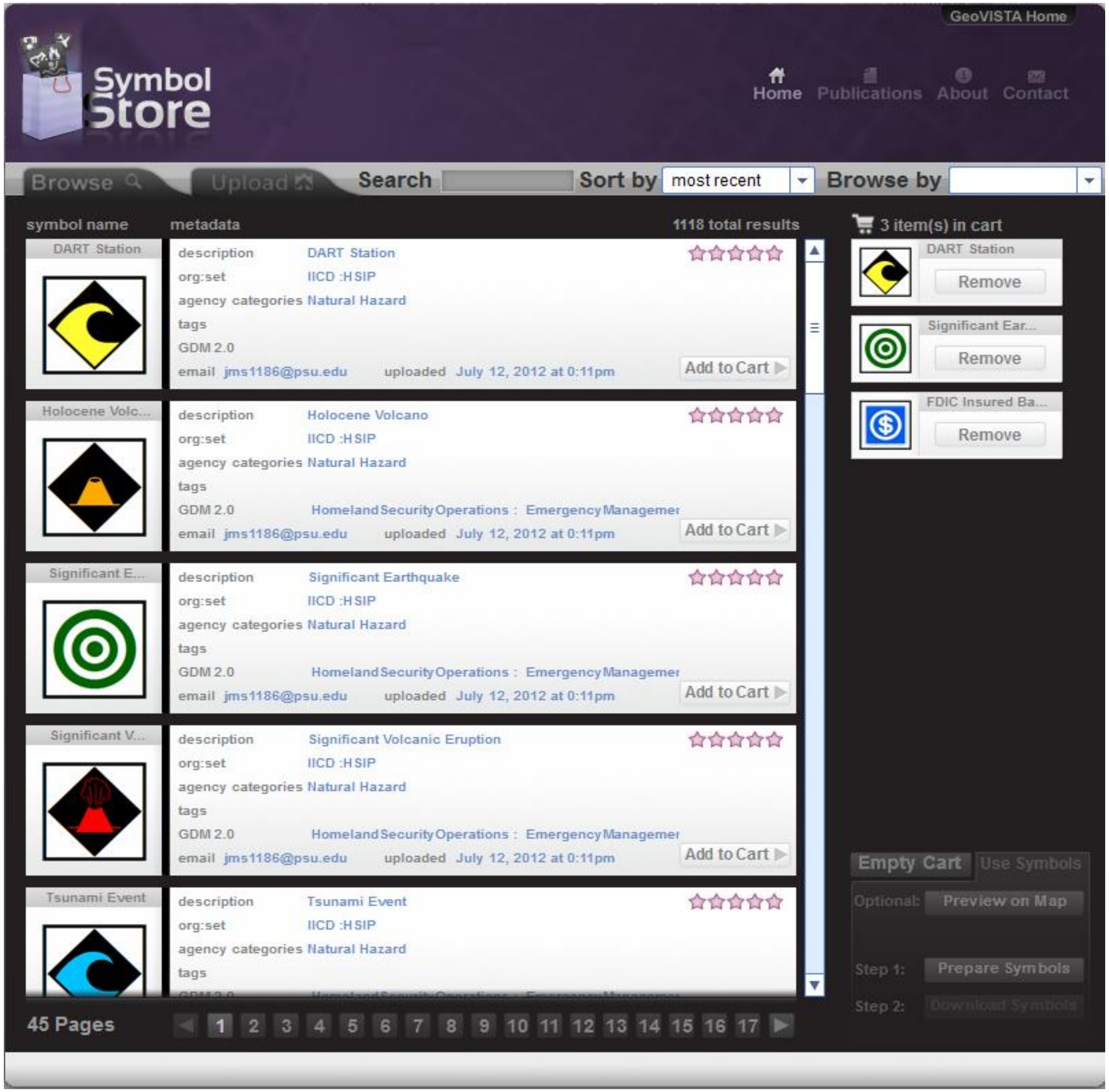


SymbolStore.org progress

Progress

- Developed interactive prototype based on static mockup + sponsor feedback (only for DHS)
- Initial evaluation with flood mapping users in California (Troedson, NACIS 2011)
- Current prototype supports all four of our initial task goals:
  - Search
  - Preview
  - Browse
  - Share
- UI Reskinned by Josh Stevens
- Public version now available at **symbolstore.org**

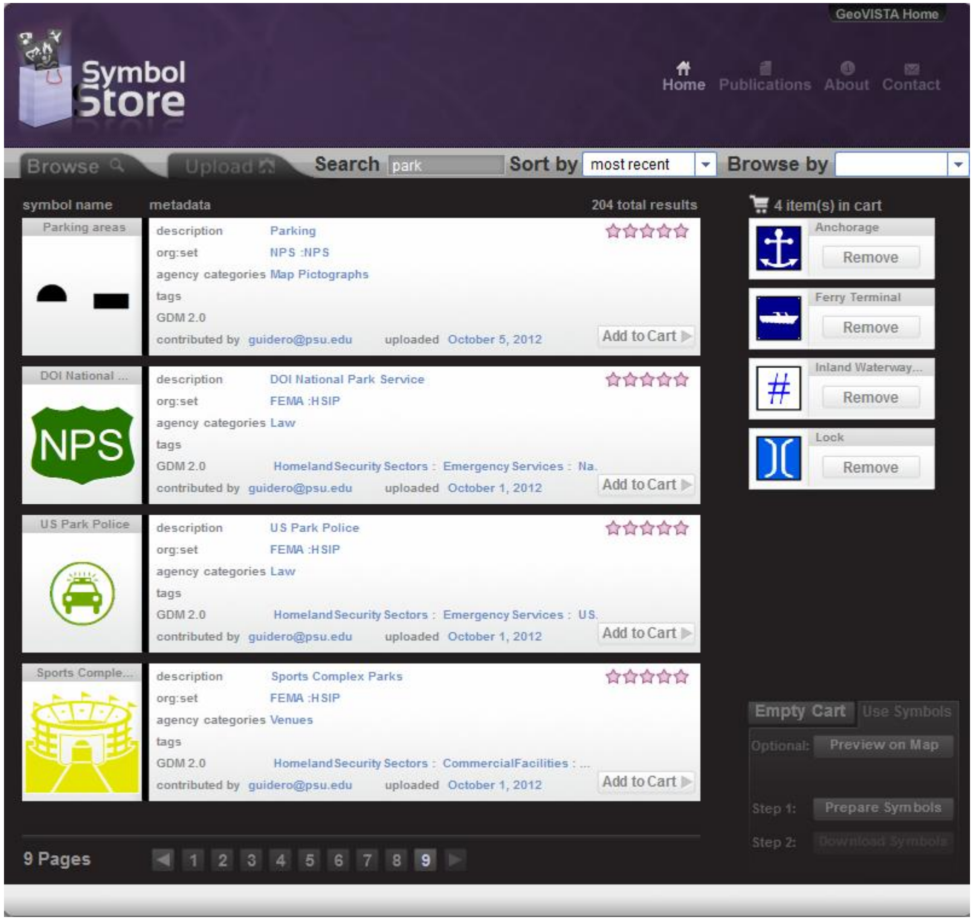
Main UI



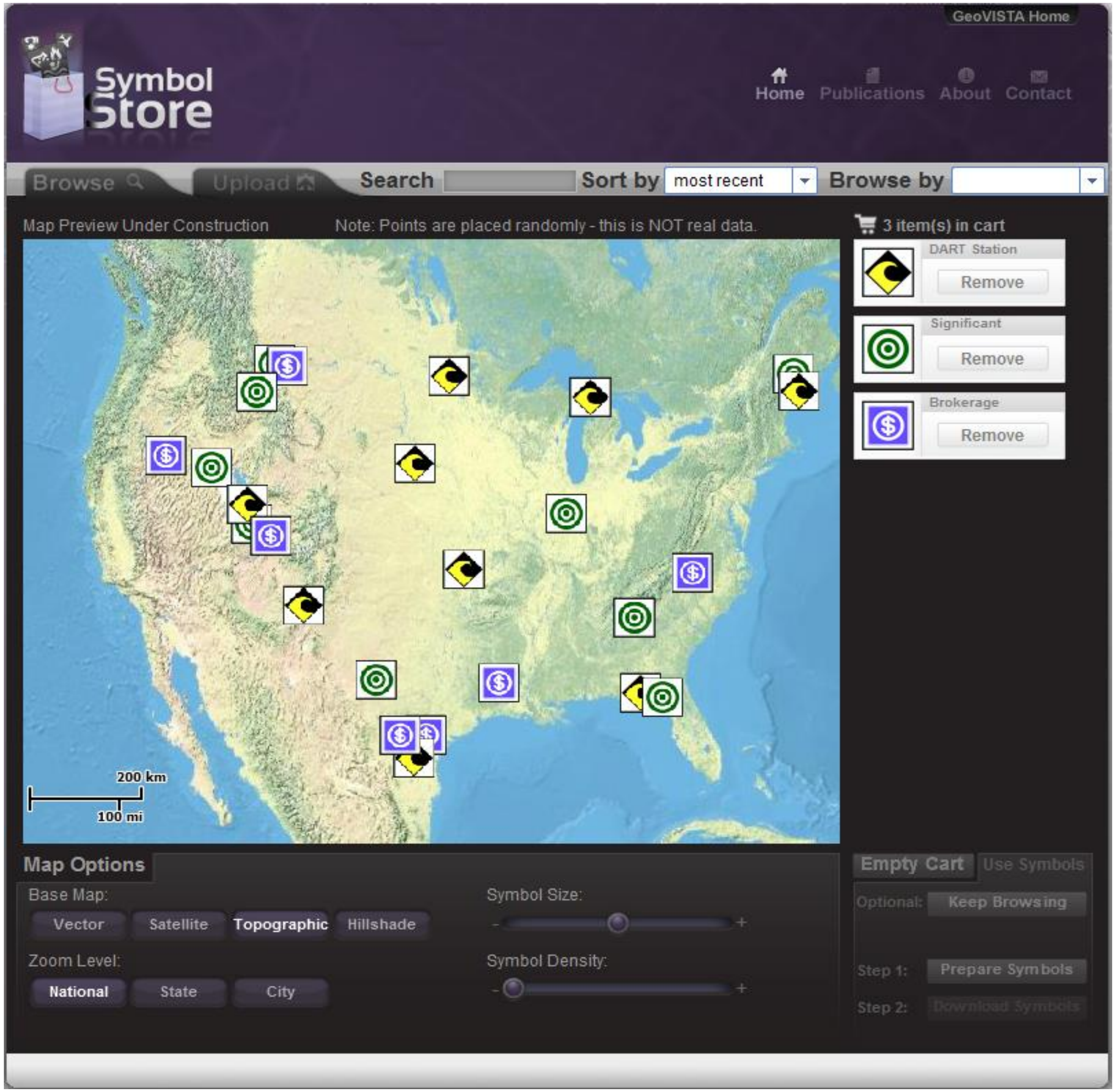
GeoVISTA Center

Search

- Lucene enables more flexible search
  - "Park" will also retrieve "Parking"
  - "Firing" will also retrieve "Fire"
- Next step involves using WordNet to expand queries automatically

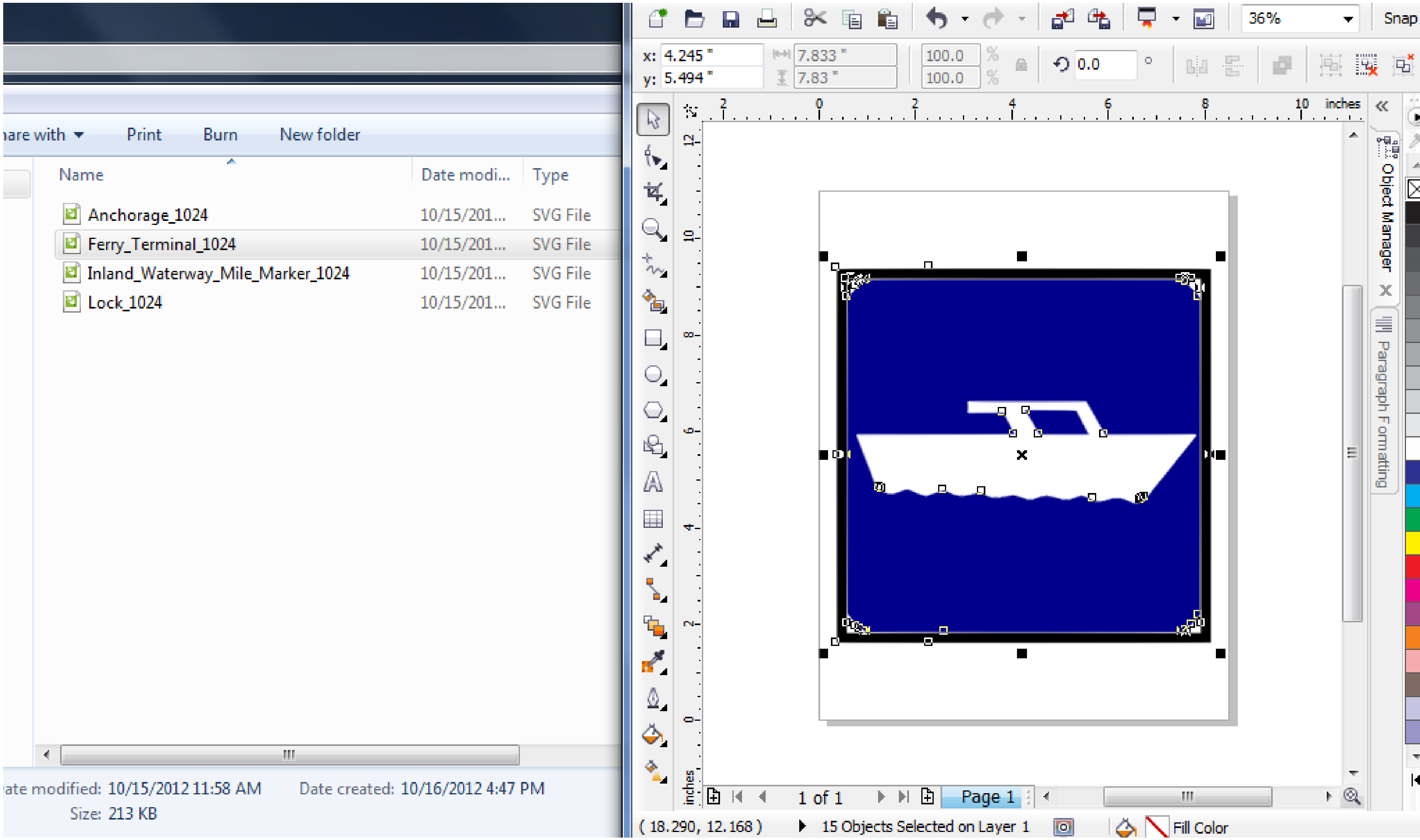


Map Preview



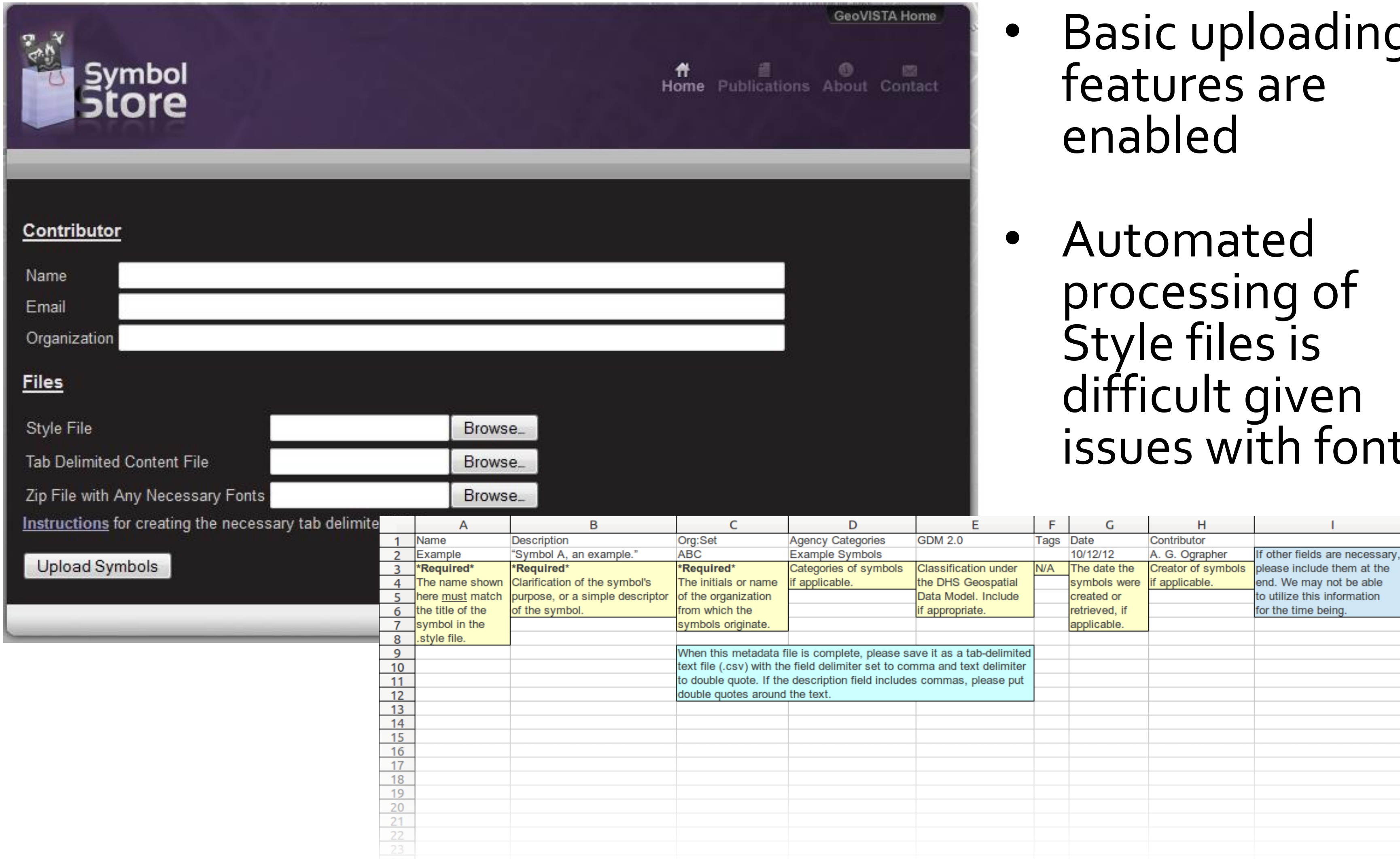
GeoVISTA Center

Symbol Download



GeoVISTA Center

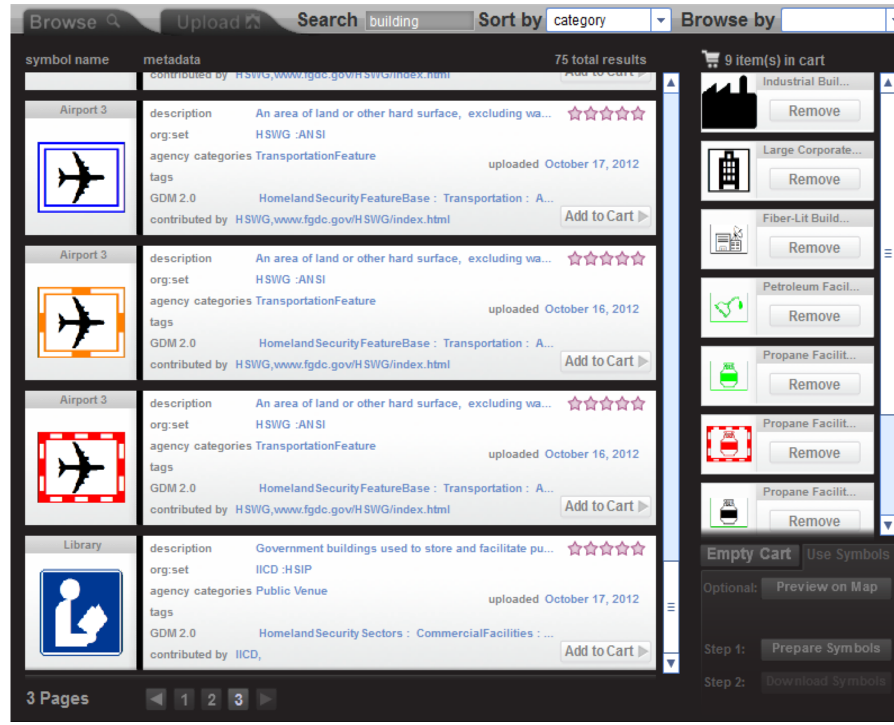
Contribute Symbols



- Basic uploading features are enabled
- Automated processing of Style files is difficult given issues with fonts

What's available

- 2480 point symbols
- Sets from
  - DHS (FEMA, HSIP, IICD…)
  - National Park Service
  - MapBox Maki
  - PSU Mobile Designs
- Support for downloading complete symbol sets coming soon



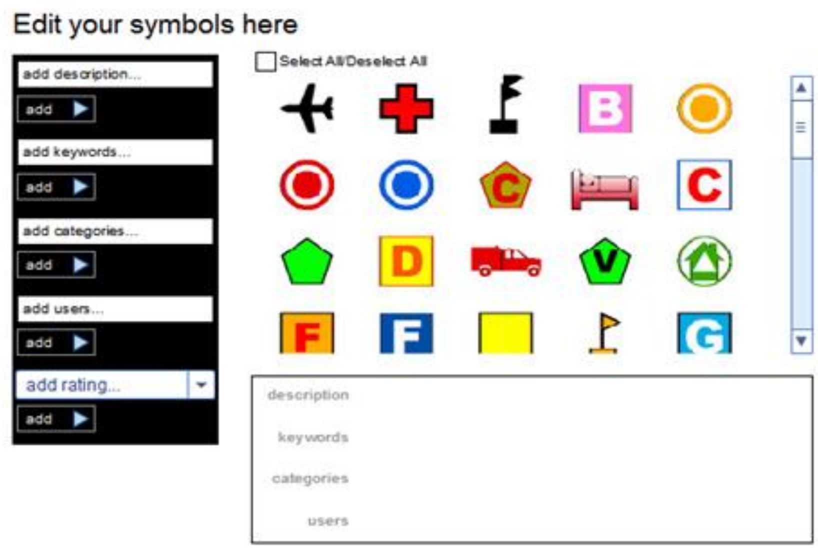
Next steps

Integration plan

- e-Symbology Portal → Symbol Store
- Support small groups of users who:
  - want to pull in symbols from the Symbol Store
  - refine them
  - contribute those changes back to the Symbol Store
- Also support individual power users who:
  - want to serve as the official curator of a symbol set
  - want to contribute metadata for large numbers of symbols
- Workshops with mapmakers to test and refine the integrated tools

Metadata contribution

- Interface and back-end procedures exist to support flexible metadata creation
- Needs to be reskinned and incorporated into the overall process
- One or more symbols can be easily selected and metadata changes applied to all of them.



GeoVISTA Center

Next Steps

- Continue progress on integration between e-Symbology Portal and the Symbol Store
- Develop automated support for other input formats (SVG)
- Collect and post more symbols
- Extend tools to handle dynamic / multi-scale point symbols
- Explore Symbol Store usage/download patterns
- Figure out what to do with duplication issues
  - same symbol, different meaning, for example

Table of Contents

for more information:

www.geovista.psu.edu/symbology/

try it now - www.symbolstore.org

email - arobinson@psu.edu

This work is supported by a contract from the U.S. Department of Homeland Security Science and Technology Directorate, Command, Control and Interoperability Division. The views and opinions expressed here are of the authors, and do not reflect the official positions of the Department of Homeland Security or the Federal Government.

GeoVISTA

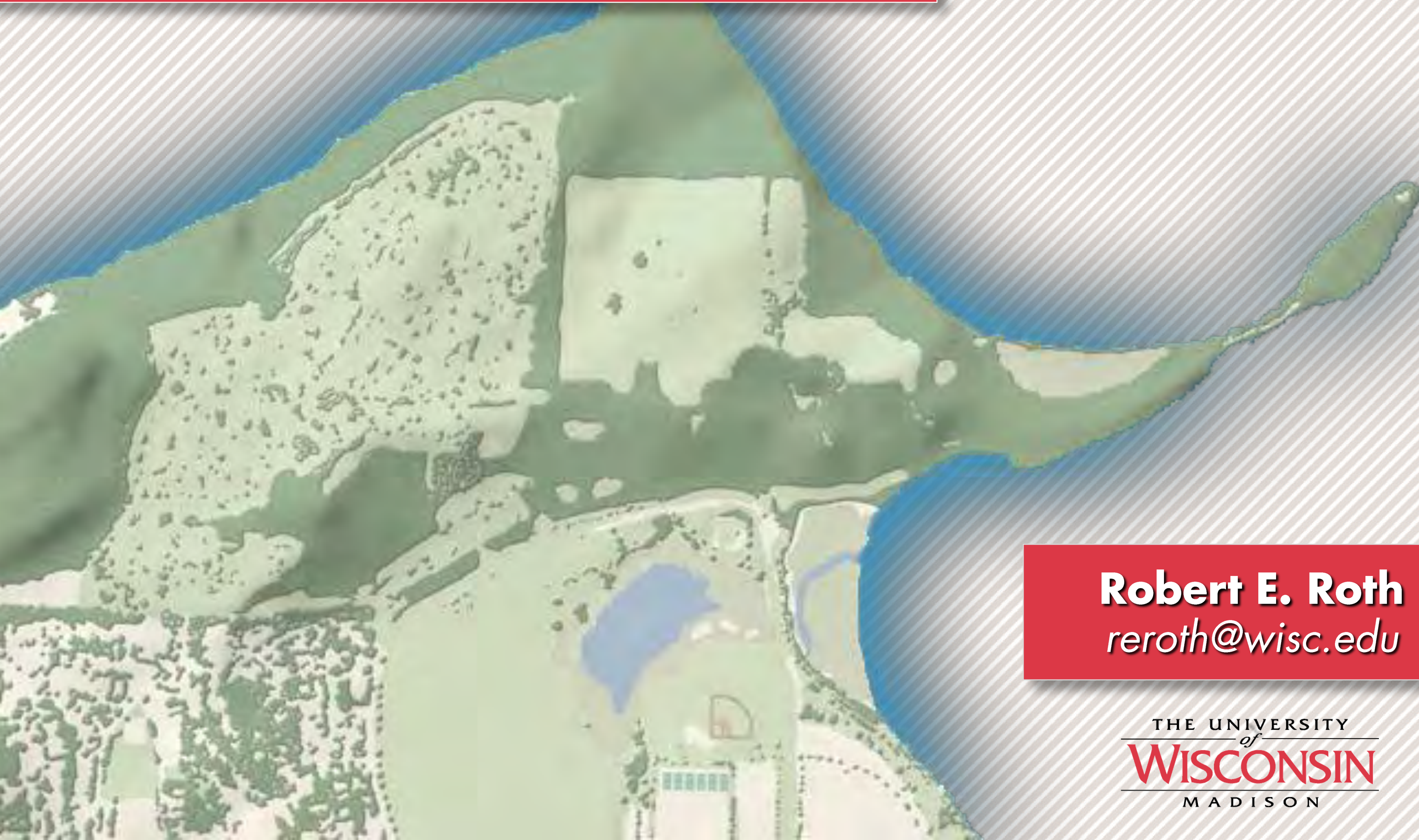
PENNSTATE

1855

The New Cartography

current states of science and technology in Interactive Cartography & Geovisualization

Table of Contents



Robert E. Roth
reroth@wisc.edu

THE UNIVERSITY of WISCONSIN MADISON

CREDIT: National Oceanic and Atmospheric Administration/Department of Commerce

Email Article
Email Editor
Discuss in Forum
Related Articles
Print Article
Free Newsletter
BOOKMARK

Search Articles | Advanced Search | Search

The Job Market

The New Cartographers

By Emily Underwood
March 18, 2013

Twenty years ago, a driver lost at night would pull his car over, take out a paper map bought at a gas station, and pore over its folds under a dim light. With luck and some critical thinking, he would eventually get where he was going. Today, he'd be more likely to swipe his finger across a smart phone screen and follow directions using Google Maps.

> *"[F]uture shortages in cartography, photogrammetry, and geodesy seem likely because the number of graduates is too small (tens to hundreds) to give NGA choices or means of meeting sudden demand." —Future U.S. Workforce for Geospatial Intelligence, a report from the National Research Council*

As maps have changed, so have mapmakers. No longer static images, maps have become active interfaces for information exchange, continuously determining where we are in relation to distant satellites and suggesting where we ought to go, says Seth Spielman, a 38-year-old geography professor at the University of Colorado, Boulder. Today, the global geoservices industry collects, shares, and analyzes data on an unprecedented scale. It's valued at as much as $270 billion per year and employs 500,000 people in the United States, according to a recent report from Google. The

http://sciencecareers.sciencemag.org/career_magazine/

1 Applications — *interactive mapping*

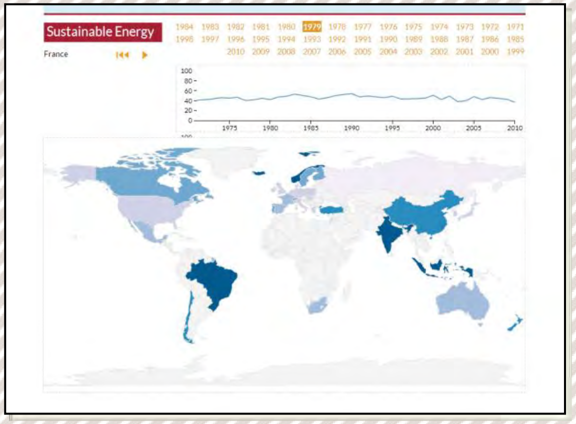
Science — *cartographic interaction* 2

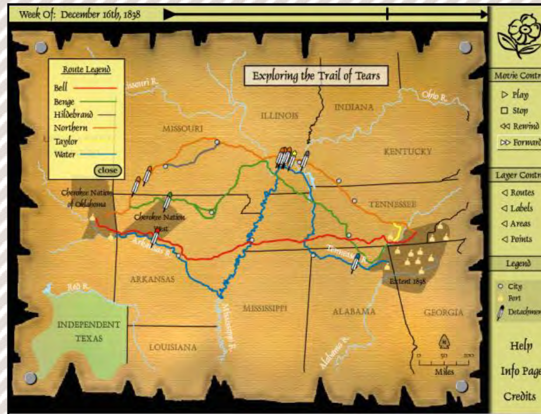
3 Technology — *emerging mapping tech*

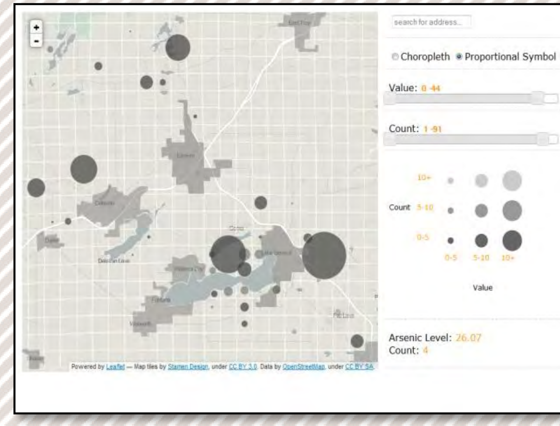
1

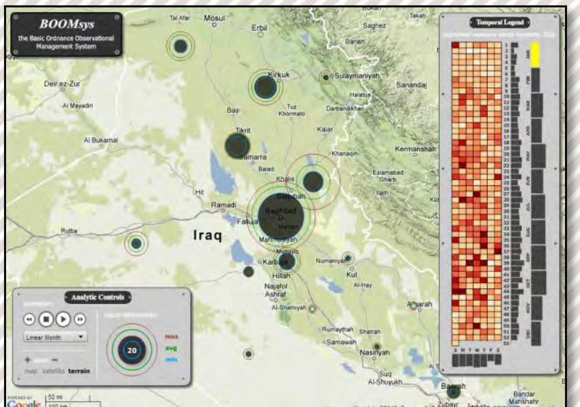
Applications

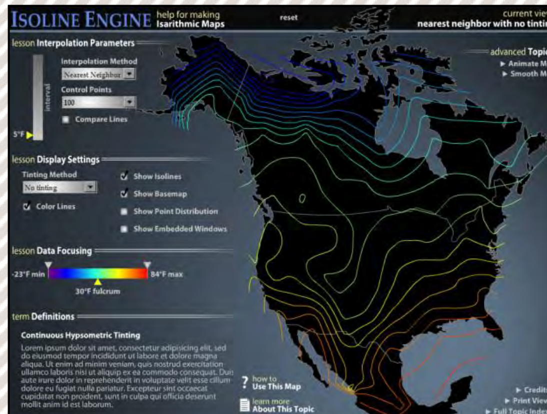
interactive cartography & geovisualization

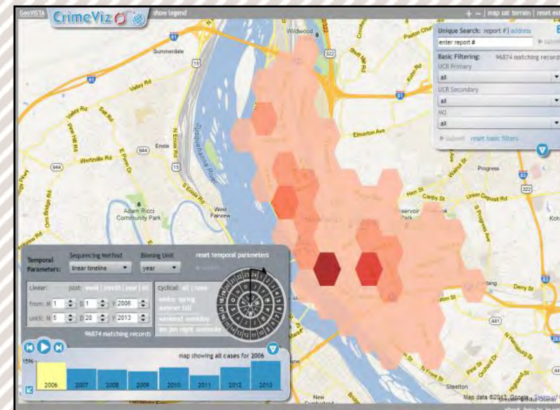

Sustainable Energy
France

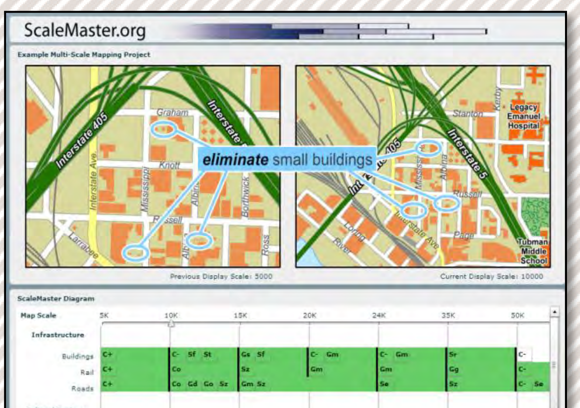

Week Of: December 16th, 1838
Exploring the Trail of Tears
Route Legend
Bell
Benge
Hildebrand
Northern
Taylor
Water
Movie Control
Play
Stop
Rewind
Forward
Layer Control
Routes
Labels
Areas
Points
Legend
Help
Info Page
Credits


Choropleth
Proportional Symbol
Value:
Count:
Arsenic Level: 26.07
Count: 4

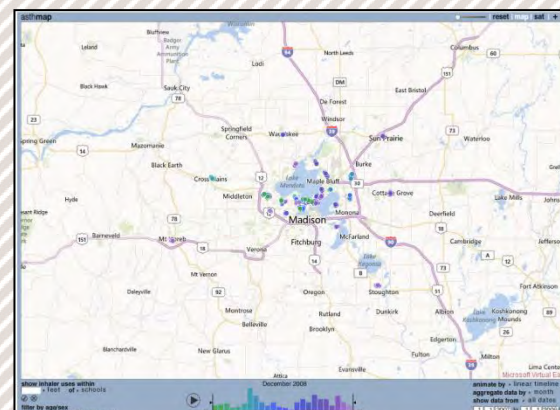

BOOMsys
the Basic Ordnance Observational Management System
Temporal Legend
Analytic Controls
Iraq


ISOLINE ENGINE
help for making isarithmic Maps
nearest neighbor with no tinting
lesson Interpolation Parameters
lesson Display Settings
lesson Data Focusing
term Definitions
Continuous Hypsometric Tinting

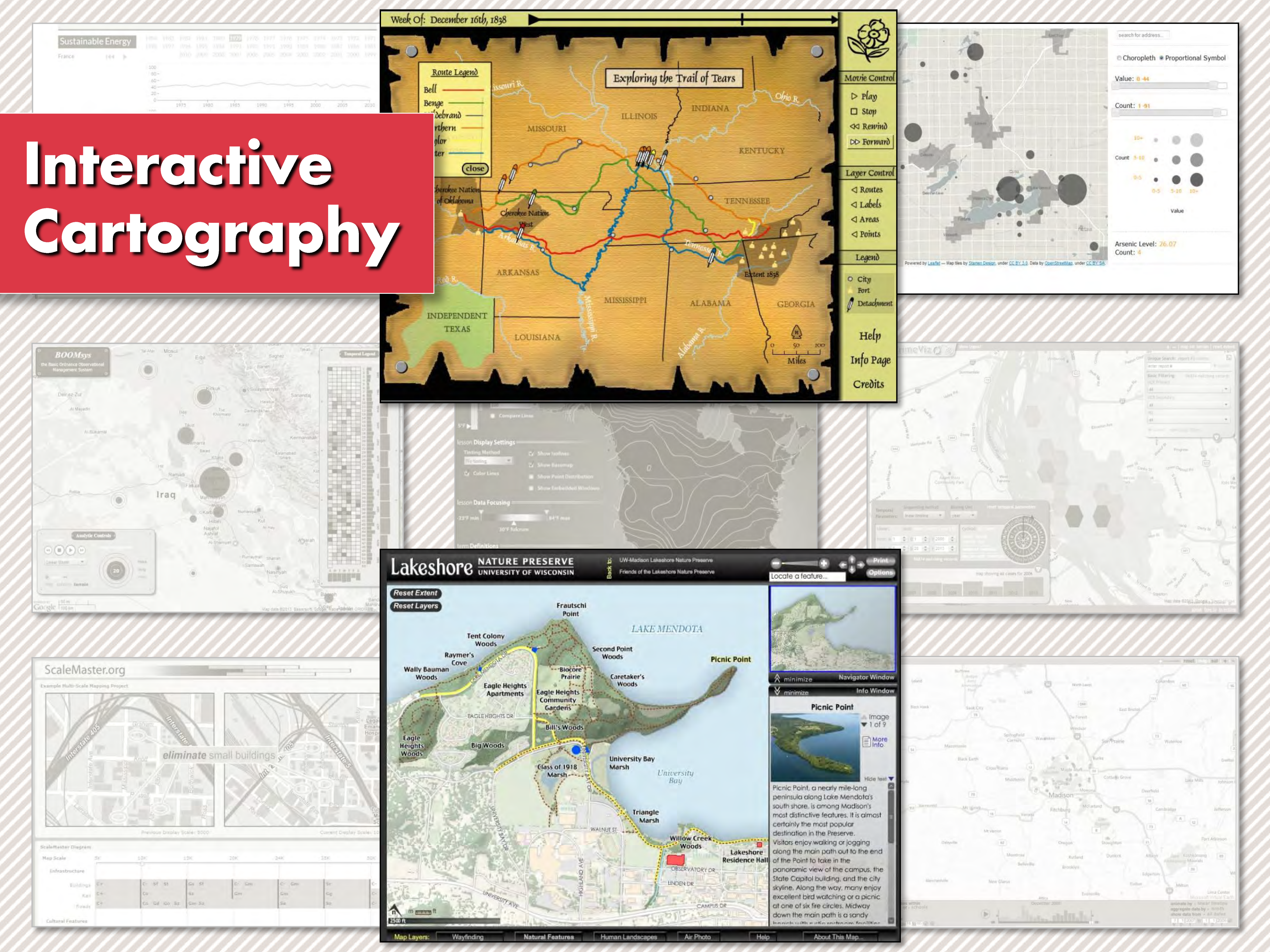

CrimeViz


ScaleMaster.org
Example Multi-Scale Mapping Project
eliminate small buildings
Previous Display Scale: 5000
Current Display Scale: 10000
ScaleMaster Diagram
Map Scale
Infrastructure
Buildings
Rail
Roads
Cultural Features


Lakeshore NATURE PRESERVE
UNIVERSITY OF WISCONSIN
Reset Extent
Reset Layers
LAKE MENDOTA
Picnic Point
Navigator Window
Info Window

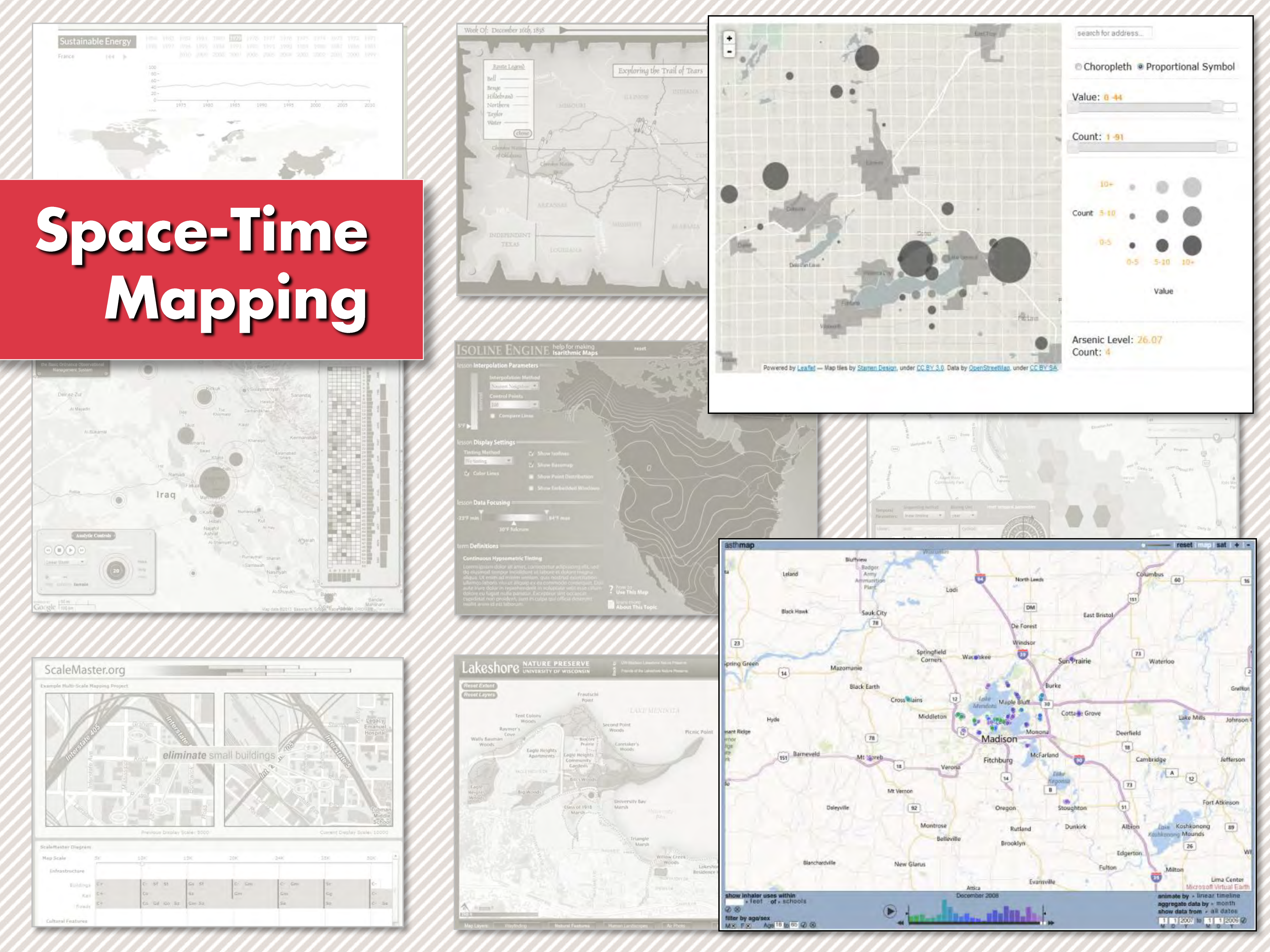

Madison

Interactive Cartography

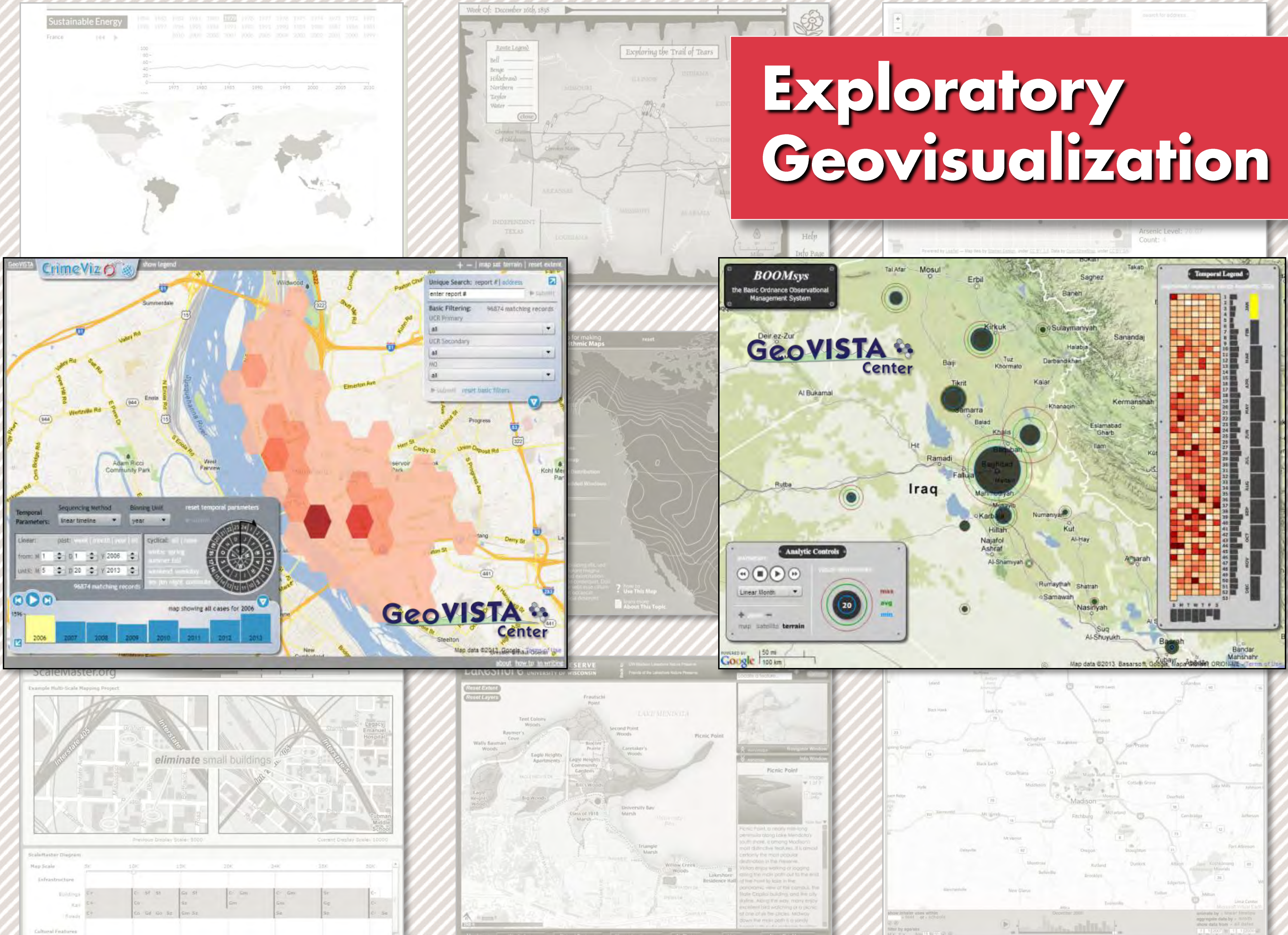


Educational Tools

Space-Time Mapping



Exploratory Geovisualization



Needs Assessment Interviews

geospatial technology unmet needs & barriers to use

Science of Cartographic Interaction (*USGIF; Doctoral*)

- **Focus:** Interactive Cartography & Geovisualization
- **Participants:** Sampling of application domains, including Emergency Response/Crisis Management (n=4) and Intelligence Analysis (n=3)

Symbology Development Best Practices (*DHS*)

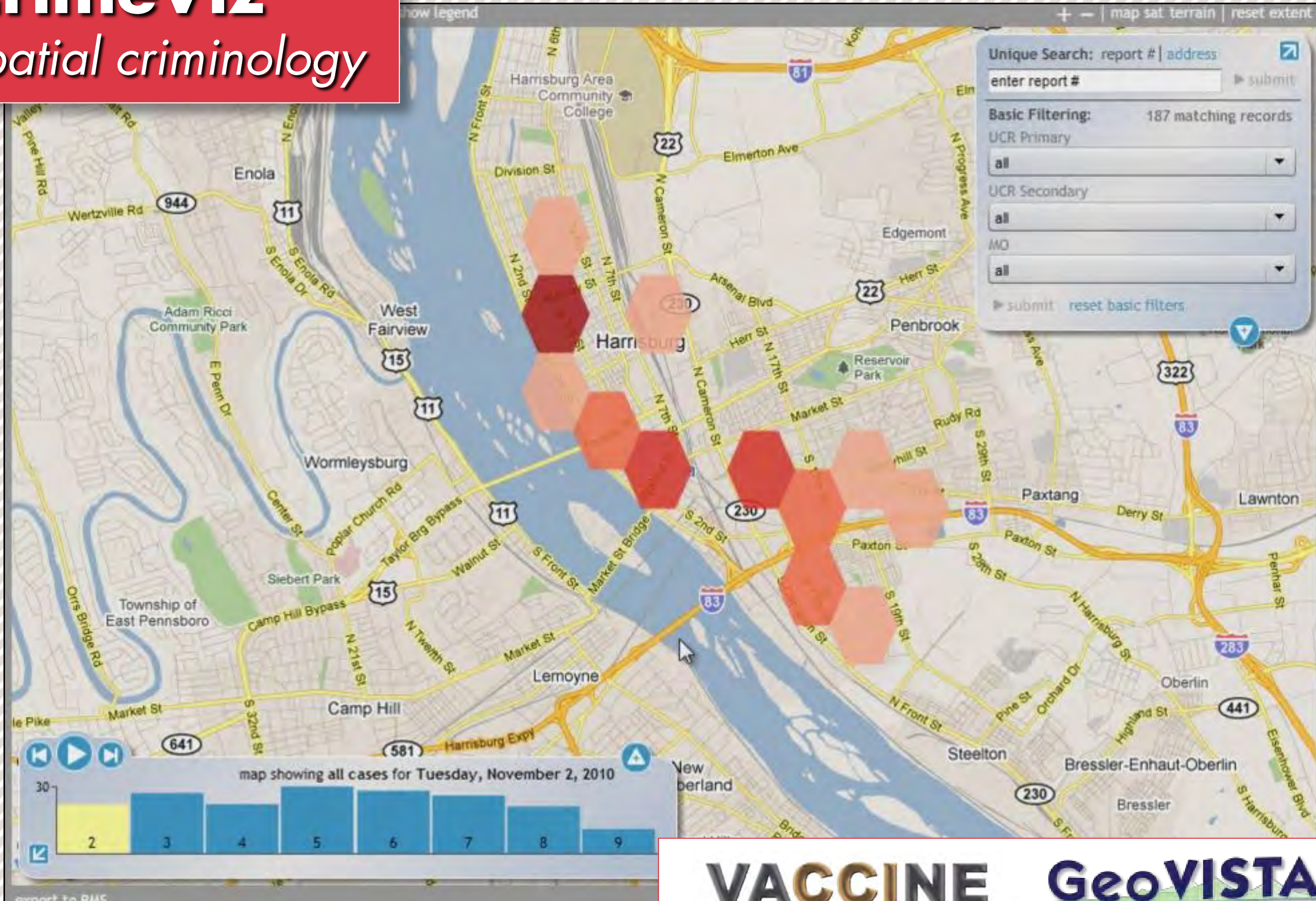
- **Focus:** ANSI INCITS 415-2006 symbol standard
- **Participants:** CBP, IICD, USCG, NOC, FEMA, USFS, DNDO (n=14)

VACCINE (*DHS Center of Excellence*)

- **Focus:** Spatial Criminology & Crime Mapping
- **Participants:** municipal law enforcement agencies, NIJ (n=9)

CrimeViz

spatial criminology



Roth et al. (2013a)

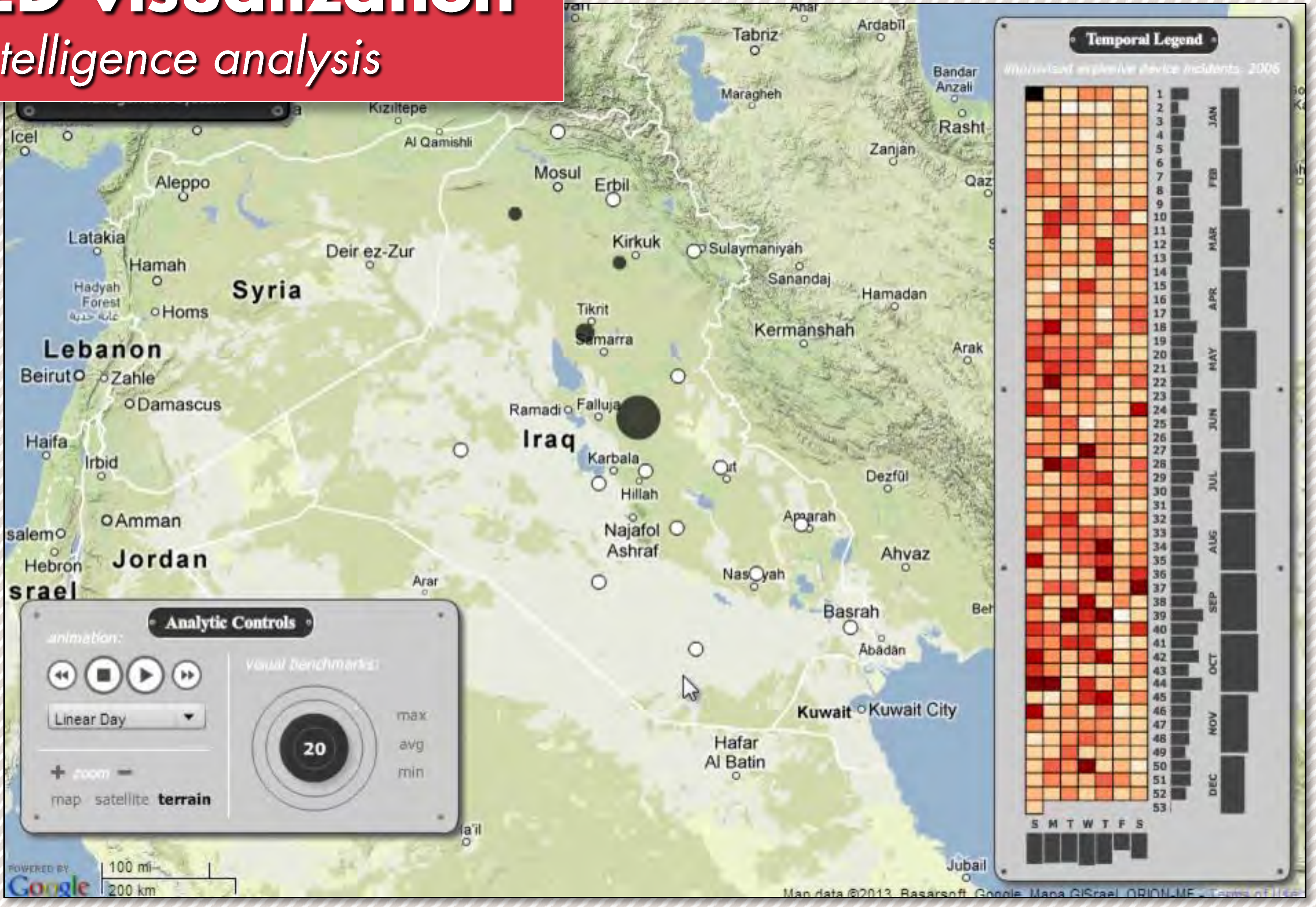
VACCINE

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

GeoVISTA Center

IED visualization

intelligence analysis



NCTC Worldwide Incident Tracking System

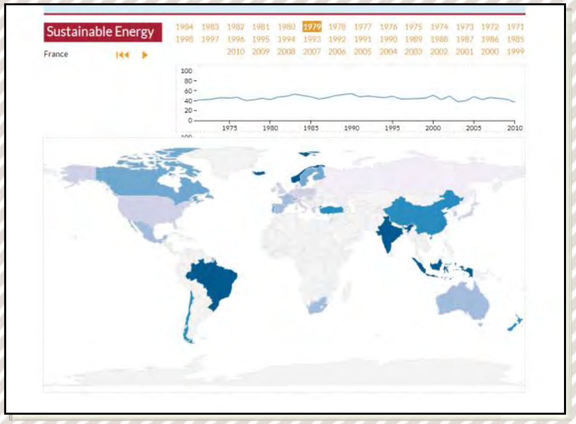
Murdock et al. (2012)

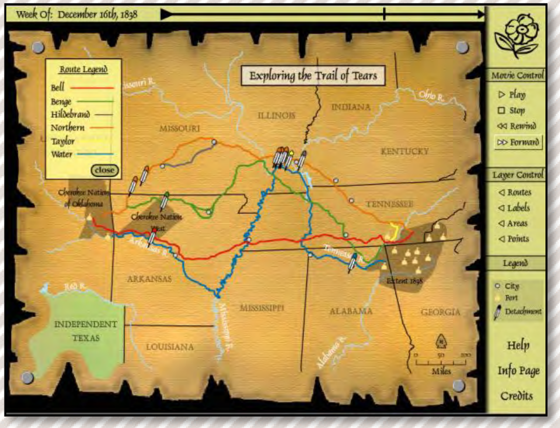
2

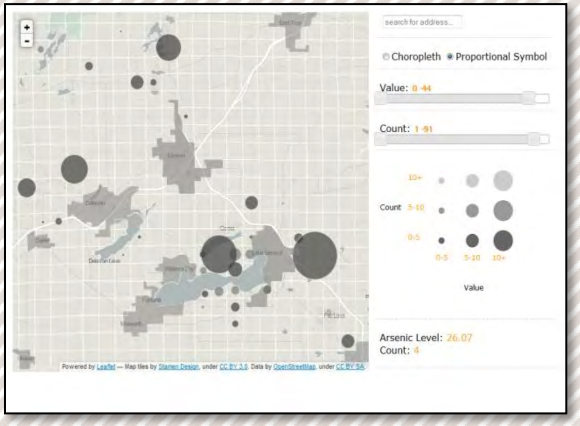
Science

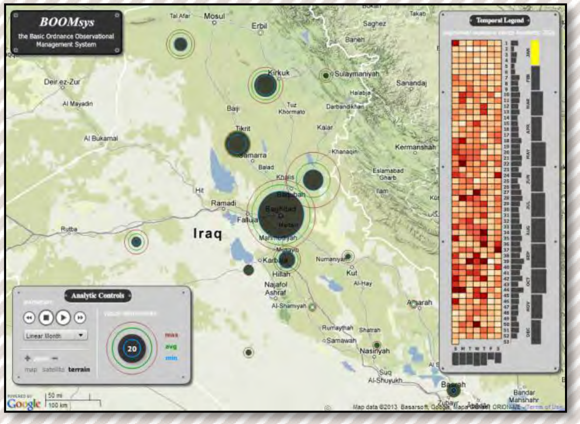
towards a science of cartographic interaction

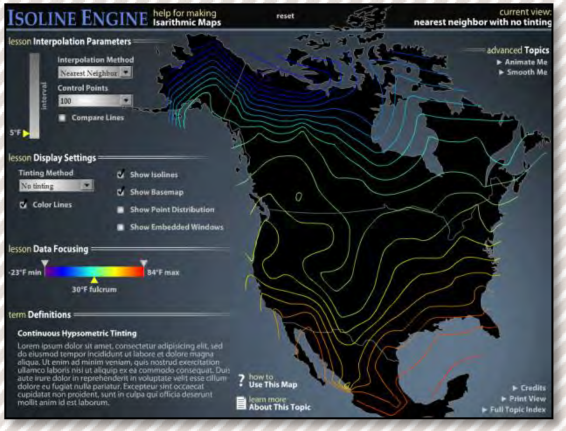
THE UNIVERSITY of WISCONSIN MADISON

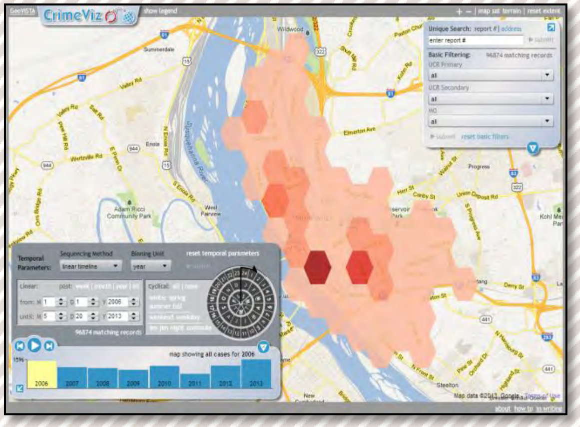

Sustainable Energy
France


Week Of: December 16th, 1838
Exploring the Trail of Tears
Route Legend
Movie Control
Layer Control
Legend
Help
Info Page
Credits


Choropleth
Proportional Symbol
Value:
Count:
Arsenic Level: 26.07
Count: 4

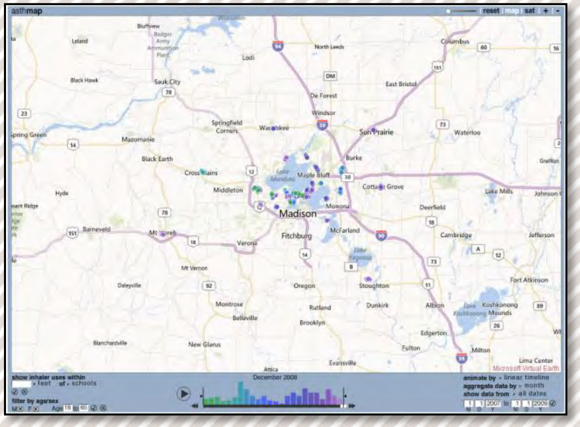

BOOMsys
the Basic Ordnance Observational Management System
Temporal Legend
Analytic Controls
Iraq


ISOLINE ENGINE
help for making Isarithmic Maps
lesson Interpolation Parameters
lesson Display Settings
lesson Data Focusing
term Definitions
Continuous Hypsometric Tinting


CrimeViz


ScaleMaster.org
Example Multi-Scale Mapping Project
eliminate small buildings
Previous Display Scale: 5000
Current Display Scale: 10000
ScaleMaster Diagram

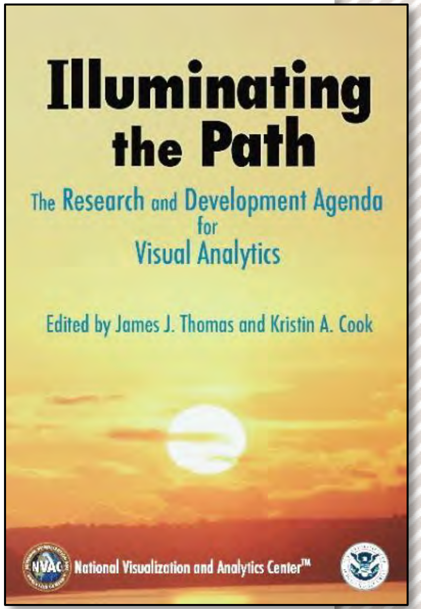

Lakeshore NATURE PRESERVE
UNIVERSITY OF WISCONSIN
LAKE MENDOTA
Picnic Point


Madison


Sustainable Energy
Exploring the Trail of Tears
Choropleth
Proportional Symbol
BOOMsys
Iraq
ISOLINE ENGINE
help for making Isarithmic Maps
Interpolation Parameters
Display Settings
Data Focusing
Definitions
CrimeViz
ScaleMaster.org
Lakeshore
NATURE PRESERVE
UNIVERSITY OF WISCONSIN
Picnic Point

The Grand Challenge
for creating a new science of interaction

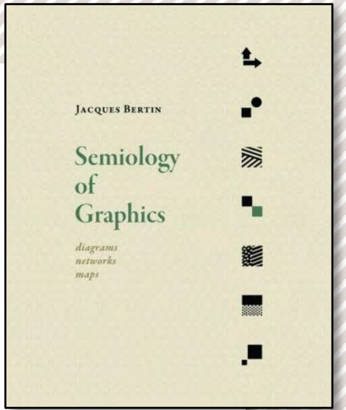
"**Recommendation 3.3:** Create a new science of interaction to support visual analytics. The grand challenge of interaction is to develop a ***taxonomy to describe the design space of interaction techniques*** that supports the science of analytic reasoning. We must characterize this design space and identify under-explored areas that are relevant to visual analytics. Then, R&D should be focused on ***expanding the repertoire of interaction techniques*** that can fill those gaps in the design space."

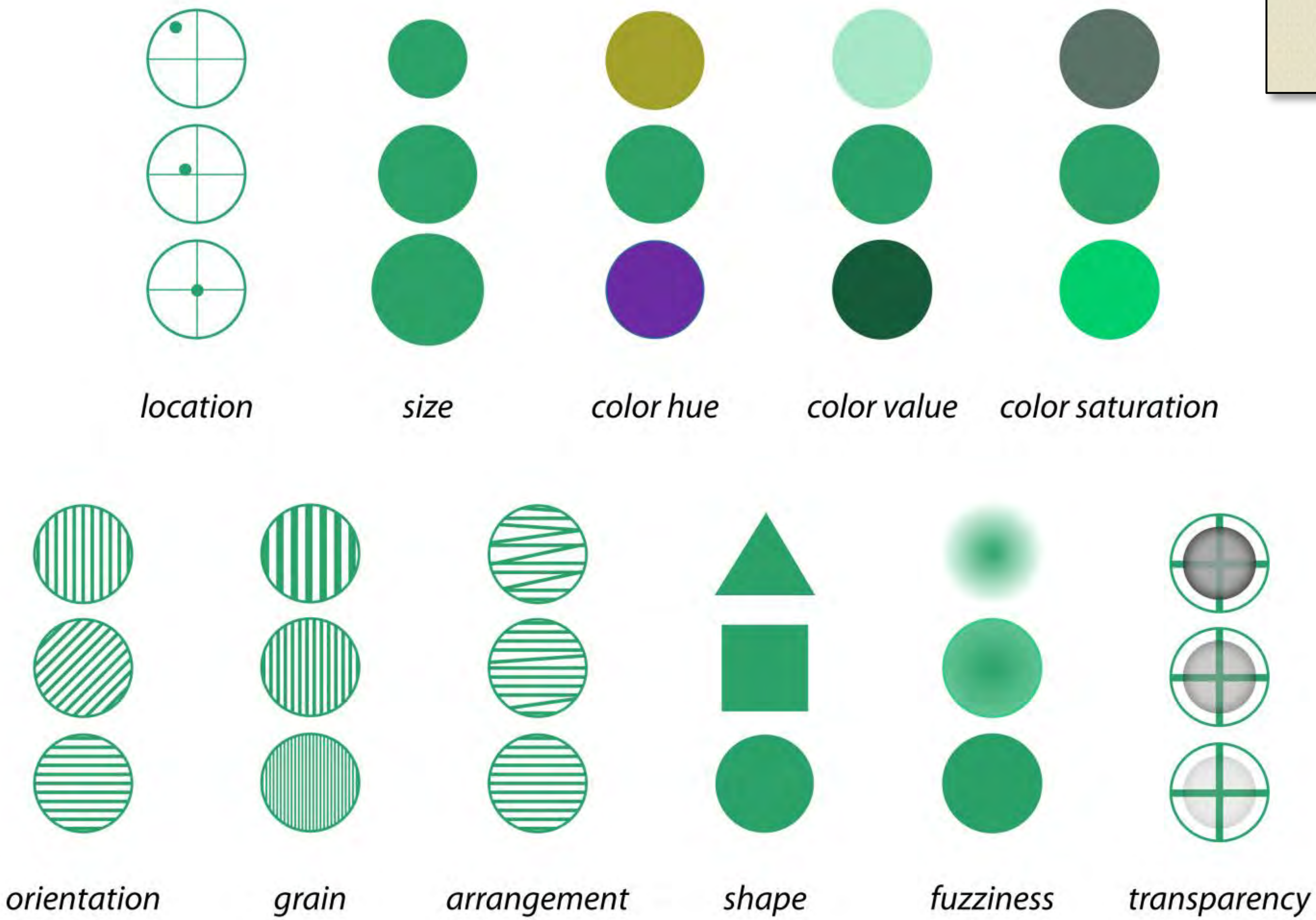


Thomas et al. (2005)

Visual Variables
how we represent

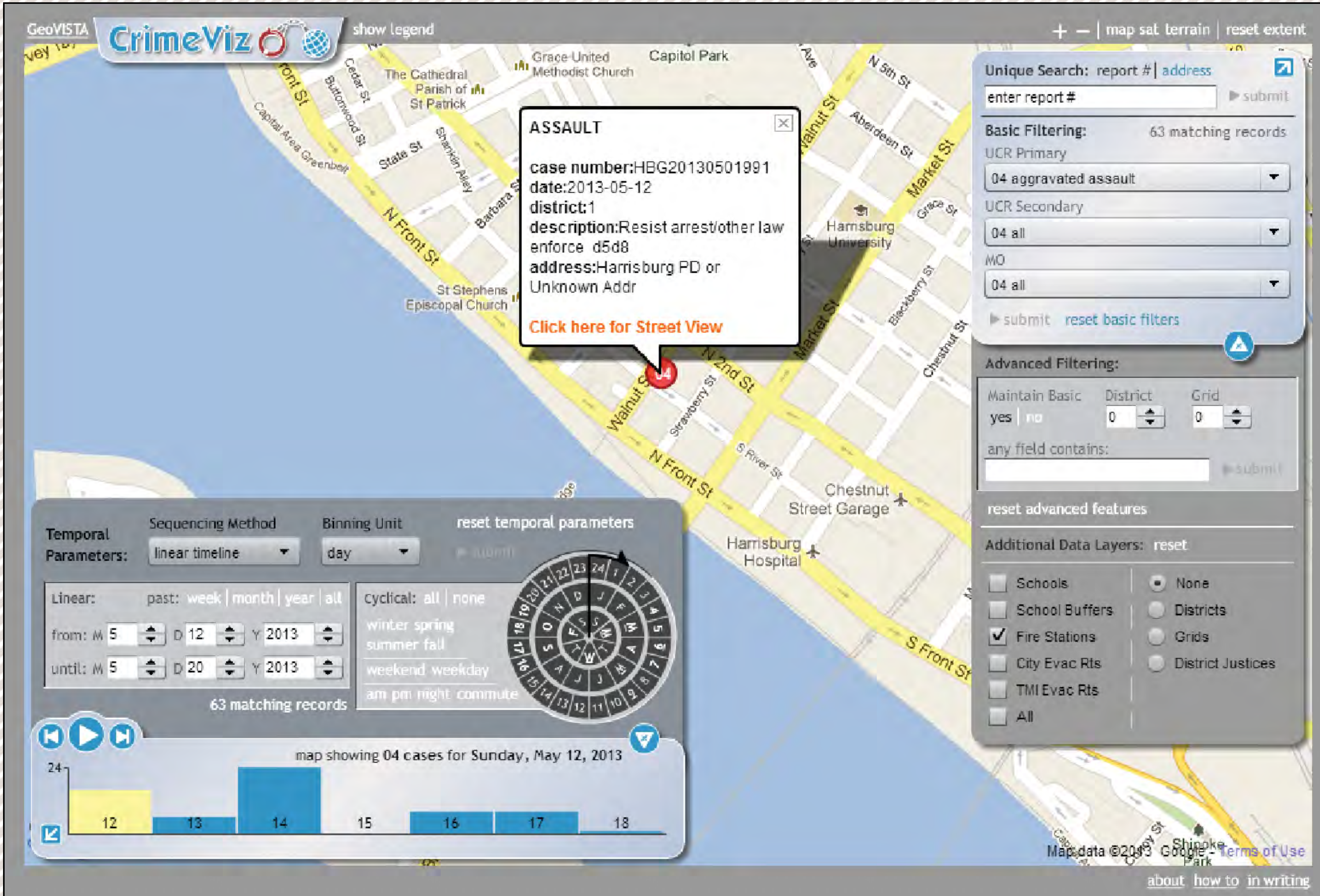
Bertin (1967)





Interaction Primitives
how we interact

Roth (2012)


GeoVISTA
CrimeViz
show legend
+ – | map sat terrain | reset extent
ASSAULT
case number:HBG20130501991
date:2013-05-12
district:1
description:Resist arrest/other law enforce d5d8
address:Harrisburg PD or Unknown Addr
Click here for Street View
Unique Search: report # | address
enter report #
submit
Basic Filtering: 63 matching records
UCR Primary
04 aggravated assault
UCR Secondary
04 all
MO
04 all
submit reset basic filters
Advanced Filtering:
Maintain Basic yes | no
District 0
Grid 0
any field contains:
submit
reset advanced features
Additional Data Layers: reset
Schools
School Buffers
Fire Stations
City Evac Rts
TMI Evac Rts
All
None
Districts
Grids
District Justices
Temporal Parameters:
Sequencing Method linear timeline
Binning Unit day
reset temporal parameters
submit
Linear: past: week | month | year | all
from: M 5 D 12 Y 2013
until: M 5 D 20 Y 2013
63 matching records
cyclical: all | none
winter spring summer fall
weekend weekday
am pm night commute
map showing 04 cases for Sunday, May 12, 2013
24
12
13
14
15
16
17
18
Capitol Park
Grace United Methodist Church
The Cathedral Parish of St Patrick
St Stephens Episcopal Church
Harrisburg University
Chestnut Street Garage
Harrisburg Hospital
N Front St
S Front St
Market St
Walnut St
N 2nd St
State St
Map data ©2013 Google - Terms of Use
about how to in writing

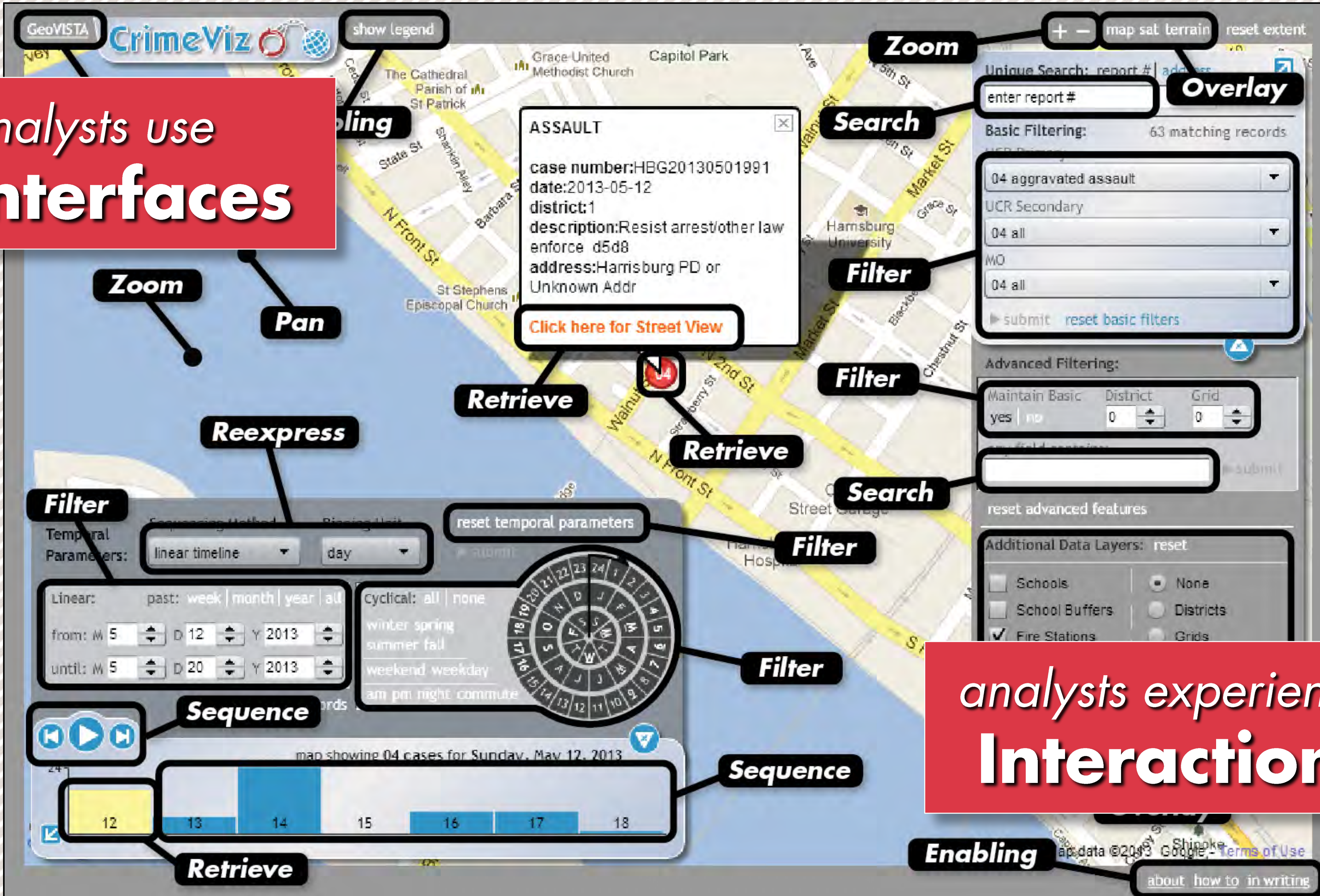
GeoVISTA
CrimeViz
show legend
Enabling
Enabling
Zoom
Pan
Reexpress
Filter
Temporal Parameters:
linear timeline
day
reset temporal parameters
Linear: past: week | month | year | all
from: M 5 D 12 Y 2013
until: M 5 D 20 Y 2013
Cyclical: all | none
winter spring summer fall
weekend weekday
am pm night commute
Sequence
map showing 04 cases for Sunday, May 12, 2013
12 13 14 15 16 17 18
Retrieve
Sequence
Filter
Filter
ASSAULT
case number:HBG20130501991
date:2013-05-12
district:1
description:Resist arrest/other law enforce d5d8
address:Harrisburg PD or Unknown Addr
Click here for Street View
Retrieve
Retrieve
Zoom
map sat terrain
reset extent
Overlay
Search
enter report #
Basic Filtering: 63 matching records
04 aggravated assault
UCR Secondary
04 all
MO
04 all
submit reset basic filters
Filter
Advanced Filtering:
Filter
Maintain Basic District Grid
yes | no
Search
reset advanced features
Additional Data Layers: reset
Schools
School Buffers
Fire Stations
City Evac Rts
TMI Evac Rts
All
None
Districts
Grids
District Justices
Overlay
Enabling
about how to in writing

analysts use

Interfaces

analysts use

Interfaces



analysts experience

Interactions

cartographic interaction primitives for

User Experience (UX) Design

objectives *(user tasks)*	**operators** *(functionality)*	**operands** *(geospatial data)*

objectives *(user tasks)*

identify
compare
rank
associate
delineate

operators *(functionality)*

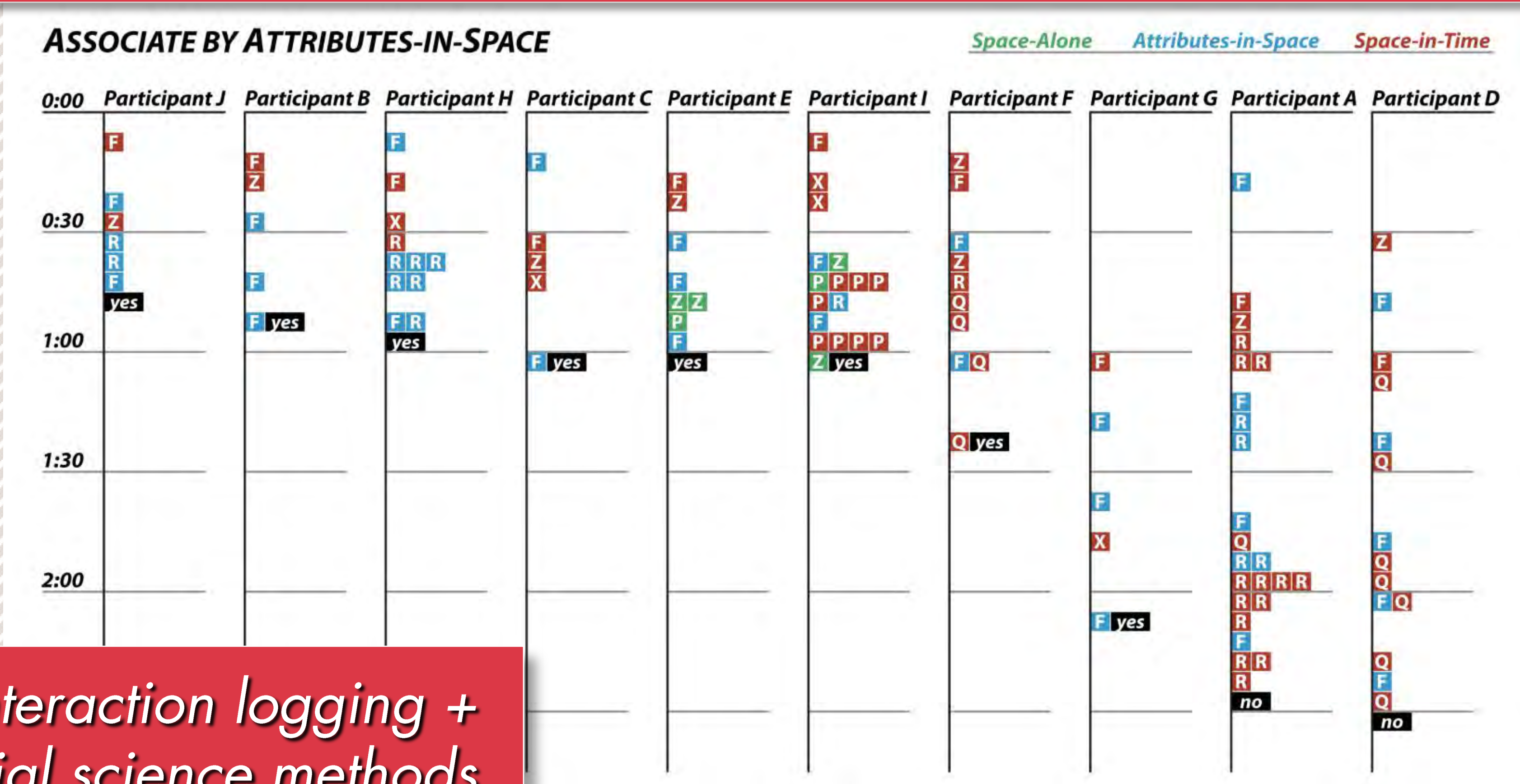
reexpress	*pan*
arrange	*zoom*
sequence	*filter*
resymbolize	*search*
overlay	*retrieve*
reproject	*calculate*

operands *(geospatial data)*

space
time
attribute

The Grand(er) Challenge

a syntactics of cartographic interaction primitives



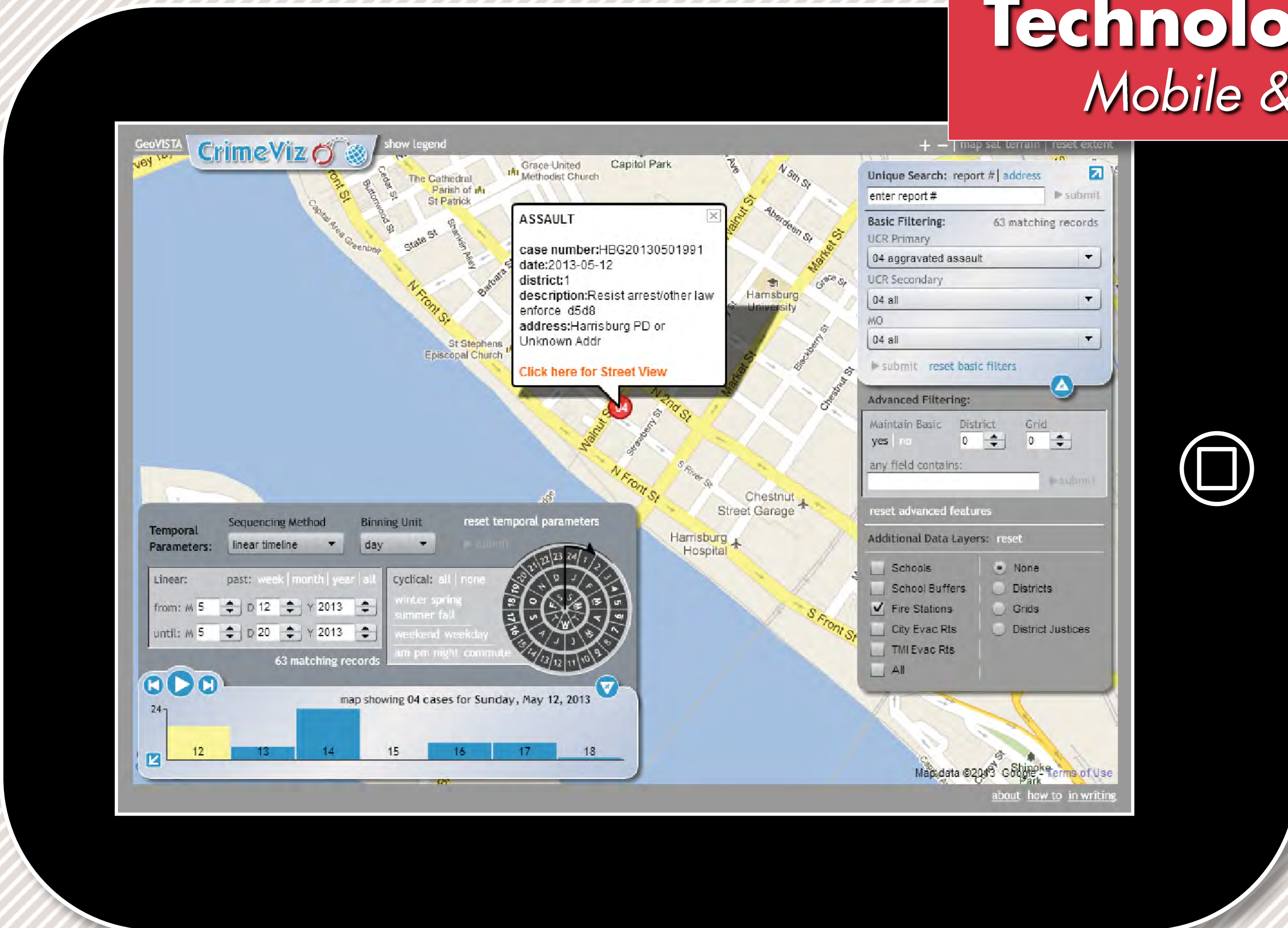
interaction logging + social science methods

Roth (forthcoming)

3 Technology

emerging technology for browser-based mapping

Technology
Mobile & LBS



Technology
Mobile & LBS

Flash Player is not available for this device

Open Source

browser-based technologies

TileMill
an open source project by MapBox

GeoMoose

Tile Stache

mapnik

MapServer
open source web mapping

NOKIA

Wax
from MapBox

Mapbender3

mapquest developers

{modest maps}

KARTOGRAPH

Processing

TILE DRAWER

Leaflet

jump

ViaMichelin

CLOUD MADE

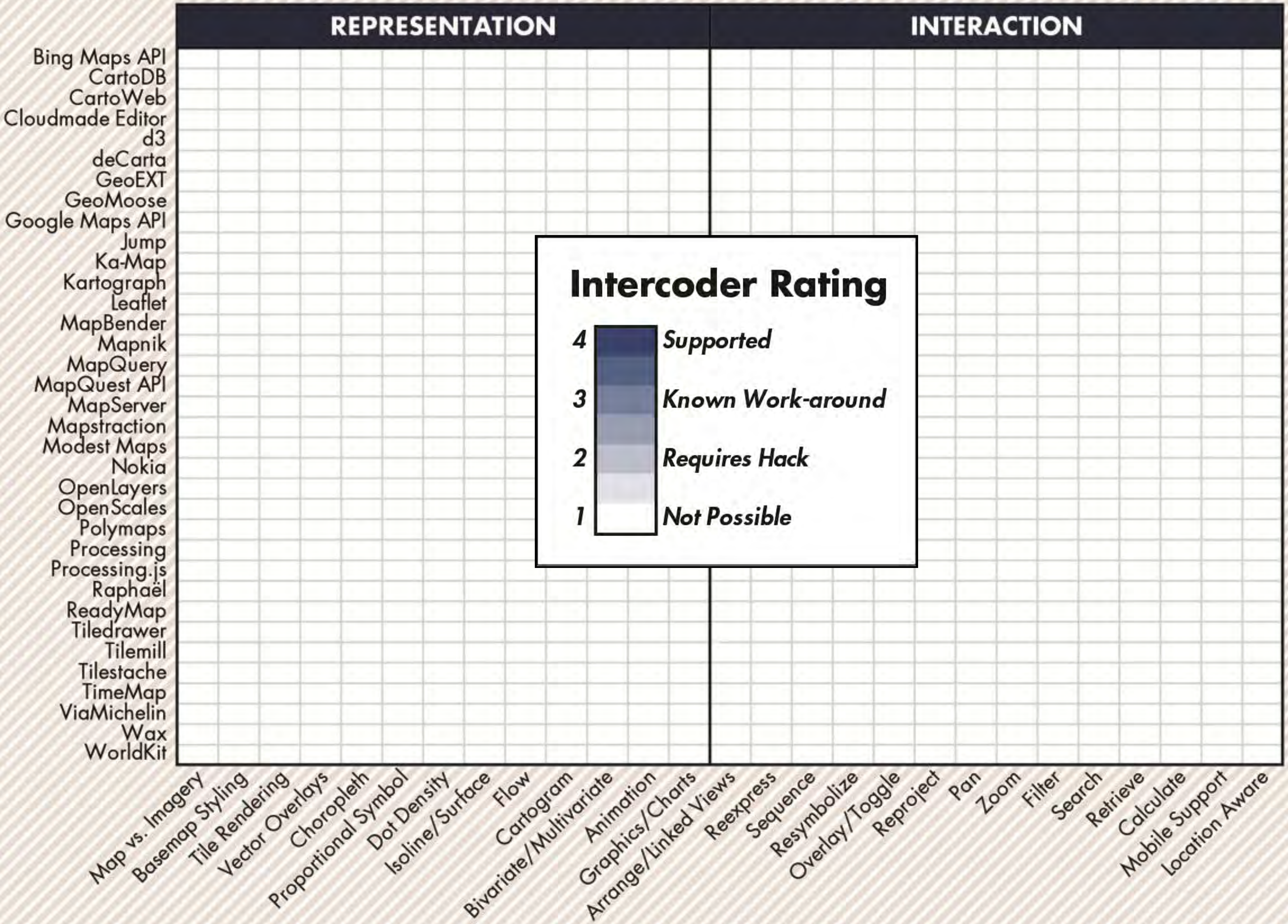
MapQuery

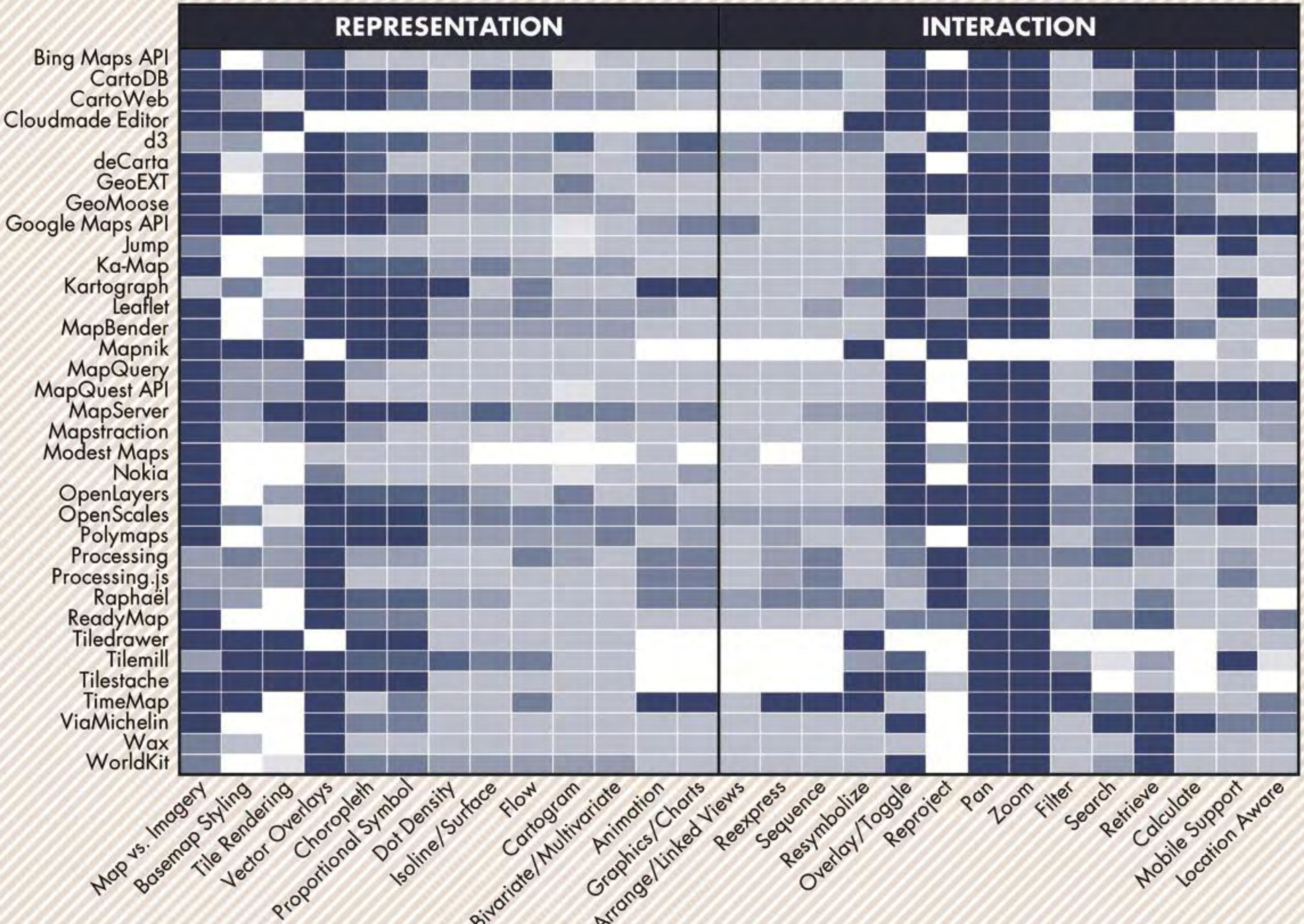
MAPSTRACTION
Javascript mapping abstraction library

OpenLayers

Bing Maps API
CartoDB
CartoWeb
Cloudmade Editor
d3
deCarta
GeoEXT
GeoMoose
Google Maps API
Jump
Ka-Map
Kartograph
Leaflet
MapBender
Mapnik
MapQuery
MapQuest API
MapServer
Mapstraction
Modest Maps
Nokia
OpenLayers
OpenScales
Polymaps
Processing
Processing.js
Raphaël
ReadyMap
Tiledrawer
Tilemill
Tilestache
TimeMap
ViaMichelin
Wax
WorldKit

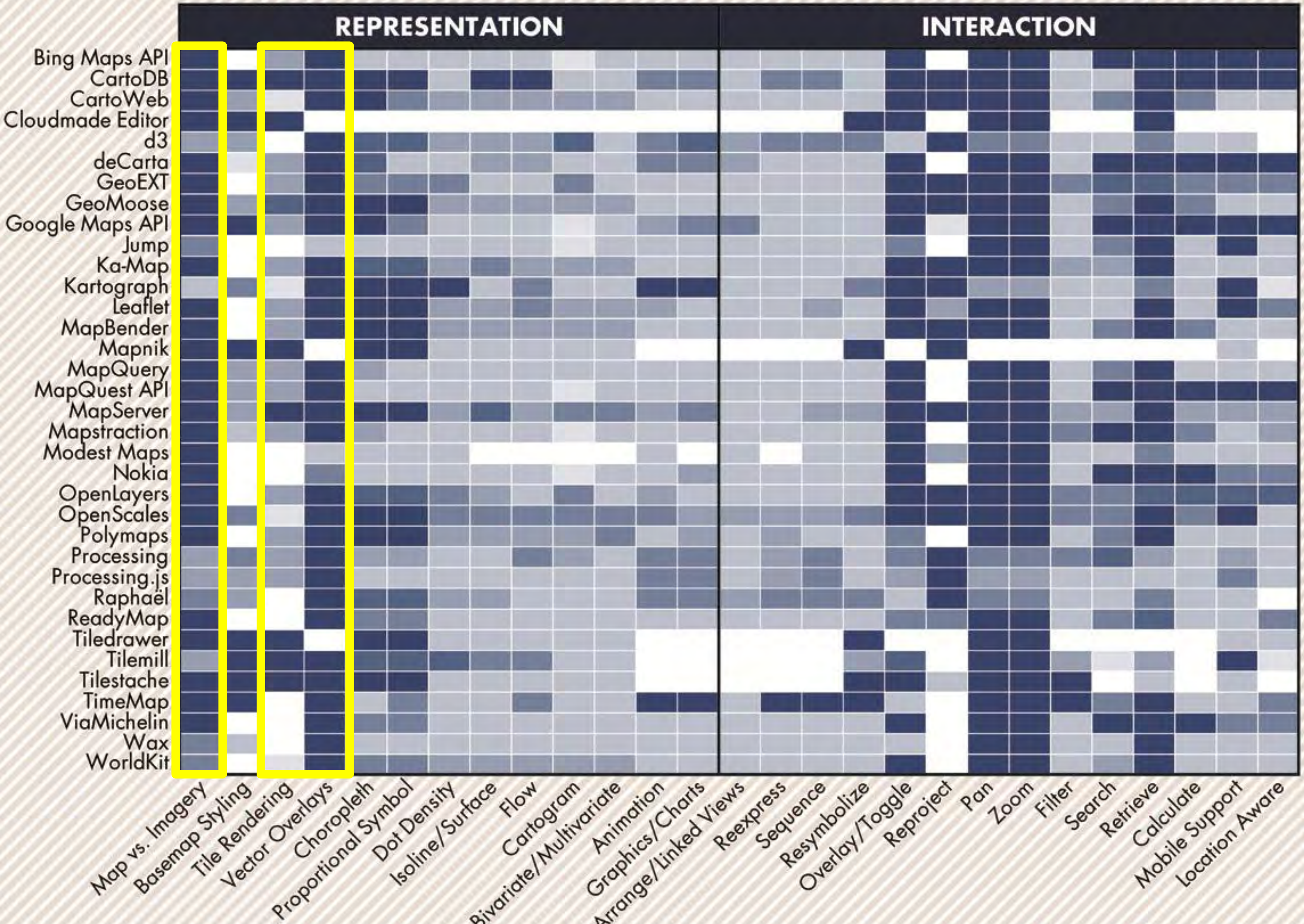
	REPRESENTATION													INTERACTION													
	Map vs. Imagery	Basemap Styling	Tile Rendering	Vector Overlays	Choropleth	Proportional Symbol	Dot Density	Isoline/Surface	Flow	Cartogram	Bivariate/Multivariate	Animation	Graphics/Charts	Arrange/Linked Views	Reexpress	Sequence	Resymbolize	Overlay/Toggle	Reproject	Pan	Zoom	Filter	Search	Retrieve	Calculate	Mobile Support	Location Aware
Bing Maps API																											
CartoDB																											
CartoWeb																											
Cloudmade Editor																											
d3																											
deCarta																											
GeoEXT																											
GeoMoose																											
Google Maps API																											
Jump																											
Ka-Map																											
Kartograph																											
Leaflet																											
MapBender																											
Mapnik																											
MapQuery																											
MapQuest API																											
MapServer																											
Mapstraction																											
Modest Maps																											
Nokia																											
OpenLayers																											
OpenScales																											
Polymaps																											
Processing																											
Processing.js																											
Raphaël																											
ReadyMap																											
Tiledrawer																											
Tilemill																											
Tilestache																											
TimeMap																											
ViaMichelin																											
Wax																											
WorldKit																											


REPRESENTATION
INTERACTION
Bing Maps API
CartoDB
CartoWeb
Cloudmade Editor
d3
deCarta
GeoEXT
GeoMoose
Google Maps API
Jump
Ka-Map
Kartograph
Leaflet
MapBender
Mapnik
MapQuery
MapQuest API
MapServer
Mapstraction
Modest Maps
Nokia
OpenLayers
OpenScales
Polymaps
Processing
Processing.js
Raphaël
ReadyMap
Tiledrawer
Tilemill
Tilestache
TimeMap
ViaMichelin
Wax
WorldKit
Map vs. Imagery
Basemap Styling
Tile Rendering
Vector Overlays
Choropleth
Proportional Symbol
Dot Density
Isoline/Surface
Flow
Cartogram
Bivariate/Multivariate
Animation
Graphics/Charts
Arrange/Linked Views
Reexpress
Sequence
Resymbolize
Overlay/Toggle
Reproject
Pan
Zoom
Filter
Search
Retrieve
Calculate
Mobile Support
Location Aware
Intercoder Rating
4 Supported
3 Known Work-around
2 Requires Hack
1 Not Possible



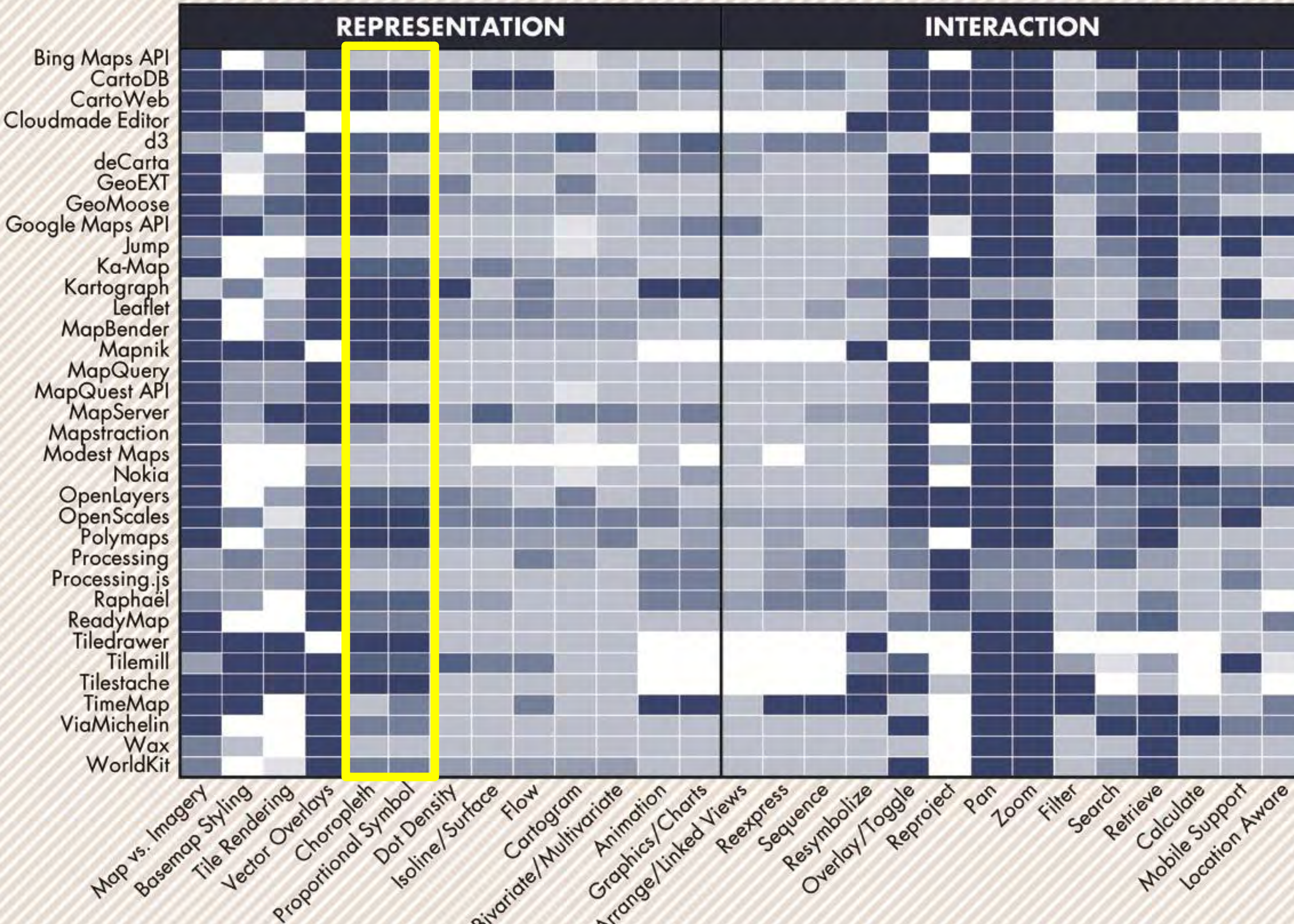
Coding Completed 2012

Roth et al. (2013b)



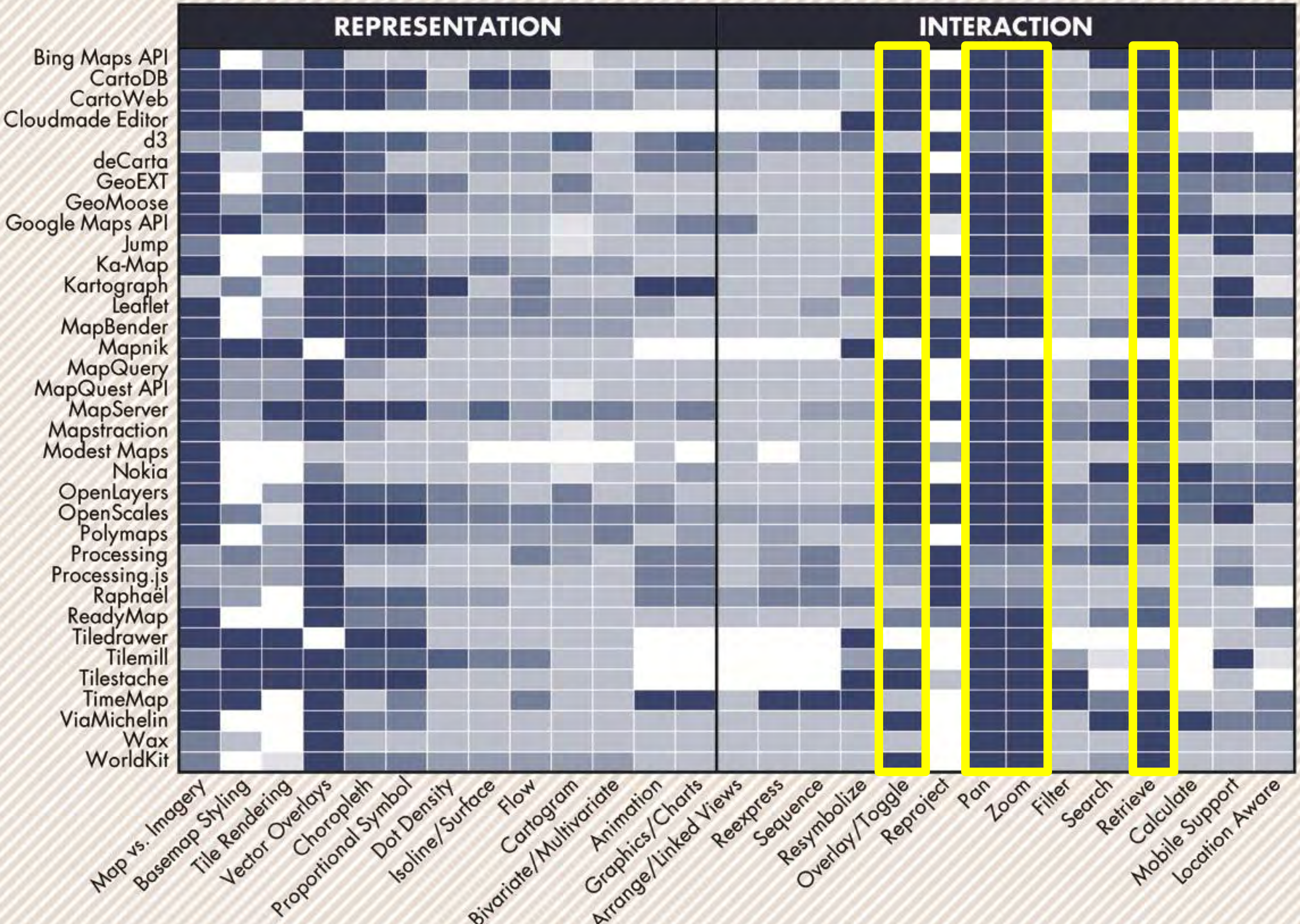
Coding Completed 2012

Roth et al. (2013b)



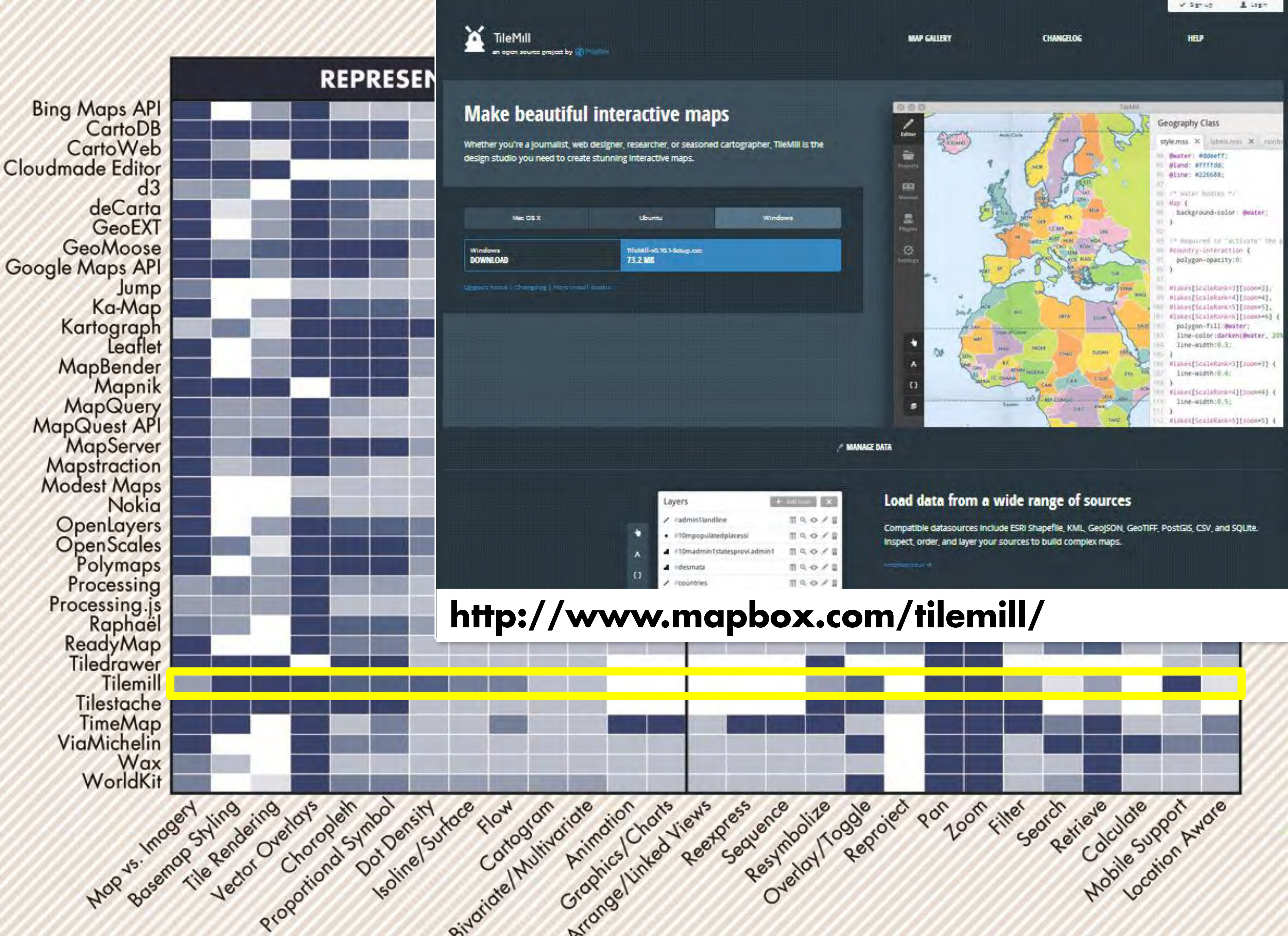
Coding Completed 2012

Roth et al. (2013b)



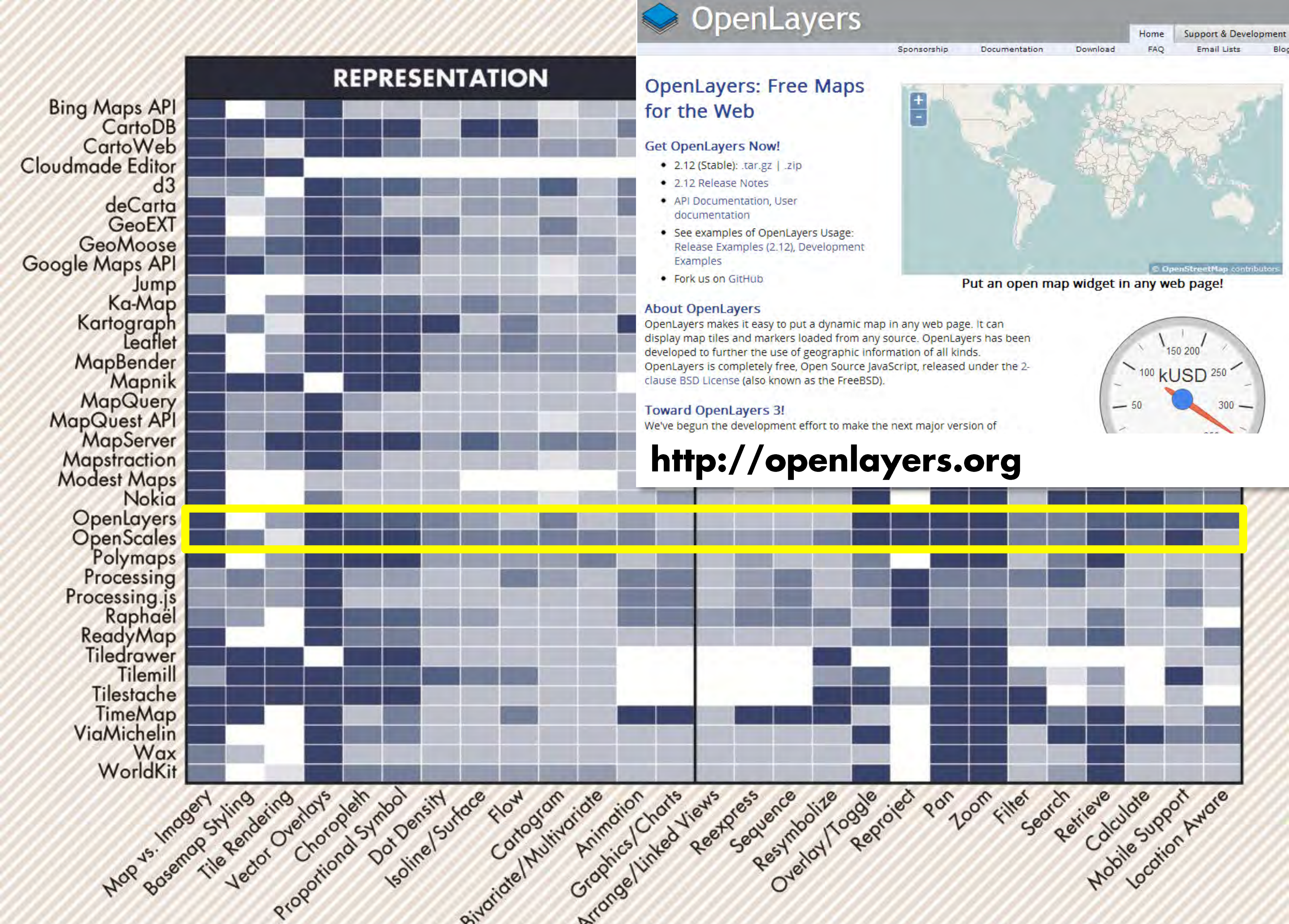
Coding Completed 2012

Roth et al. (2013b)



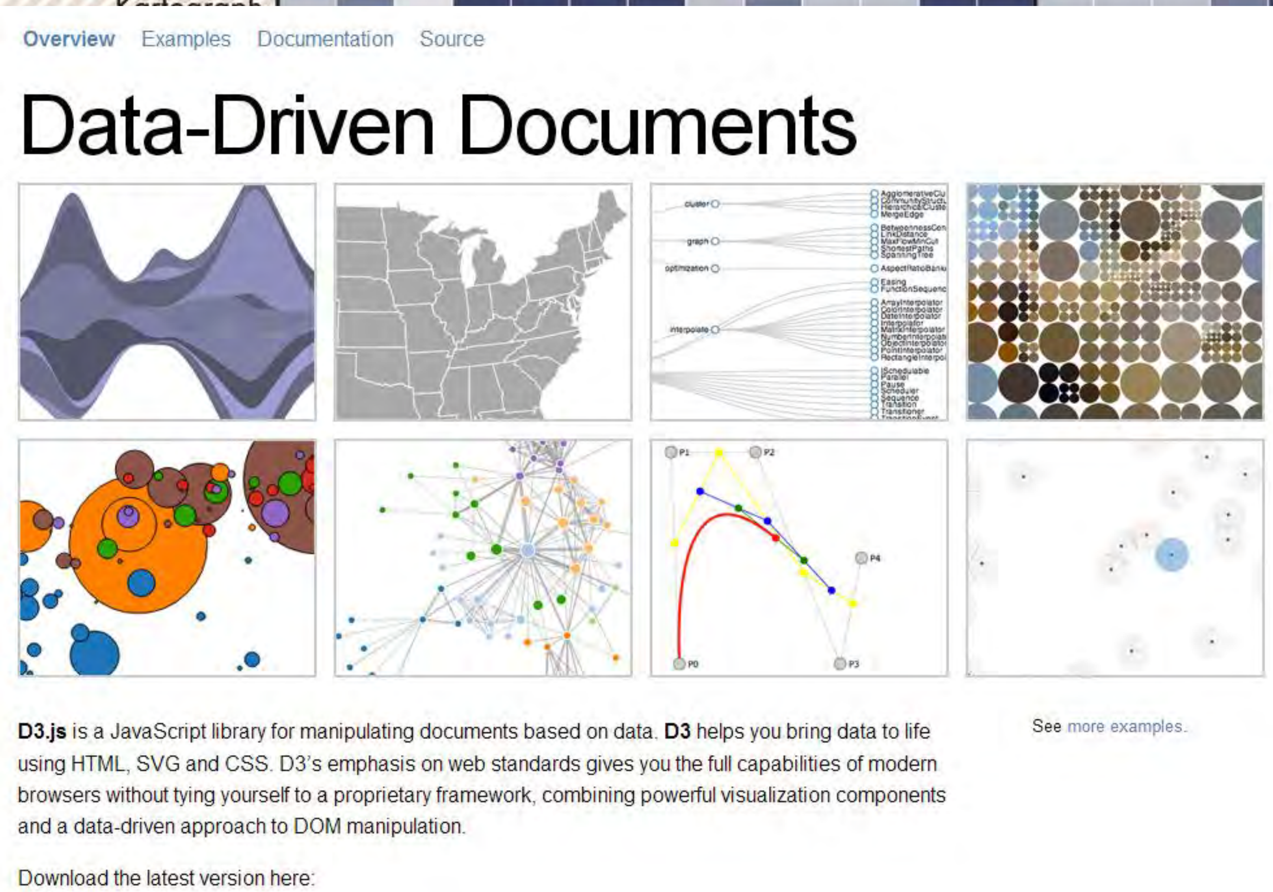
Coding Completed 2012

Roth et al. (2013b)



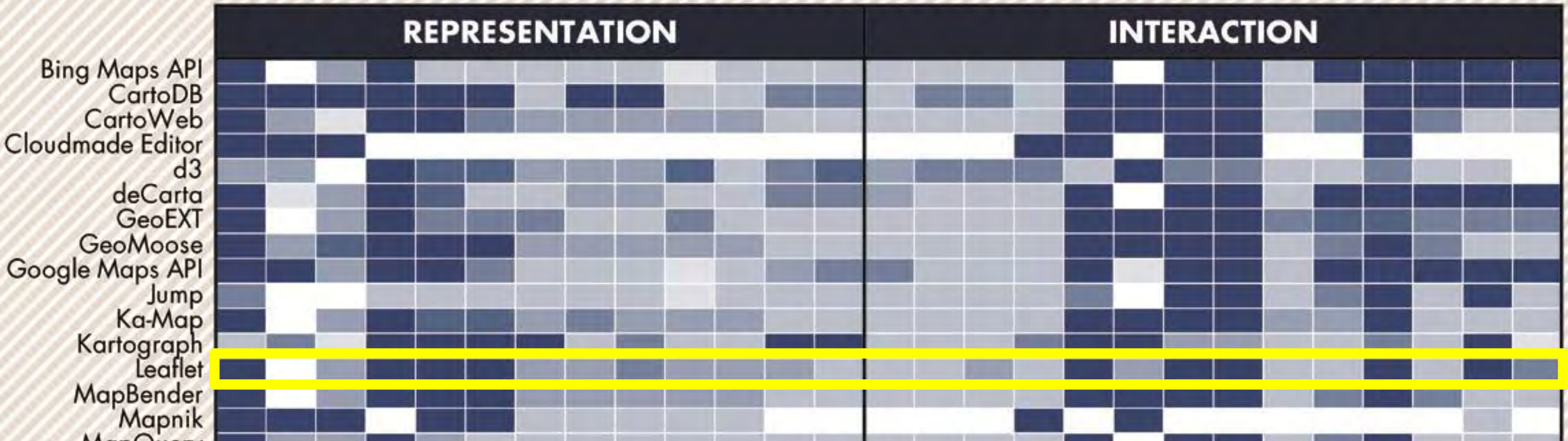
Coding Completed 2012

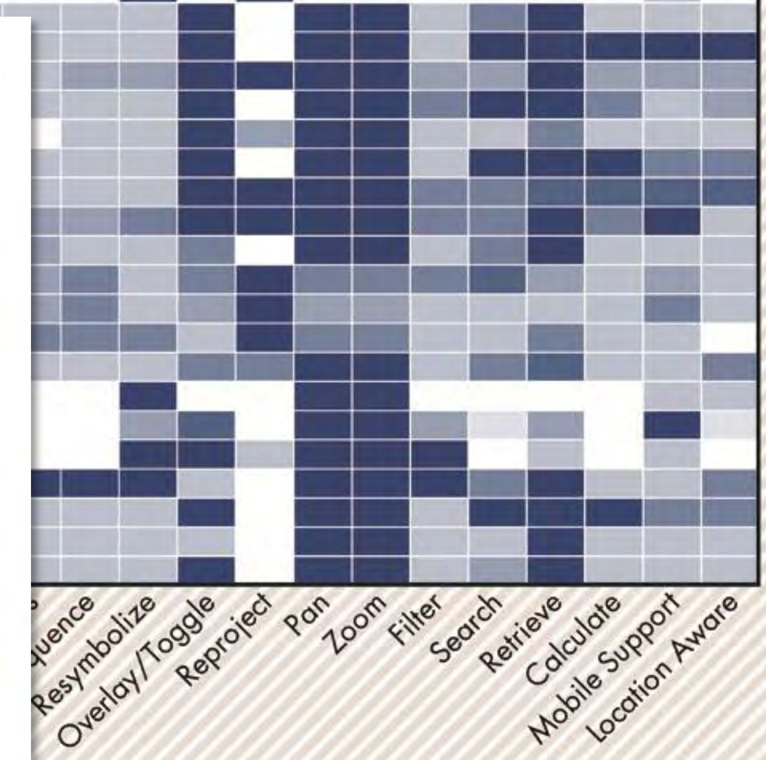
Roth et al. (2013b)

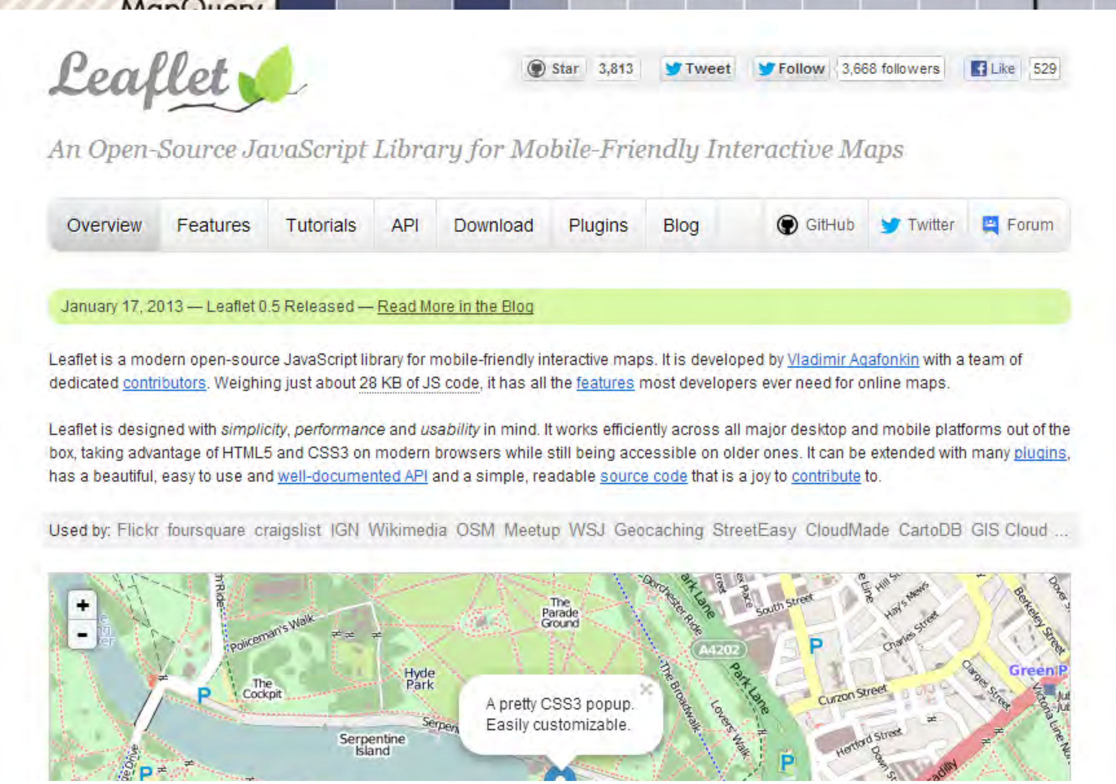


http://d3js.org

Roth et al. (2013b)

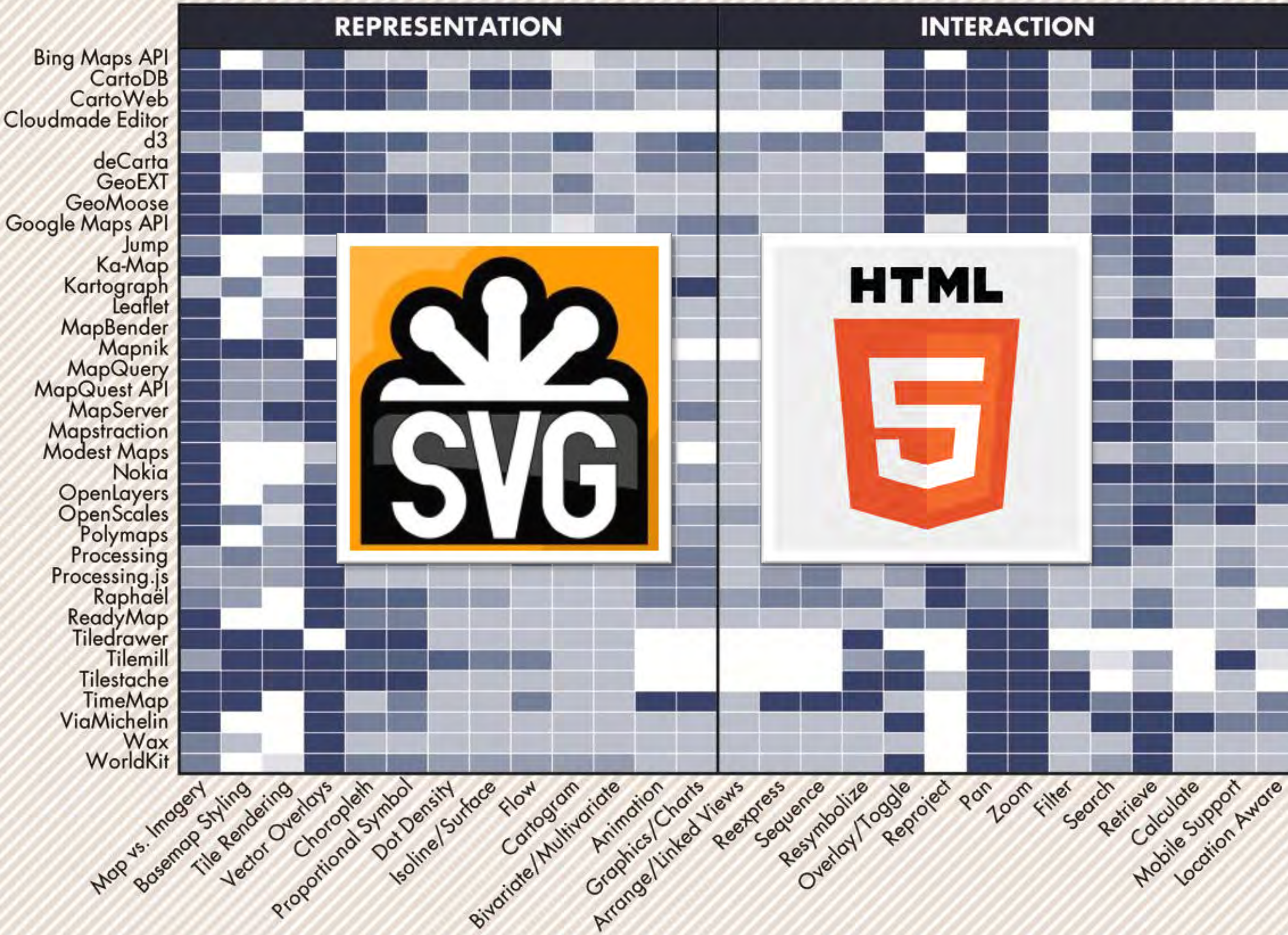


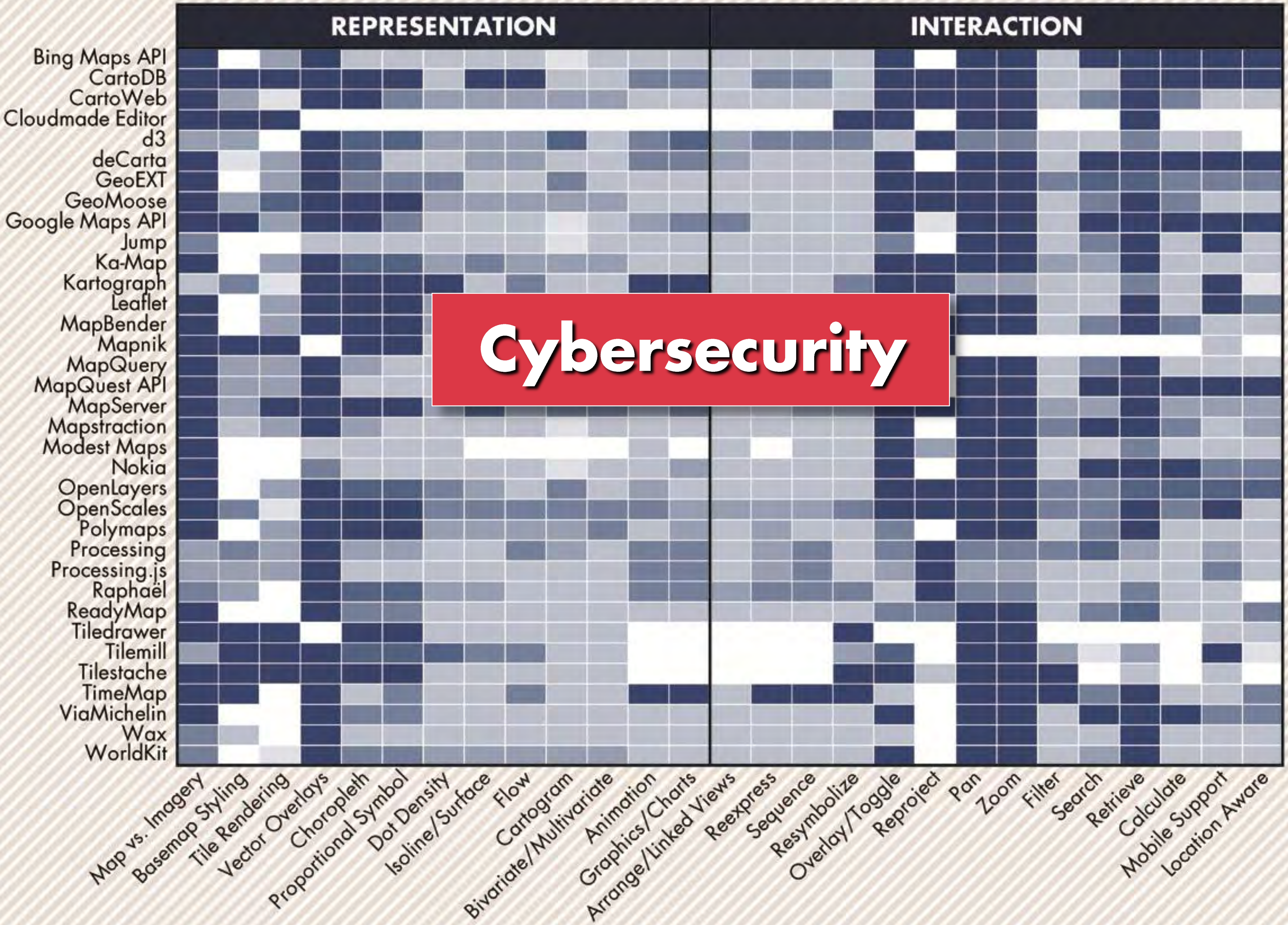




http://leafletjs.com

Roth et al. (2013b)


REPRESENTATION
INTERACTION
Bing Maps API
CartoDB
CartoWeb
Cloudmade Editor
d3
deCarta
GeoEXT
GeoMoose
Google Maps API
Jump
Ka-Map
Kartograph
Leaflet
MapBender
Mapnik
MapQuery
MapQuest API
MapServer
Mapstraction
Modest Maps
Nokia
OpenLayers
OpenScales
Polymaps
Processing
Processing.js
Raphaël
ReadyMap
Tiledrawer
Tilemill
Tilestache
TimeMap
ViaMichelin
Wax
WorldKit
SVG
HTML
Map vs. Imagery
Basemap Styling
Tile Rendering
Vector Overlays
Choropleth
Proportional Symbol
Dot Density
Isoline/Surface
Flow
Cartogram
Bivariate/Multivariate
Animation
Graphics/Charts
Arrange/Linked Views
Reexpress
Sequence
Resymbolize
Overlay/Toggle
Reproject
Pan
Zoom
Filter
Search
Retrieve
Calculate
Mobile Support
Location Aware


Cybersecurity
REPRESENTATION
INTERACTION
Bing Maps API
CartoDB
CartoWeb
Cloudmade Editor
d3
deCarta
GeoEXT
GeoMoose
Google Maps API
Jump
Ka-Map
Kartograph
Leaflet
MapBender
Mapnik
MapQuery
MapQuest API
MapServer
Mapstraction
Modest Maps
Nokia
OpenLayers
OpenScales
Polymaps
Processing
Processing.js
Raphaël
ReadyMap
Tiledrawer
Tilemill
Tilestache
TimeMap
ViaMichelin
Wax
WorldKit
Map vs. Imagery
Basemap Styling
Tile Rendering
Vector Overlays
Choropleth
Proportional Symbol
Dot Density
Isoline/Surface
Flow
Cartogram
Bivariate/Multivariate
Animation
Graphics/Charts
Arrange/Linked Views
Reexpress
Sequence
Resymbolize
Overlay/Toggle
Reproject
Pan
Zoom
Filter
Search
Retrieve
Calculate
Mobile Support
Location Aware

To Conclude

think differently about…

THE UNIVERSITY *of* WISCONSIN MADISON

think differently about your

Geospatial Data

Natural Earth

Free vector and raste... 1:10m, 1:50m, and 1:...

Home | Features | Downloads | Blog | Forums | Corrections | About

Map Gallery

Natural Earth is a public domain map dataset available at 1:10m, 1:50m, and 1:110 million scales. Featuring tightly integrated vector and raster data, with Natural Earth you can make a variety of visually pleasing, well-crafted maps with cartography or GIS software.

Natural Earth was built through a collaboration of many volunteers and is supported by NACIS (North American Cartographic Information Society), and is free for use in any type of project (see our Terms of Use page for more information).

Get the Data

Convenience

Natural Earth solves a problem: finding suitable data for making small-scale maps. In a time when the web is awash in geospatial data, cartographers are forced to waste time sifting through confusing tangles of poorly attributed data to make clean, legible maps. Because your time is

Neatness Counts

The carefully generalized linework maintains consistent, recognizable geographic shapes at 1:10m, 1:50m, and 1:110m scales. Natural Earth was built from the ground up so you will find that all data layers align precisely with one another. For example, where rivers and

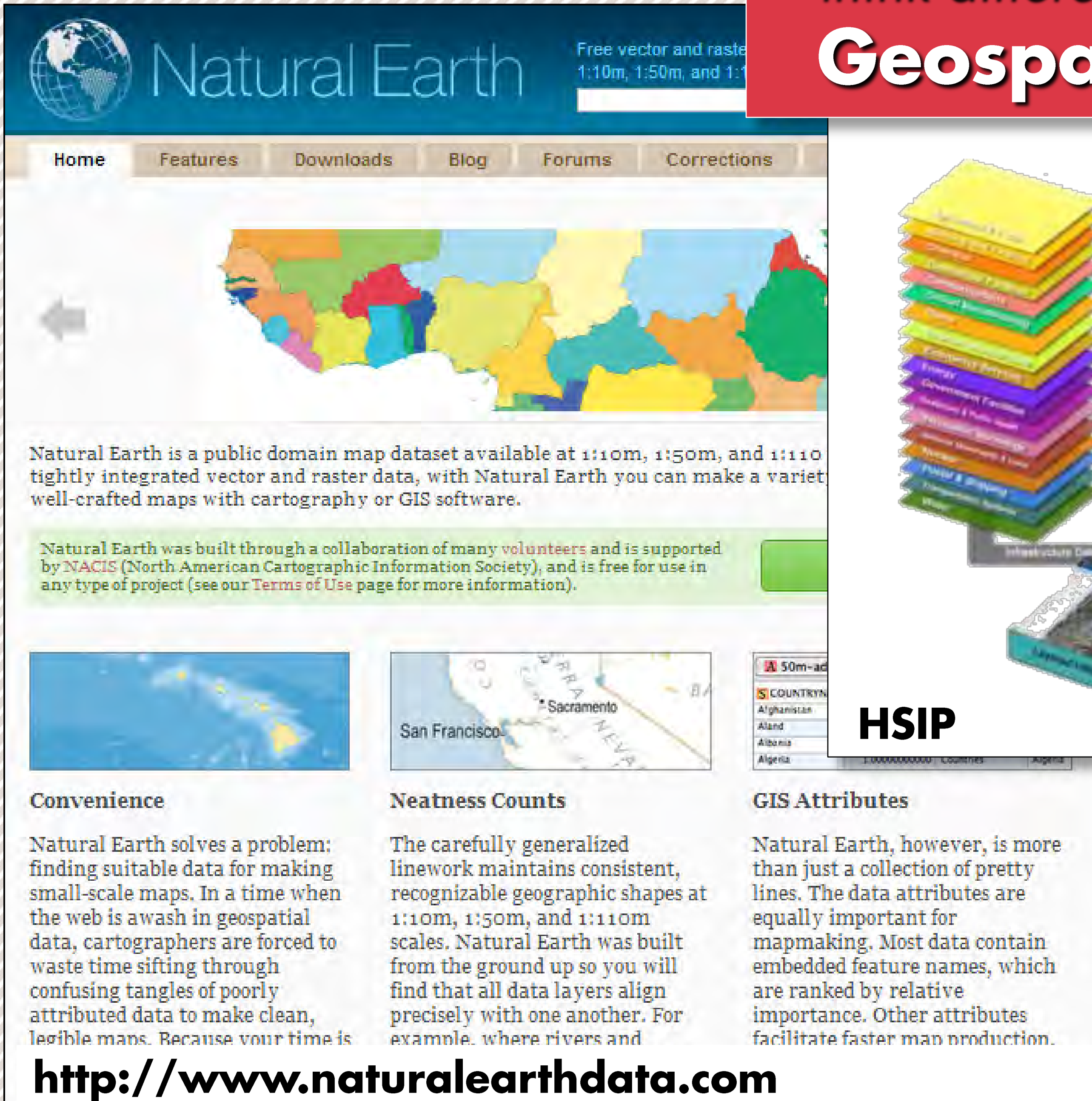
GIS Attributes

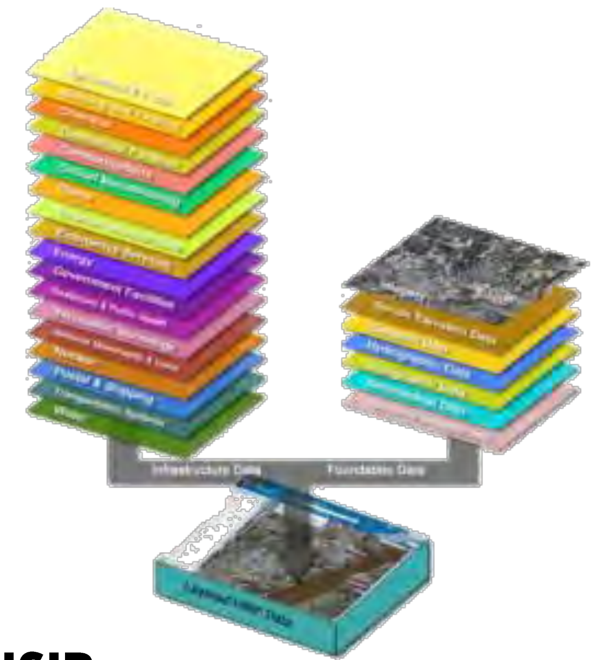
Natural Earth, however, is more than just a collection of pretty lines. The data attributes are equally important for mapmaking. Most data contain embedded feature names, which are ranked by relative importance. Other attributes facilitate faster map production.

http://www.naturalearthdata.com

think differently about your

Geospatial Data



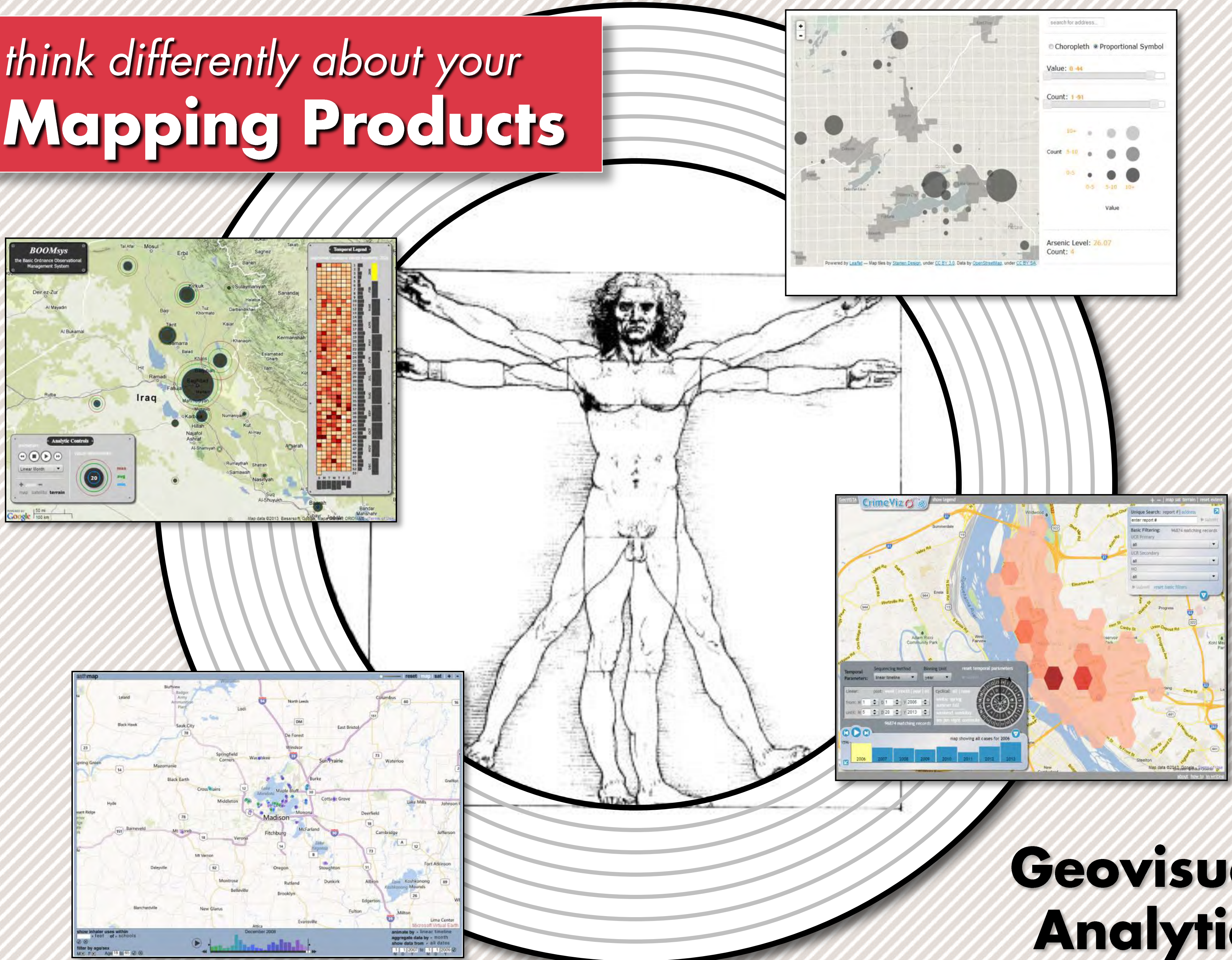


HSIP

http://www.naturalearthdata.com

think differently about your

Mapping Products



Geovisual Analytics

think differently about your Cartographers

CREDIT: National Oceanic and Atmospheric Administration/Department of Commerce

Email Article
Email Editor
Discuss in Forum
Related Articles
Print Article
Free Newsletter
BOOKMARK

Search Articles | Advanced Search | Search

The Job Market

The New Cartographers

By Emily Underwood
March 18, 2013

Twenty years ago, a driver lost at night would pull his car over, take out a paper map bought at a gas station, and pore over its folds under a dim light. With luck and some critical thinking, he would eventually get where he was going. Today, he'd be more likely to swipe his finger across a smart phone screen and follow directions using Google Maps.

"[F]uture shortages in cartography, photogrammetry, and geodesy seem likely because the number of graduates is too small (tens to hundreds) to give NGA choices or means of meeting sudden demand." —Future U.S. Workforce for Geospatial Intelligence, a report from the National Research Council

As maps have changed, so have mapmakers. No longer static images, maps have become active interfaces for information exchange, continuously determining where we are in relation to distant satellites and suggesting where we ought to go, says Seth Spielman, a 38-year-old geography professor at the University of Colorado, Boulder. Today, the global geoservices industry collects, shares, and analyzes data on an unprecedented scale. It's valued at as much as $270 billion per year and employs 500,000 people in the United States, according to a recent report from Google. The rapid transformation, which Spielman equates with a "renaissance" in the field, has overturned traditional ideas of what a geographer does.

Thank You!

Robert E. Roth | reroth@wisc.edu

Further Reading:
(*open access articles at http://www.geography.wisc.edu/faculty/roth*

Roth RE. (forthcoming) Interactive Maps: What we know and what we need to know. *Journal of Spatial Information Science.* [open access]

Roth RE, KS Ross, BG Finch, W Luo, & AM MacEachren. 2013a. Spatiotemporal crime analysis in US law enforcement agencies: Current practices and unmet needs. *Government Information Quarterly*. 30(3).

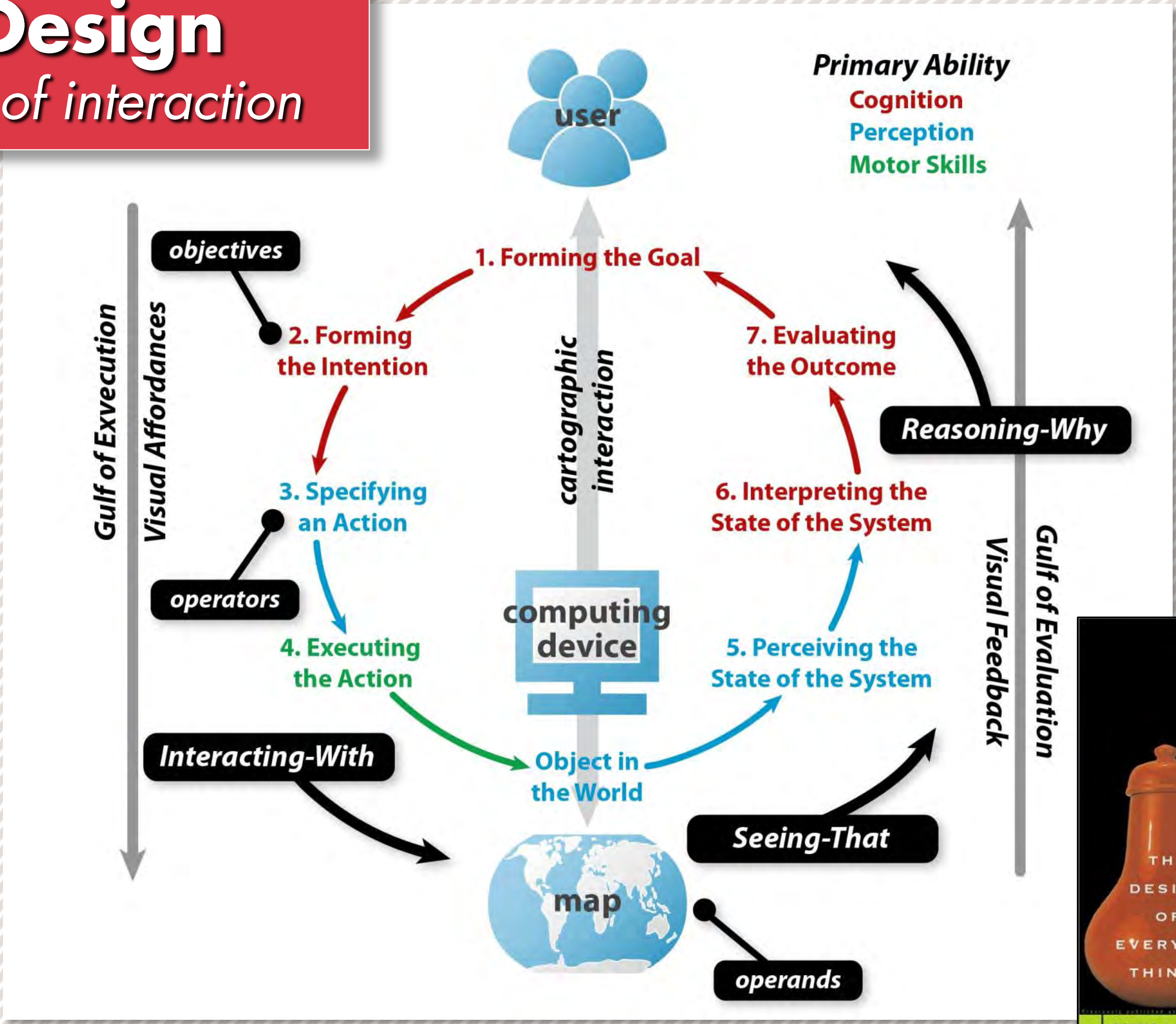
Roth RE, RG Donohue, CM Sack, TR Wallace, & TMA Buckingham. 2013b. A process for assessing emergent web mapping technologies. In: *Proceedings of the 26th International Cartographic Conference.* Dresden, Germany. [open access]

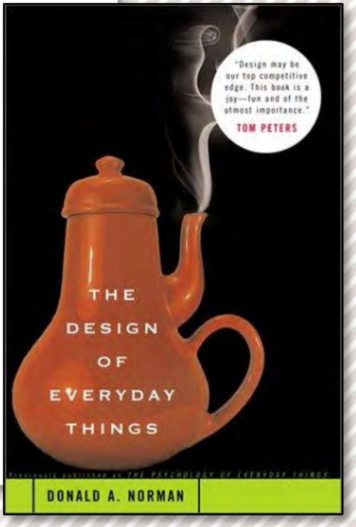
Roth RE. 2012. Cartographic interaction primitives: Framework & Synthesis. *The Cartographic Journal*. 49(4): 376-395.

Murdock MJ, RE Roth, and NV Maziekas. 2012. The Basic Ordnance Observational Management System: Geovisual analytics of improvised explosive device incidents. *The Journal of Maps*. 8(1): 120-124. [open access]

Roth RE. 2011. Interacting with maps: The science and practice of cartographic interaction. PhD Dissertation in Geography. The Pennsylvania State University. University Park, PA. [open access]

UX Design
stages of interaction





Norman (1985); Roth (2011)

Research Agenda

towards a science of cartographic interaction

Table of Contents

How?	*interaction primitives • stages of interaction • interaction strategies • objectives, operators, operands • interface styles • interface design*
What?	*interaction vs. interface • digital vs. analog • stages of interaction • system response time • interactive maps vs. mapping systems vs. map mashups*
Why?	*visual thinking • insight • exploration/confirmation/synthesis/presentation • discussion / debate • decision making*
When?	*interface complexity • interface freedom & flexibility • convention & standardization • work productivity • efficiency and effectiveness*
Who?	*user-centered design • user ability • user expertise • user motivation • adaptive cartography • role-based interfaces • geocollaboration*
Where?	*input capabilities • bandwidth size • processing power • display capabilities • interoperability • mobile mapping • open source development*

Roth (forthcoming)

Symbol Store

Table of Contents

The SymbolStore: Expanding to Provide a Social Forum for the Creation, Sharing, and Evaluation of Symbols

Ryan S Mullins
Scott Pezanowski
Anthony C. Robinson, PhD
Alan M MacEachren, PhD

GeoVISTA Center
Department of Geography
The Pennsylvania State University

GeoVISTA

PENNSTATE 1855

DEPARTMENT OF GEOGRAPHY
COLLEGE OF EARTH AND MINERAL SCIENCES

Outline

SymbolStore and Reviewer

Social Symbol Creation & Refinement

Features

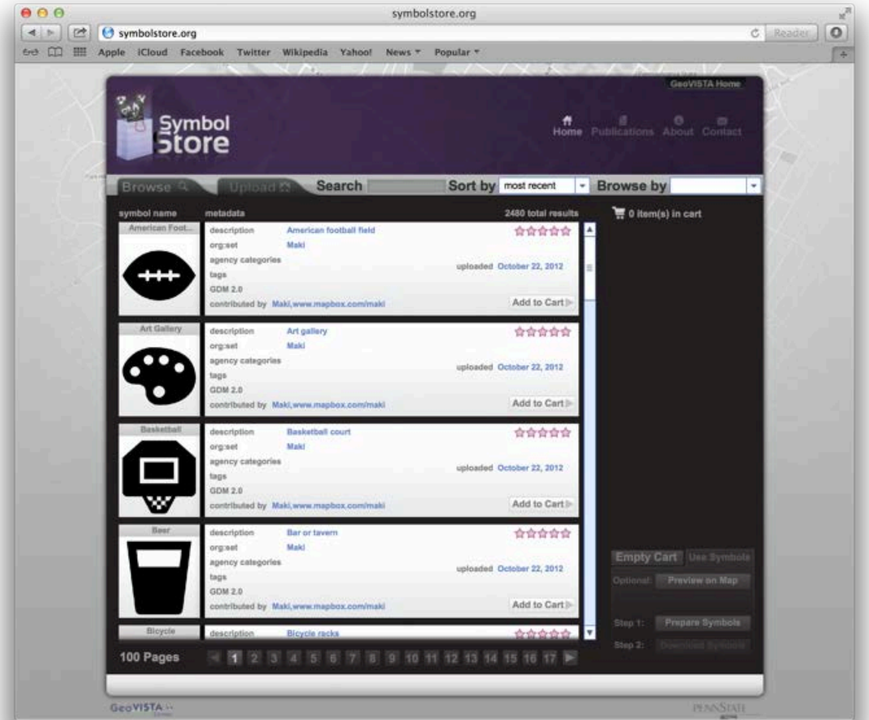
Interface

Moving Forward

GeoVISTA Center

The SymbolStore

- Search and browse symbols
  - By keyword, organization, category, etc.
- Preview symbols on realistic maps
  - Similar to ColorBrewer and others
- Download symbols for use
- Share symbols with others



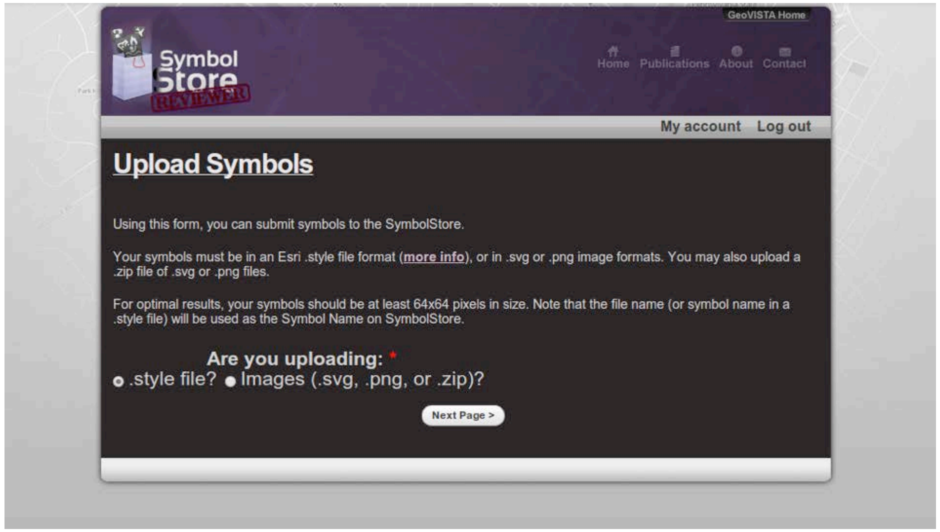
http://SymbolStore.org

GeoVISTA Center

The SymbolStore Reviewer

- Web-based activities via custom Drupal site

- Review Cycle:
  - Phase 1: Needs Assessment
  - Phase 2: Initial Standard Development
  - Phase 3: Standard Refinement
  - Phase 4: Implementation & Quality Control



- Methods
  - Round-based discussion
  - Voting (modified Delphi)

Robinson, Anthony C., et al. "Developing map symbol standards through an iterative collaboration process." Environment and Planning-Part B 39.6 (2012): 1034.

Lingering Questions

- How do we create a request for a new symbol, or modifications to an existing symbol?
- How do we track changes to symbols?
- Do we provide a means for discussion of designs, or will that create too many problems?
- How can we leverage community members to help the SymbolStore, and its content, grow?

Social Symbol Creation and Refinement

Target Audience

- Proactive cartographers from participating organizations
- Estimated 5–10% of users

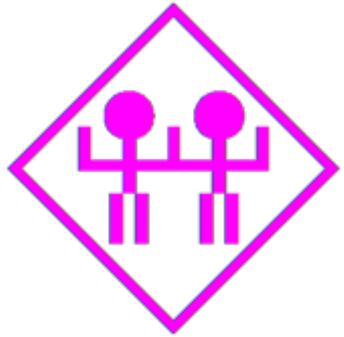
Needs

- Textual description of issues/needs
- Revision control
- Visual comparison across versions
- Discussion of
- Interfaces for:
  - Requesting changes
  - Contributing new versions
  - Examining and discussing

GeoVISTA Center

Features: Defining the Problem

- Sets:
  - Problem symbols
    - Redundant – Two symbols, same meaning
    - Confusing – Visual does not describe referent
    - Multiple meanings – One symbol, multiple meanings
    - Poorly designed – Displays poorly on digital/print media
  - Missing symbols



- Symbols:
  - Track problems with symbol individually
  - Similar to bug trackers for source code

- Universally:
  - Detailed description of set/symbol purpose

Features: Revision Control

- Revision control similar to those used for source code control
- Utilize popular, existing tools (e.g. Git)
- Allow comments for change logs
- Link to individual issues defined in request
- Review/merge functionality for symbol contributor

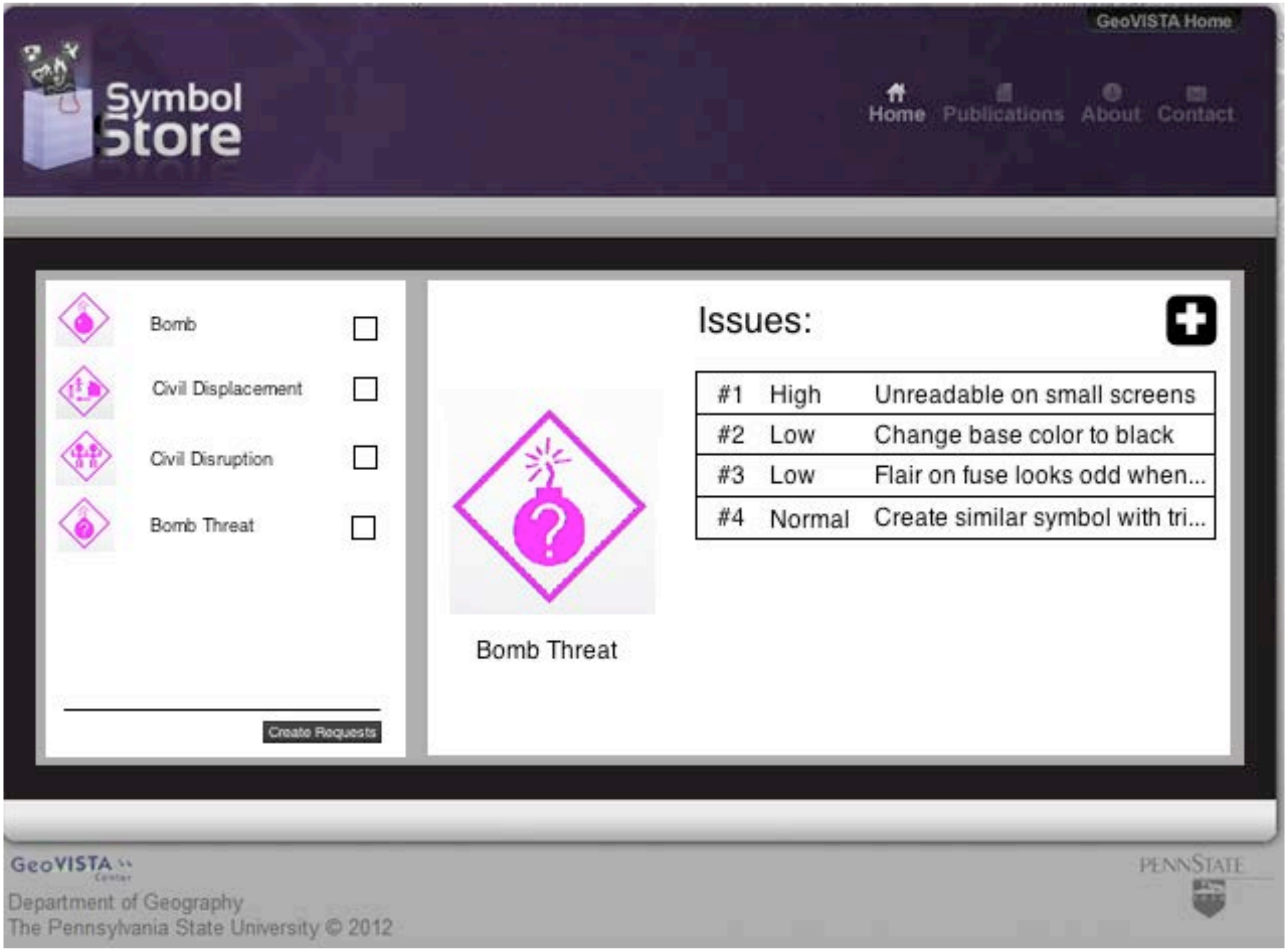
GeoVISTA Center

Features: Discussion

- Forum to record discussion about sets/symbols
- Threaded replies to topic
  - Toggle to show/hide
  - Vote up or down based on relevance
  - Filter by issue, symbol, etc.
- (Semi-)Automated moderation

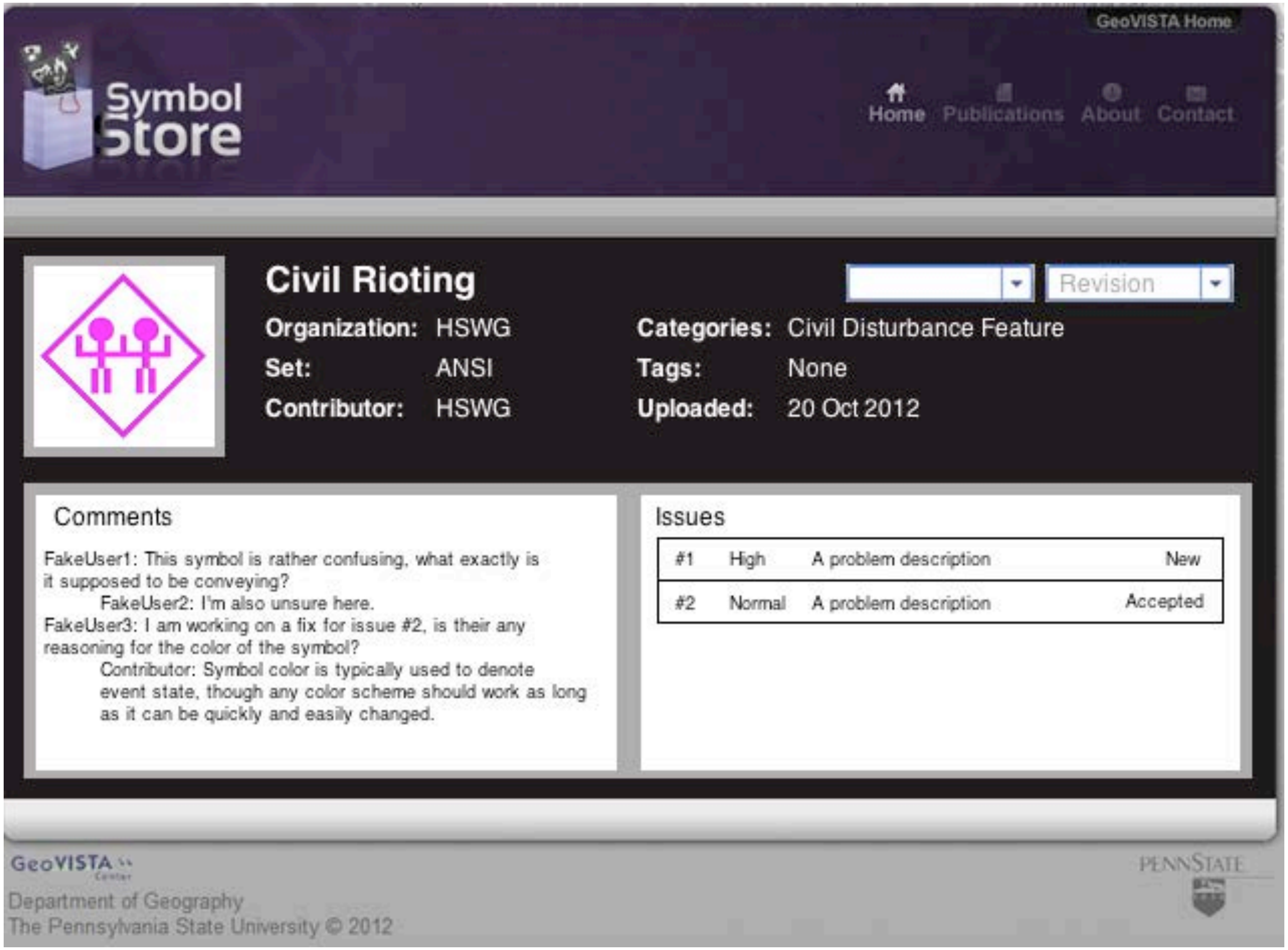
GeoVISTA Center

Interface Concept: Requests



GeoVISTA Center

Interface Concept: Examining



Moving Forward: Challenges

- Automated generation of symbols in supported formats
- Scalability of hardware/software
- Linking revision control and issue tracking

GeoVISTA Center

Table of Contents

For more information:

Try it now – www.symbolstore.org

RyanMullins@psu.edu
@RyanMullins

This work is supported by a contract from the U.S. Department of Homeland Security Science and Technology Directorate, Command, Control and Interoperability Division. The views and opinions expressed here are of the authors, and do not reflect the official positions of the Department of Homeland Security or the Federal Government.

GeoVISTA

PENNSTATE 1855

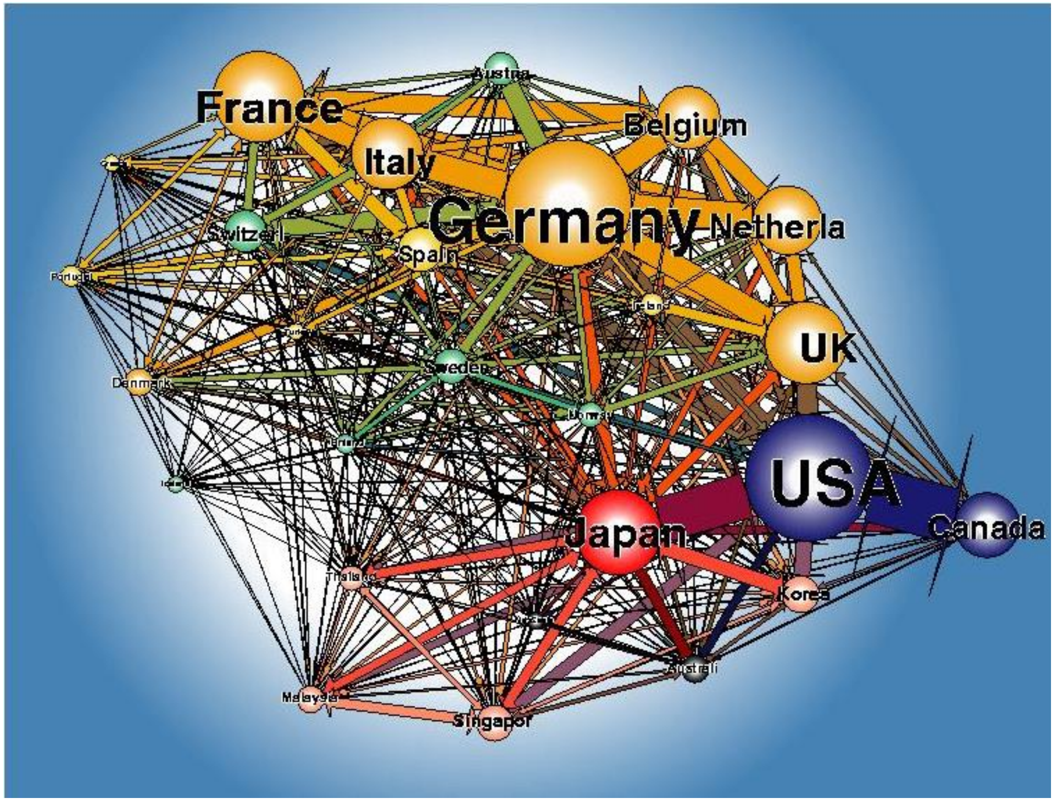
DEPARTMENT OF GEOGRAPHY
COLLEGE OF EARTH AND MINERAL SCIENCES

Table of Contents

Understanding Spatial and Social Relationships in International Trade Network: A Geovisual Analytic Approach

Wei Luo
Peifeng Yin
Frank Hardisty
Alan M. MacEachren
GeoVISTA Center, Department of Geography
Penn State University

Motivation



World trade among some countries in 1992 (Krempel and Pluemper 1999)

The role of geographical proximity in shaping the International Trade Network (ITN) has not been explored, especially across different regions.

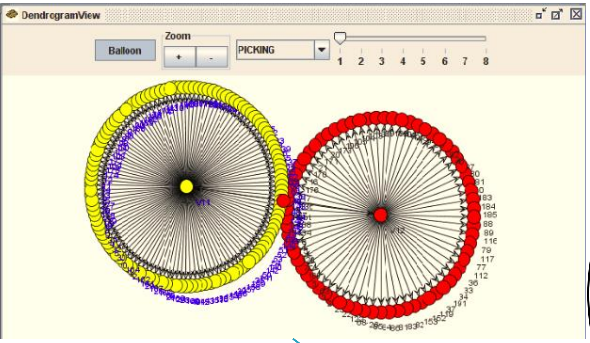
Research Questions

- What regions in the world are more influenced by the geographical proximity relationship?
- What regions in the world are more influenced by the international trade network relationship?
- What regions in the world on which both relationships have the approximately similar influence?

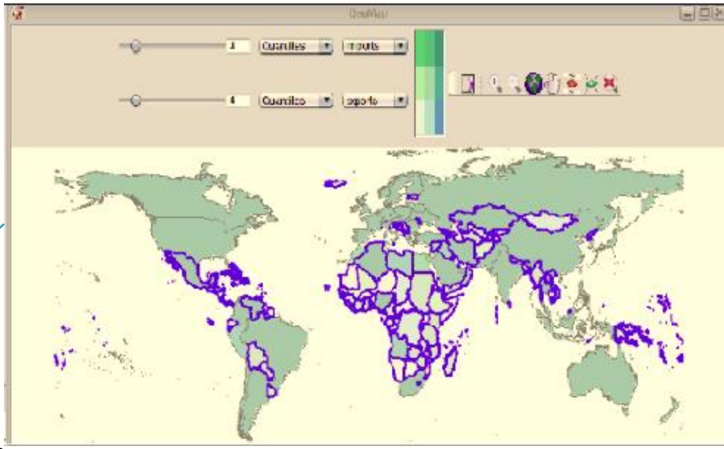
Challenge

- It is a challenge to distinguish the role of the social relationship and spatial relationship in the ITN, because both relationships always interact with each other.

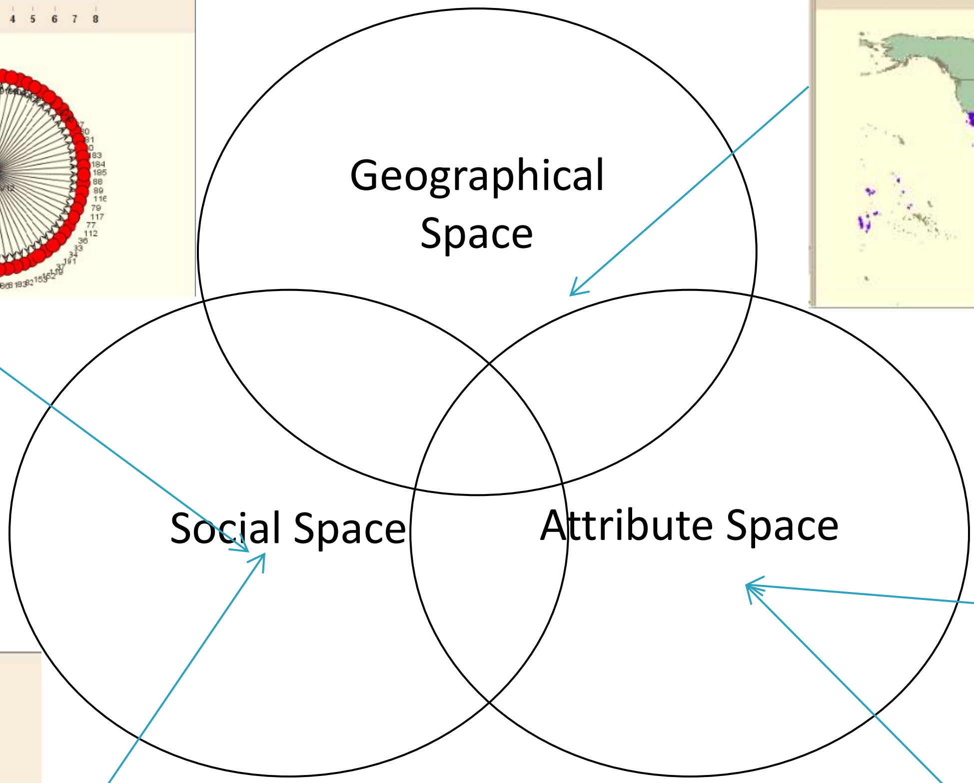
GeoSocialApp

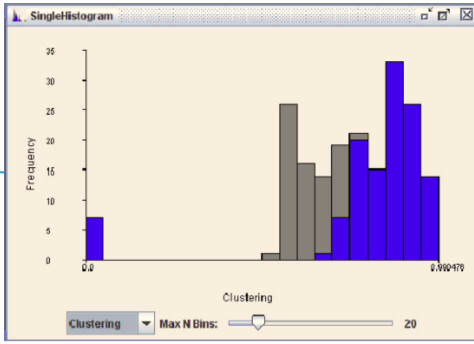


Dendrogram View

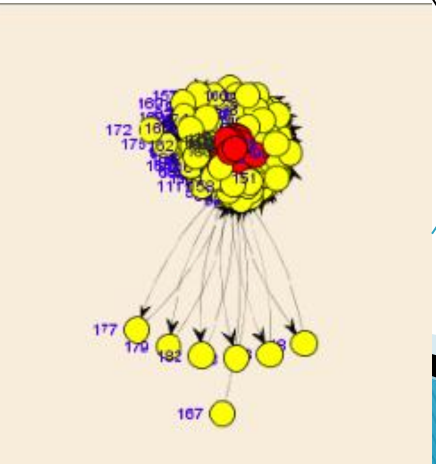


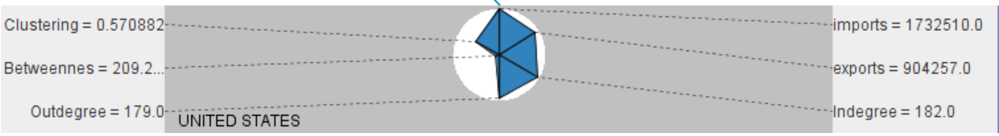
Chroploth Map





Modified from (Hess 2004)





http://www.geovista.psu.edu/GeoSocialApp/

Data: Network

Imports-exports relationship

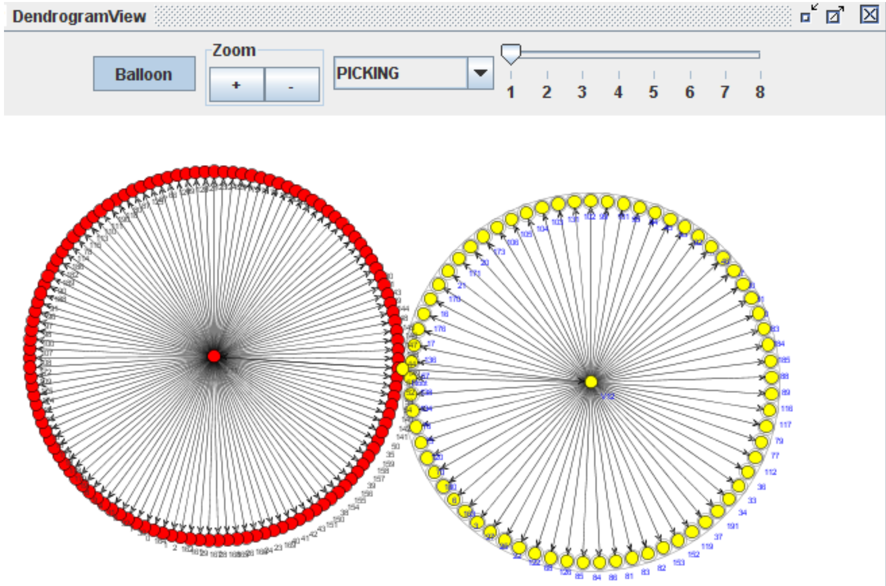
Years	importer1	importer2	flow1	flow2
2005	United States	Canada	291944	195151
2005	United States	Bahamas	726.3	1945.79
2005	United States	Cuba	0	397.87
2005	United States	Haiti	458.5	756.91
2005	United States	Dominican R	4721.4	5179.24
2005	United States	Jamaica	410.9	1962.2
2005	United States	Trinidad and	8342.2	1583.01
2005	United States	Barbados	33.4	595.28
2005	United States	Dominica	3.8	67.43
2005	United States	Grenada	6	90.64
2005	United States	St. Lucia	33.6	148.06
2005	United States	St. Vincent a	15.7	79.98
2005	United States	Antigua & Ba	4.6	189.9
2005	United States	St. Kitts and	55.8	138.82

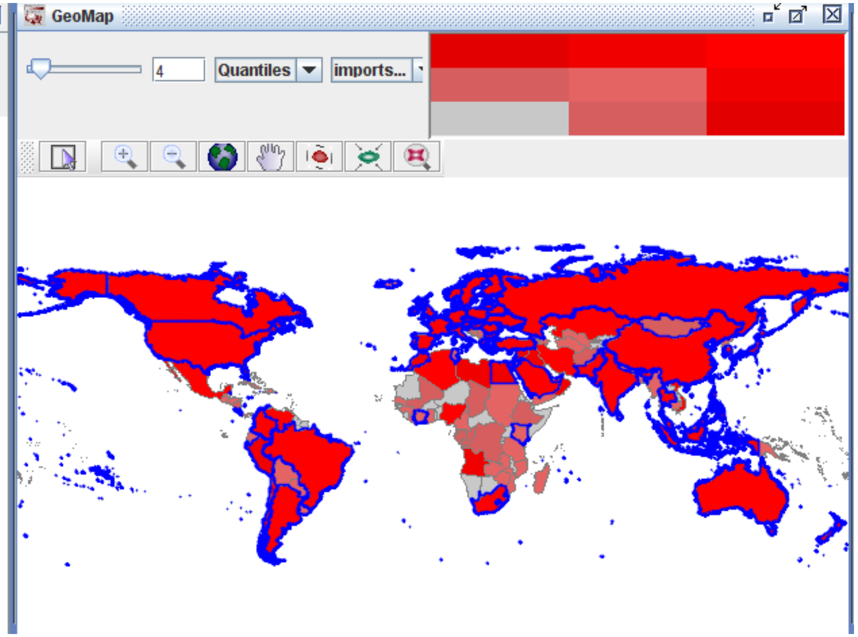
Binary matrix for 0% Threshold in 2005

	AFGHANISTA	ALBANIA	ALGERIA	AMERICA	ANDORRA	ANGOLA	ANTIGUA	ARGENTIN	ARMENIA	AUSTRALI
AFGHANISTAN	0	0	0	0	0	0	0	1	0	1
ALBANIA	0	0	1	0	0	0	0	1	1	1
ALGERIA	1	0	0	0	0	1	0	1	0	1
AMERICAN SAM	0	0	0	0	0	0	0	1	0	1
ANDORRA	0	0	0	0	0	0	0	0	0	0
ANGOLA	0	0	1	0	0	0	0	1	0	1
ANTIGUA AND E	0	0	0	0	0	0	0	1	0	1
ARGENTINA	1	1	1	0	0	0	0	0	0	1
ARMENIA	0	0	0	0	0	0	0	1	0	1
AUSTRALIA	1	1	1	1	0	1	1	1	1	0

CorrelatesOfWar Database

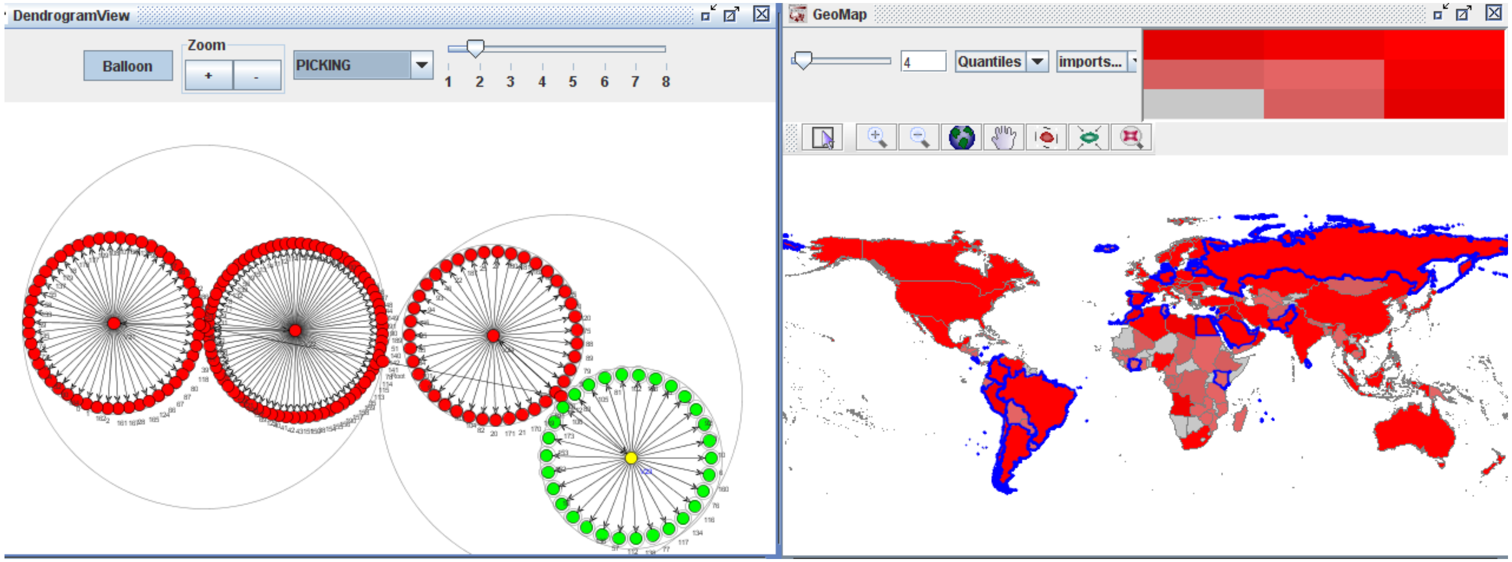
Results: The First Split (A Core-Periphery Structure)

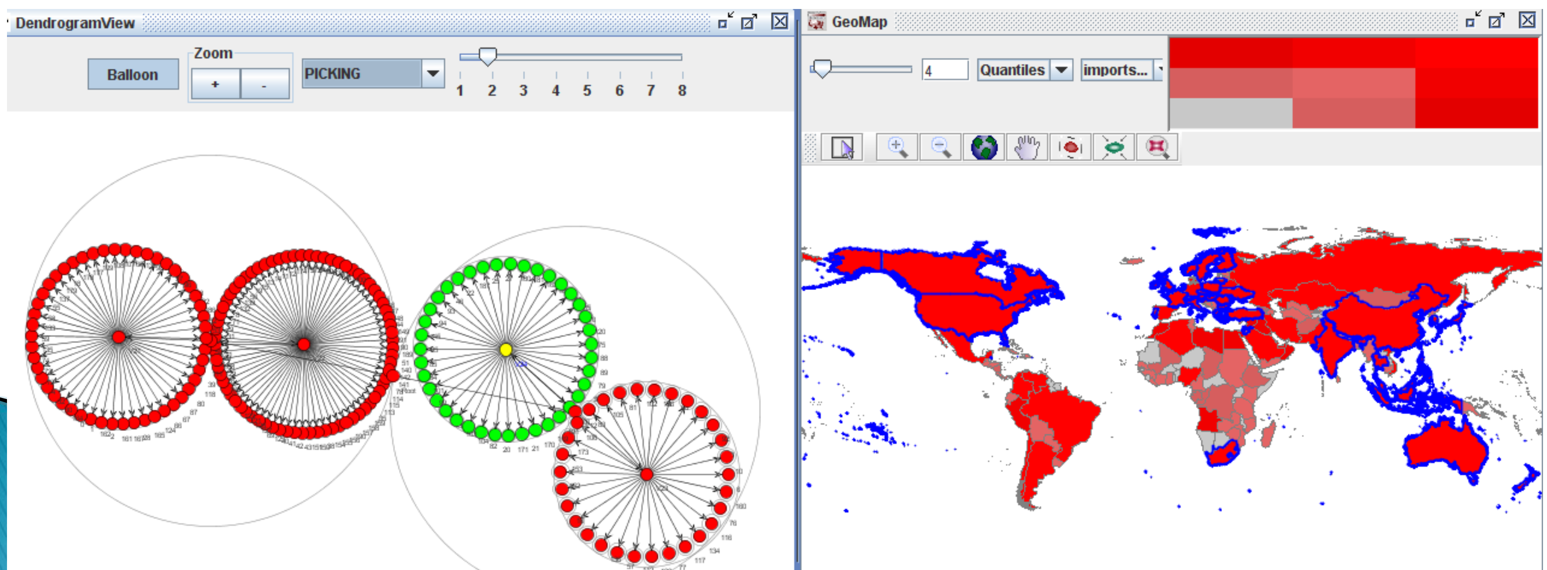




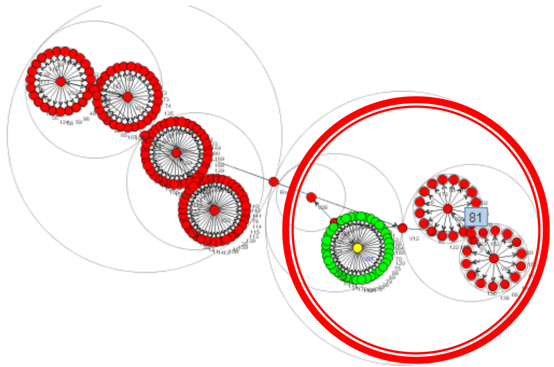
The convergence of the iterated correlations (CONCOR) algorithm (Breiger et al. 1975).

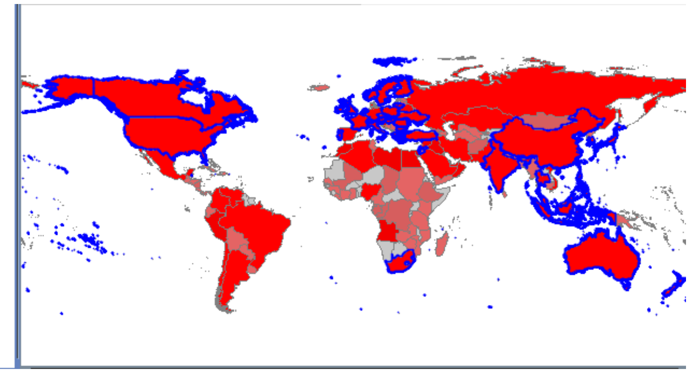
The Second Split (Core Group)

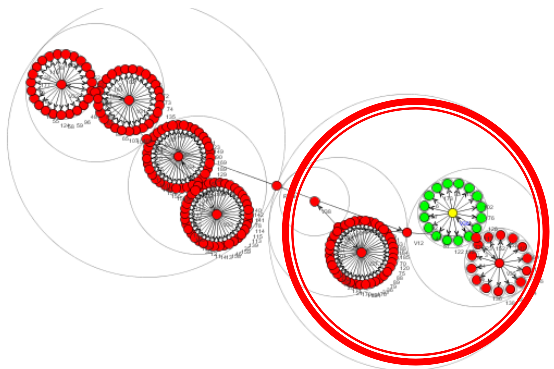


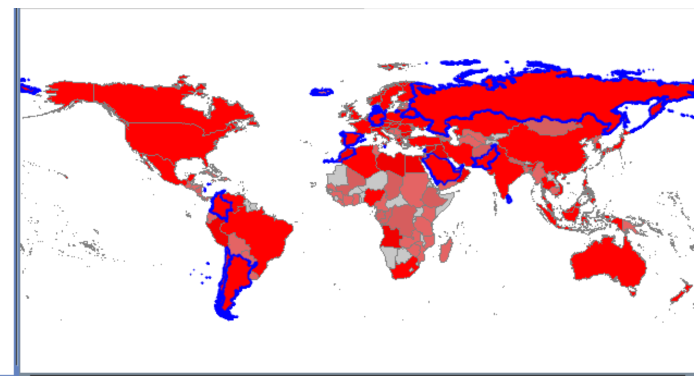


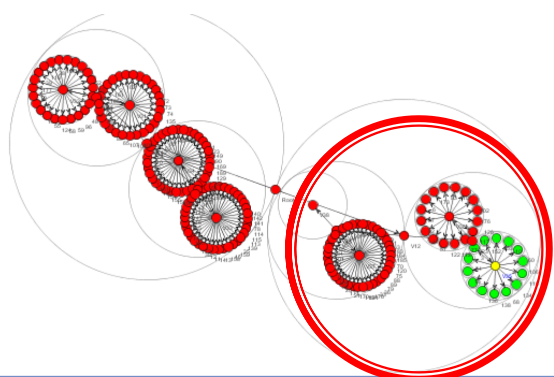
The Third Split (Core Group)

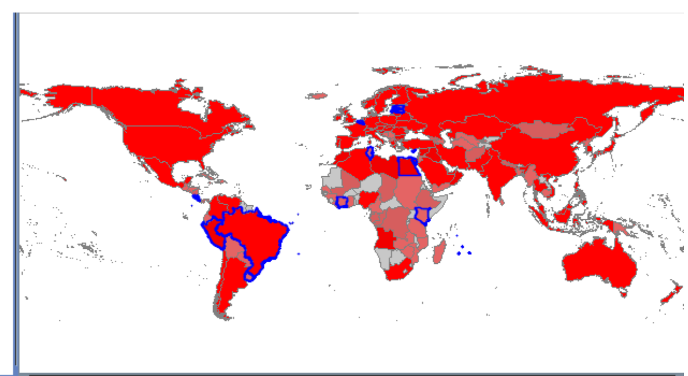




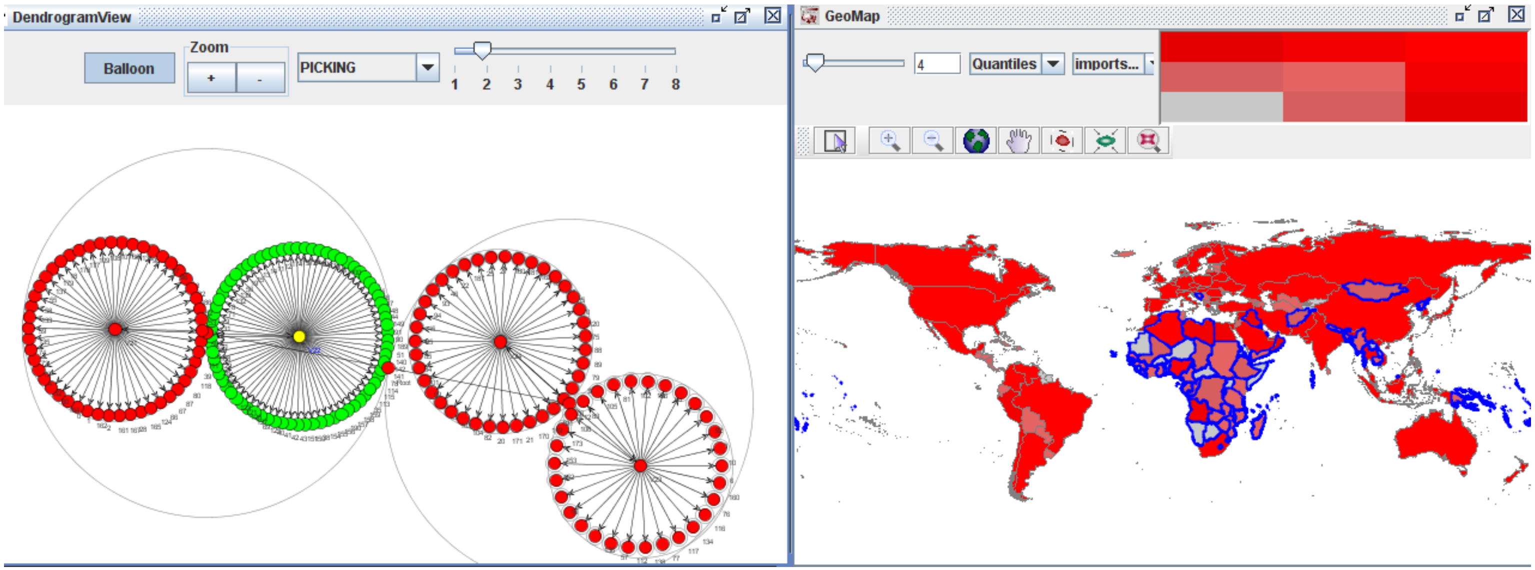


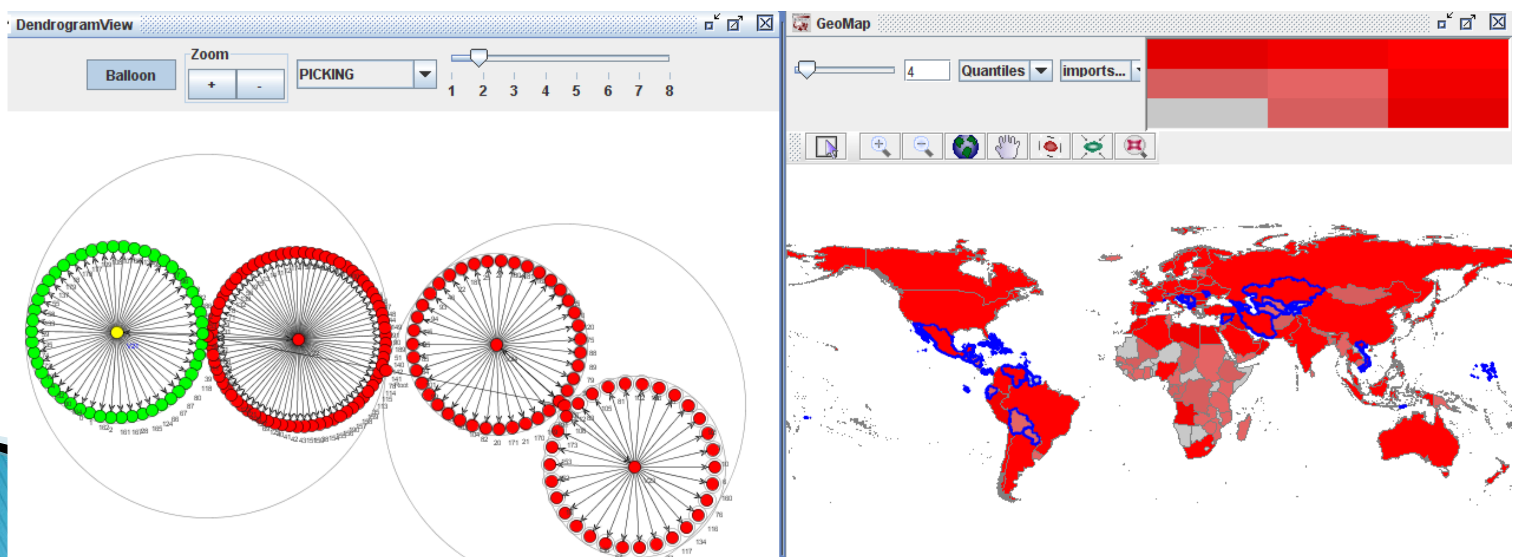




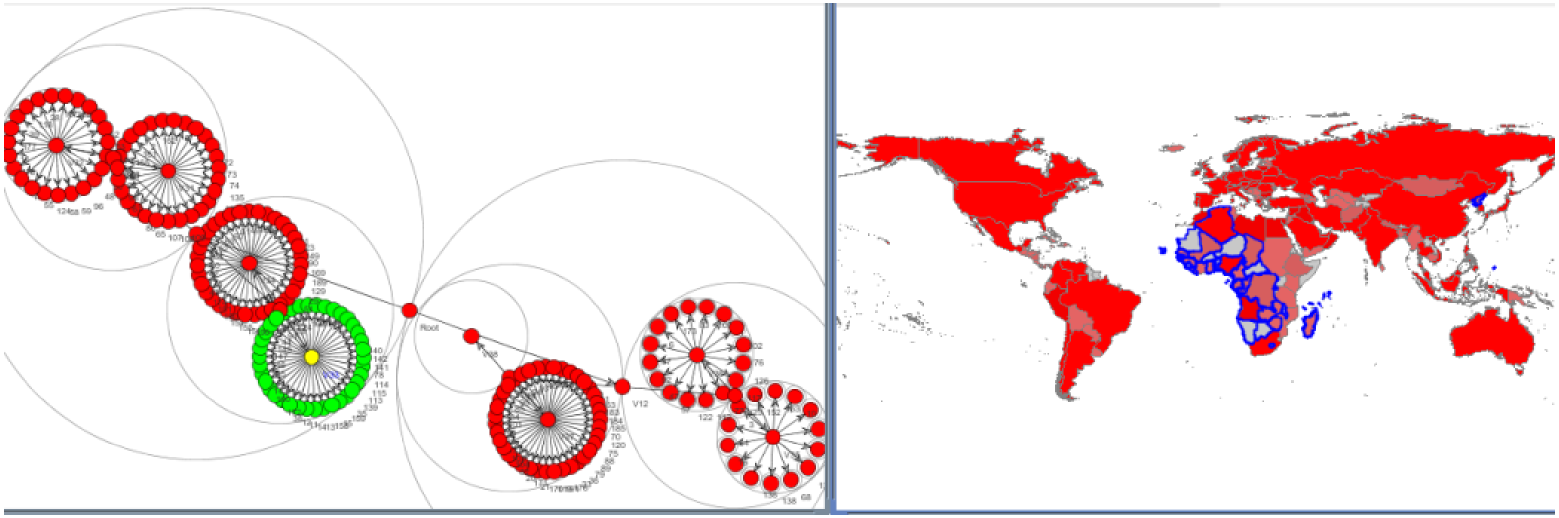


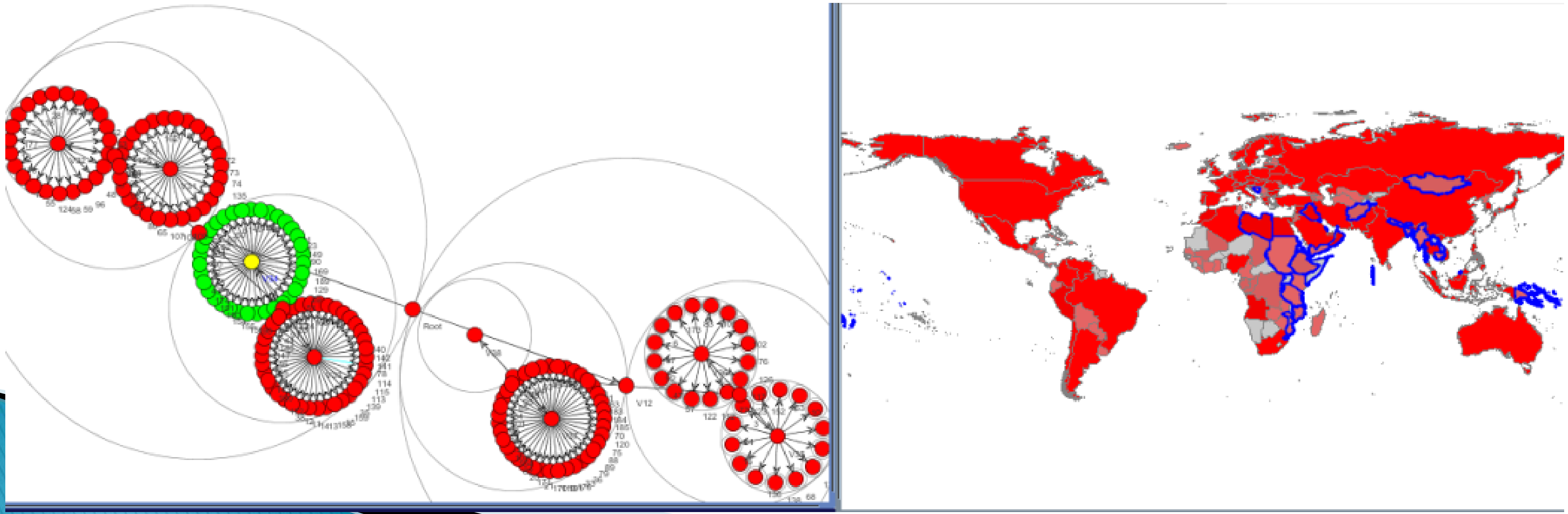
The Second Split (Periphery Group)



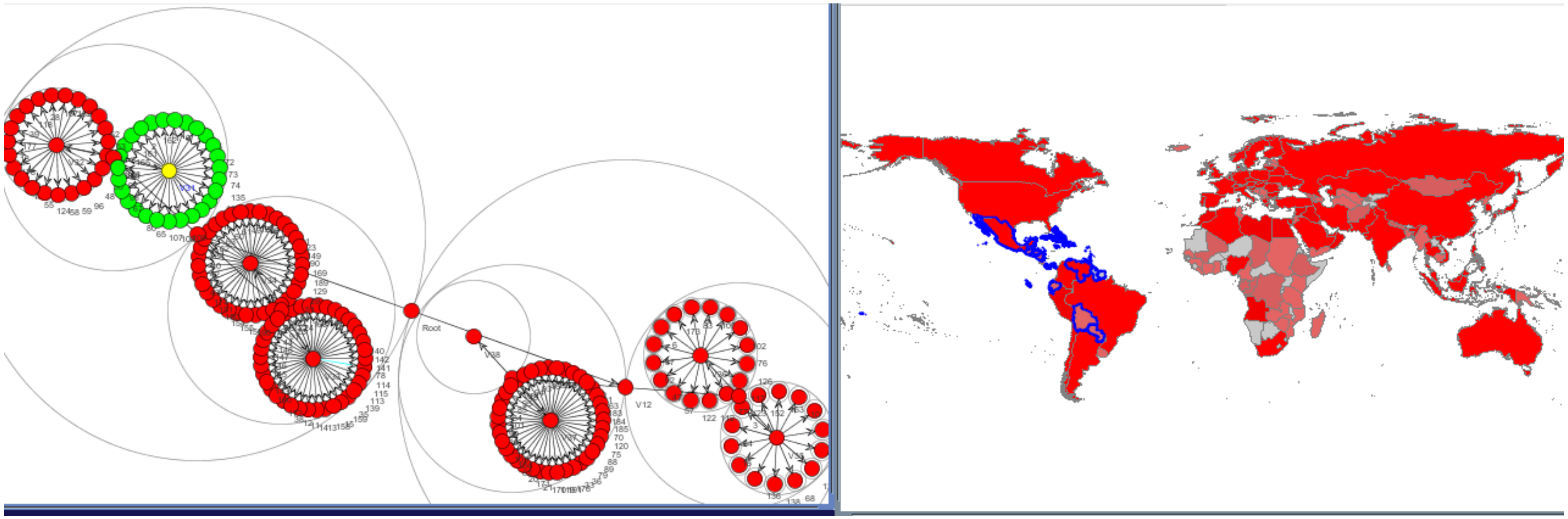


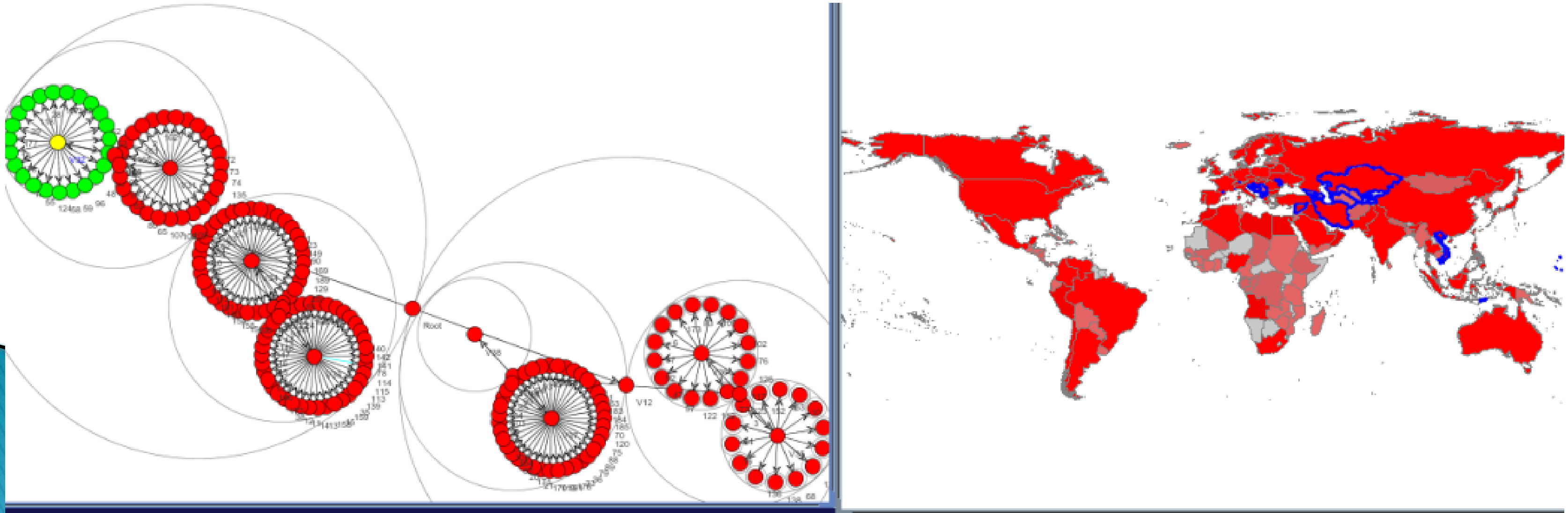
The Third Split (Periphery Group)





The Third Split (Periphery Group)





Summary

- Global trade is hierarchical with a core-periphery structure in terms of international trade network.
- The complex interaction between the social relationship and spatial relationship always exists in the ITN, but the spatial relationship are more influential to developing countries, and the social relationship are more influential to developed countries.
- The interaction of spatial proximity and social closeness in terms of the ITN cross different regions also indicates that the development of countries experiences a process of loosened spatial constraints (Temporal implications).

GeoVISTA Center

Conclusion

- Spatial analysis and social network analysis can give insights on the same datasets from different perspectives, so it is necessary to consider simultaneously.
- New approaches should be developed to study the interaction between social relationships and spatial relationships.

Future Work

- Use the tool to explore the international trade data for multiple years.

- Questions?

References

Table of Contents

- Andrienko, G., N. Andrienko, D. Keim, A. M. MacEachren & S. Wrobel (2011) Challenging Problems of Geospatial Visual Analytics (editorial introduction). *Journal of Visual Languages & Computing, 22, **251-256.***
- Barbieri, K., Keshk, O. M. G. and Pollins, B. 2008. Correlates of war project trade data set codebook, Version 2.0.
- DiBiase, D. (1990) Visualization in the earth sciences. Earth and Mineral Sciences, 59, 13-18.
- Goodchild, M. & D. Janelle (2010) Toward critical spatial thinking in the social sciences and humanities. *GeoJournal, 75, **3-13.***
- Hardisty, F. and Robinson, A. 2010. The geoviz toolkit: using component-oriented coordination methods for geographic visualization and analysis. International Journal of Geographical Information Science, 25, 191-210.
- Hess, M. (2004) "Spatial" relationships? Towards a reconceptualization of embeddedness. Progress in human geography, 28, 165-186.
- Kali, R., Méndez, F. and Reyes, J. 2007. Trade structure and economic growth. *Journal of International Trade Economic Development*, 16, 2, 245-269.
- Kali, R. and Reyes, J. 2007. The architecture of globalization: a network approach to international economic integration. Journal of International Business Studies, 38, 4, 595-620.
- Keim, D., F. Mansmann & J. Thomas (2010) Visual analytics: how much visualization and how much analytics? ACM SIGKDD Explorations Newsletter, 11, 5-8.
- Keim, D. 2011. The Role of Visualization in Visual Document Analysis. In IEEE Workshop on Interactive Visual Text Analytics for Decision Making. Providence R.I.
- Perer, A. and Shneiderman, B. 2008. Integrating statistics and visualization: case studies of gaining clarity during exploratory data analysis. In Proceedings of the Proc. SIGCHI conference on Human factors in computing systems (2008). ACM.
- Thomas, J. & K. Cook. 2005. Illuminating the path: The research and development agenda for visual analytics. IEEE Computer Society.
- Vespignani, A. (2009) Predicting the behavior of techno-social systems. Science, 325, 425-428.
- Anselin, L., I. Syabri & Y. Kho (2006) GeoDa: An introduction to spatial data analysis. Geographical Analysis, 38, 5-22.
- Krempel, L. & T. Pluemper (1999) International division of labor and global economic processes: an analysis of the international trade in automobiles. Journal of World-Systems Research, 5, 390-402.
- Luo, W., A. M. MacEachren, P. Yin & F. Hardisty. 2011. Spatial-Social Network Visualization for Exploratory Data Analysis. In SIGSPATIAL International Workshop on Location-Based Social Networks (LBSN 2011) in conjunction with the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS) 2011. November 1-4, 2011, Chicago, Illinois.: ACM.

Table of Contents

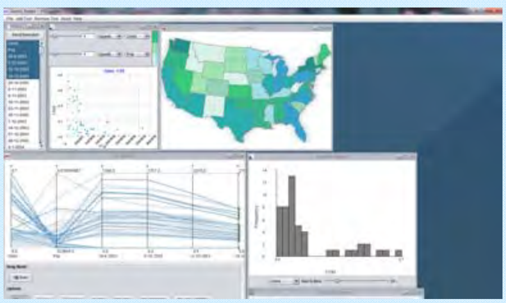
Visualizing Spatial, Temporal and Social Graph Information With the GeoViz Toolkit

Frank Hardisty
GeoVISTA Center, Penn State

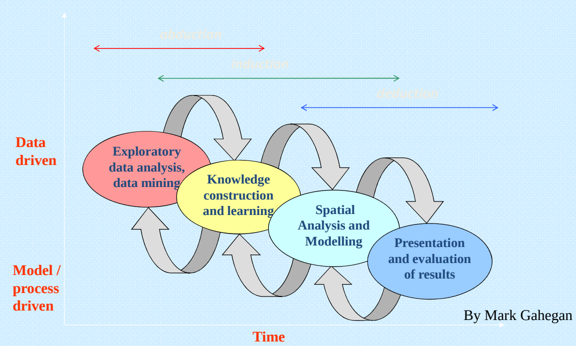
Outline

- GeoViz Toolkit
- STempo
- GeoTxt.org
- GeoViz.js
- Cloud and Server GIS

GeoViz Toolkit



The Process of Inference



Starplot Overview

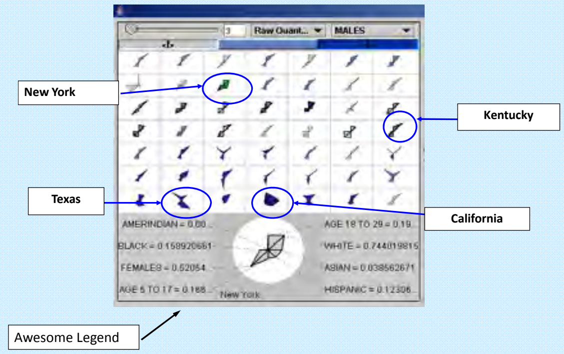
- Starplots
  - What and why of Starplots
  - My innovations

Why Starplots?

- Multivariate Exploratory Mapping (Geovisualization)
- Good for: a large number of attributes
  - Think of students who struggle to make a thematic map with two variables
- Bad for: a large number of observations

What are Starplots?

What are Starplots?

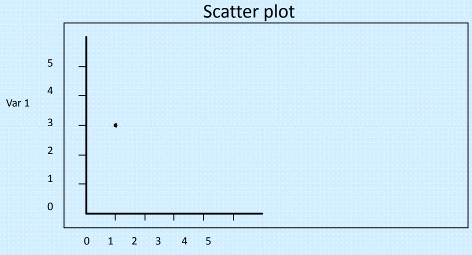


New thing: Starplot Maps

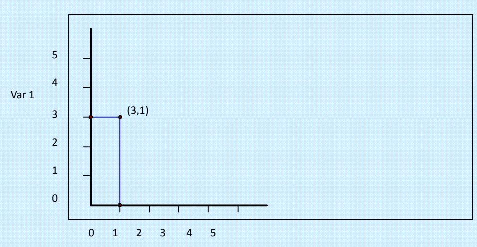
Demo

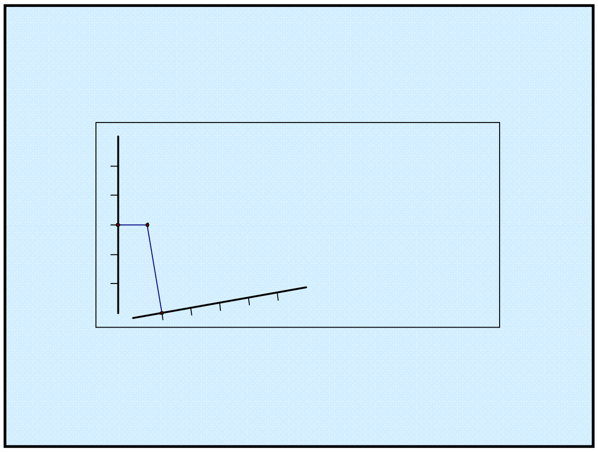
What is a parallel coordinate plot?

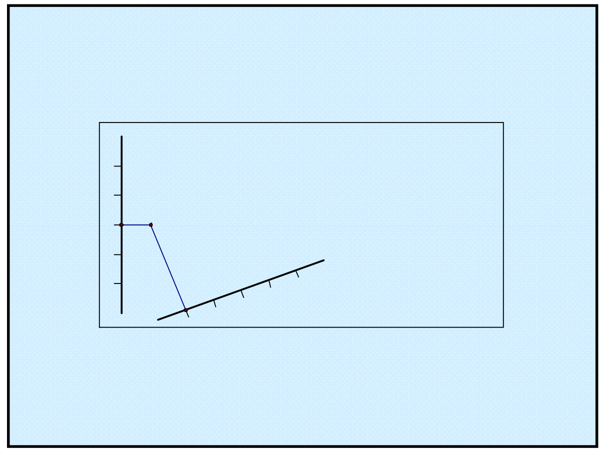
Scatter plot

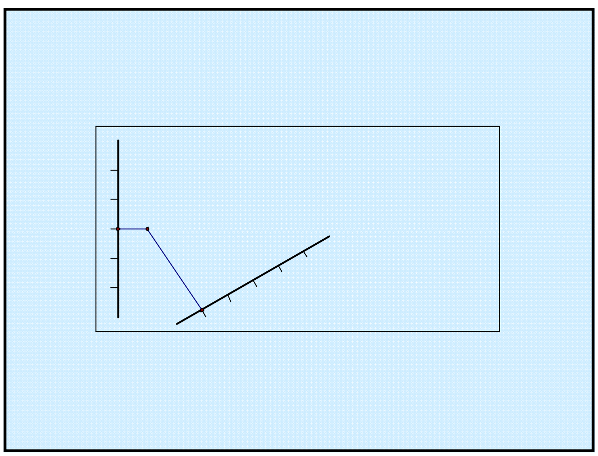


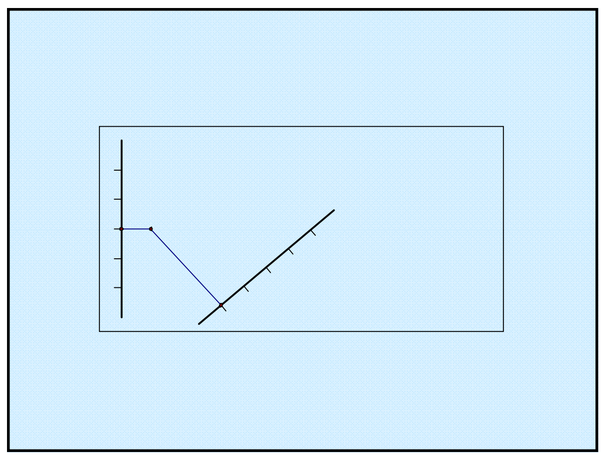
Slides by Linda Pickle at NCI

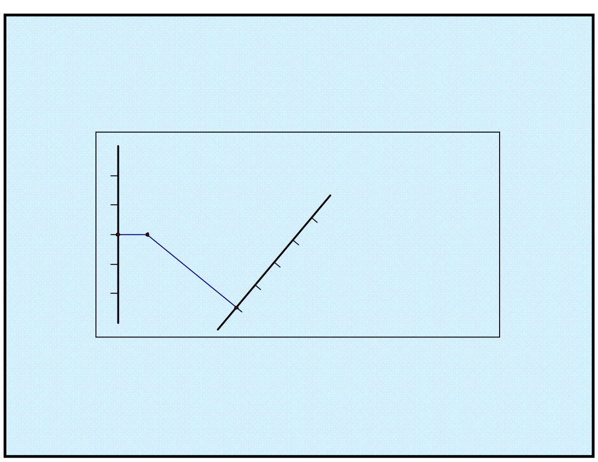


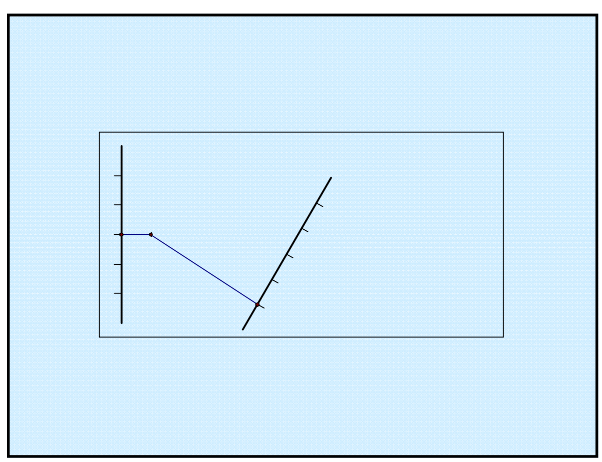


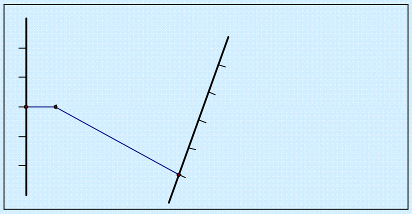


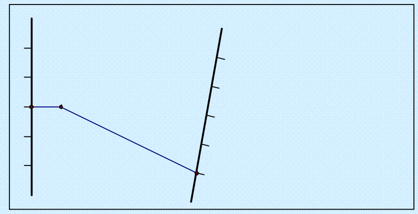


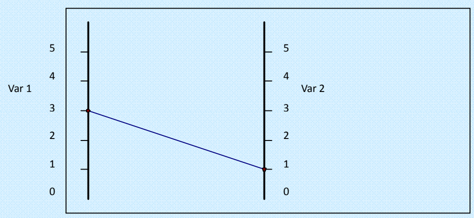


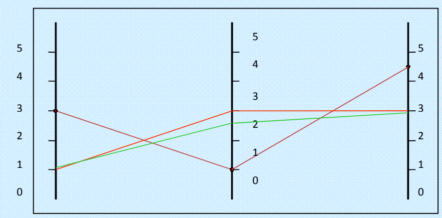




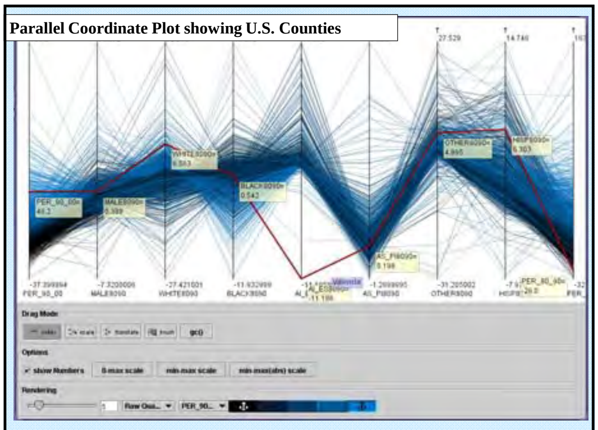






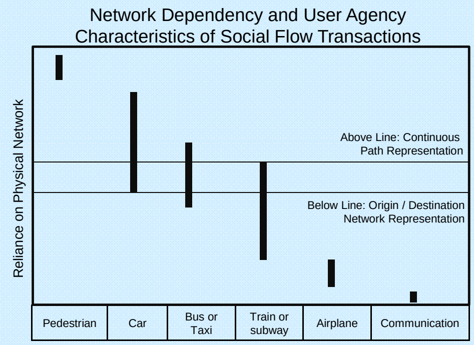


Parallel Coordinate Plot showing U.S. Counties

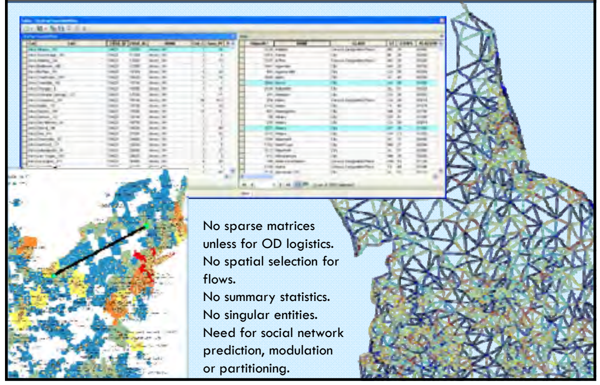


Demo

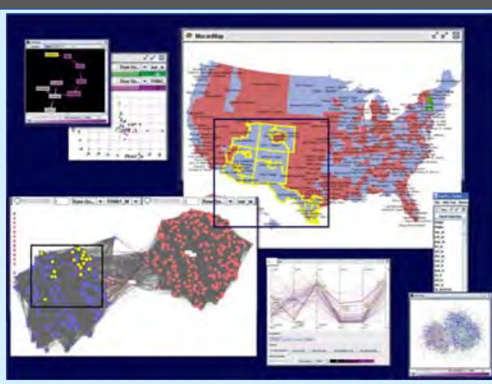
Need for Flow Theory: Edge Assignment



Lack of GIS Infrastructure



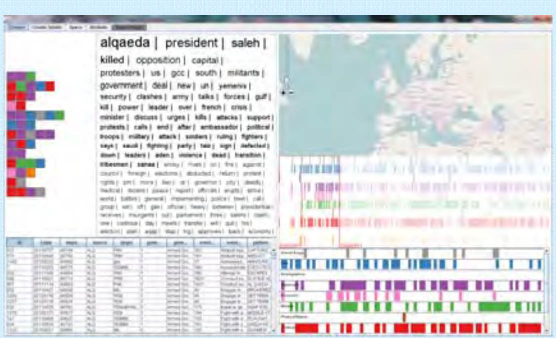
Need for ESDA Tools



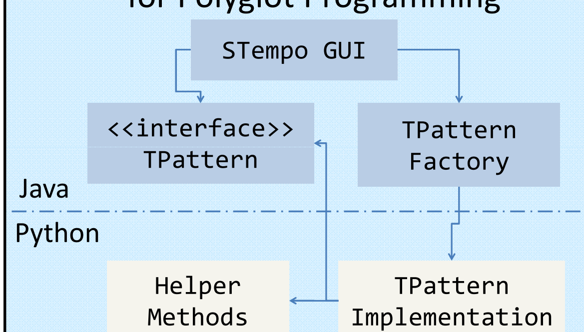
Open Alternatives

- R
- Build your own with D3.js
- Geometry: JTS, GEOS, Shapely

STempo



Factory Pattern in STempo for Polyglot Programming

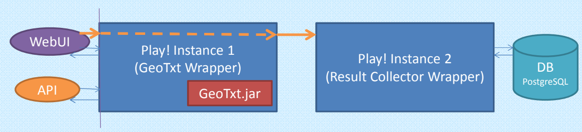


Demo

Open Alternatives

- Same as GeoViz Toolkit

GeoTxt.org

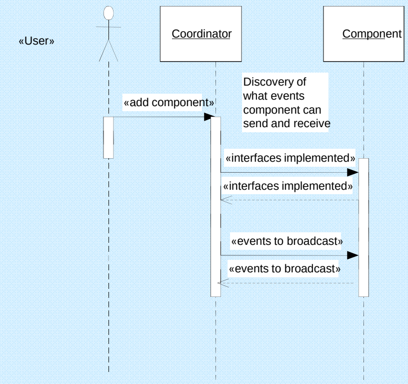


Demo

Open Alternatives

- Two steps: NER and Geocoding
- NER
  - Stanford NLP
  - GATE
- Geocoding
  - GeoNames local
  - GeoNames webservice
- NOT Google, Yahoo APIs if you need to respect TOS

GeoViz.js



SensePlace2

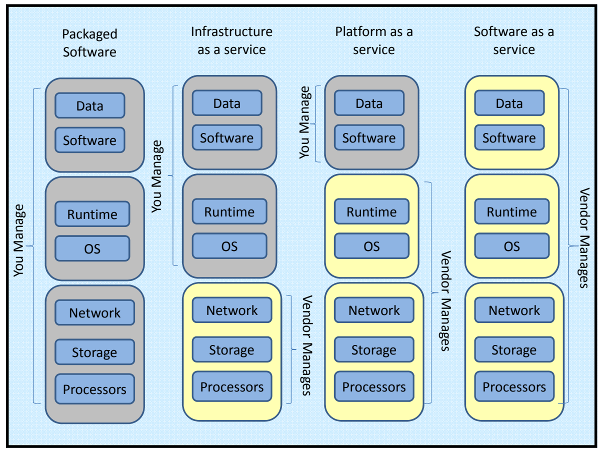
Demo

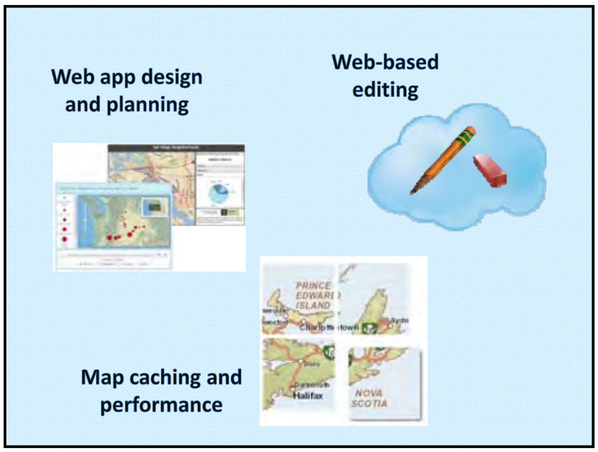
Open Alternatives

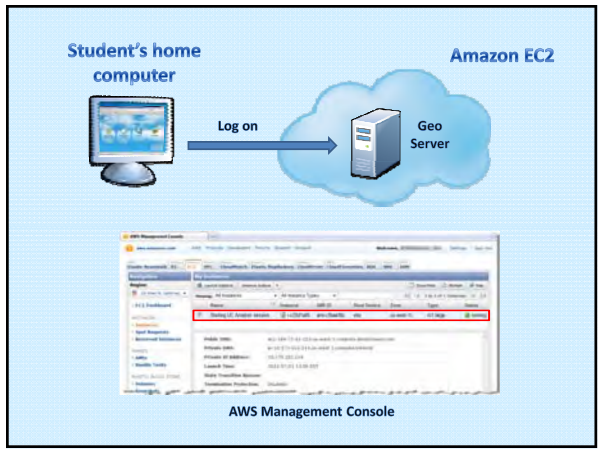
- Tons of JavaScript frameworks that do coordination of some kind
  - jQuery, Backbone, Ember, Angular, Tangle….
- Write your own

GEOG 897C: Cloud and Server GIS

Network accessibility
On-demand service
Resource pooling
Elasticity
Metered service







Demo

Open Alternatives

- Cloud providers
  - Amazon
  - Heroku
- MOOCs rock the world
  - Coursera
  - edX
  - Udacity
- Coursera Demo

Table of Contents

Pep Talk

Spatial Data is Hot!

Collaborators and Sponsor

Research:	Teaching:	Sponsors:
Donna Peuquet	Sterling Quinn	NGA -- STempo
Alan MacEachren	Ryan Baxter	DHS -- VACCINE
Alex Klippel	Jim Sloan	Amazon
Anthony Robinson	Jim Detwiler	
	Anthony Robinson	
Jim Kroon	Marty Gutowski	
Sam Stehle	Mark Wherley	
	Wonderful students	

Thank you!

Table of Contents

Visualizing Uncertainty and Decision-Making

Alan M. MacEachren

GeoVISTA Center
Department of Geography
Penn State University
maceachren@psu.edu

University of Zurich, March 21, 2013

Outline

- What is uncertainty (in the context of geographical information)?
- How *can* we depict uncertainty visually (in static and interactive displays)?
- How *should* we depict uncertainty? What do we know about which methods work?
- What happens when we depict uncertainty? What are the implications for reasoning and decisions?

Outline

- What is uncertainty (in the context of geographical information)?

Categorizing Uncertainty: GIScience

Data Quality \ Data Type	Positional Accuracy	Attribute Accuracy	Logical Consistency	Completeness	Lineage
Discrete Points and Lines	Size Shape (Error ellipses) (Epsilon bands)	Value Color Saturation (Feature code checks)	Color mixing Redundancy by overprinting Slivers by solid fills (Topological cleaning)	Mapping Technique Density traces Marginalia Generalization algorithm Mapping tolerance Buffer size	Mapping Technique Minimum Bounding Rectangles
Categorical Aggregation & Overlay (Tesselation, tiling, Areal coverages)	Texture Value (Certainty of boundary location)	Color mixing (Attribute code checks) (Topographic classifier)	lack error models	Mapping Technique Missing values Logical adjacency surface Marginalia Discrete model weights	
Partitioning & Enumeration (Metric class breaks)	not meaningful	Size = height (Blanket of error)	Size = height (Maximum likelihood prism maps)	Mapping Technique Missing values Missclassification matrix Classing scheme OAL/TAI	Marginalia Source of data Scale/Resolution Date Geometry
Continuous Interpolation (Surfaces and volumes)	no clear distinction b/w the two Value Color Saturation (Continuous tone vignettes) (Continuous tone isopleths)		Size = line wt Color Shape = compactness (TIN links)	not possible by definition Mapping Technique Surface of search attenuation Marginalia Interpolation algorithm	

Graphical Syntax ←→ Graphical/Lexical Syntax

after: Buttenfield, B.P. and Weibel., R. 1988: Visualizing the Quality of Cartographic Data., Presented at Third International Geographic Information Systems Symposium (GIS/LIS 88), San Antonio, Texas..

Categorizing Uncertainty: Cartography

	Accuracy	*Precision*
Location	***position*** of vegetation boundaries	***state*** birth rate
Attributes	***total*** HIV positive caes/county	soil ***order***
Time	***date*** of the last glacier	mean ***monthly*** rainfull

after: MacEachren, Alan M. (1992) Visualizing uncertain information. *Cartographic Perspectives*, 13:10-19.

Categorizing Uncertainty

Category	Components		
	Space	**Time**	**Attributes**
Accuracy/ error	coordinates., buildings	+/- 1 day	counts, magnitudes
Precision	1 degree	once per day	nearest 1000
Lineage	geographic sources/transforms	time sources/transforms	attribute sources/transforms
Consistency	from / for a place	5 say Mon; 2 say Tues	multiple classifiers
Currency/ timing	age of maps	C = Tpresent - Tinfo	census data
Credibility	knowledge of place	reliability of model	U.S. analyst vs. informant
Subjectivity	local ←→ outsider	expert ←→ trainee	fact ←→ guess
Interrelatedness	source independence	source independence	source independence

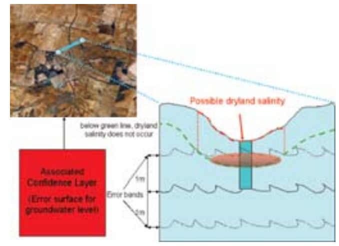
after: MacEachren, A.M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M. and Hetzler, E. 2005: Visualizing Geospatial Information Uncertainty: What we know and what we need to know. *Cartography & Geographic Info. Science* 32, 139-160.

Outline

- What is uncertainty (in the context of geographical information)?
- How *can* we depict uncertainty visually (in static and interactive displays)?

error bounds to depict uncertainty

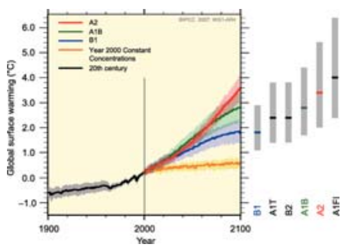
"Visualisation of uncertainty using a cross-sectional transect (Lowell et al. 2007). The confidence layer depicted here can be represented by the statistical 90% confidence interval"



Benke, K.K., Pettit, C.J. and Lowell, K.E. 2011: Visualisation of spatial uncertainty in hydrological modelling. Journal of Spatial Science 56, 73-88. {figure reproduced from: Lowell, K.E., Benke, K.K., & Pettit, C. (2007) Spatial considerations for uncertainty in hydrological modelling and implications for visualisation. SSC2007: Biennial Internationsl Conference of the Spatial Sciences Institute. Hobart, Tasmania.}

depicting error range and agreement:
IPCC Fourth Assessment Report: Climate Change 2007

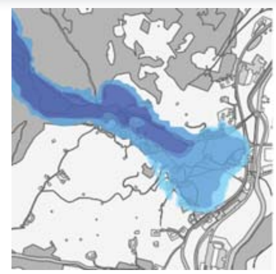
Figure SPM.5. Solid lines are multi-model global averages of surface warming (relative to 1980–1999) for the scenarios A2, A1B and B1, shown as continuations of the 20th century simulations. Shading denotes the ±1 standard deviation range of individual model annual averages. The orange line is for the experiment where concentrations were held constant at year 2000 values. The grey bars at right indicate the best estimate (solid line within each bar) and the likely range assessed for the six SRES marker scenarios. The assessment of the best estimate and likely ranges in the grey bars includes the AOGCMs in the left part of the figure, as well as results from a hierarchy of independent models and observational constraints. {Figures 10.4 and 10.29}



http://www.ipcc.ch/publications_and_data/ar4/wg1/en/figure-spm-5.html

Uncertainty via visual variables: transparency/color value

Uncertainty visualization of snow avalanche intensities (buffer of 10%)

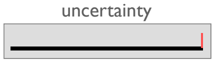


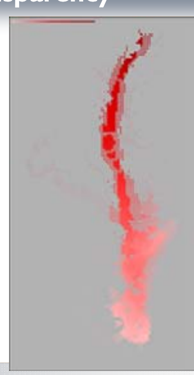
Kunz, M., Lienert, C. and Hurni, L. 2009: Uncertainty visualizations in hazard maps for consistent decision making. *Cartography and Geoinformatics for Early Warning and Emergency Management: Towards Better Solutions*, Prague, Czech Republic.

Uncertainty via visual variables: size/transparency

- dynamic depiction of data (dissolved inorganic nitrogen) and data uncertainty

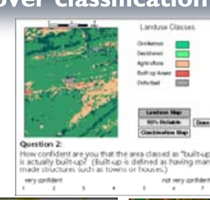
uncertainty

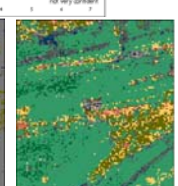




Uncertainty via visual and dynamic variables: Landcover classification

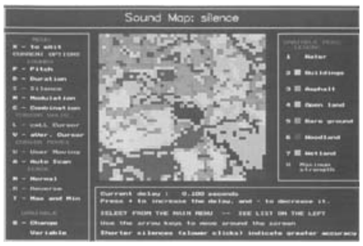
- Evans, B.J. 1997: Dynamic display of spatial data-reliability: does it benefit the user? *Computers & Geosciences, special issue on Exploratory Cartographic Visualization 23, 409-422.*





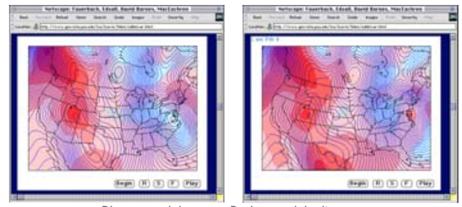
Uncertainty via sonic variables: Landcover classification

- Fisher, P. (1994). "Hearing the reliability in classified remotely sensed images." Cartography and Geographic Information Systems 21(1): 31-36.



Uncertainty as a data layer: Meteo. forecast uncertainty

Fauerbach, E., Edsall, R., Barnes, D. and MacEachren, A. 1996: Visualization of uncertainty in meteorological forecast models. In Kraak, M.J. and Molenaar, M., editors, *Proceedings of the International Symposium on Spatial Data Handling, Delft, The Netherlands, August 12-16: Taylor & Francis, 465-476.*



Blue = models agree; Red = models disagree

http://www.geovista.psu.edu/sites/icavis/icavis/febm/sdhbivar.html

Visualizing uncertainty & ensemble data

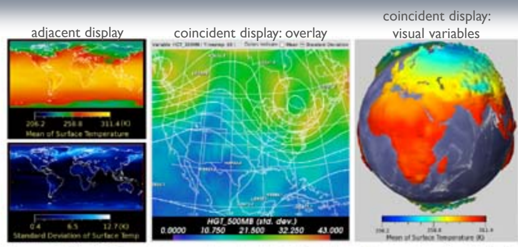
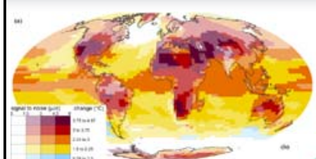


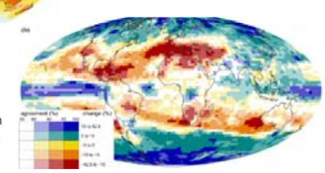
Figure 3. Visualizations of mean and standard deviation. (Left) Mean and standard deviation are visualized independently using color maps. (Center) Mean is presented through a color map, and standard deviation is shown as an overlaid contour. (Right) Standard deviation is mapped to a height field and mean is color mapped.

Potter, K., Wilson, A., Bremer, P.-T., Williams, D., Doutriaux, C., Pascucci, V. and Johhson, C. 2009: Visualization of uncertainty and ensemble data: Exploration of climate modeling and weather forecast data with integrated ViSUS-CDAT systems. *Journal of Physics: Conference Series 180, 012089.*

Visualizing uncertainty & ensemble data coincident display: visual variables



Estiamted temperature change between 1961–1990 and 2070–2099 for the mean of the IPCC AR4 ensemble based on the SRES A1B scenario



Estiamted precipitation change between 1961–1990 and 2070–2099 for the mean of the IPCC AR4 ensemble based on the SRES A1B scenario

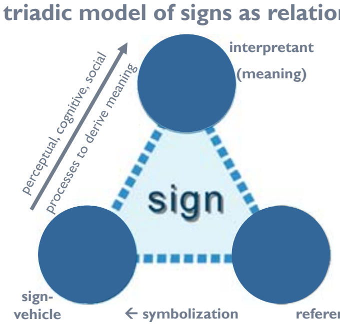
Kaye, N., Hartley, A. and Hemming, D. 2011: Mapping the climate: guidance on appropriate techniques to map climate variables and their uncertainty. Geoscientific Model Development Discussions 4, 1875-1906.

Outline

- What is uncertainty (in the context of geographical information)?
- How *can* we depict uncertainty visually (in static and interactive displays)?
- How *should* we depict uncertainty? What do we know about which methods work?

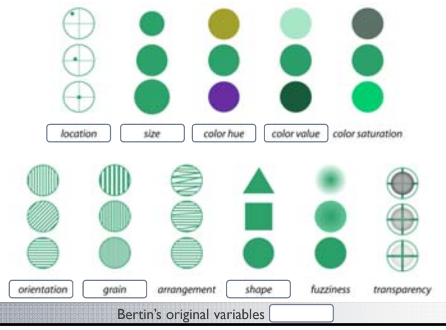
Semiotics: the science of "signs" and sign systems

triadic model of signs as relations



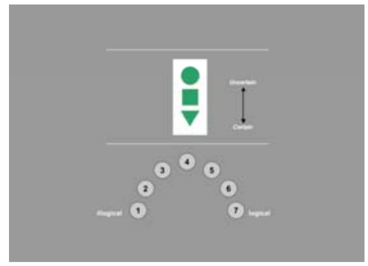
Visual Semiotics: which visual variable imply uncertainty?

- Visual variables

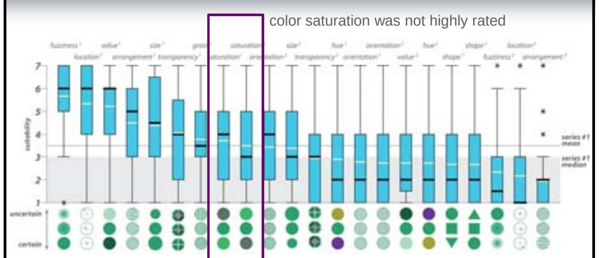


Experiment 1: Assessing intuitiveness

- Participants (31 in pilot + 72 in main study) rate the logic (intuitiveness) of the symbol set



Series 1: Results



fuzziness, *location*, and *value* rate high; *arrangement* is marginally successful

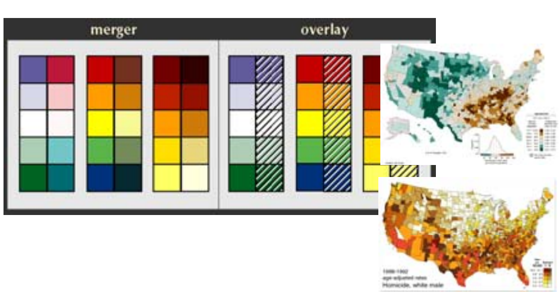
color hue and *shape* rate low (as expected), as do *inverted* *fuzziness*, *location*, *value*

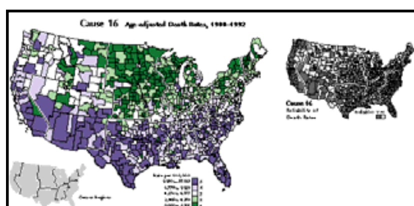
NCHS atlas color/reliability schemes

Cindy Brewer and **Alan MacEachren** researched the color and reliability map symbols for the *Atlas of United States Mortality* by the National Center for Health Statistics (CDC).

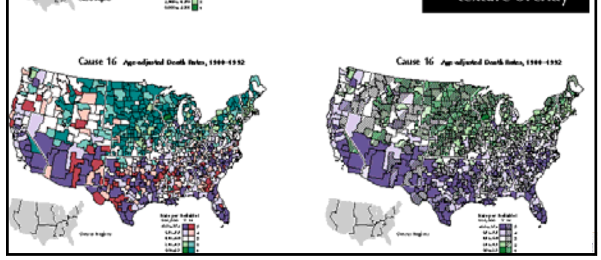
www.cdc.gov/nchswww/products/pubs/pubd/other/atlas/atlas.htm

colors/textures to represent reliability

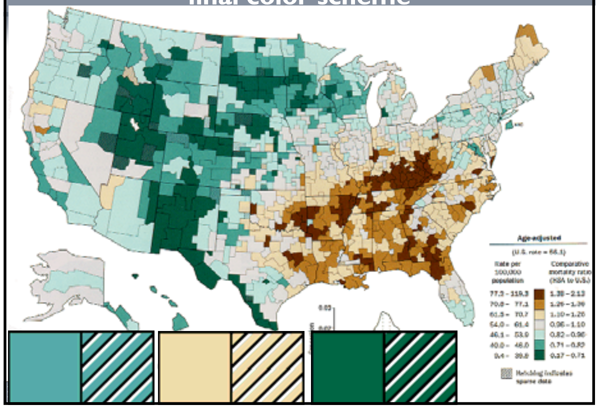




Sample Test Maps:
- cause: 16
- scale: quarter
- color scheme: purple-green
- reliability schemes: map pairs color change texture overlay



final color scheme



Outline

- What is uncertainty (in the context of geographical information)?
- How can we depict uncertainty visually (in static and interactive displays)?
- How should we depict uncertainty? What do we know about which methods work?
- What happens when we depict uncertainty? What are the implications for reasoning and decisions?

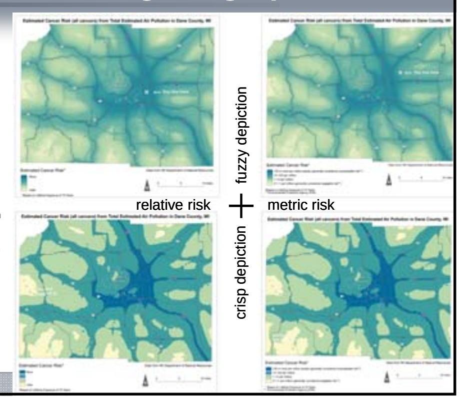
Decision-making

- The impacts of uncertainty on analytical reasoning and decision-making are not well understood. Only a few studies have been carried out, e.g.
  - Deitrick and Edsall (2006) present evidence that uncertainty information has an influence on decision making.
  - Leitner and Buttenfield (2000) found that uncertainty representation for facility location can clarify the map and lead to more accurate decisions.
  - Aerts, et al (2003) report (for urban planning) that uncertainty information can improve decision-making efficiency.
  - Hope and Hunter (2007) found evidence for "ambiguity aversion" and associated irrational decisions when uncertainty was depicted on maps. In complementary research, different methods for depicting positional uncertainty resulted in very different decision behaviors related to travel risks (Hope and Hunter, 2007).
  - Severtson and Meyers (2012), found that geo-uncertainty representation factors ("focus" / crispness of map contours and verbal-relative versus numeric risk expression) and personal characteristics (prior beliefs and numeracy) interacted to influence risk beliefs related to modeled cancer risk from air pollution.
  - Cliburn, et al (2002) "Decisions cannot be made from uncertain data; it only leads decision-makers to discount the results. Unfortunately, not considering uncertainty may lead to inappropriate decisions. A potential collaborator, who viewed the application in its later stages, suggested incorporating a reasoning network of potential actions to problems presented by the visualizations."

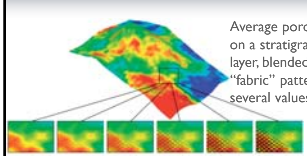
Communicating ambiguity in risk

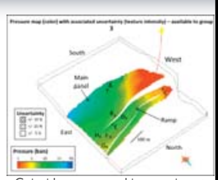
"Overall, results indicate incremental shading effectively conveys a dose-response message, and contour focus and risk expression show promise for conveying information uncertainty."

Severtson, D.J. and Myers, J.D. 2012: The Influence of Uncertain Map Features on Risk Beliefs and Perceived Ambiguity for Maps of Modeled Cancer Risk from Air Pollution. *Risk Analysis, online pre-print.*



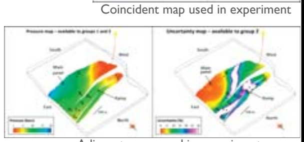
Coincident versus Adjacent display & Decision-making





Coincident map used in experiment

Findings:
- uncertainty influences decision-making
- coincident display better for complex tasks
- uncertainty via coincident depiction does not decrease data interpretation accuracy
- no performance difference for 2.5D vs 2D data-uncertainty depictions

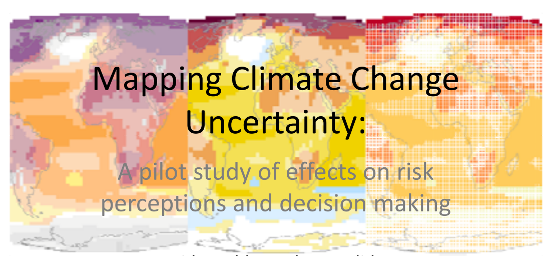


Adjacent maps used in experiment

Viard, T., Caumon, G. and Lévy, B. 2011: Adjacent versus coincident representations of geospatial uncertainty: Which promote better decisions? Computers & Geosciences 37, 511-520.

PENNSTATE

Department of Geography



Mapping Climate Change Uncertainty:

A pilot study of effects on risk perceptions and decision making

David Retchless, PhD candidate

dpr173@psu.edu

for details, see: Retchless, D. 2012: Mapping Climate Change Uncertainty: Effects on Risk Perceptions and Decision Making, Paper #GC43B-1024. *AGU Fall Meeting, San Francisco, CA. fallmeeting.agu.org/2012/eposters/eposter/gc43b-1024*

Research Questions

- Which maps of temperature change and uncertainty are easiest to understand?
- Do map users combine magnitude and certainty of change when:
  - assessing risks?
  - making decisions?
- Which maps do users prefer?

Methods

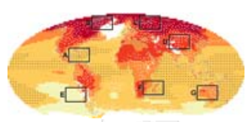
- Survey: Mechanical Turk & Survey Gizmo
- 4 Ranking Questions, each ranking 7 map regions
  - Temperature change
  - Certainty (signal/noise, described as precision)

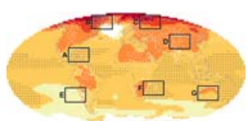
  UNDERSTANDING TEMPERATURE AND CERTAINTY
  - Harm to environment
  - Suitability for nature reserve (given temperature requirements)

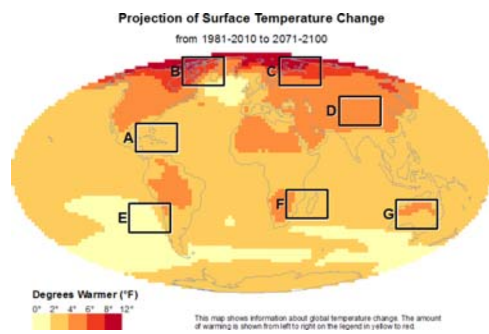
  COMBINING TEMPERATURE AND CERTAINTY

Methods

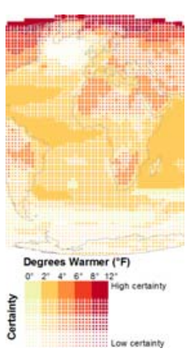
- 274 respondents, randomly assigned 1 of 20 maps:
  - 10 types of maps
  - 2 emissions scenarios (high & low)
  - ~14 respondents/map
  - Data from CMIP5


HIGH

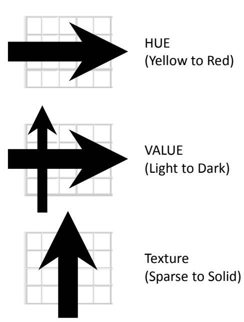

LOW



Control (Temperature Only)



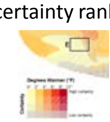
Texture – Squares

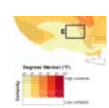


Conclusions: Understanding Magnitude & Certainty

- Temperature ranking was easy, uncertainty ranking was hard.
- Consistent with MacEachren et al. (1998), texture outperformed color for uncertainty ranking.
- Best maps for uncertainty ranking:


Control with small map


Texture – Lines


Texture – Spots

- Magnitude was primary driver of risk assessment and decisions.

Table of Contents

Challenges

- understand components of uncertainty and their relationships to domains, users, and information needs
- develop representation methods for depicting multiple kinds of uncertainty
- develop methods and tools for interacting with uncertainty depictions
- understand how (or whether) uncertainty visualization aids exploratory analysis
- understand how knowledge of uncertainty influences information analysis, decision making, and outcomes
- develop methods to capture and encode analysts' or decision makers' uncertainty
- assess usability and utility of uncertainty capture, representation, and interaction methods and tools.

after: MacEachren, A.M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M. and Hetzler, E. 2005: Visualizing Geospatial Information Uncertainty: What we know and what we need to know. *Cartography & Geographic Info. Science* 32, 139-160.

That's all: Thanks

maceachren@psu.edu
www.geovista.psu.edu

Acknowledgements: Parts of the work represented here were supported by funds from: the Dept of Homeland Security Visual Analytics for Command, Control, and Interoperability Environments Center of Excellence. A large number of people contributed to work represented. These include Mark Gahegan, Anthony Robinson, Robert Roth, …

Table of Contents

Gang Graffiti Automatic Recognition and Interpretation (GARI)

Edward J. Delp

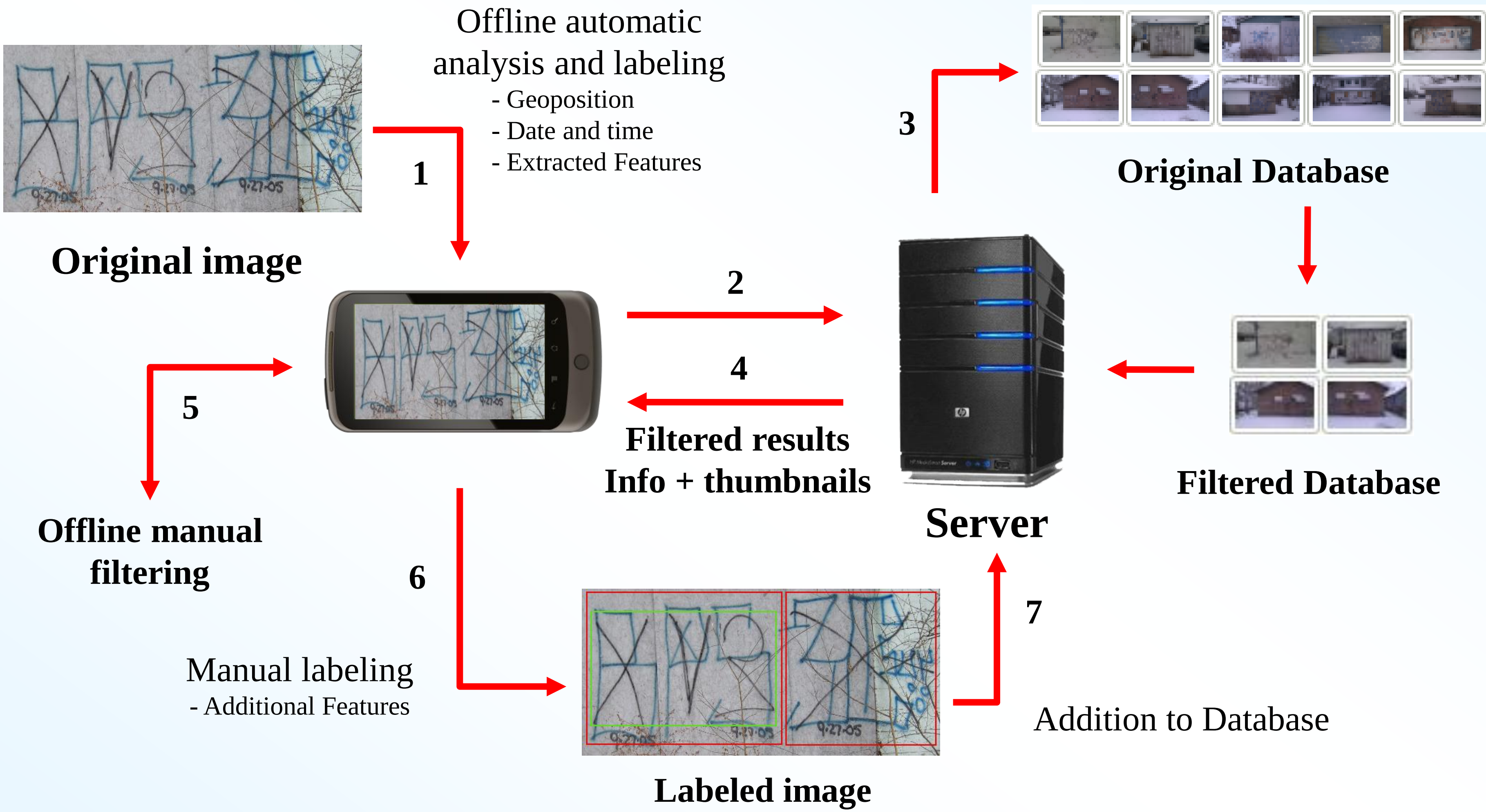
Video and Image Processing Laboratory
School of Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA
ace@ecn.purdue.edu

Acknowledgements

- This work is funded by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001

- Indiana Gang Intelligence Network (INGangNetwork)

GARI System Overview



GARI System Goals

- System that allows first responders to:
  - Collect gang graffiti and gang tattoos images
  - Analyze and interpret the images
  - Image retrieval
  - Interact with a database of images
  - Browse database using an interactive map

- Platforms
  - Android application
  - Web-based interface
  - iPhone

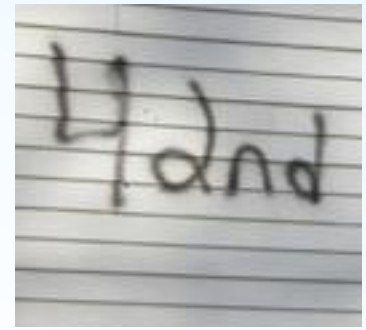
Gang Graffiti

Gang Tattoos

Gang Graffiti Interpretation

SHAPE

SYMBOLS

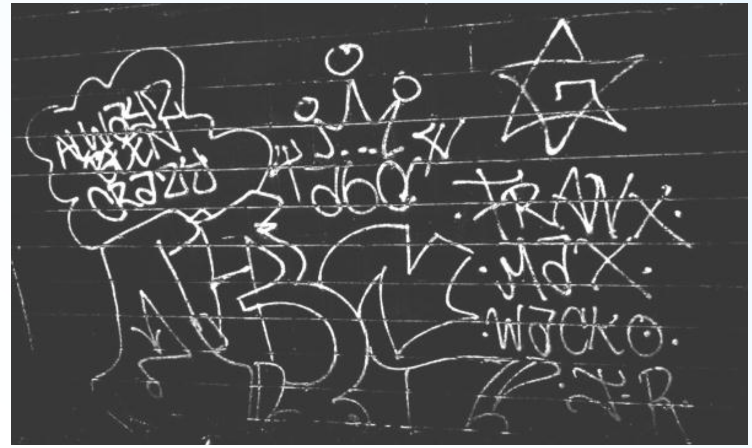


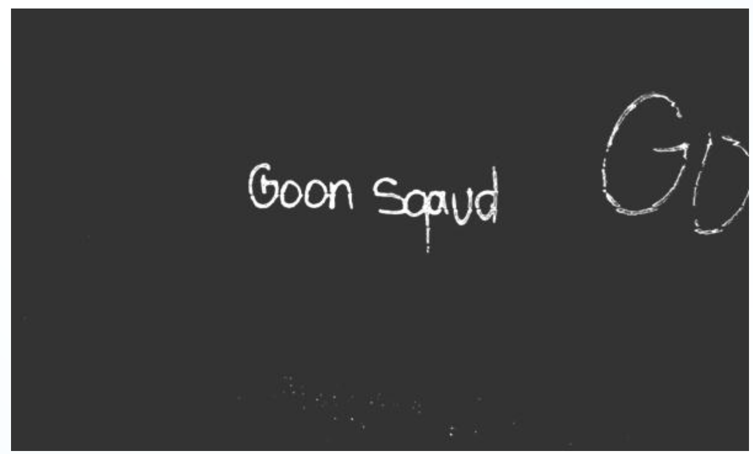
NUMBERS/LETTERS

TIME

COLOR

Image Analysis






2-5
Hillside
Bitch
2500


2-5
Hillside
Bitch


2-5
25
7 6 4
G D

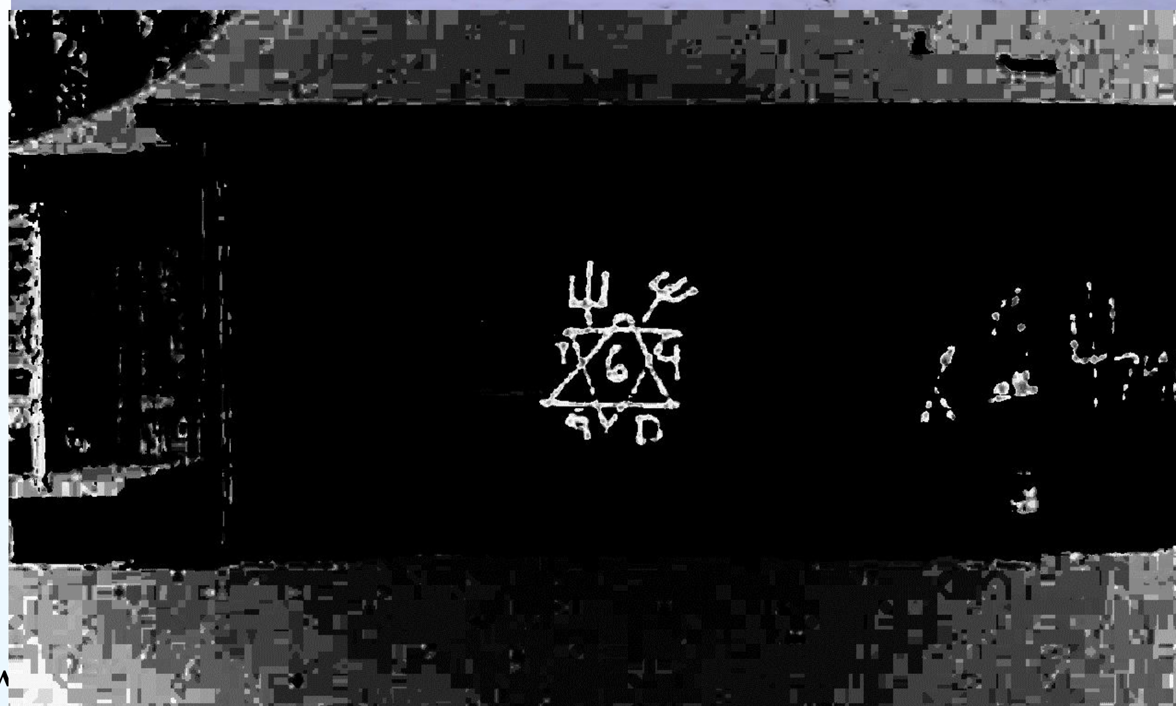


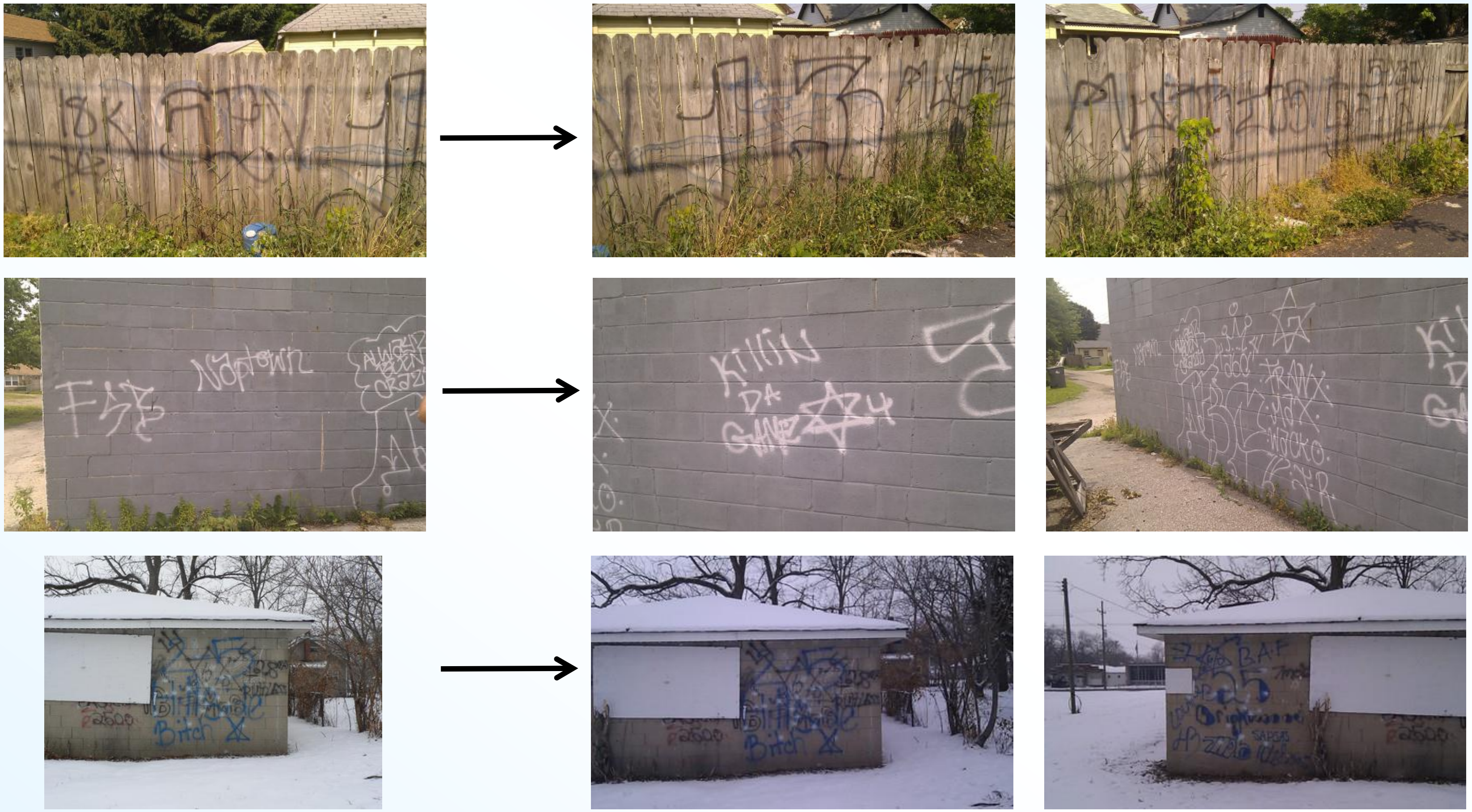
Image Analysis

GANG RIVALRY TRACKING

- **Date:** 08/19/2010
- **Time:** 3.25 PM
- **Geo:** 41.387917, 2.169919

- **Date:** 01/03/2011
- **Time:** 5.11 PM
- **Geo:** 41.387917, 2.169919

Image Retrieval



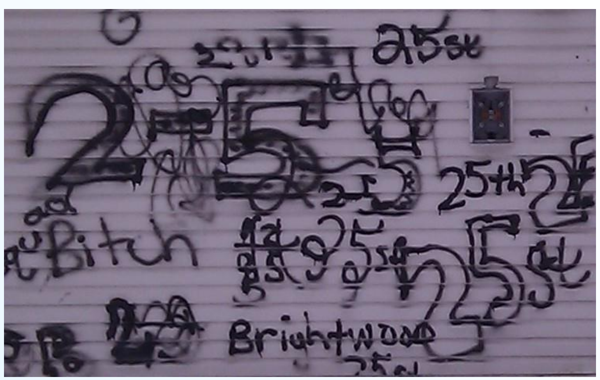
User Filtering

KNOWN GANG

18th Street Gang

CLEAR TOKENS

Ruthless



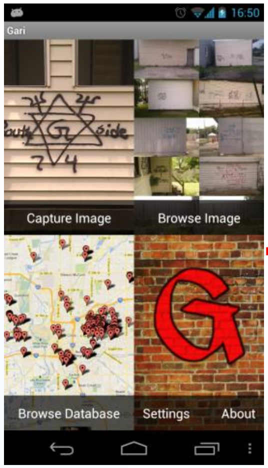
SPECIAL FONT TYPE

KNOWN SYMBOLS

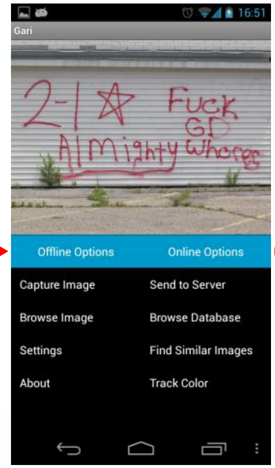
Down arrows

Mobile App

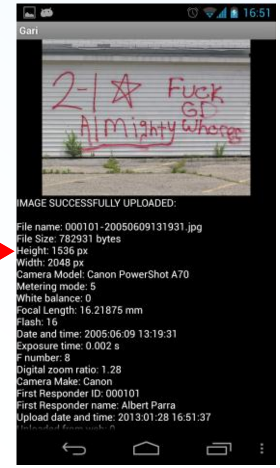
Main Screen



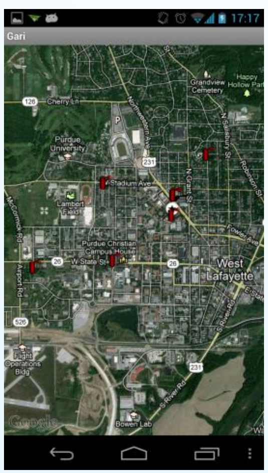
Capture Image

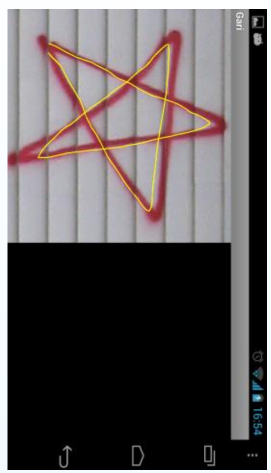


Send to Server



Possible Options





Browse Database **Trace Color** **Find Similar Images**

Web Interface

INDIANA INTELLIGENCE FUSION CENTER

VACCINE

VIPER

Video and Image Processing Laboratory

Main

Internal

Gang Graffiti Recognition and Analysis Using a Mobile Telephone

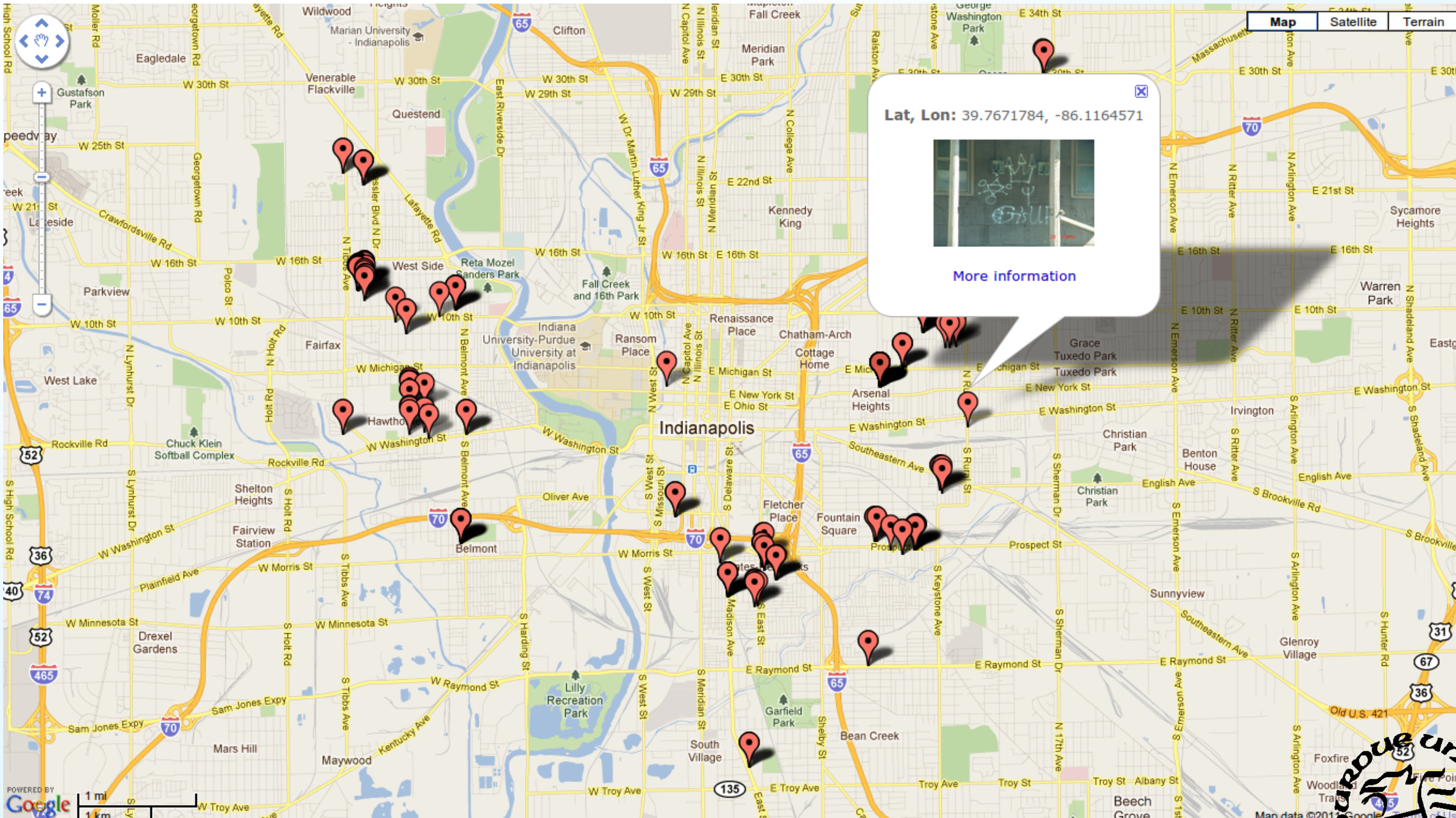
Gangs are a serious threat to public safety throughout the United States. Gang members are continuously migrating from urban cities to suburban areas. They are responsible for an increasing percentage of crime and violence in many communities.

According to the National Gang Threat Assessment, approximately 1 million gang members belonging to more than 20,000 gangs were criminally active within all 50 states and the District of Columbia as of September 2008. Criminal gangs commit as much as 80 percent of the crime in many communities according to law enforcement officials throughout the nation.

Street gang graffiti is their most common way to communicate messages, including challenges, warnings or intimidation to rival gangs. It is, however, an excellent way to track gang affiliation and growth, or even sometimes to obtain membership information.

Web-Based Interface

Browse Database

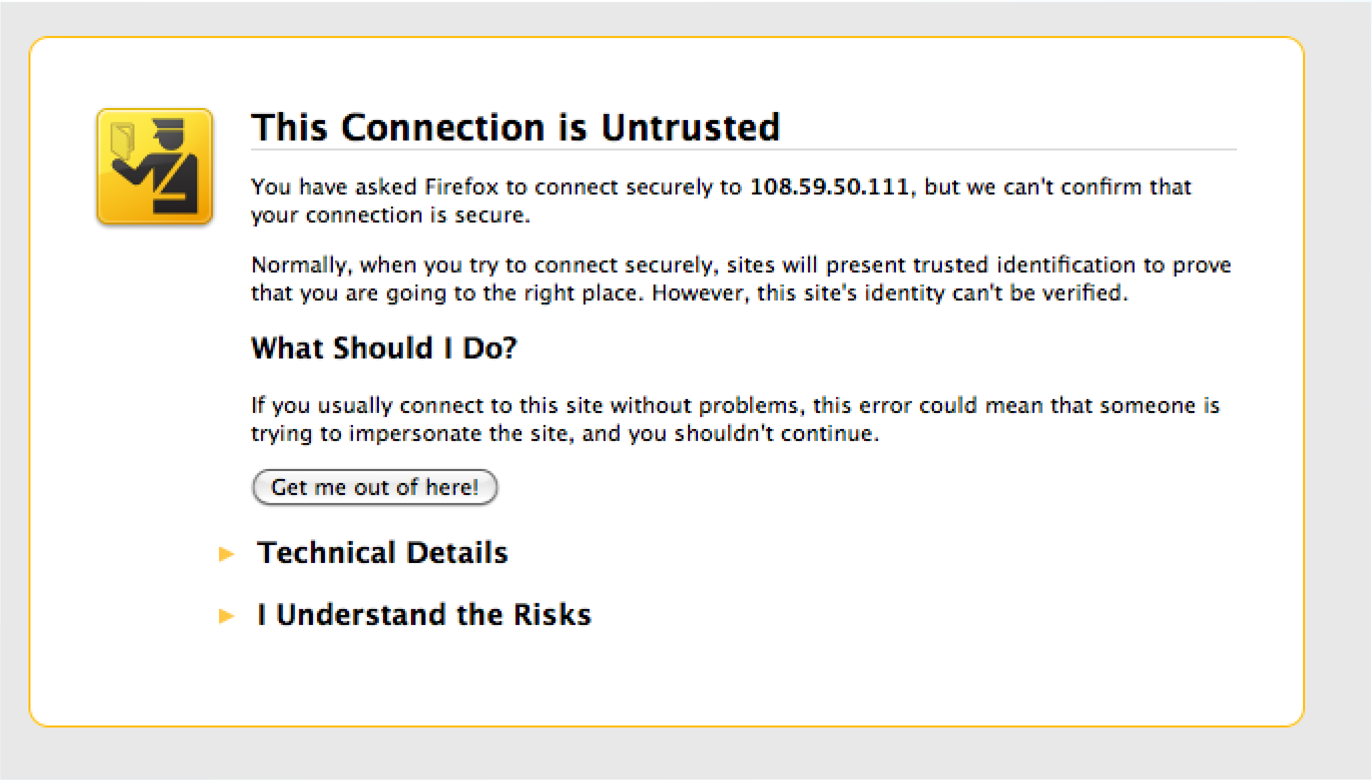


Installation Process

- **Your identity is authenticated by InGang you will receive a "welcome" email from *gari2ind@ecn.purdue.edu***
  - **email contains your GARI credentials (id and password)**
  - **Need to change password on first use (you must also accept the Disclaimer)**
- **Android app – need Gmail email associated with you're the phone – the app is then emailed to you (should be supplied when you register for access)**
- **iPhone – the link is in the "welcome" email (you MUST open the email on the iPhone)**
- **Questions or problems – *gari2ind@ecn.purdue.edu***

One Issue/Bug

- **Web-Based Interface**
  - **https://108.59.50.111/~gari/**



Gang Graffiti Automatic Recognition and Interpretation (GARI)

Edward J. Delp

Video and Image Processing Laboratory
School of Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Automatic Recognition and Interpretation of Gang Graffiti (GARI)

Edward J. Delp

Video and Image Processing Laboratory

School of Electrical and Computer Engineering

Purdue University

West Lafayette, Indiana USA

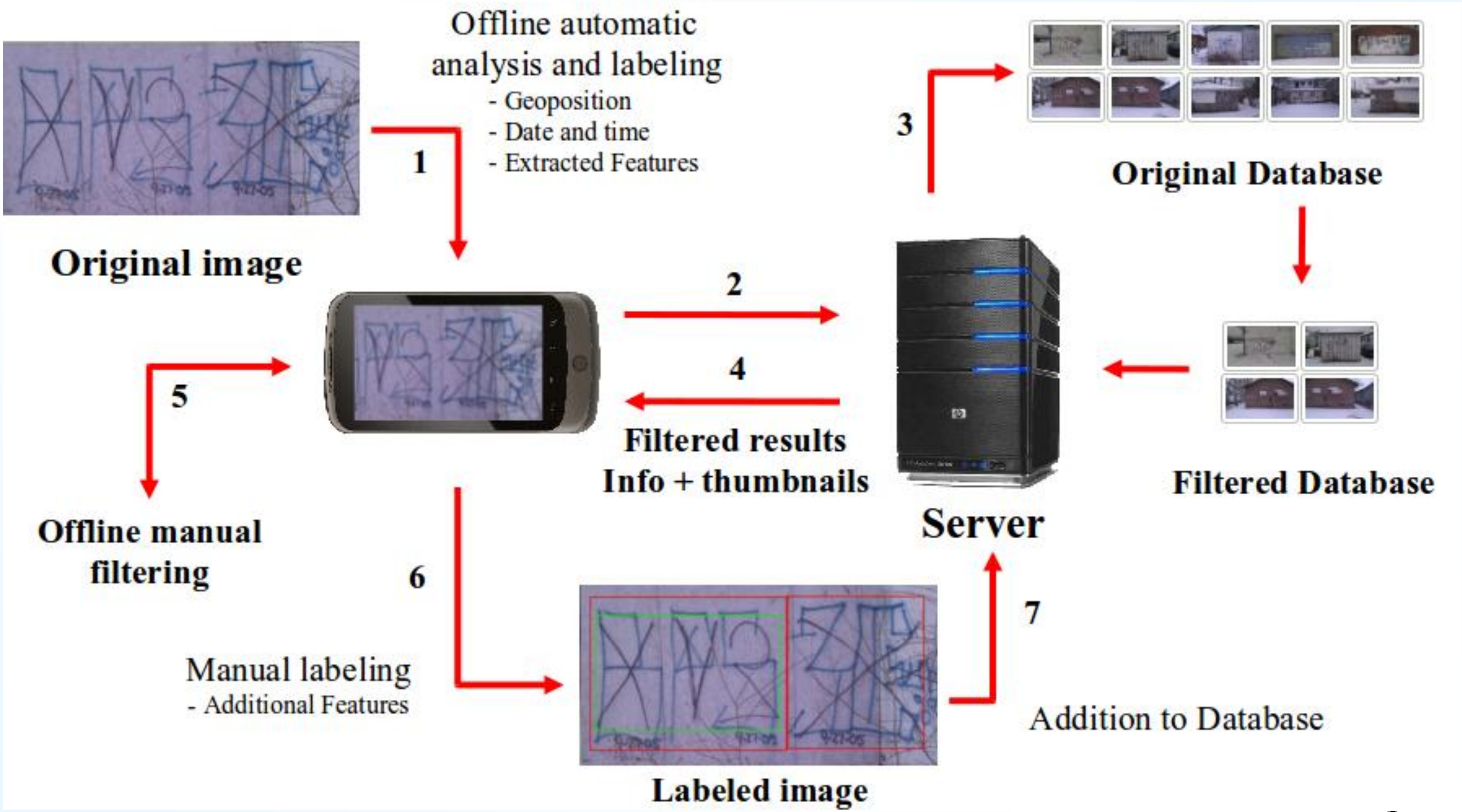
ace@ecn.purdue.edu

Acknowledgements

- This work is funded by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001

- Faculty and graduate students involved have included:
  - Professor Mireille Boutin
  - Andrew Haddad
  - Albert Parra
  - Kevin Lorenz
  - Bin Zhao
  - Joonsoo Kim

GARI System Overview



Project Status

- Deployed on Android Phone
- Main phone application is image database system and browsing (acquisition tool) with some limited analysis capabilities
- Desktop backend
  - Browse by radius
  - Upload images
  - View and edit image details
  - Interact with map
  - Searching for similar images
  - Integration with social media (Facebook and Twitter)

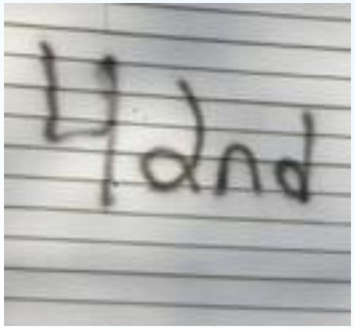
System Goals

- **Network connected phone:**
  - **Image Analysis**
  - **User input**
  - **Interacts with a database of gang symbols**

Gang Graffiti Interpretation

SHAPE

Simple, Straightforward



NUMBERS

42nd street gang

SYMBOLS

6-point star, pitchforks

COLOR

Goon Squad: Red/Black

Gang Graffiti Interpretation

LETTERS

East Side Gang

TIME

Black: 18 ST (18th Street Gang)

Red: 13 SUR (Sureños 13)

POSITION/ALIGNMENT

Letters at star points

Numbers in the middle

Letters at the bottom

Pitchforks upright

Image Analysis – Color Recognition

MEXICANOS
MALDITOS SUREÑOS 13

18 STREET GANG

SUREÑOS 13

Image Analysis – Scene Analysis

GANG RIVALRY TRACKING

- **Date:** 08/19/2010
- **Time:** 3.25 PM
- **Geo:** 41.387917, 2.169919

- **Date:** 01/03/2011
- **Time:** 5.11 PM
- **Geo:** 41.387917, 2.169919

Database Structure Descriptions

Database Fields: automatically populated

- Camera EXIF data
  - Make
  - Model
  - Focal length
  - ISO equivalent speed
  - Exposure time
  - Exposure bias
  - Exposure mode
  - Aperture
  - X/Y resolution
  - White balance
  - Compressed bit per pixel
  - Metering mode
  - Flash
  - F number
  - Interoperability offset
  - Sensing method
  - Custom rendered
  - Digital zoom ratio
  - YCbCr positioning

Database Structure Descriptions

Database Fields: automatically

- **Image properties**
  - **File size**
  - **Resolution**

Database Structure Descriptions

Database Fields: automatically/manually populated

- **Geoposition**
  - **Country**
  - **State**
  - **County**
  - **City**
  - **ZIP code**
  - **Address**

- **Date and time**

Database Structure Descriptions

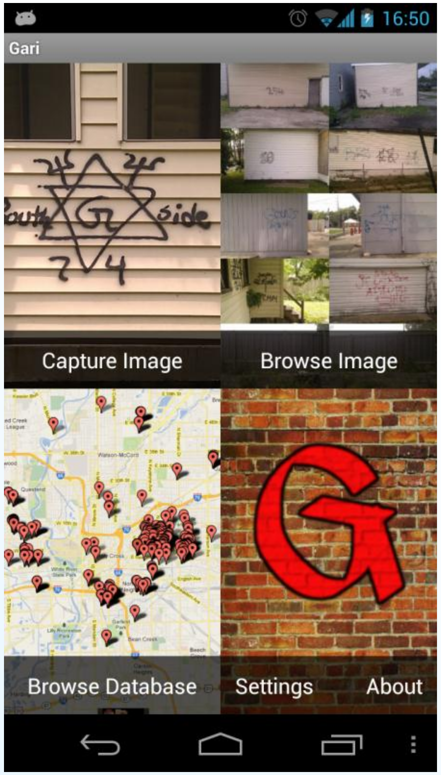
Database Fields: manually populated

- First responder
  - Name/ID
  - Affiliation
- Gang information
  - Gang Name
  - Gang Member
- Tattoo (only GARI)
  - Tattoo?
  - Gang?
  - Prison

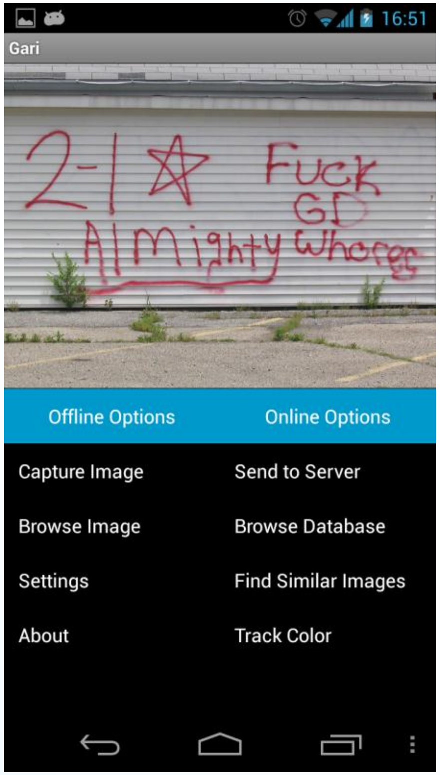
GARI App Features

Gang Graffiti –User Interface

App Version 2.1



Main menu

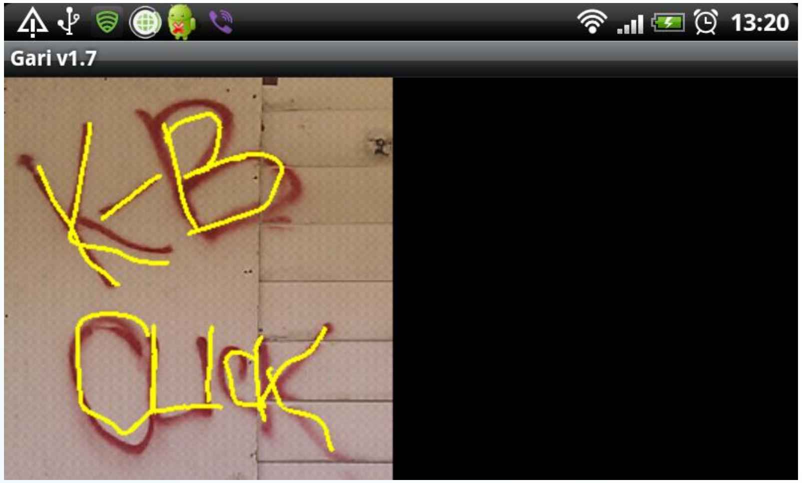


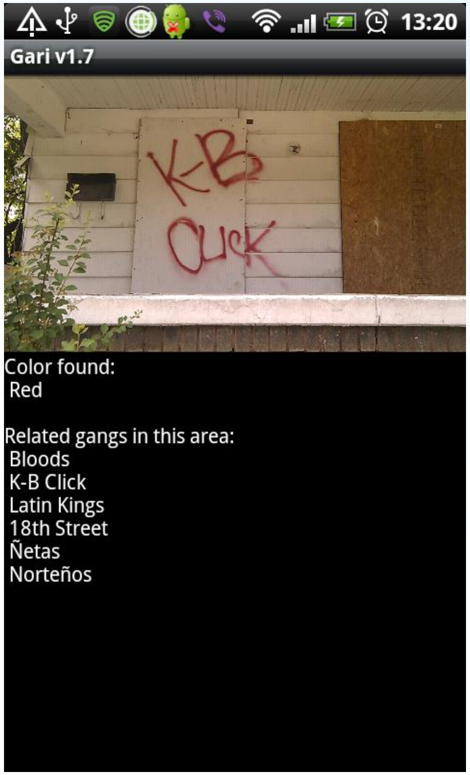
User options

Analysis results

Image matching

Gang Graffiti –User Interface





Color tracking

Gang Graffiti – Browse Database

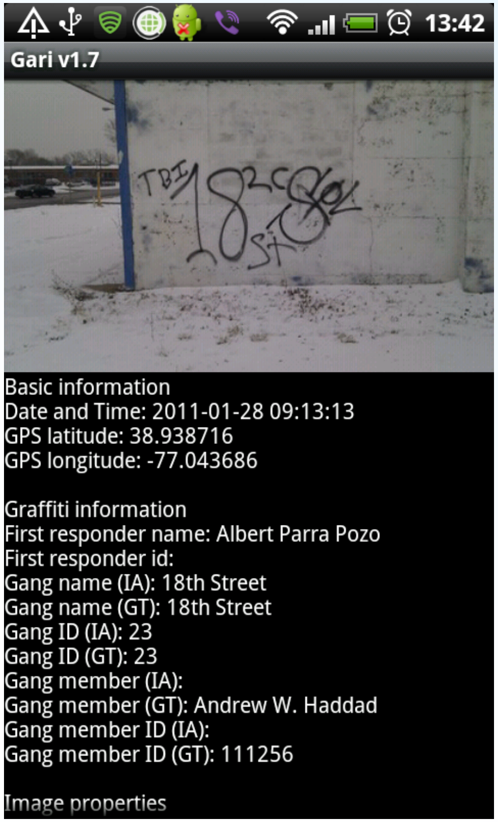
- Browse by radius from current position
- Download images and information from server (image EXIF tags, gang related information)
  - Compare results
  - Track graffiti
- Browse map (Google Maps API)

- Network connection required

Gang Graffiti – Browse Database

Browse general results

Show graffiti on map



Inspect specific graffiti

Gang Graffiti – Browse Database

Desktop Version

Gang Graffiti – Browse Database

Desktop Version

VIPER

Video and Image Processing Laboratory

VACCINE

Main

Archive

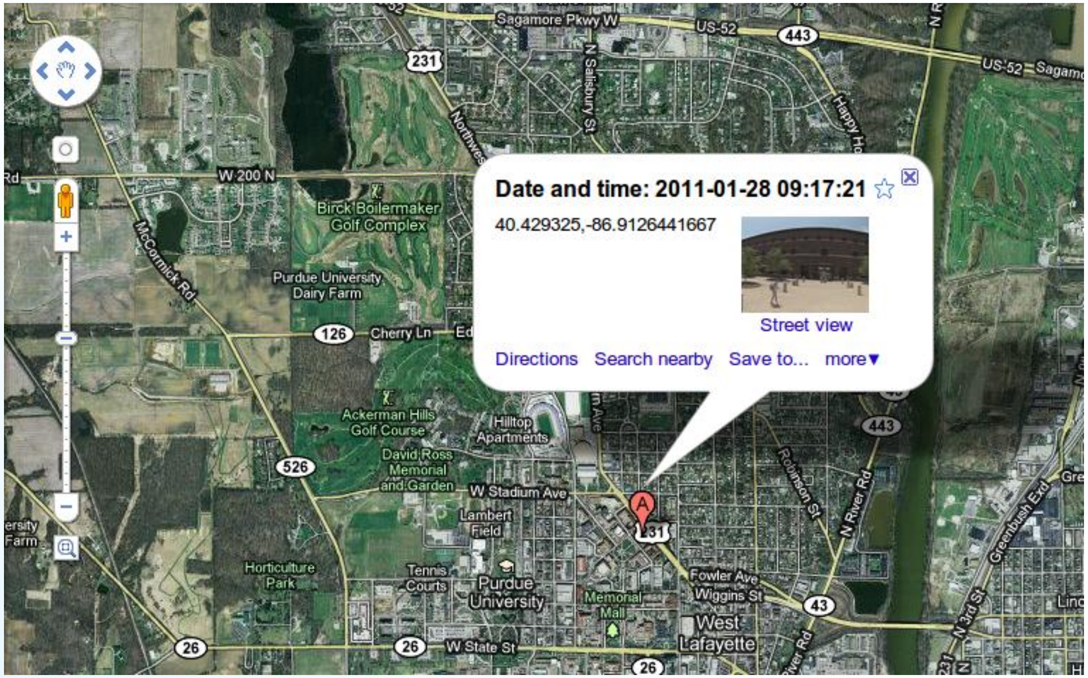
Browse Database

Date and Time: 2011-01-28 09:17:21
GPS latitude: 40.429325
GPS longitude: -86.9126441667
File Size: 664501 bytes
Height: 1552 px
Width: 2592 px
Focal length: 4.31 mm
Camera make: HTC
Camera model: HTC Desire
Image id: 1111
Show in map

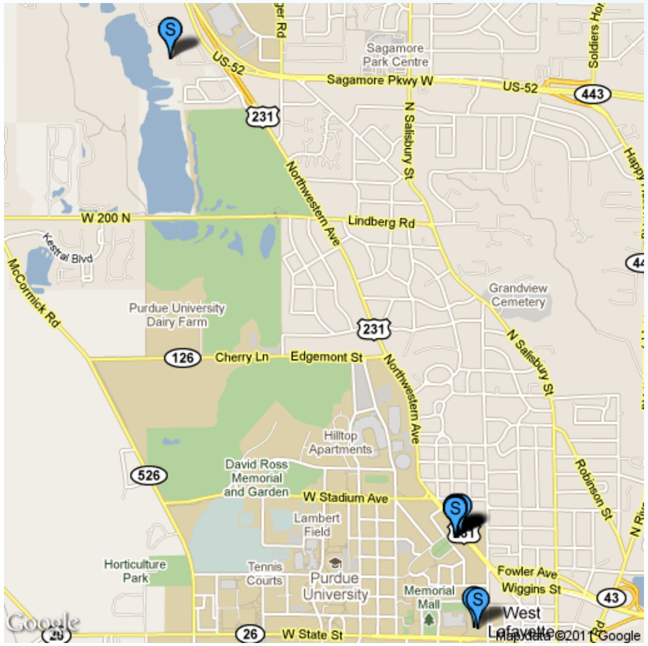
Gang Graffiti – Browse Database

Desktop Version

Google maps



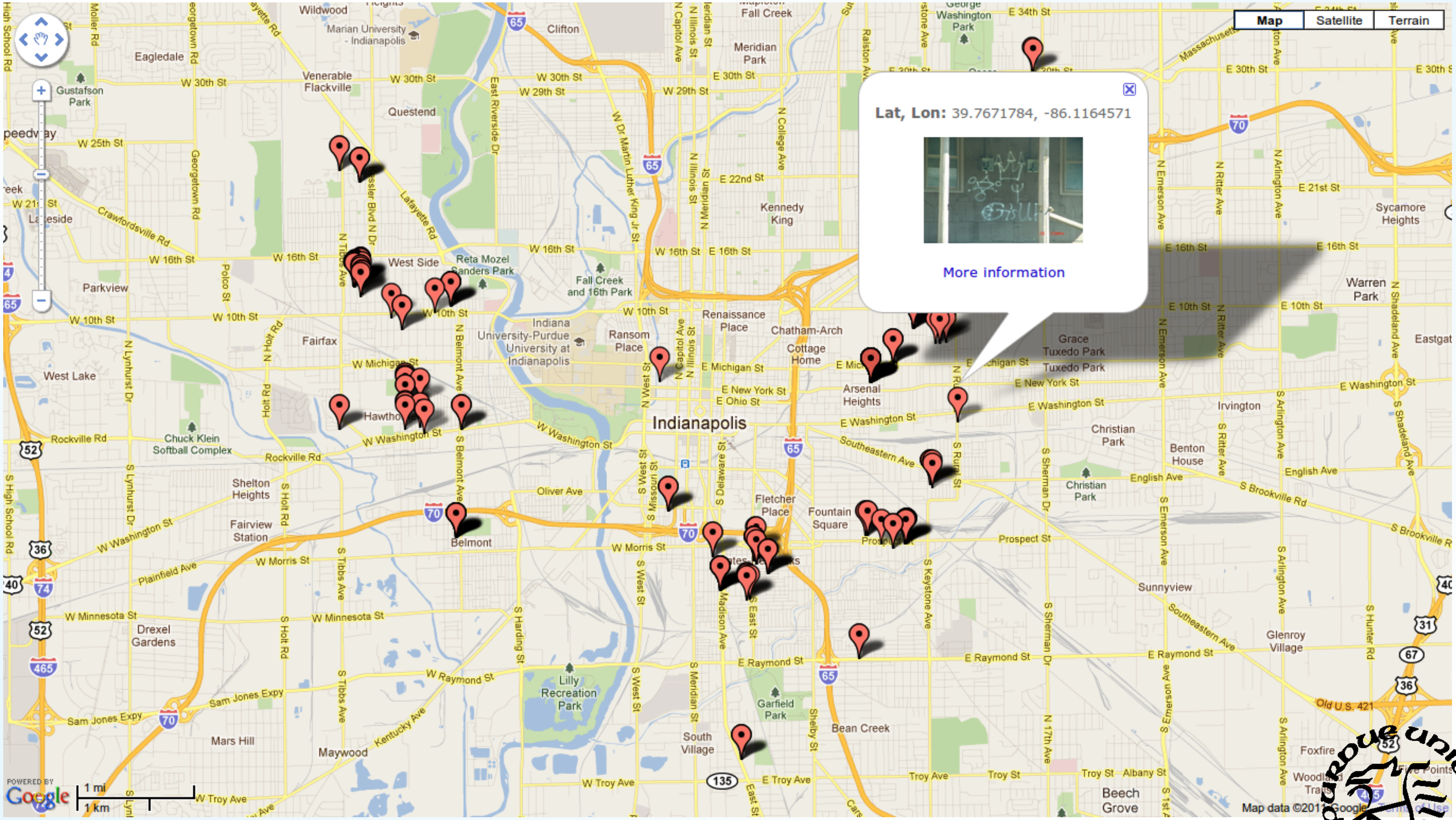
Specific image



All images in radius

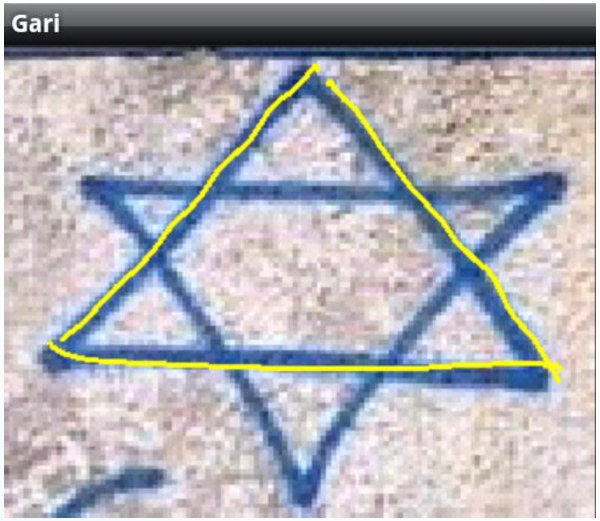
Gang Graffiti – Browse Database

Desktop Version



Gang Graffiti – Color Recognition

Device



Finger tracking

Color detection

Project Status

- Android App Version
  - Browse by radius
  - Upload images
  - View image details
  - Find similar images
  - Track color
  - Interact with map
  - Uses SSL security
  - Password protected

Project Status

- Desktop Database Version
  - Browse by radius/date/address
  - Upload images
  - View and edit image details
  - Interact with map
  - Uses SSL security
  - Password protected

How to Get The App?

- Contact gari@ecn.purdue.edu
  - It requires Android phone (app version 2.1)
  - We assign user IDs and initial passwords
  - Database server

System Requirements for GARI server

- Hardware
  - Intel Core i7 Processor (or similar)
  - Multithreading
  - 6M Cache
  - 2.20 GHz
  - at least 8 GB RAM

- Image storage
  - Taken with phone: 500 KB/image x 30,000 = 15 GB
  - Taken with camera: 5 MB/image x 30,000 = 150 GB

Table of Contents

System Requirements for GARI server

- **Web-server software**
  - **Server version: Apache/2.2.9 (Debian)**
  - **PHP Version: 5.2.6-1+lenny13**
  - **OpenSSH_5.1p1 Debian-5, OpenSSL 0.9.8g 19 Oct 2007**

- **Database management system**
  - **psql (PostgreSQL) 8.1.11**
  - **contains support for command-line editing**

Table of Contents

CVADA

Center for Visual Analytics and Data Analytics

A U.S. Department of Homeland Security Center of Excellence

Predictive Visual Analytics

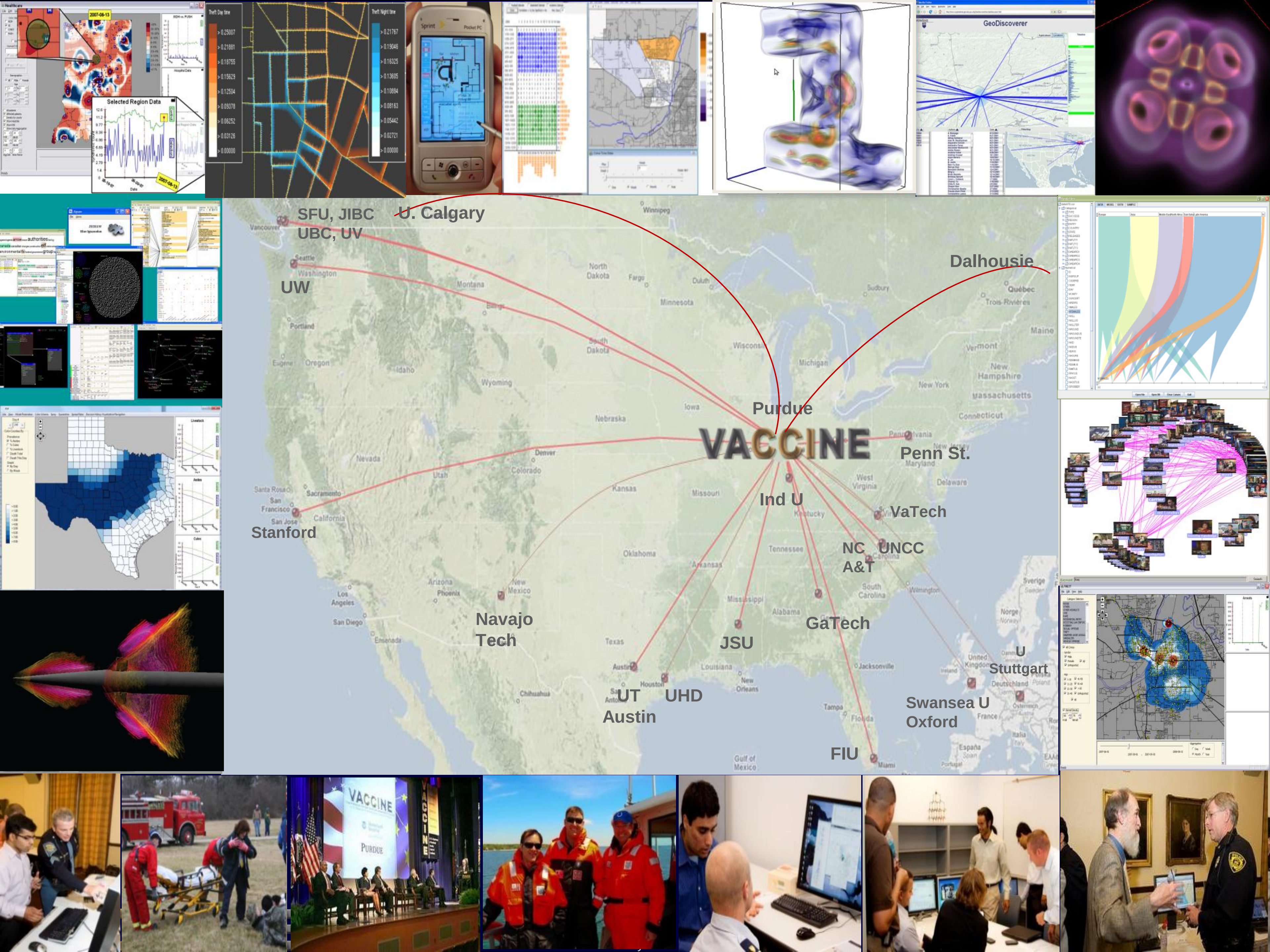
David S. Ebert

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

February 2013



SFU, JIBC
UBC, UV
U. Calgary
UW
Dalhousie
Purdue
VACCINE
Penn St.
Ind U
VaTech
NC
A&T
UNCC
Stanford
Navajo
Tech
GaTech
JSU
U
Stuttgart
UT
Austin
UHD
Swansea U
Oxford
FIU
GeoDiscoverer
Selected Region Data

Who We Are:
International Team of Experts
- 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

VACCINE Multidisciplinary Team

- Experts in computer science, image analysis, signal processing, cognitive science, decision theory, management science, geography and GIS, HCI, visualization, HPC, statistics, political science …

- Partners in nursing, public health, economics, databases, command and control, emergency management

VACCINE's Value

We enable users to be more effective through innovative interactive visualization, analysis, and decision making tools

- Provide the right information, in the right format within the right time to solve the problem
- Turn data deluge into a pool of relevant, actionable knowledge
- Enable user to be more effective from planning to detection to response to recovery
- Enable effective communication of information

Approach: Partner-driven solutions and research

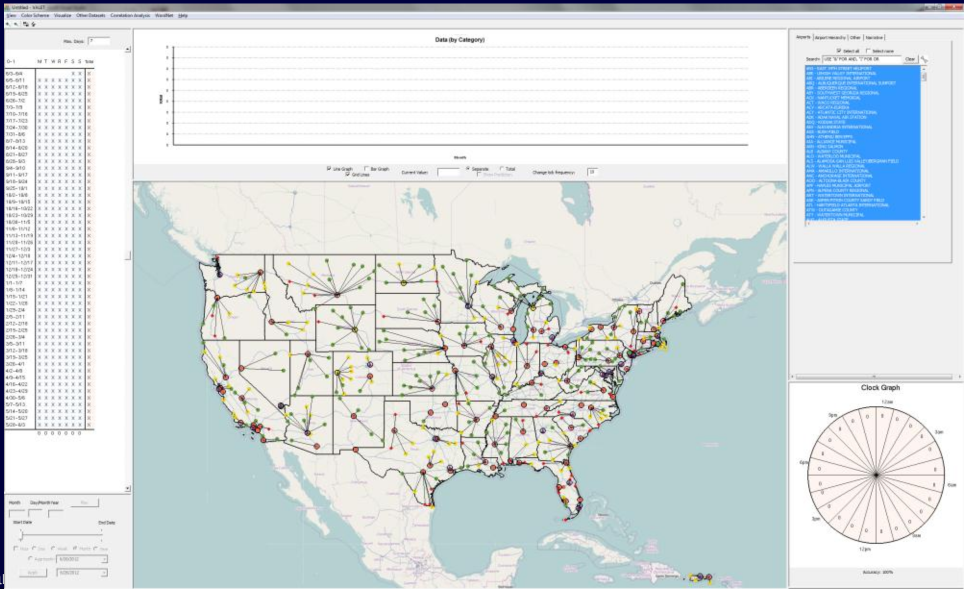
Engaged End-Users

- Federal Operating Compoents:
  - US Coast Guard
  - US Transportation Security Agency
  - US Federal Emergency Management Agency
  - US Customs and Border Patrol
  - US CERT
- Law Enforcement
  - Over 40 local and state agencies
- Fusion Centers
  - Ohio and Indiana

VACCINE

Example Projects

- Predictive Analytics
- Financial Visual Analytics
- TSA
  - Integrating staffing, financial, performance metric, legal data into a unified VA system
- Law Enforcement
- USCG – Risk-based decision making

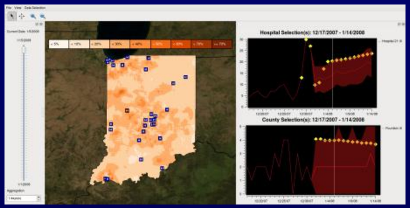


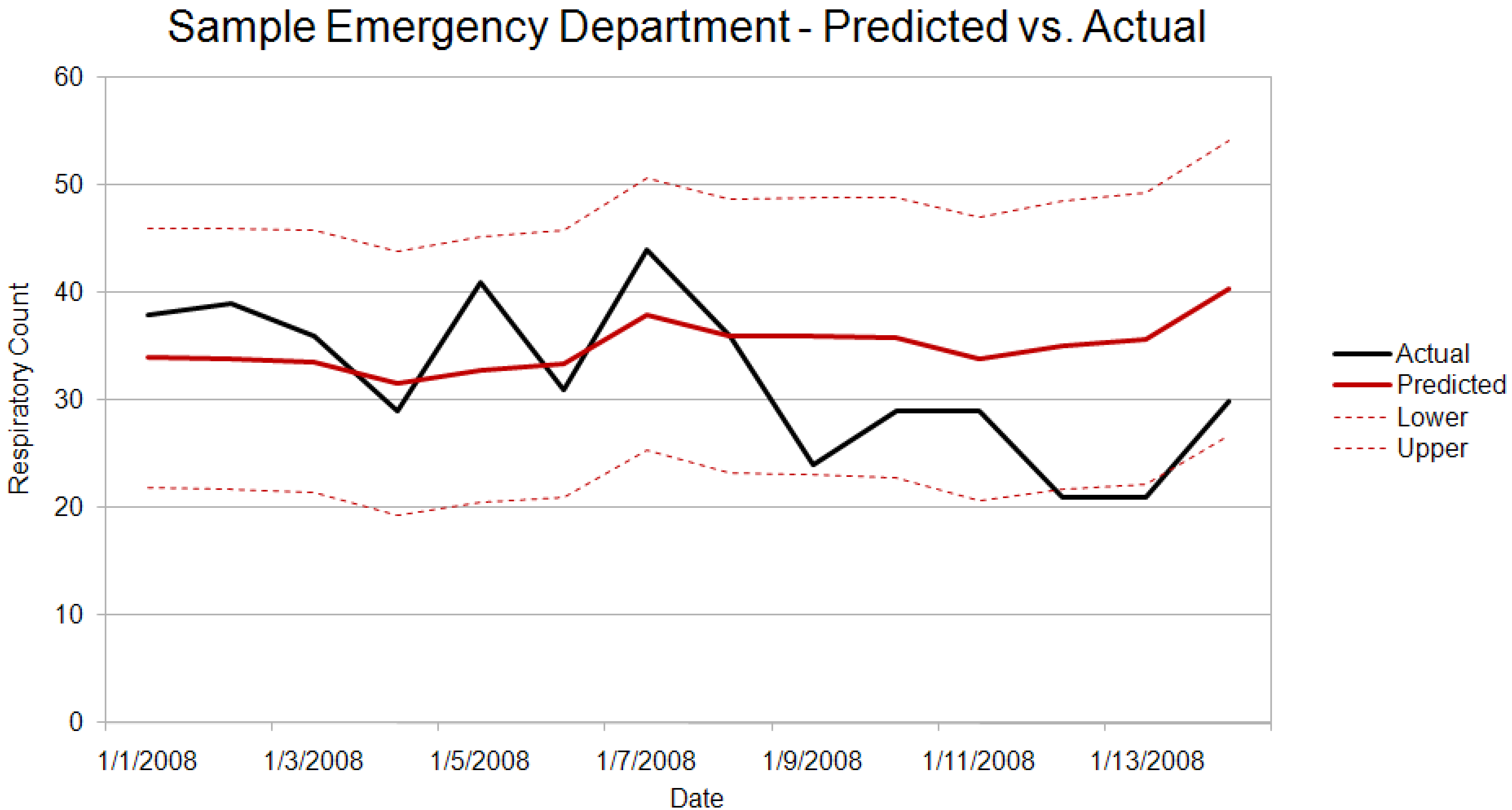
TSA System

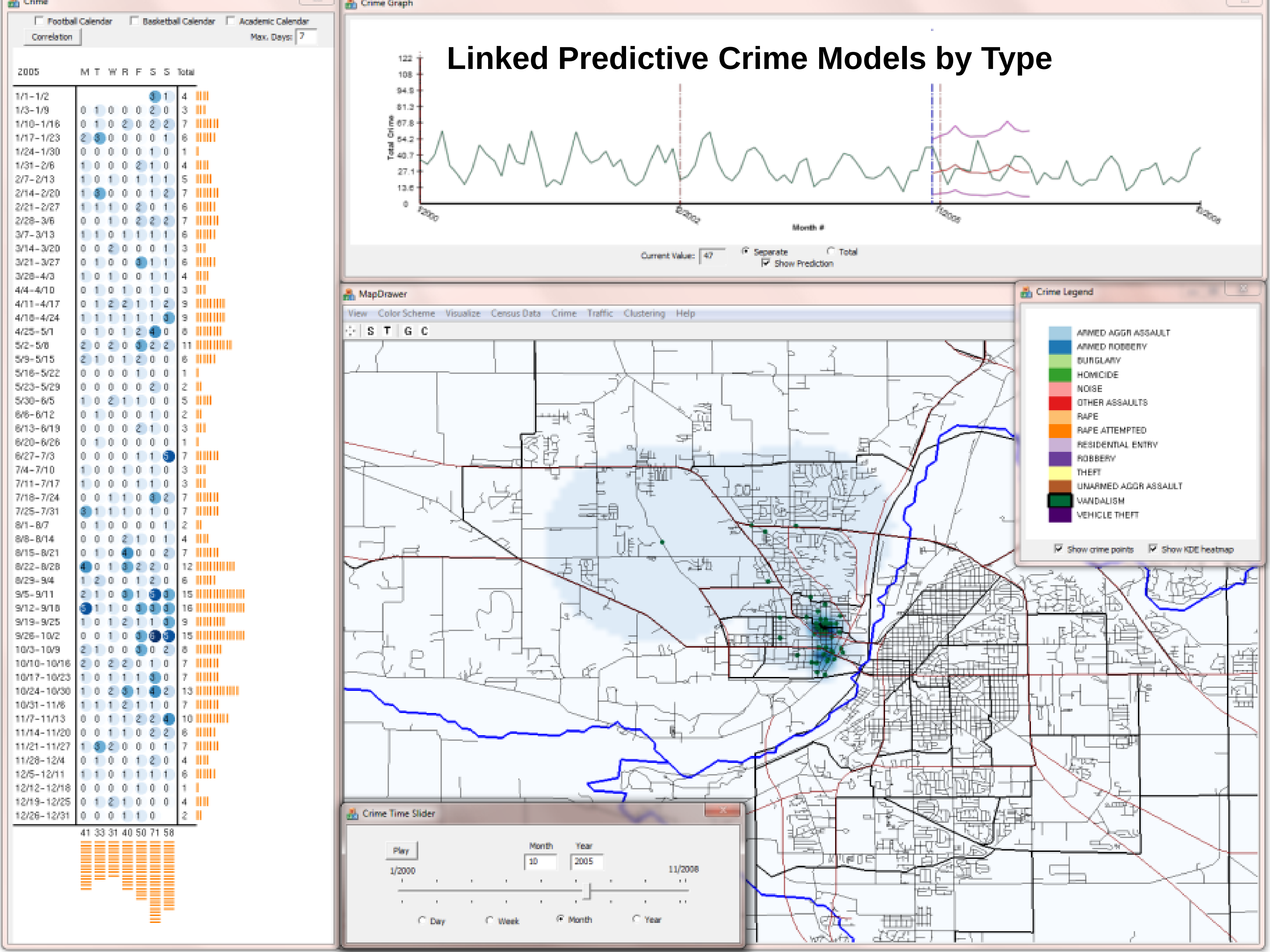
Predictive Analytics

- Use STL per data item to develop predicted values with confidence intervals
- Applied to health data and crime data
- Adapting to appropriate spatial resolution for most accurate predictions
- Spatial prediction based on population adapted kernel density estimation

Predictive Visual Analytics



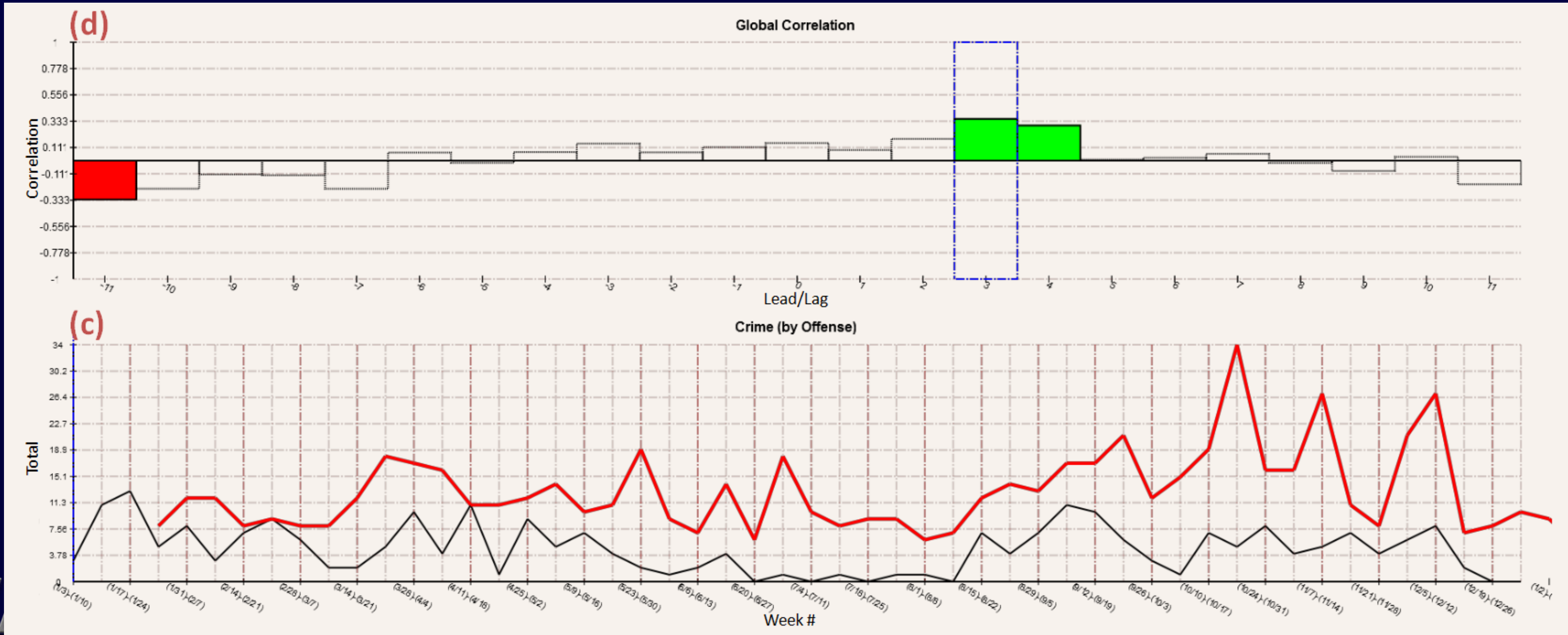




Linked Predictive Crime Models by Type
Crime
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
Max. Days:
2005
M T W R F S S Total
Crime Graph
Total Crime
Month #
Current Value:
Separate
Total
Show Prediction
MapDrawer
View
Color Scheme
Visualize
Census Data
Crime
Traffic
Clustering
Help
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Show crime points
Show KDE heatmap
Crime Time Slider
Play
Month
Year
1/2000
11/2008
Day
Week
Month
Year

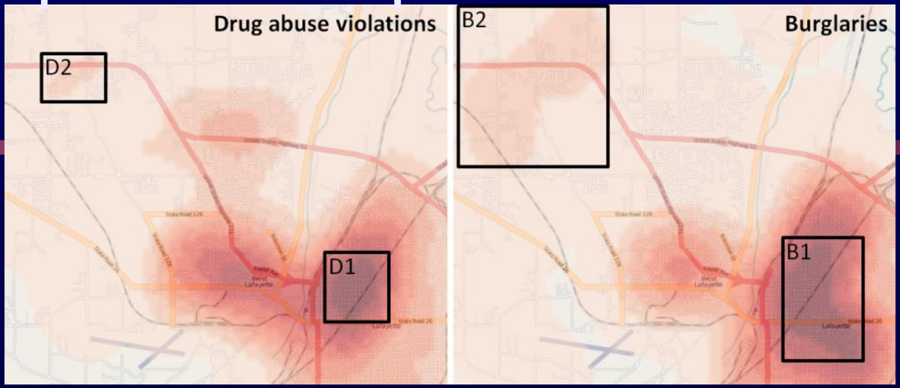
Multivariate Correlative Predictive Analytics: Three Approaches

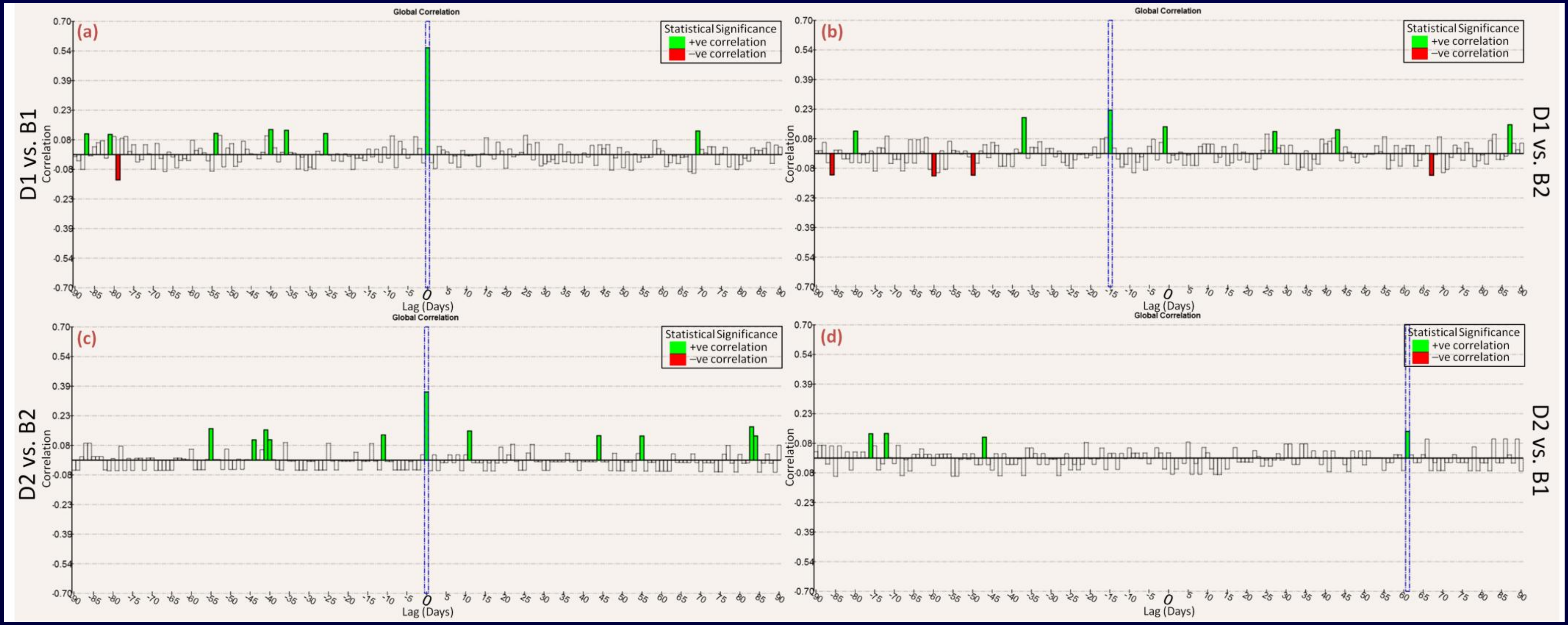
- Automatic correlation computation against lead/lags
- Temporal and spatial windowing
- Data category parameter space



Example: Drug Abuse Violations Vs. Burglaries – Focus on Geospatial Hotspots

- High positive correlation in same neighborhood with zero lag

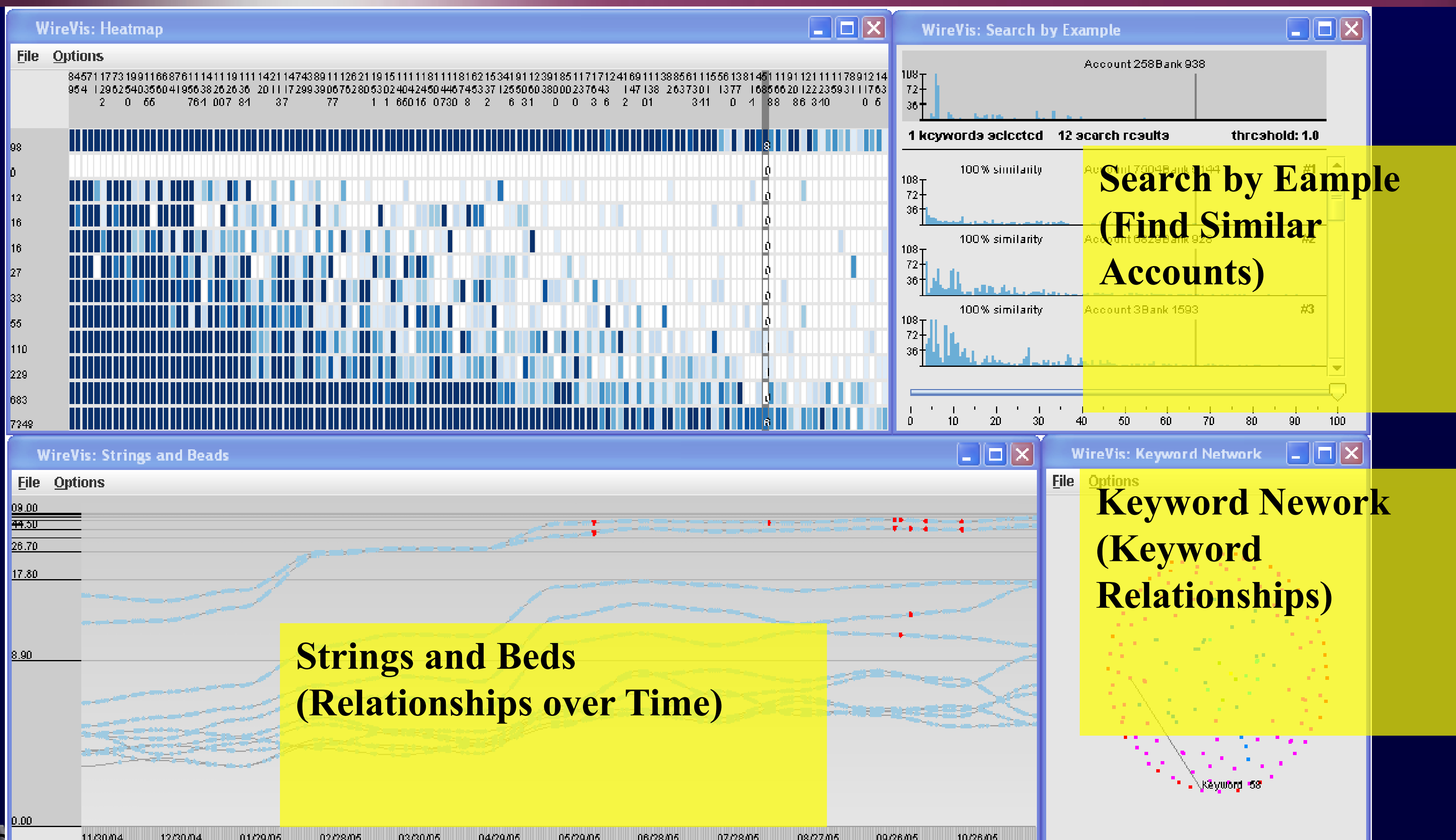




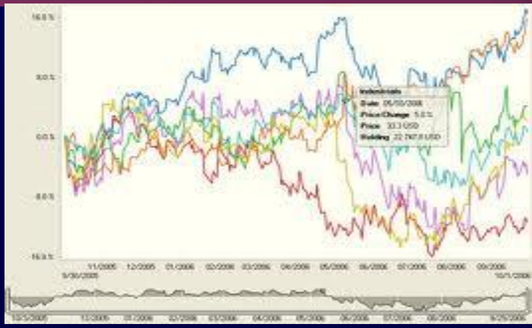
VACCINE

WireVis – Streaming Data - Multiple Linked Views

- Temporal, geospatial, theme, cluster, list views with association linkages between views

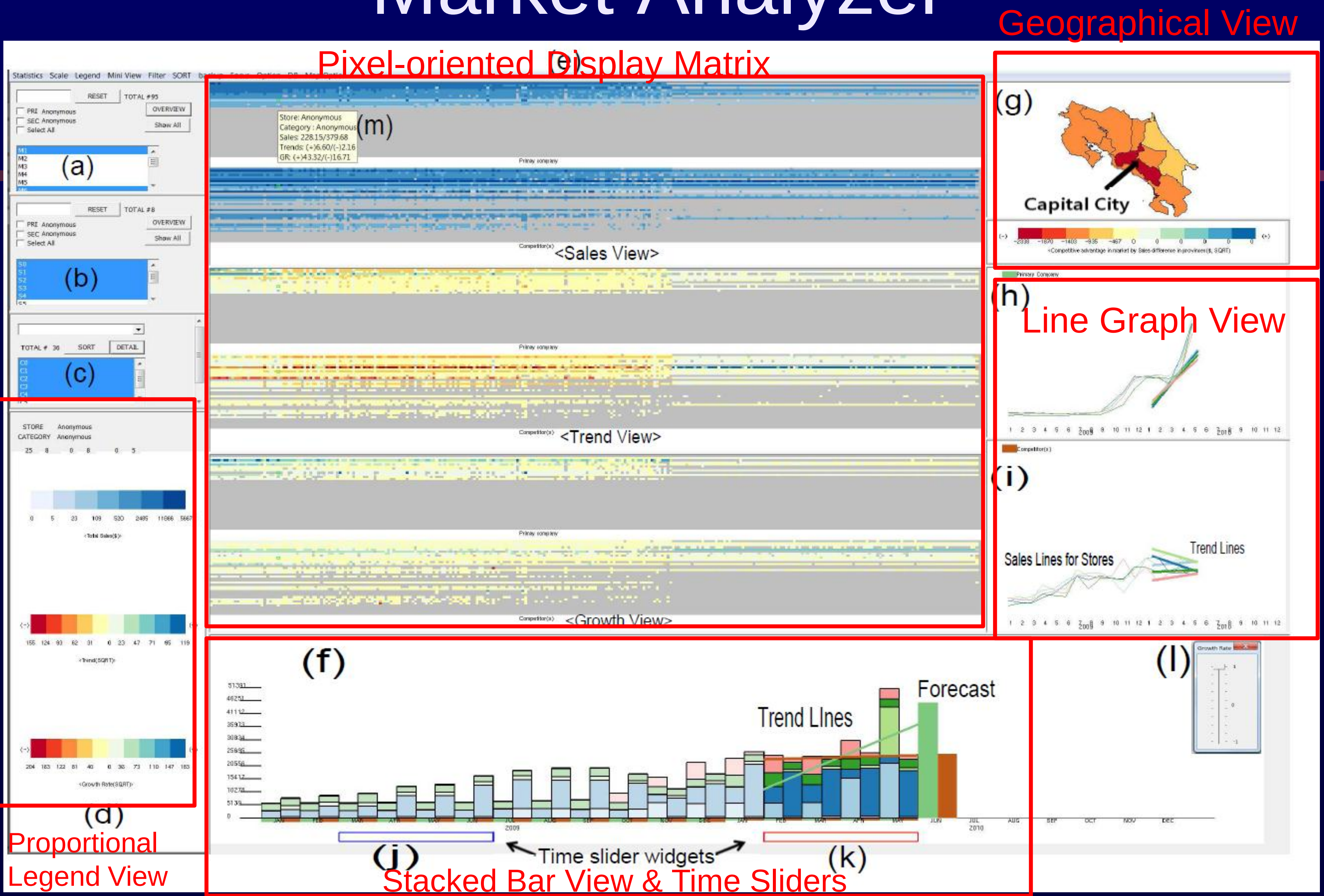


Competitive Business Intelligence Based On Point of Sales Data



- Characteristics of point of sale data
  - Multivariate
    - Large # of dimensions
      – E.g., 38 categories(products) in 288 stores
  - Temporal : 18 months
  - Spatial
    - Stores located all over Costa Rica
  - Requirements
    - Supporting easy comparison among companies
      – E.g., visualize all data at once, sorting by importance
      – Enable geographical comparison
    - Easy recognition of any change in sales
      – E.g., proportional legends
    - Forecast for decision-making

Market Analyzer

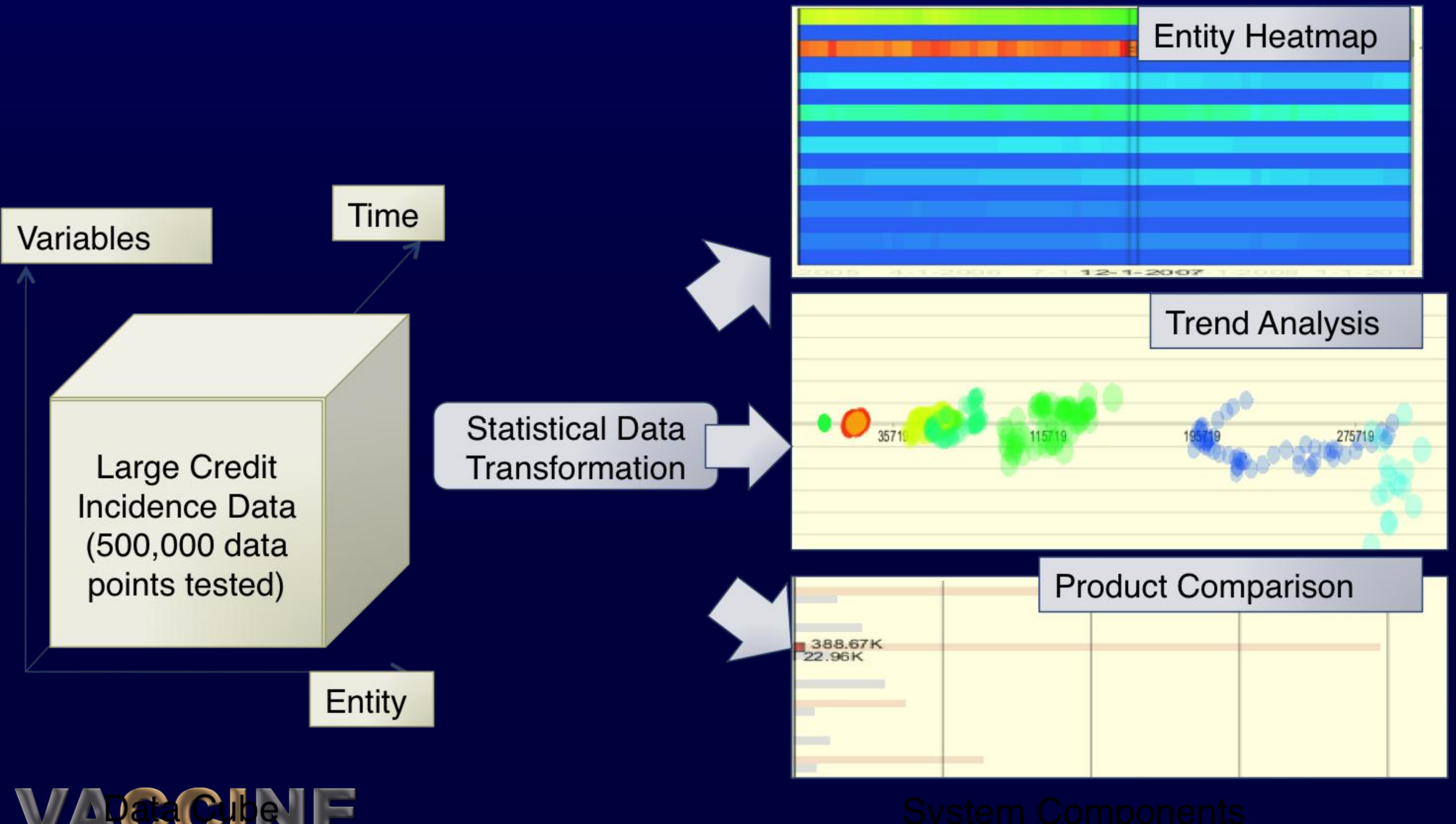


VACCINE

RiskVA: Key Credit Risk Analytics Challenges

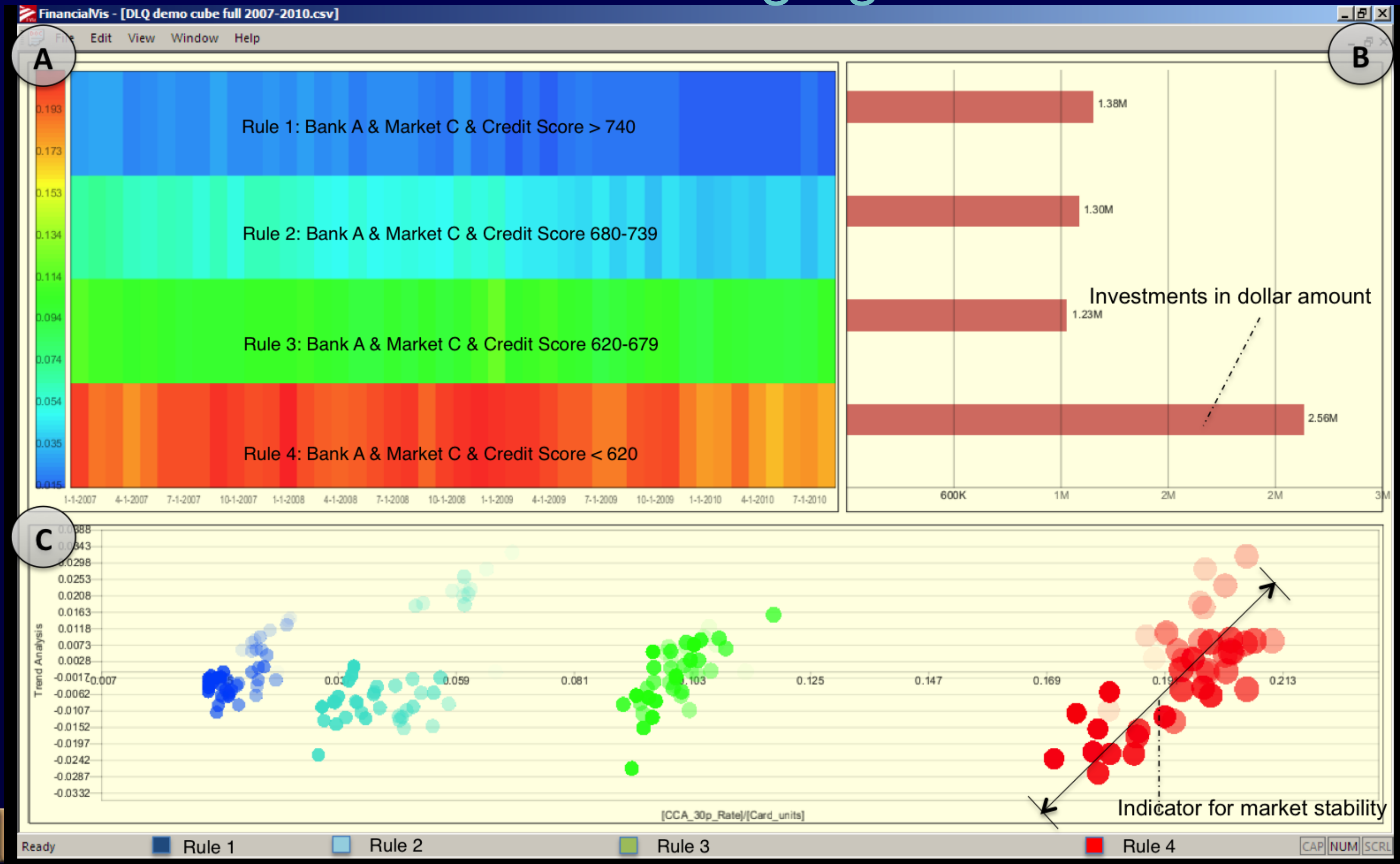
- Consumer credit data is large, temporal, and related across multiple investments and financial markets.
- The data are heterogeneous, not clean, have missing values, may be misleading and inefficient to explore.
- The data contains important behaviors and relations/groupings that change over time.
- Analysts need to have customer-centric interactive systems to achieve full analysis of risks

RiskVA Data Integration



RiskVA Overview: Interactive Exploratory Visualizations

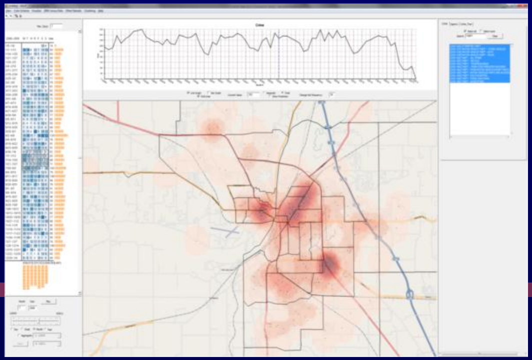
Identification of Emerging Risks

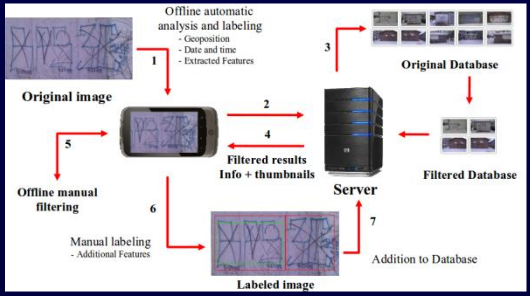


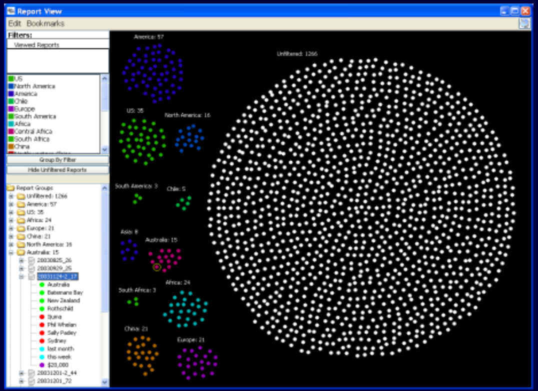
VACCI

Law Enforcement

- Law enforcement visual analytics
  - VALET, iVALET (iPhone/iPad), CrimeViz
    - Visual Analytics Law Enforcement Toolkit
    - Analyzing crime patterns and time of day problems
- Gang activity analytics
  - GARI
    - Gang Graffiti Recognition and Interpretation using a mobile telephone
    - Allowing police to catalog and analyze gang graffiti images, better track and determine gang activities
- Document visual analytics
  - JIGSAW
    - Visualization for investigative analysis
    - Discovery of hidden relationship and threats across documents



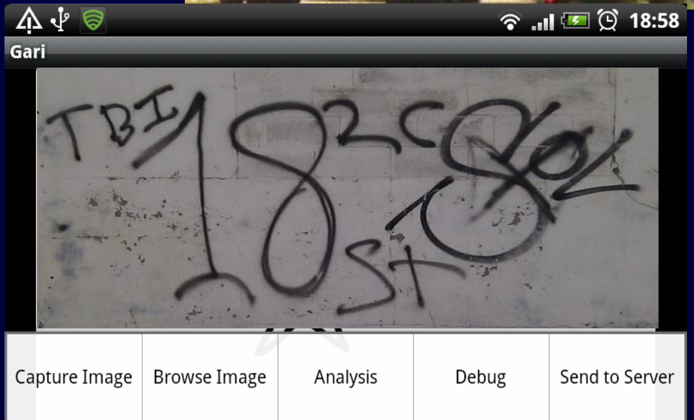




We Have an App For That!

Our Mobile Tools

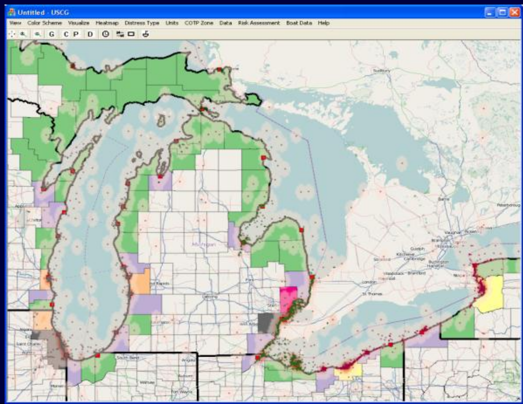
- VALET
- Evacuation Planning
- Rosetta Phone
- Hazmat app
- Gang graffiti app
- Tatoo app



VACCINE

Visual Analytics Uses for Risk-Based Decision Making

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



USCG: Effective Risk-based Decision Making and Resource Allocation Visual Analytics

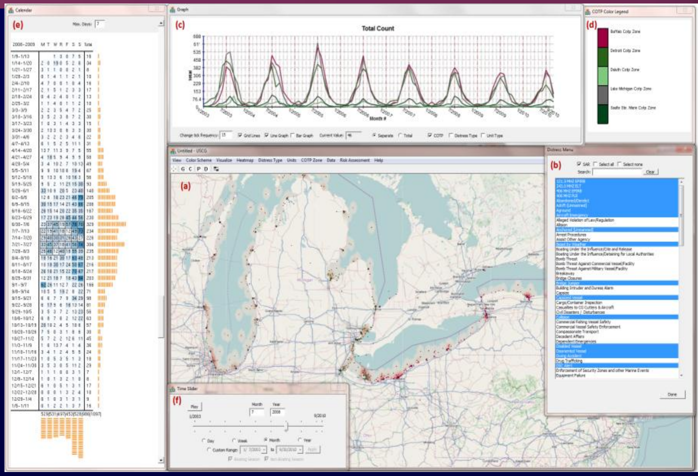
- Evaluate current and historical mission area:
  - Demands
  - Risks (total, mitigated, residual)
  - Resource allocation
  - Return on investment
- Evaluate courses of action
- Evaluate above at both Strategic and Tactical/Operational level

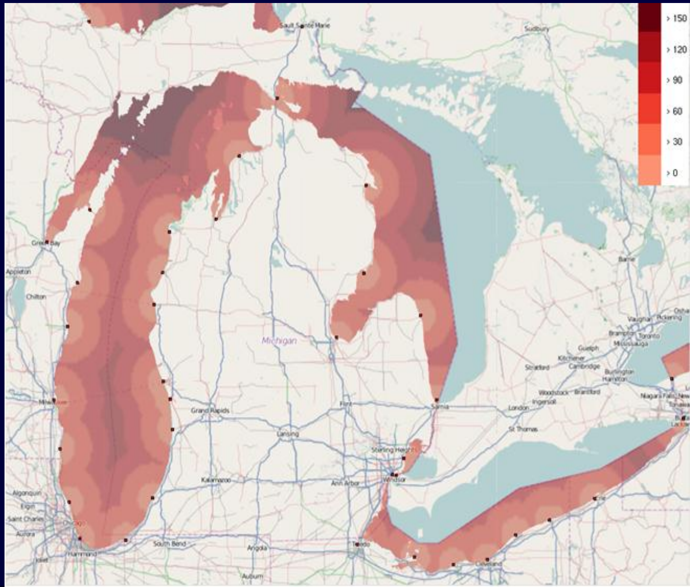
U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7, USCG D9, USCG D5, USCG HQ 771

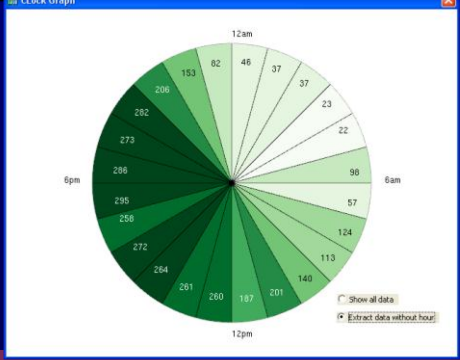
IMPACTS:

- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Superstorm Sandy: Used for resource allocation in response and in rebuilding
- Used for Hurricane Irene resource allocation decisions
- Informed Commandant's budget testimony and recommendations to Congress
- Key component of USCG D9 reallocation plan for 2011-2012 based on decreased budget
- Key component of Coastal Operations Allocation Suite of Tools (COAST) – USCG HQ





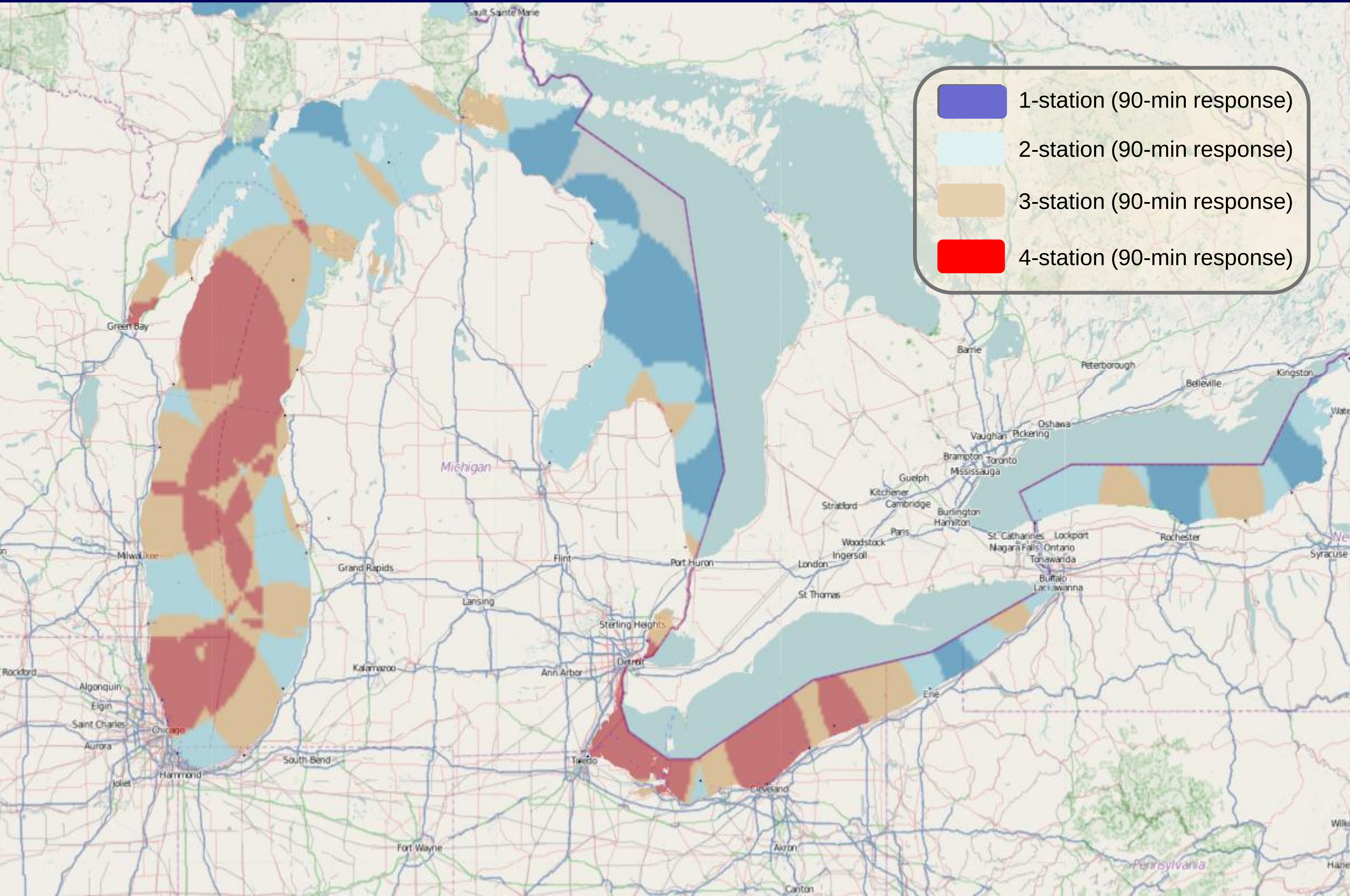
Risk-Based Allocations



- Comparative visual analysis of mission cases/hours vs. staffing hours
- Comparative visualization of resources vs. risk
- Trend visual analytics
  - Increase/decrease in resource allocation
  - Increase/decrease in risk (total, mitigated, residual)
  - Increase/decrease in incidents
- Exploration of alternatives and effect on risk
- Predictive analytics based on historical data (STL and EWMA)

VACCINE

Response Efficiency – Potential Future Assets



For Further Information

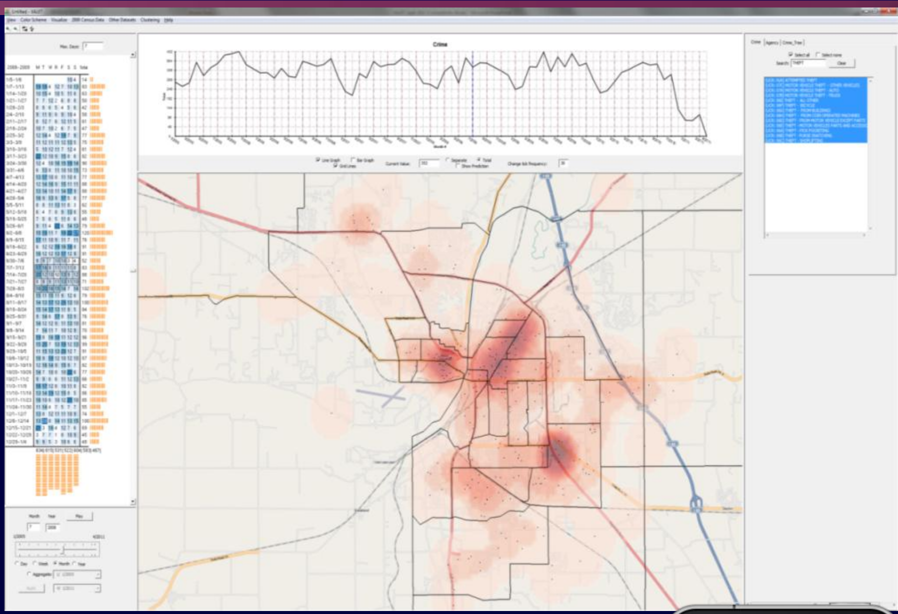
www.VisualAnalytics-CCI.org

vaccine@purdue.edu

ebertd@purdue.edu

VACCINE

Visual Analytics Law Enforcement Toolkit (VALET, iVALET)



Impacts:

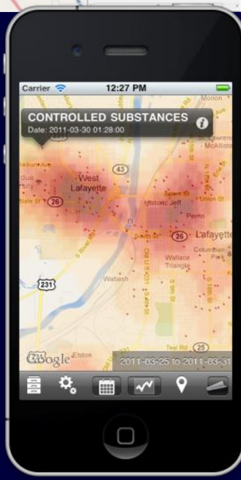
- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (September 2012) for community-based policing
- Investigating correlation of bus routes and crime, street lights and crime
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning

VALET delivered:

- Spring 2011: WL, Lafayette Police

iVALET delivered:

- October 2011: Purdue, WL Police



VALET Overview

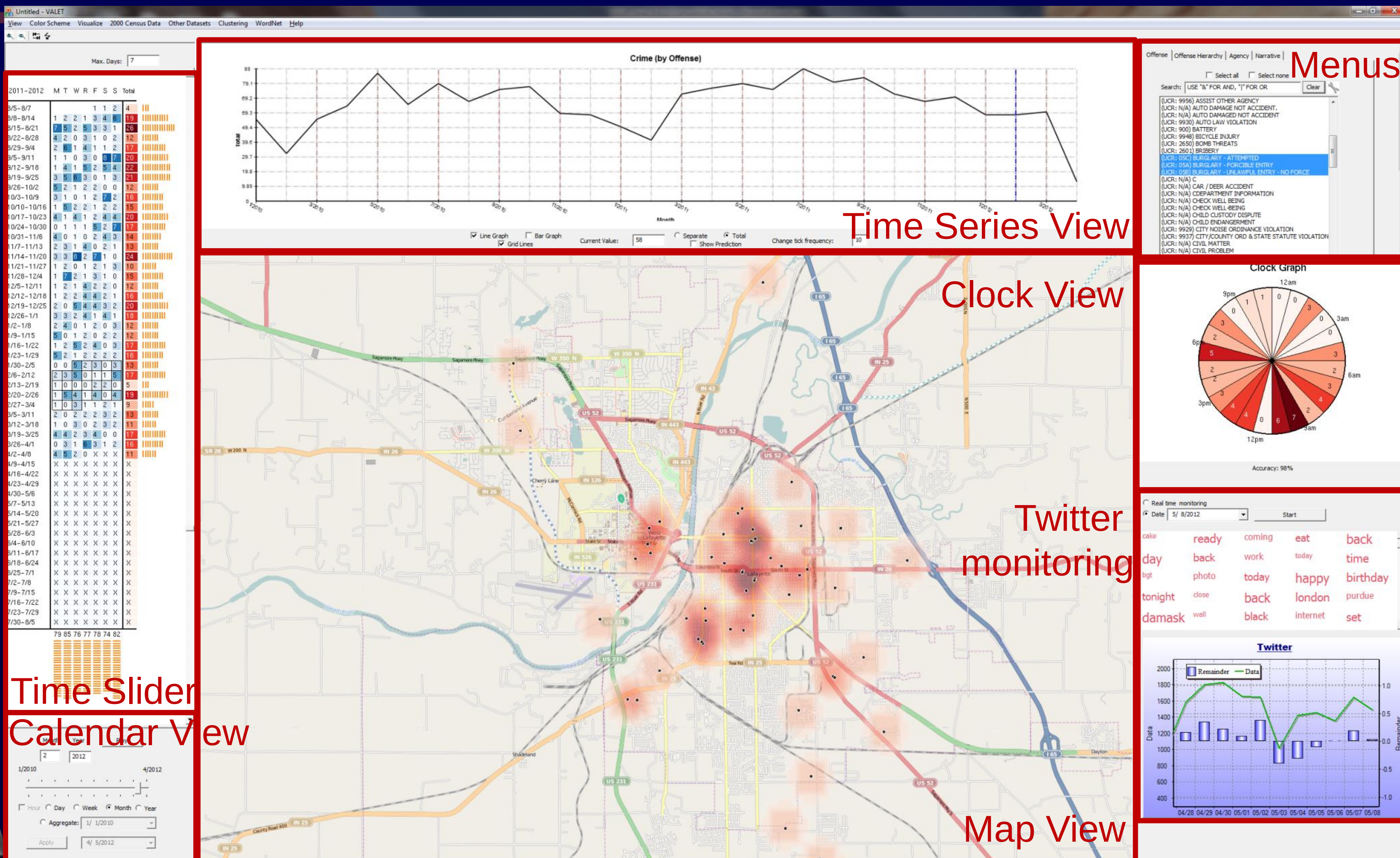



Table of Contents
Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
11/14/2010
01/01/2011
02/13/2011
03/27/2011
Calendar View
M T W Th F Sa S
1/1/10 36 46 29
1/4/10 38 48 44 43 42 32 30
1/11/10 62 53 53 56 59 45 34
1/18/10 62 49 53 72 61 52 49
1/25/10 98 66 57 66 64 55 49
2/1/10 58 63 50 51 73 51 42
2/8/10 50 46 48 63 52 39 35
2/15/10 48 65 48 54 55 67 51
2/22/10 51 66 48 56 57 46 47
3/1/10 50 59 42 54 82 45 41
3/8/10 50 57 64 78 63 36 30
Clock View
12Hrs
10:00
Accuracy: 64%
Choro
Plume
Moran
Region
Route
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20
©2012 Google

Table of Contents

Leave Blank

Fundamentals of Visual Analytics

Ross Maciejewski and David S. Ebert

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

ASU Ira A. Fulton school of engineering
ARIZONA STATE UNIVERSITY

The Ongoing Data Deluge

Leave Blank

- Since 2003, digital information has accounted for 90% of all information produced[1]
- In 2009, drones from Iraq and Afghanistan recorded 24 years of video footage
- In 2010, the amount of information added annually to the digital universe was estimated to be nearly 1 ZB
- Wal-mart process > 1 million transactions per hour
- By 2013 Cisco estimates the annual internet traffic will be 667 EBs

VACCINE

1 – C. R. Johnson, R. Moorhead, T. Munzner, H. Pfister, P. Rheingans, and T. S. Yoo, Eds., *NIH-NSF Visualization Research Challenges Report*, IEEE Press, ISBN 0-7695-2733-7, 2006

ASU

The Ongoing Data Deluge

Leave Blank

- FedEx's ships more than 8.5 million packages per day (http://www.fedex.com/us/about/today/companies/corporation/facts.html)
- Consumers carry more than 1 billion Visa cards worldwide. More than 450 million of those cards are in the United States (http://www.creditcards.com/credit-card-news/credit-card-industry-facts-personal-debt-statistics-1276.php)

Data Overload

Leave Blank

- Opportunity: Huge amounts of data available in digital form and ready for analysis!
  - But, how can we make sense of this data?
  - How can we harness this data in the decision making process?
  - How do we avoid being overwhelmed by all of this data?

The Power of Visualization

Leave Blank

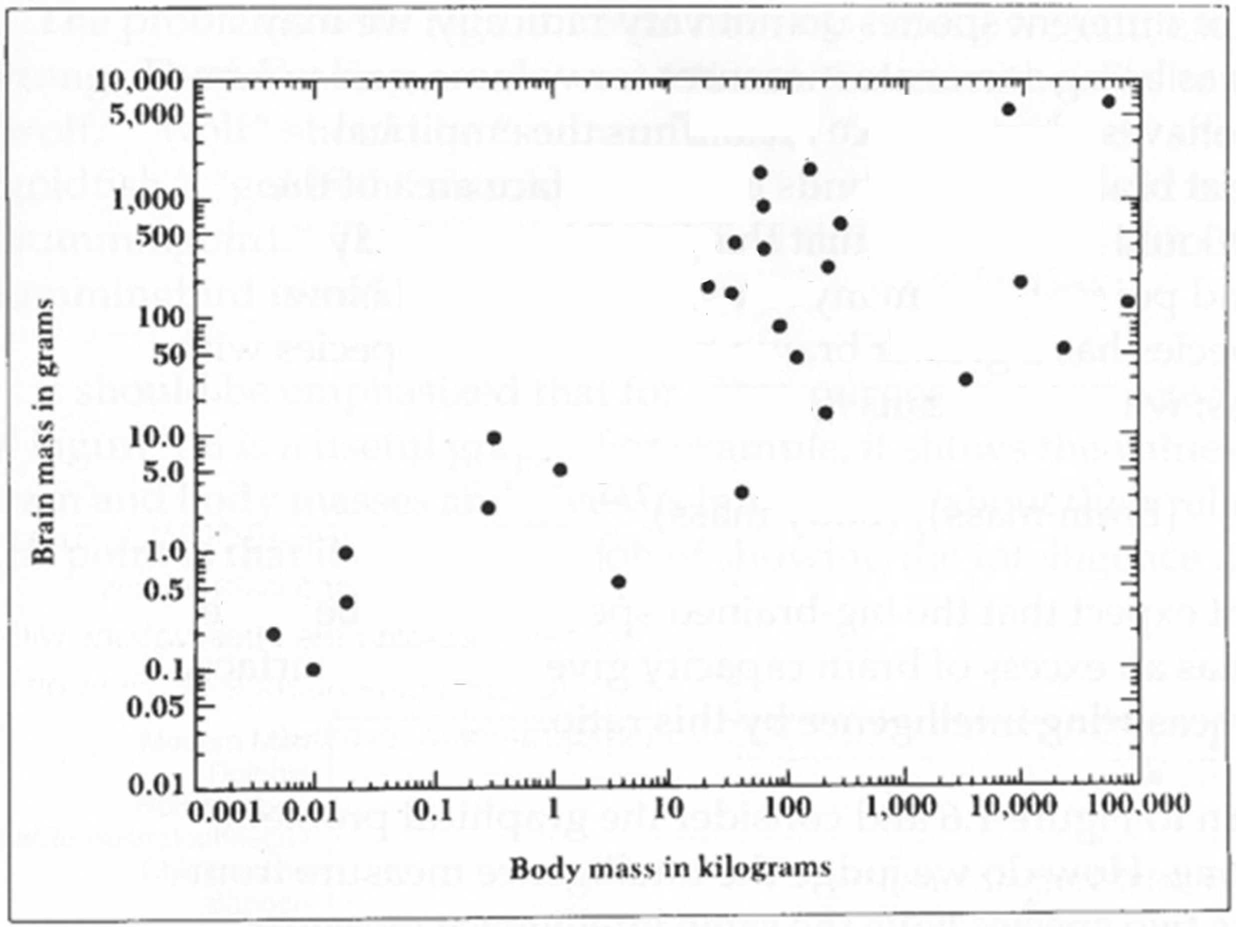
- The goal is to take all of this *data* and transform it into *information*
- How many terabytes of data we have collected doesn't matter, it's how many petaflops of *insights* we can generate from this data
- "*The purpose of computing is insight, not numbers.*" Dedication to his book: RW Hamming (1971). *Introduction to Applied Numerical Analysis*. McGraw Hill.
- We need to make the data understandable to people and a key way of doing this is through *visualization*

	A	B	C	D
1	Animal	Brain Weight (kg)	Body Weight (kg)	
2	Mountain beav	1.35	465	
3	Cow	465	423	
4	Grey wolf	36.33	119.5	
5	Goat	27.66	115	
6	Guinea pig	1.04	5.5	
7	Dipliodocus	11700	50	
8	Asian elephant	2547	4603	
9	Donkey	187.1	419	
10	Horse	521	655	
11	Potar monkey	10	115	
12	Cat	3.3	25.6	
13	Giraffe	529	680	
14	Gorilla	207	406	
15	Human	62	1320	
16	African elephar	6654	5712	
17	Triceratops	9400	70	
18	Rhesus monke	6.8	179	
19	Kangaroo	35	56	
20	Golden hamste	0.12	1	
21	Mouse	0.023	0.4	

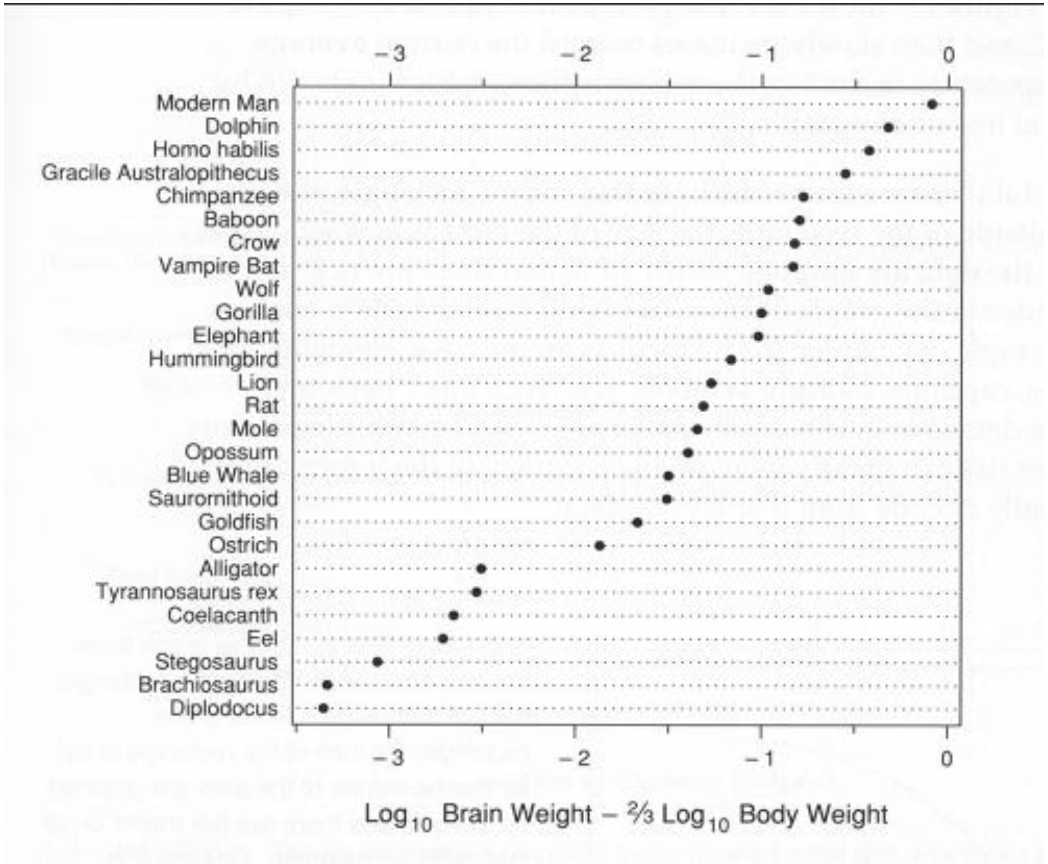
Leave Blank

Which animal weights the least/most?
Is there a relationship between brain weight and body weight?
If so, are there any outliers?

Leave Blank


10,000
5,000
1,000
500
100
50
10.0
5.0
1.0
0.5
0.1
0.05
0.01
Brain mass in grams
0.001
0.01
0.1
1
10
100
1,000
10,000
100,000
Body mass in kilograms

Leave Blank



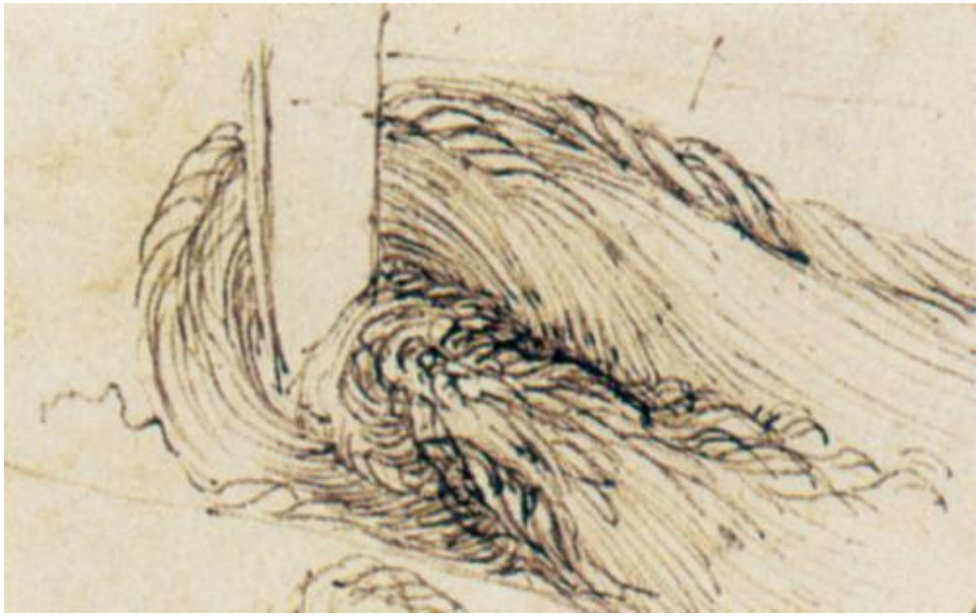
WS Cleveland – The Elements of Graphing Data, 1994

VACCINE

Visualization Through the Ages

Leave Blank

- Hand-drawn illustration[1] of water by Leonardo da Vinci from his studies to determine the processes underlying water flow (1510)

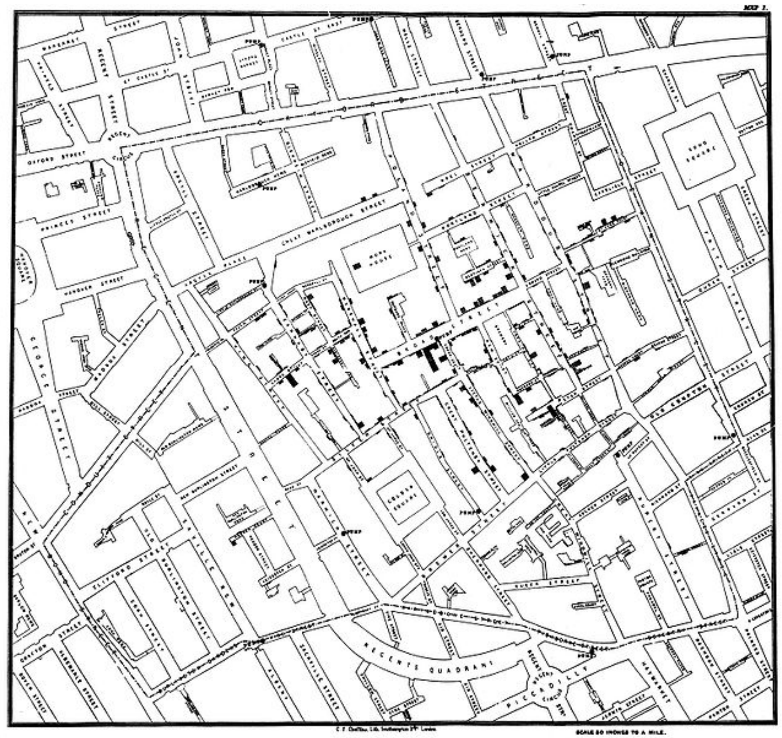


1 – Leonardo da Vinci. *Old Man Seated on Rocky Outcrop, Seen in Profile to the Right, with Water Studies*. Windsor Castle, Royal Liberty, RL 12579r, c. 1510-1513.

Visualization Through the Ages

Leave Blank

- John Snow's map of Cholera cases in the London epidemic of 1854[1]

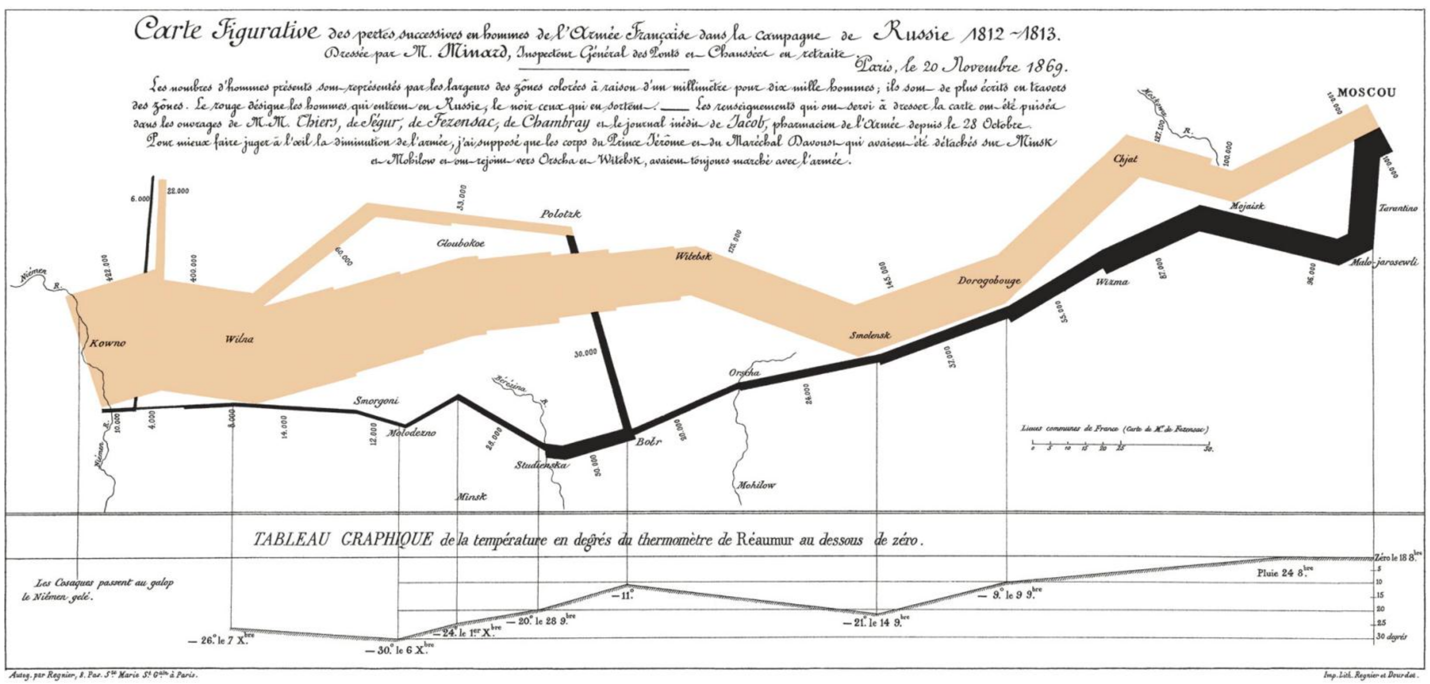


1 – The Visual Display of Quantitative Information, E. R. Tufte, 1983

Visualization Through the Ages

Leave Blank

- Charles Minard's 1869 flow map[1] showing Napoleon's disastrous Russian campaign of 1812.



1 – Charles Joseph Minard: Mapping Napoleon's March, 1861 by John Corbett, Center for Spatially Integrated Social Science

Visualization Through the Ages

Leave Blank

- James Maxwell's[1] thermodynamic surface sculpture (1874)
- Three-dimensional plot of the states of a fictitious water-like substance
- Coordinates are volume (x), entropy (y), and energy (z)

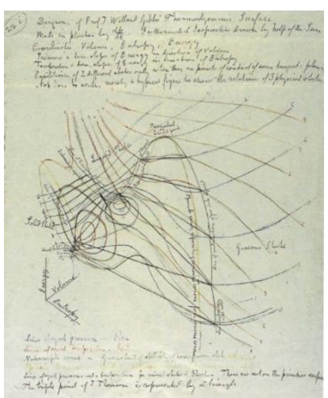


Plate IV Diagram of the lines on Gibbs' thermodynamic surface, 8 July 1875 (Number 564).

1 – James Clerk Maxwell and P. M. Harman (2002), The Scientific Letters and Papers of James Clerk Maxwell, Volume 3: 1874-1879, Cambridge University Press

Why is Visualization Helpful?

Leave Blank

- It utilizes the high bandwidth of the human visual systems
- The human mind is fast and parallel
- Humans are great at visual pattern recognition
- We have pre-attentive visual phenomena
- Visual tools can extend memory and cognitive capacity
- We think visually!
- "A picture is worth a thousand words." – *Printers Ink,* pp. 96-97 December 8, 1921

Why Is Visualization So Helpful?

Leave Blank

- Amplifies cognition
  - Expands our working memory by allowing us to offload results
  - Reduces the search time
  - Pattern detection and recognition can be improved through perceptual cues and inference
  - Visualizations can be designed to control attention interaction for improving cognition

Readings in Information Visualization: Using Vision to Think, SK Card, J Mackinlay and B.Shneiderman, 1999

What is Visualization?

Leave Blank

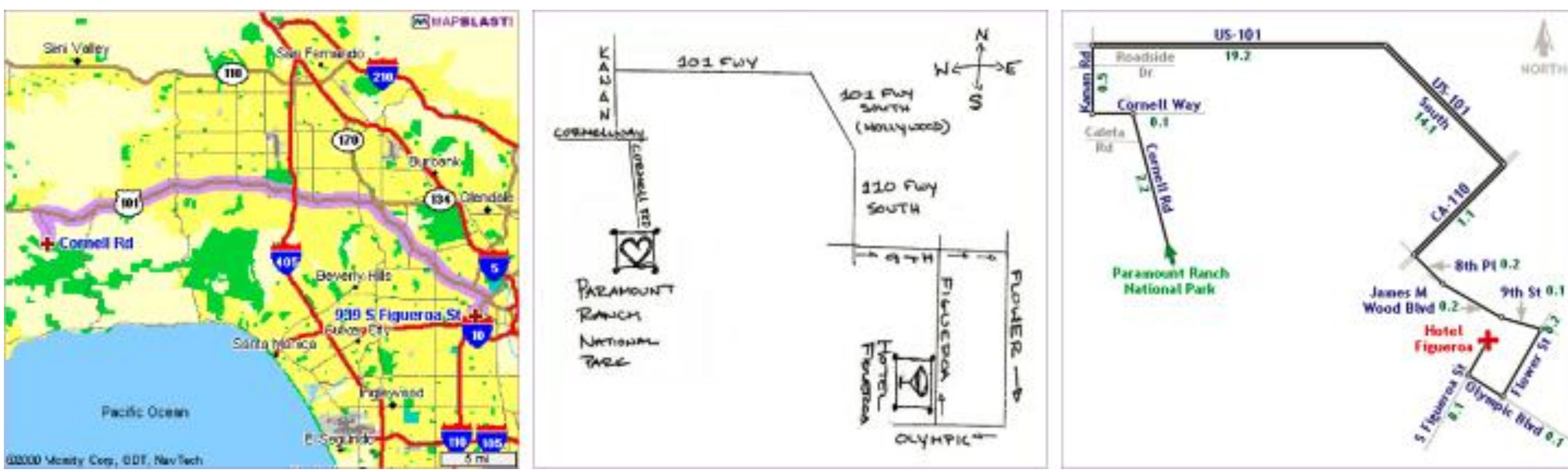
- "The use of computer-supported, interactive visual representations of data to amplify cognition."[1]
- This is not simply the process of making a graphic or an image, the goal is to create insight, not pretty pictures
- We want to help people form a mental image of something and internalize their own understanding
- We want to promote discovery, decision making and explanations

1-Readings in Information Visualization: Using Vision to Think, SK Card, J Mackinlay and B.Shneiderman, 1999

What is Visualization

Leave Blank

- We want to find and utilize cognitive and perceptual principles
- We want to optimize our visualizations and our interactions with the visualization according to these principles



Maneesh Agrawala, Chris Stolte: Rendering effective route maps: improving usability through generalization. SIGGRAPH 2001: 241-249

Scientific Visualization

Leave Blank

- Primarily relates to and represents something physical or geometrics
- The structure of the data is typically defined or given
- Examples
  - Air flow over a wing
  - Stresses on a girder
  - Organs in the human body
  - Molecular bonding

1 – Descriptions on this slide are borrowed from John Stasko's "InfoVis Overview" lecture: http://www.cc.gatech.edu/~stasko/7450/Notes/overview.pdf

Information Visualization

Leave Blank

- Primarily relates to data that does not have a direct physical correspondence
- Notion of the data is abstract
- Examples
  - Baseball statistics
  - Stock trends
  - My social network

1 – Descriptions on this slide are borrowed from John Stasko's "InfoVis Overview" lecture: http://www.cc.gatech.edu/~stasko/7450/Notes/overview.pdf

Purpose of Visualization

Leave Blank

- Analysis – Understand your data better and act upon that understanding
  - Given a data set, compare, contrast, assess, evaluate
  - Solve a problem!
- Presentation – Communicate and inform others more effectively
- Visualization is most useful in **exploratory data analysis**[1]

1 – J. W. Tukey. Exploratory Data Analysis

The Information Seeking Mantra

Leave Blank

Overview first,

then zoom and filter,

details on demand

1 - B. Shneiderman. *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations*. In the Proceedings of the IEEE Symposium on Visual Languages, pp. 336-343, 1996.

Tasks for Visualization

Leave Blank

- **Searching and browsing** - find a specific piece of information, inspect data, seek information
- **Analyze** – do comparisons, differences, look for outliers, extrema and patterns
- **Monitor** – look for changes and trends

Leave Blank

"Contained within the data of any investigation is information that can yield conclusions to questions not even originally asked. That is, there can be surprises in the data ... To regularly miss surprises by failing to probe thoroughly with visualization tools is terribly inefficient because the cost of intensive data analysis is typically very small compared with the cost of data collection."

W. S. Cleveland

The Elements of Graphing Data

Is This the Total Solution?

Leave Blank

- Traditional visualization misses several key factors in how people solve difficult problems
- Visual analytics aims to overcome these shortfalls and create an interactive human-computer exploration and decision making environment

Visual Analytics

Leave Blank

- Visualization is good for exploring data, but we can do more than just explore
- “Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces”[1]
- Visual analytics combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets.[2]
- “A graphic display has many purposes but it achieves its highest value when it forces us to see what we were not expecting.”[3]

1 - J. J. Thomas and K. A. Cook (Ed.). *Illuminating the Path: The R&D Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.
2 – Keim et al. chapter in *Information Visualization: Human-Centered Issues and Perspectives*, 2008.
3 – H. Wainer. *A Graphic Discovery: A Trout in the Milk and Other Visual Adventures*.

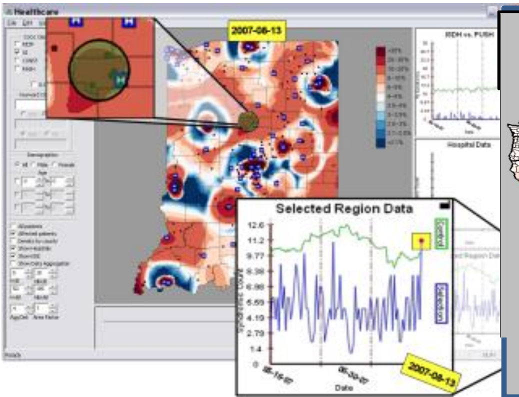
Example Systems And Applications

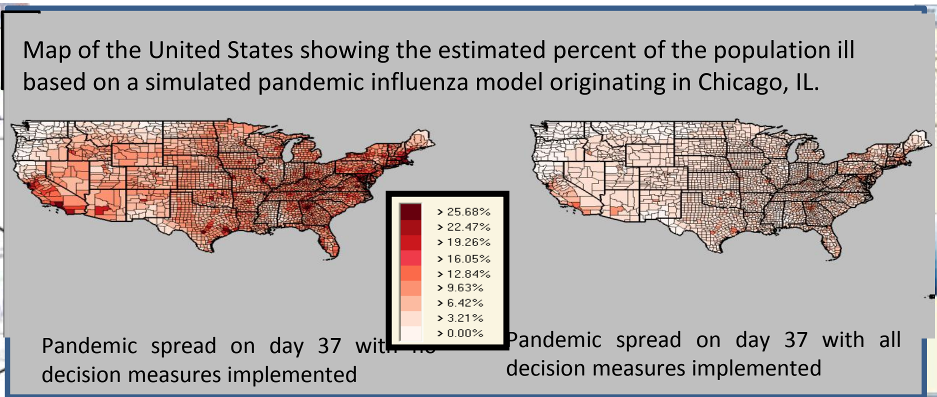
Leave Blank

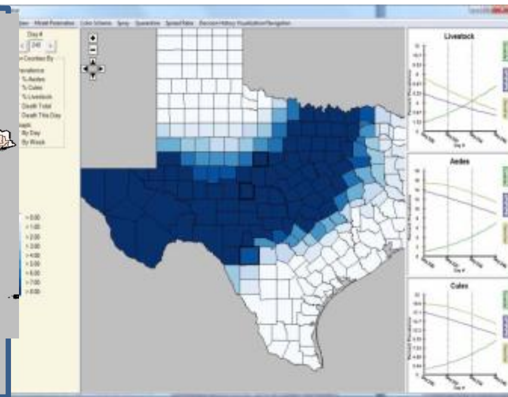
VACCINE

Leave Blank

Example: Public Health Visual Analytics





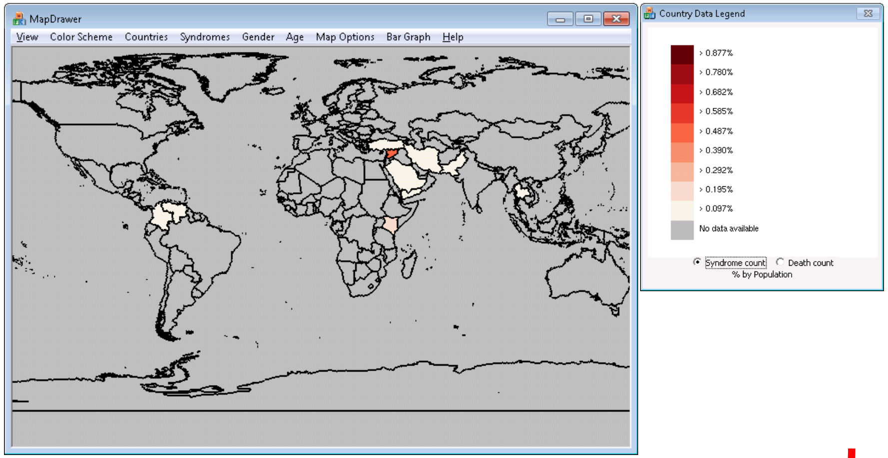


Syndromic Surveillance

Leave Blank

– Syndromic surveillance is the detection of adverse health events focusing on pre-diagnosis information to improve response time
– Pre-diagnosis information can consist of multiple data sources:
  - Over the counter medicine sales
  - News reports on emerging diseases
  - Pro-med news feeds
  - Emergency department chief complaints

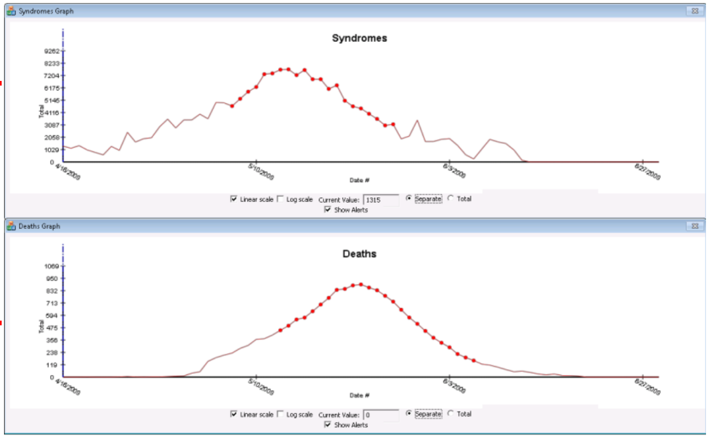
Syndromic Surveillance



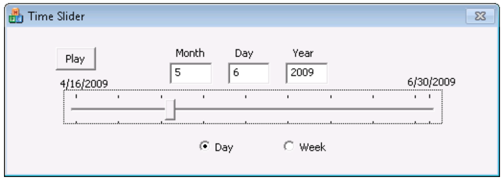
Leave Blank

Patient ID	Date	Chief Complaints	Location
9398	4/16/09	Ear pain	Karachi
10816	4/16/09	Stuffy nose	Lebanon
1491	4/16/09	Fever	Allepo
16237	4/16/09	Head bleed	Yemen

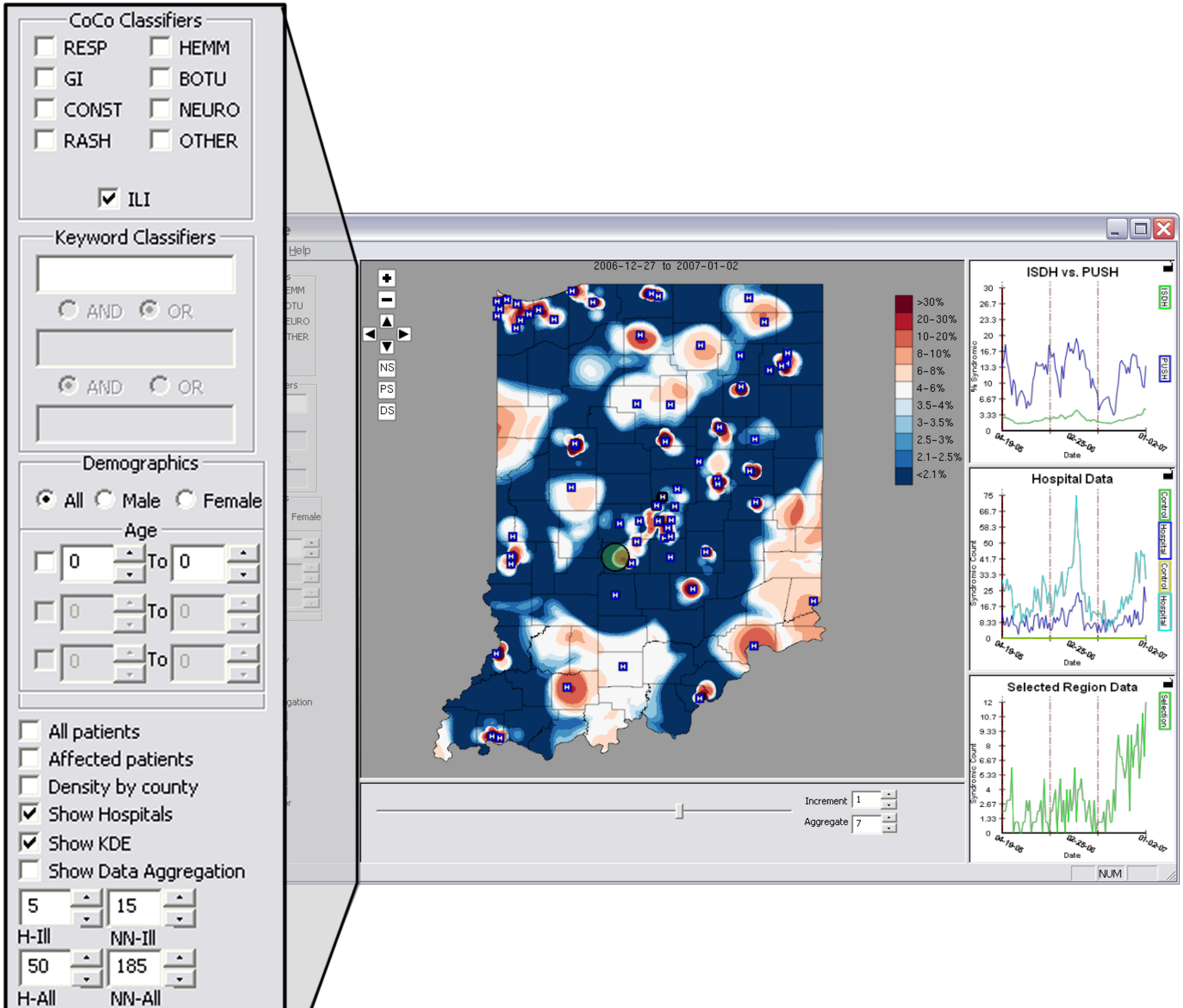
CoCo Classifier



1) **Respiratory**
2) **Gastro-intestinal**
3) **Neurological**
4) **Botulinic**
5) **Constitutional**
6) **Rash**
7) **Hemmoraghic**
8) **Other**



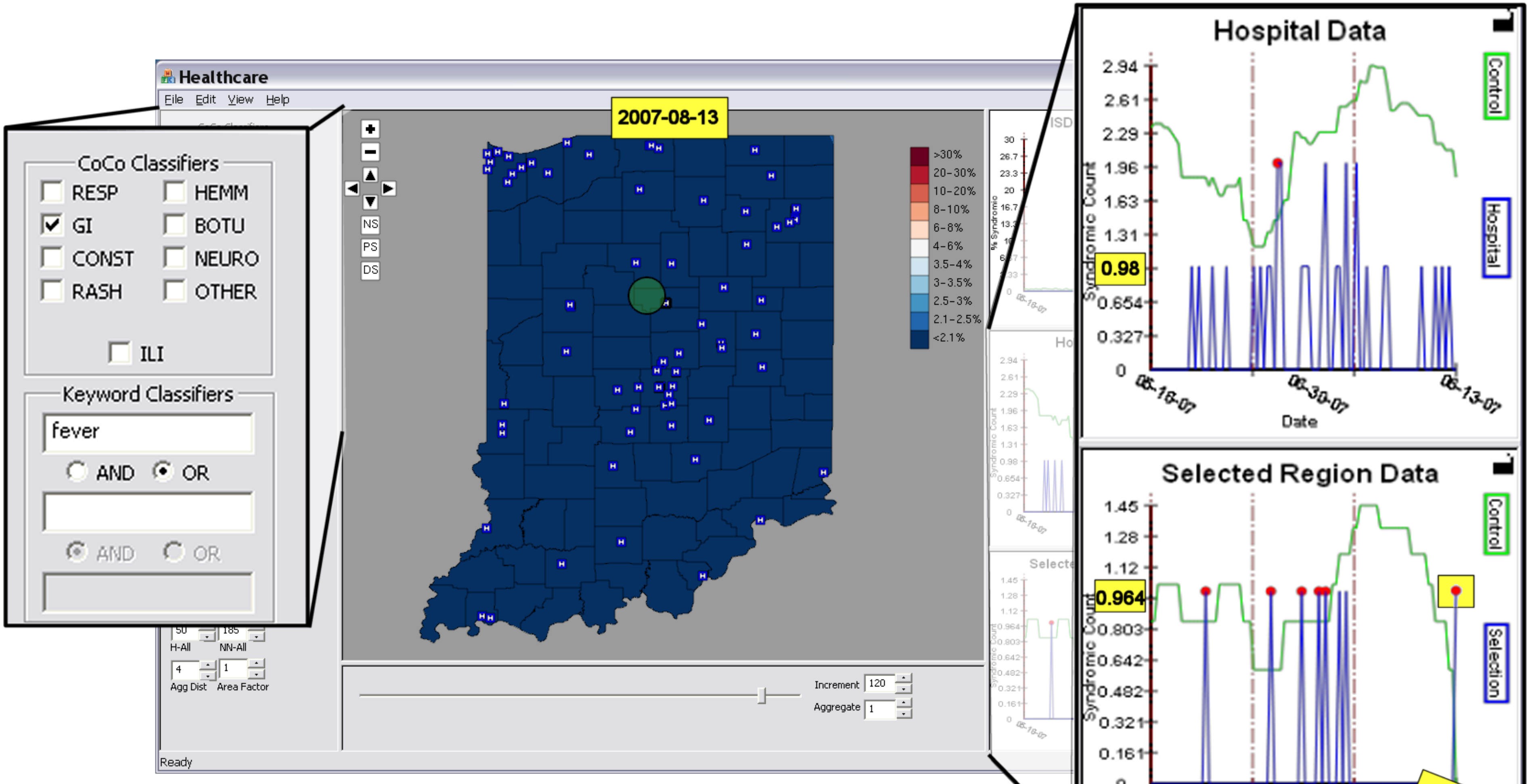
VAST 2010 Challenge Award: Support for Future Detection, IEEE Conference on Visual Analytics Science and Technology, October 2010 for "VACCINATED," (Malik, A., Afzal, S., Hodgess, E., Ebert, D., Maciejewski, R.).



Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W., Grannis, S., Ebert, D., "A Visual Analytics Approach to Understanding Spatiotemporal Hotspots," *IEEE Transactions on Visualization and Computer Graphics*, 16(2): 205-220, 2010.

Hypothesis Generation and Exploration

Leave Blank

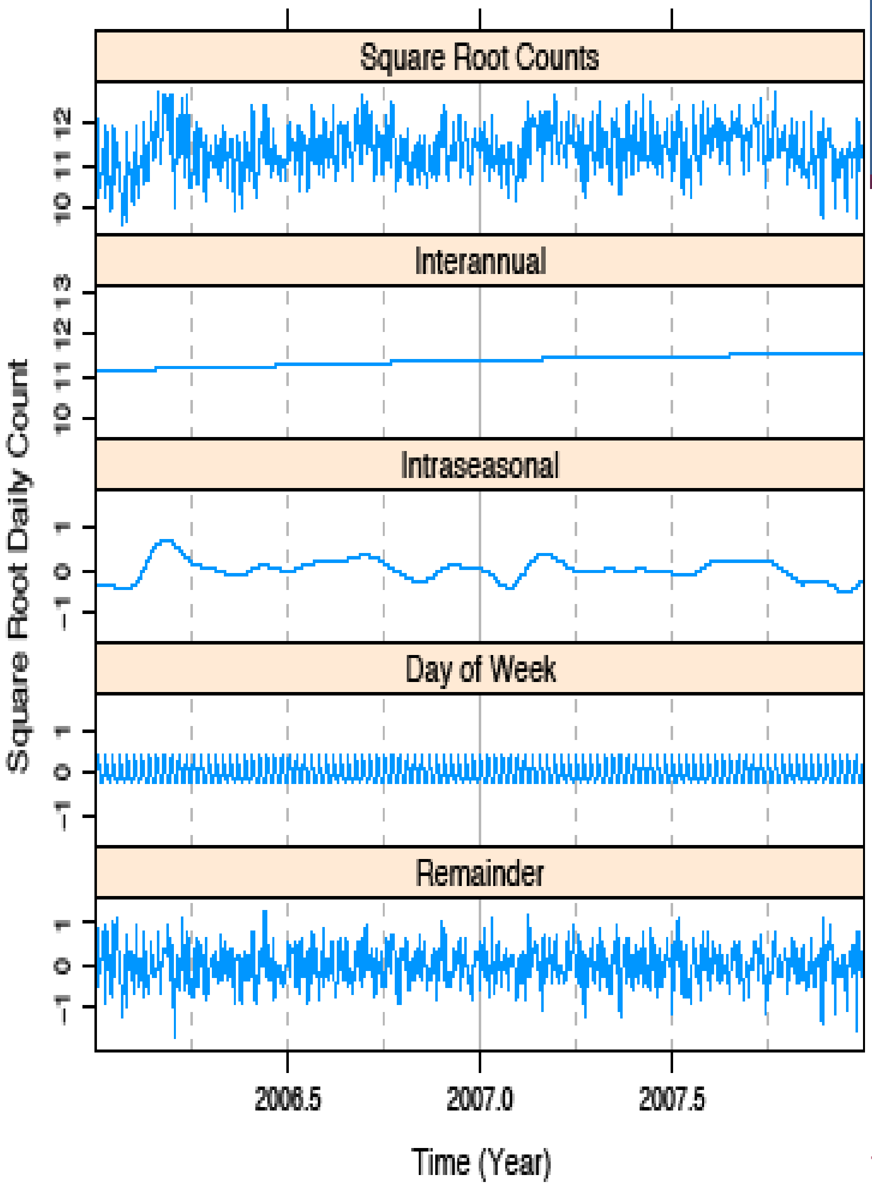


Best Paper Nominee, IEEE Symposium on Visual Analytics Science and Technology (VAST), October 2008, for "Understanding Syndromic Hotspots – A Visual Analytics Approach," (Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W., Grannis, S., Wade, M., Ebert, D.).

Time Series Modeling

Leave Blank

- Seasonal-Trend Decomposition Based on Loess
  - Time series can be viewed as the sum of multiple trend components
  - For each data signal, components are extracted
  - Can then analyze correlation between components

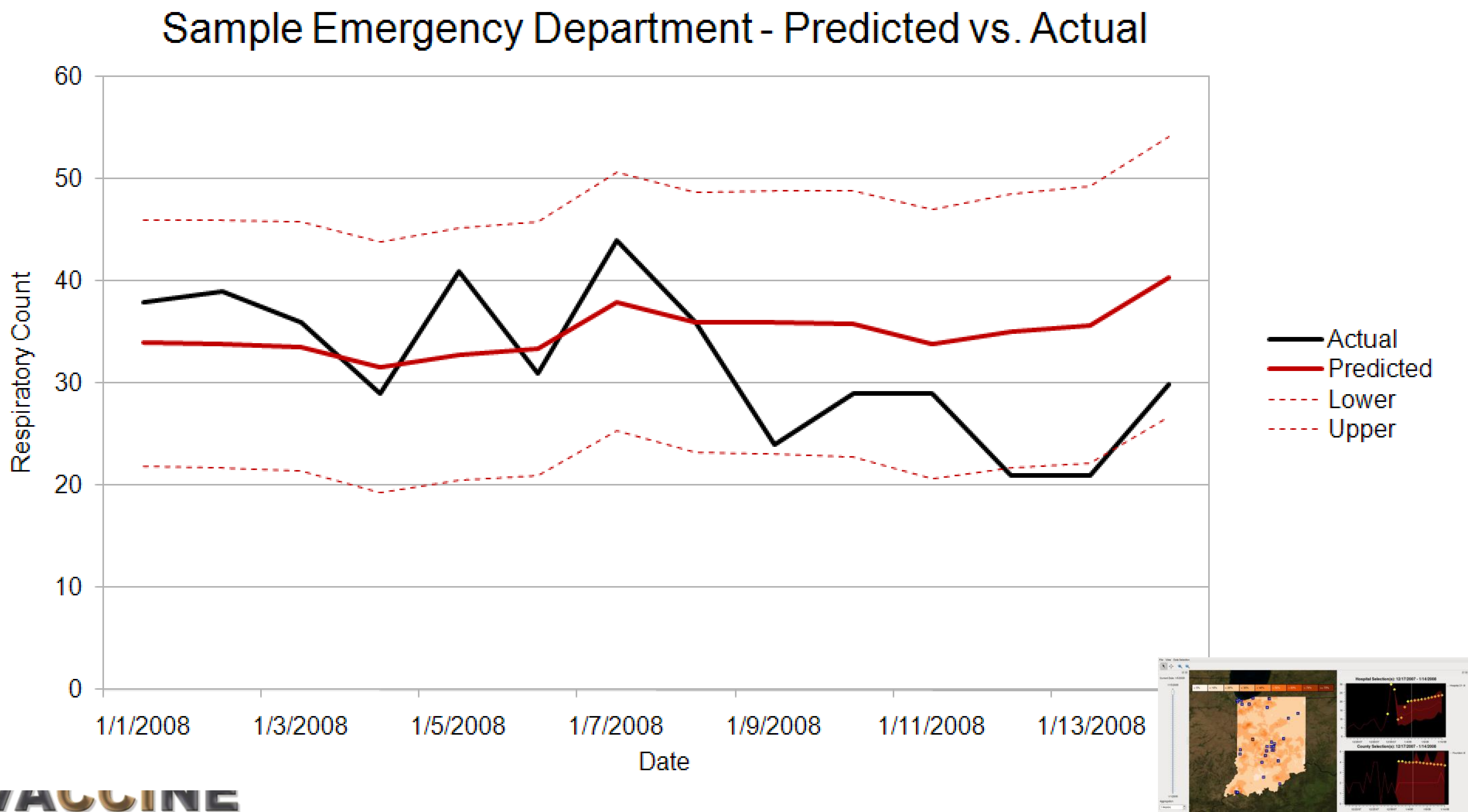


Hafen, R., Anderson, D., Cleveland, W., Maciejewski, R., Ebert, D., Abusalah, A., Yakout
Monitoring Disease Counts," *BMC Medical Informatics and Decision Making*, 9(21), 20

VACCINE

Predictive Visual Analytics

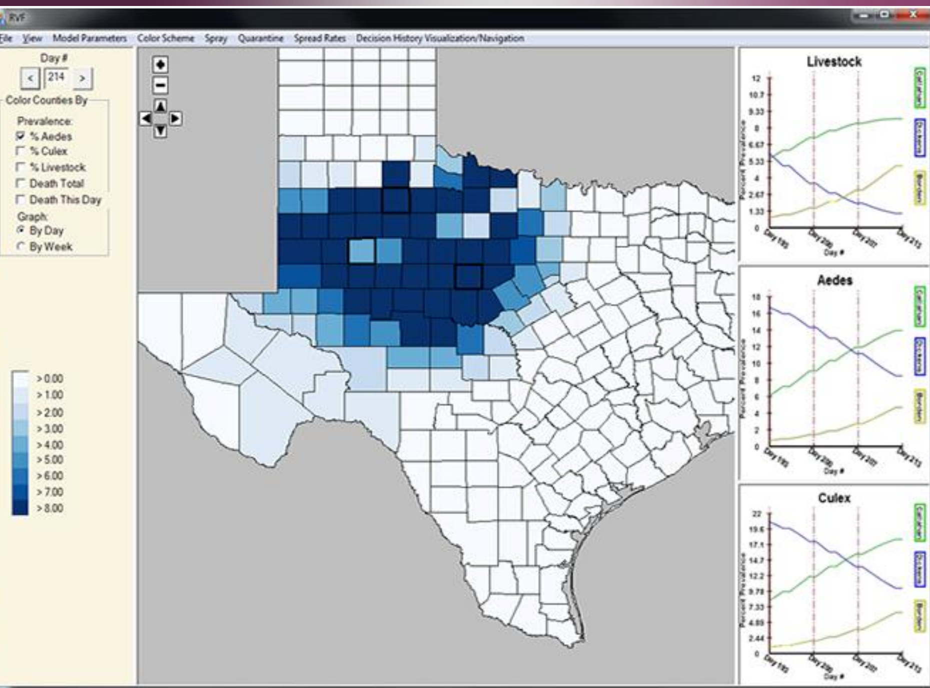
Leave Blank

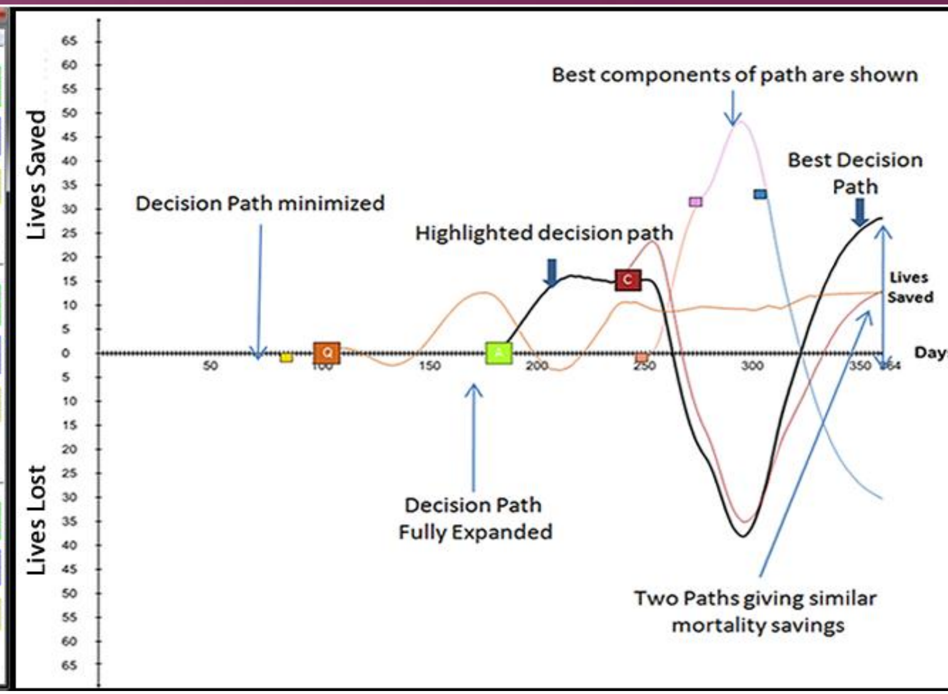


VACCINE

Advanced Decision Support Tools: Rift Valley Fever

Leave Blank

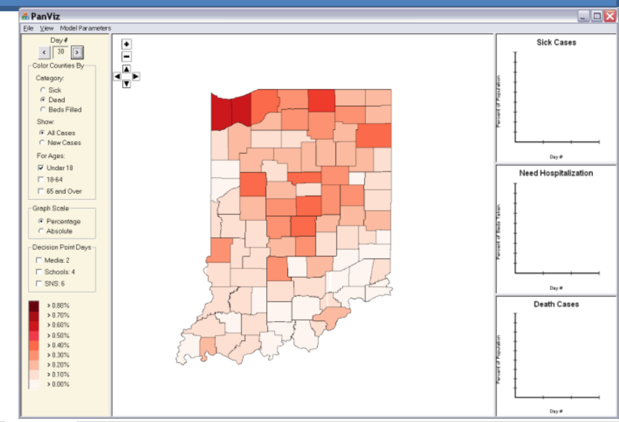




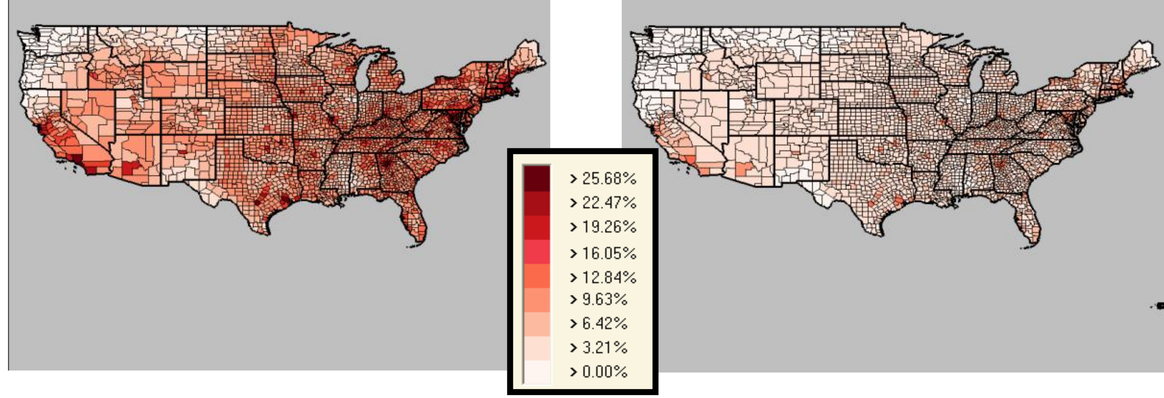
Modeling a Pandemic

Leave Blank

- Pandemic Influenza Planning Tool
- Models user specified:
  - Pandemic influenza characteristics
  - County population, demographics, hospital beds
- Decision measures
  - Strategic National Stockpile deployment
  - School Closures
  - Media Alerts



Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.



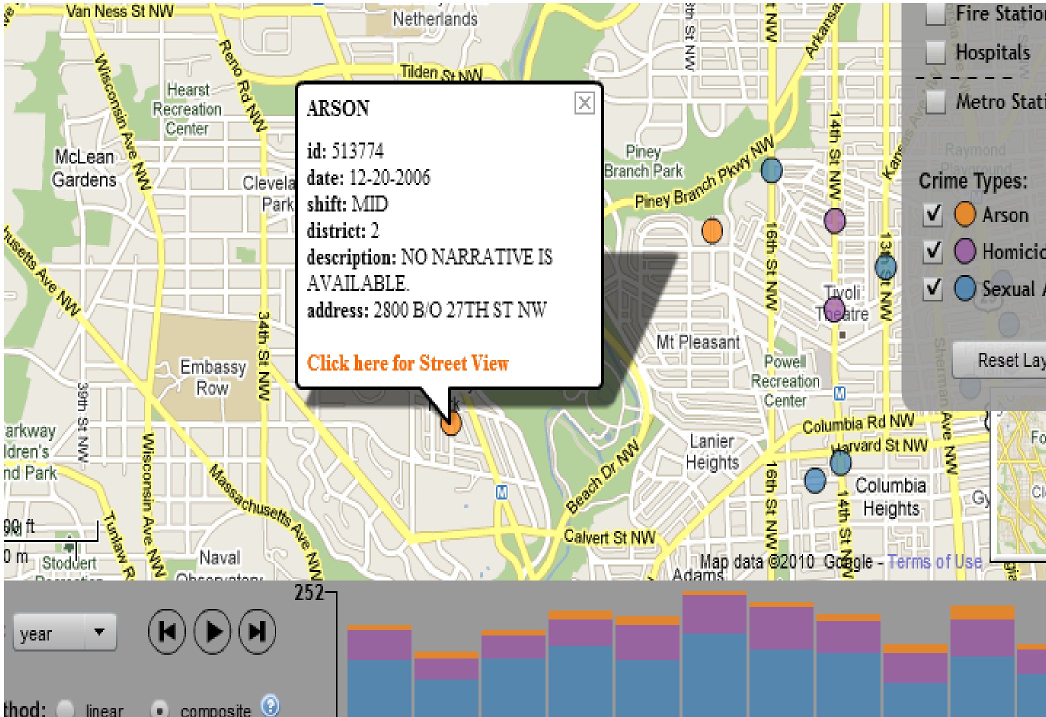
Pandemic spread on day 37 with no decision measures implemented

Pandemic spread on day 37 with all decision measures implemented

Visual Analytics Uses for Crisis Management and Response

Leave Blank

- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis
- Purpose:
  - Planning for resiliency
  - Long-term analysis
  - Predictive analytics
  - Training
  - Detection
  - Investigation
  - Response
  - Recovery, remediation

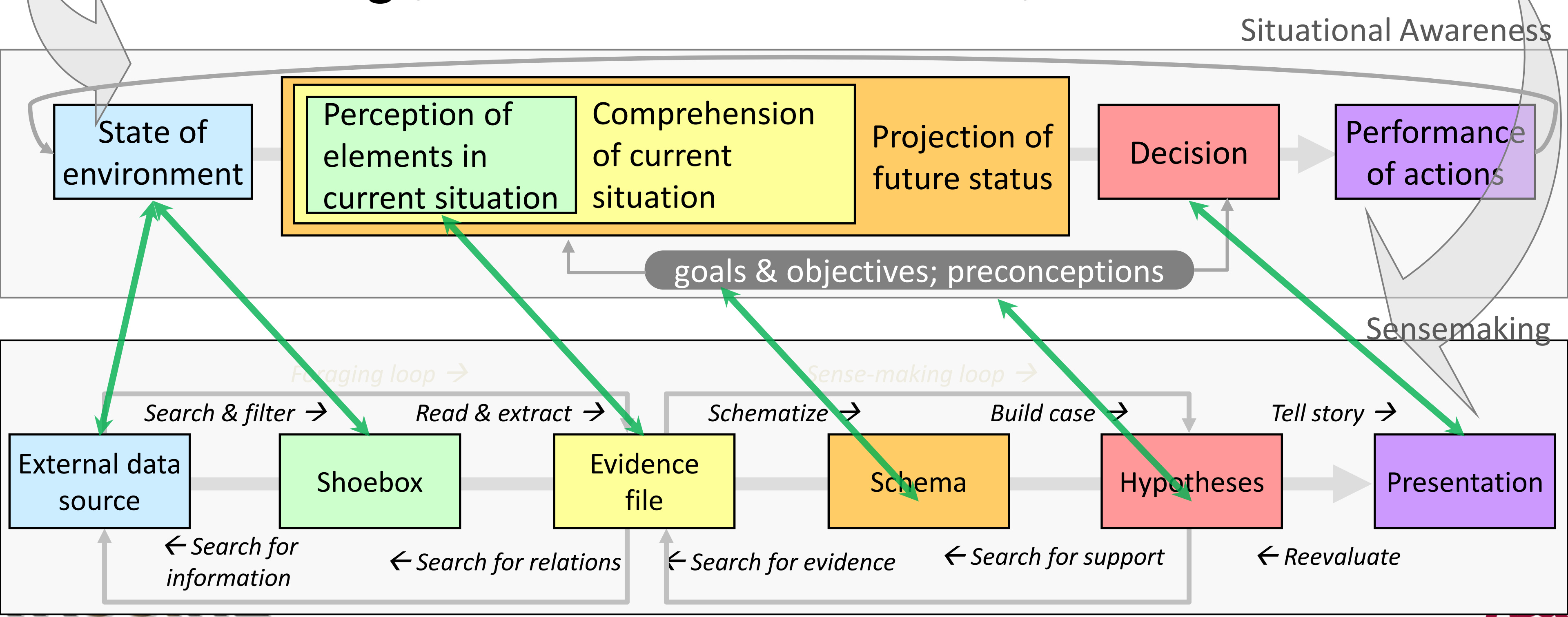


VACCINE

Visual Analytics Situational Awareness←→ Sensemaking (courtesy Alan MacEachren)

Leave Blank

Situational awareness model (after Endsley, 1995; top) compared to model of sense making (after Pirolli & Card, 2005; bottom)



Situational Awareness, Planning, Investigation, Response Visual Analytics

Leave Blank

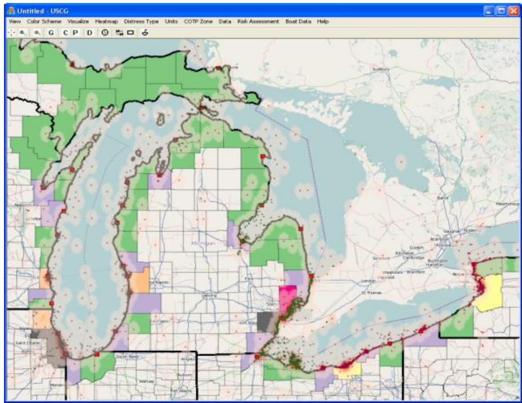
- Situational Awareness and Assessment from massive data
  - In-field and desktop historical and current incident/crime information
  - Innovative spatiotemporal visualization, analysis, prediction
  - Image analysis tools
  - Integrated statistical forecasting and simulation models
  - Advanced social media analytics
- Tools for Intelligence analysis
- Planning and response support
  - Integrated simulation, resource allocation, census, weather data

VACCINE

ASU

Visual Analytics Uses for Public Safety

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia massive data integration & analysis



VACCINE

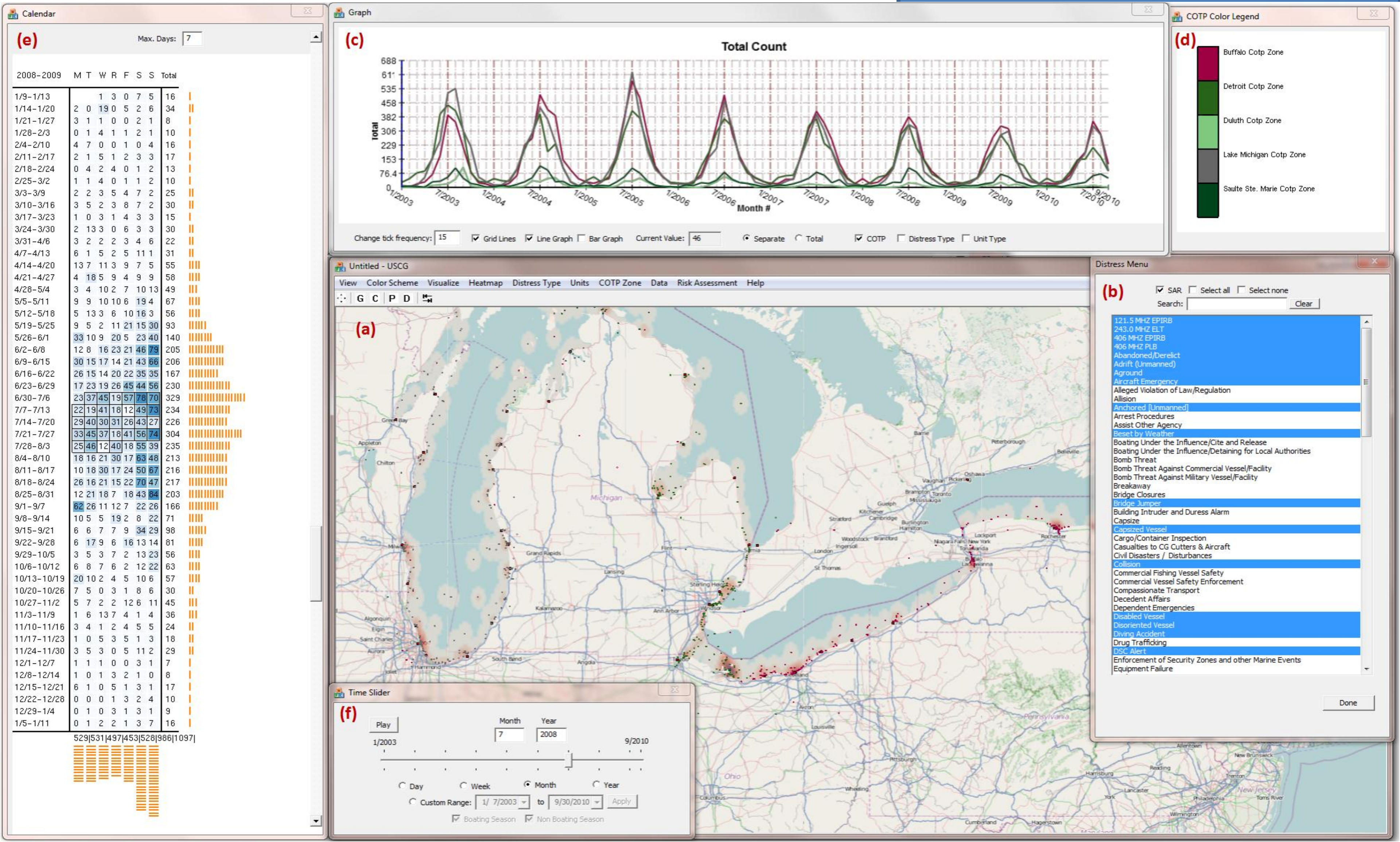
Example Maritime Safety Visual Analytics Environment

Leave Blank

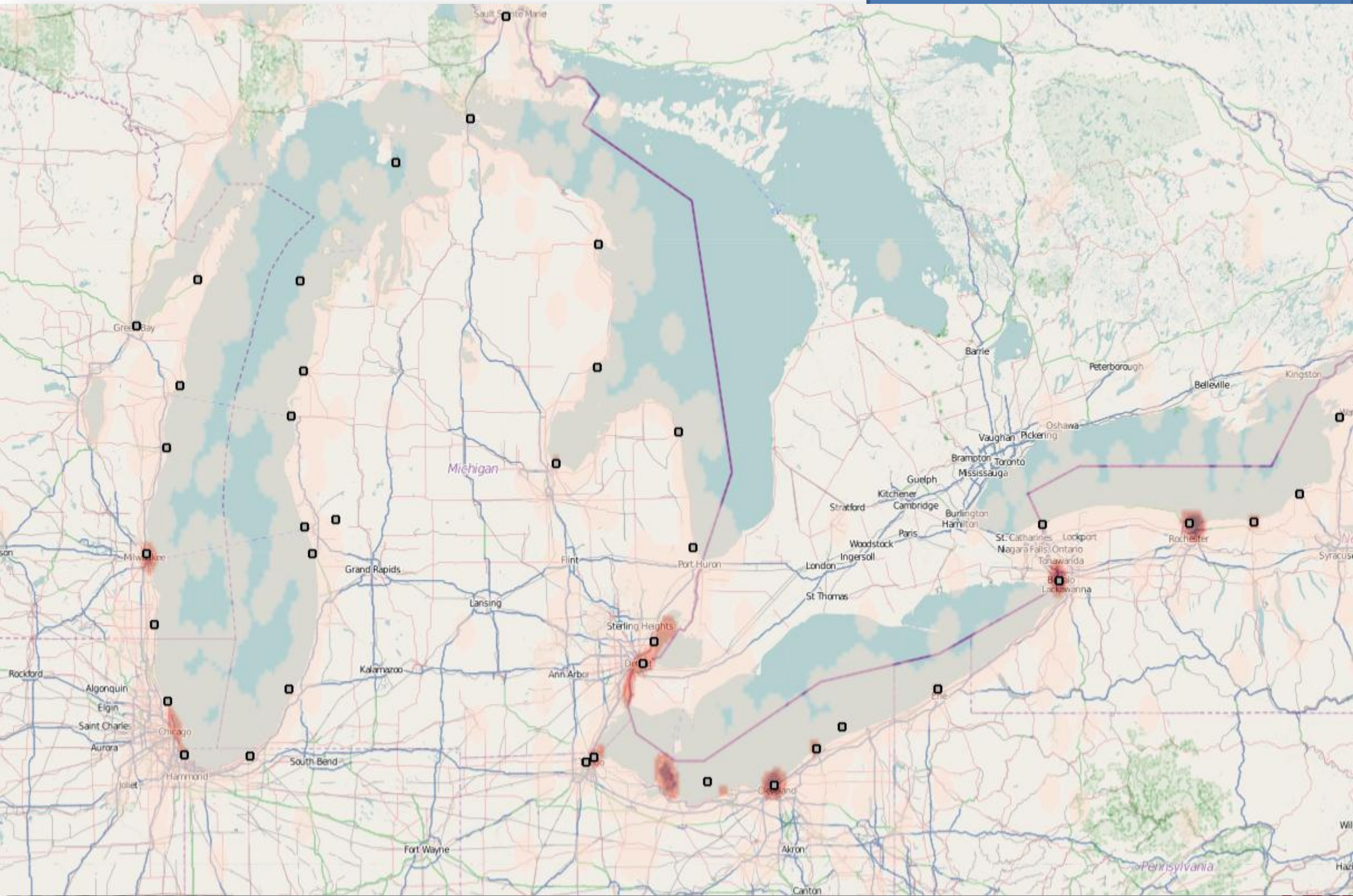
- Supports decision making and risk assessment
- Interactive exploration and analysis of trends, patterns and anomalies
- Allows analysis of risks associated with
  - Closing one or more Coast Guard stations
    - Find optimal stations that absorb work load of the closing station
  - Allocating new resources
    - Impact on safety and efficiency of operation
- Currently being used by analysts at the U.S. Ninth District, Fifth District, HQ, and Atlantic Commands

Example: USCG D9 Search & Rescue Operational Analysis

Leave Blank

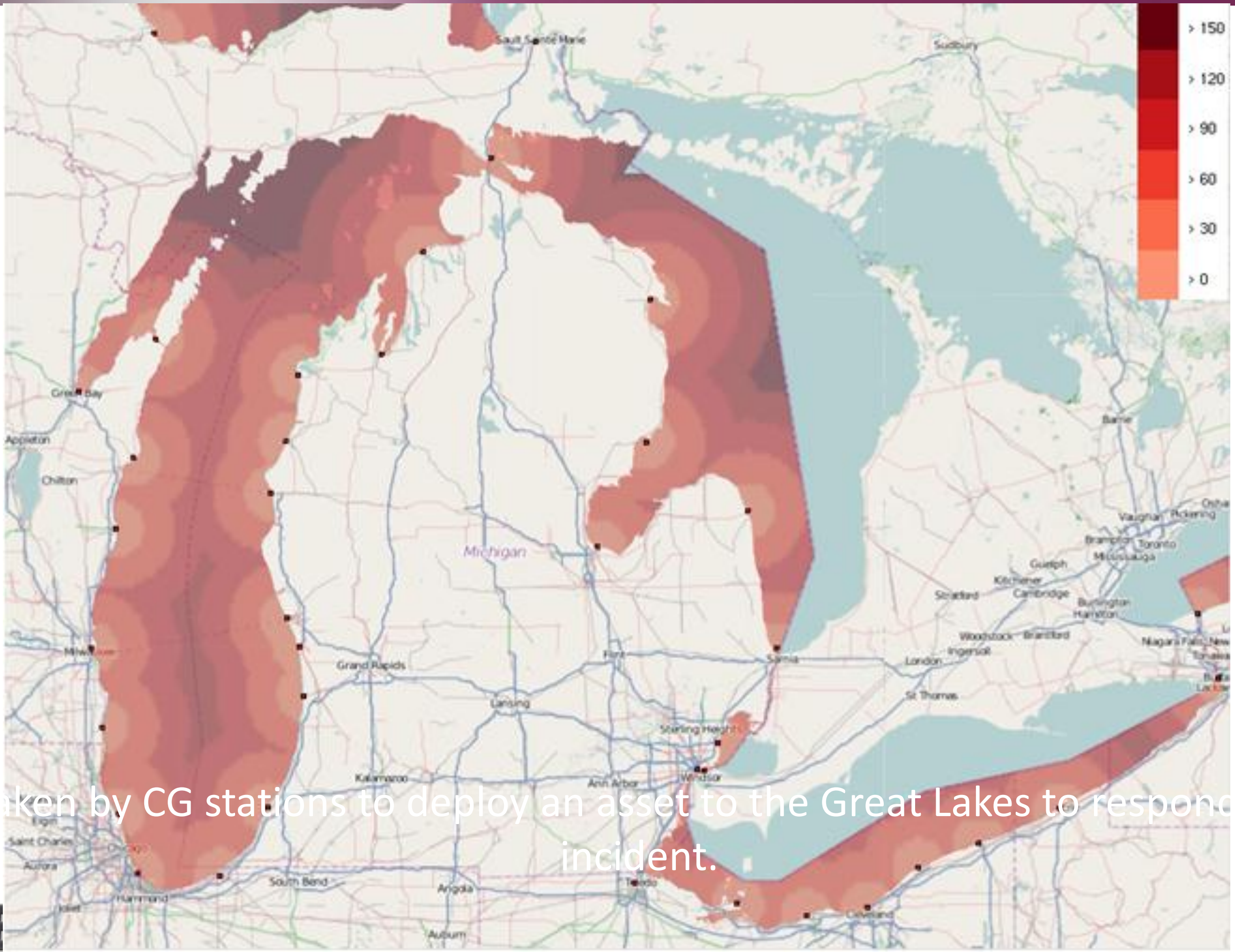


Total Risk Based on SAR Cases



System Features: SAR Risk Profile

Leave Blank

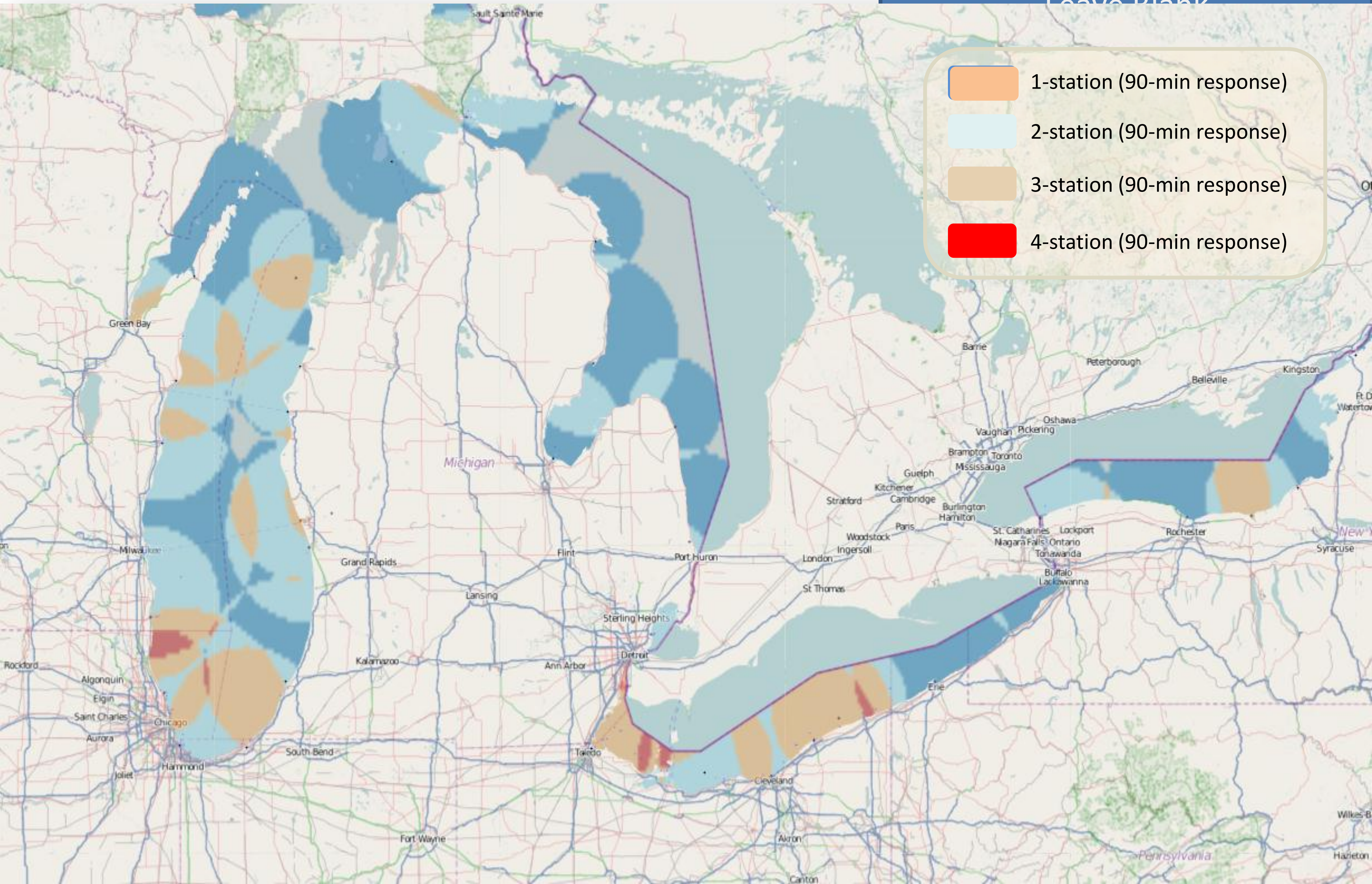


…ken by CG stations to deploy an asset to the Great Lakes to respond… incident.

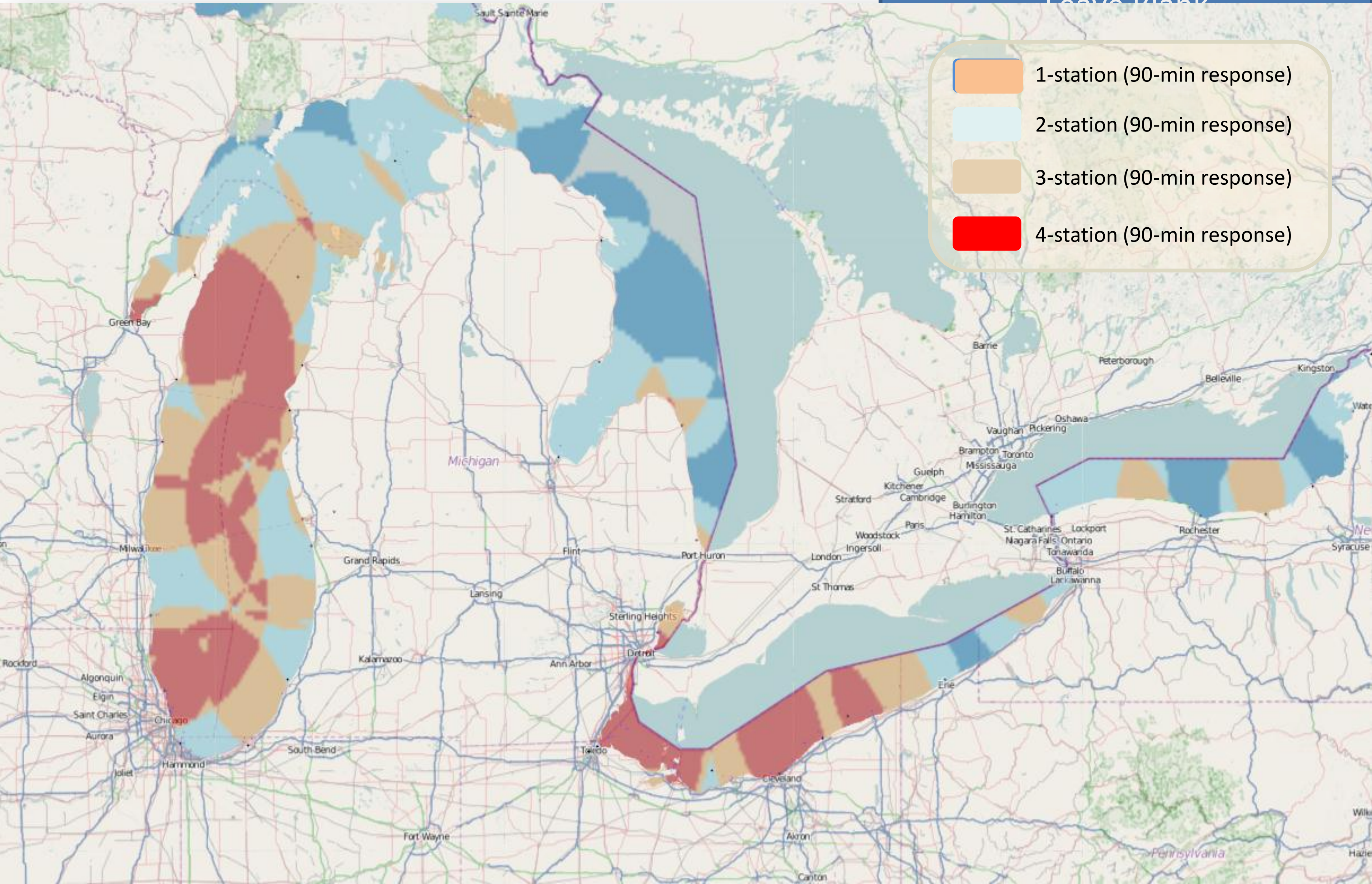
VACCINE

ASU

Response Efficiency – Current Assets

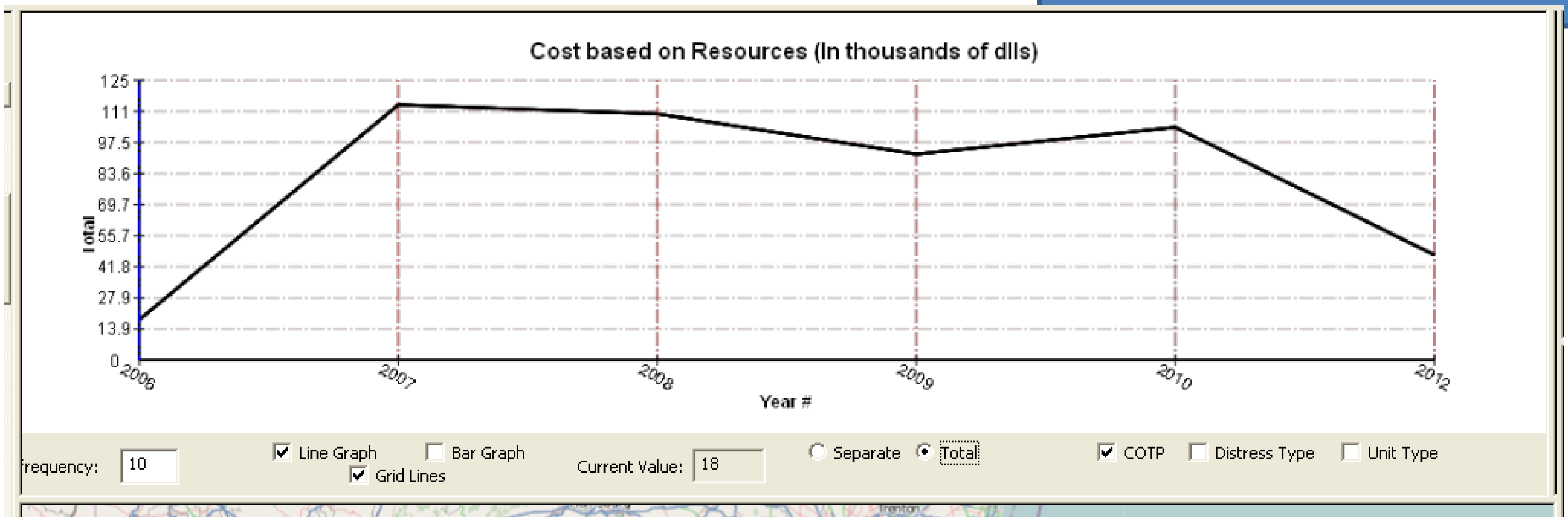


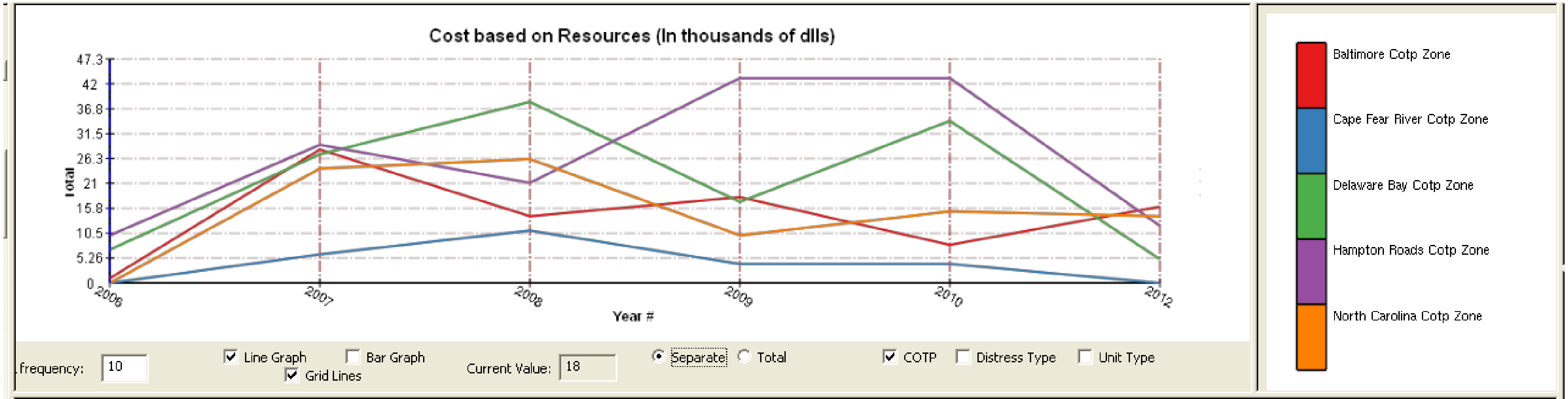
Response Efficiency – Potential Future Assets



n-Suspended Sorties over

Leave Blank

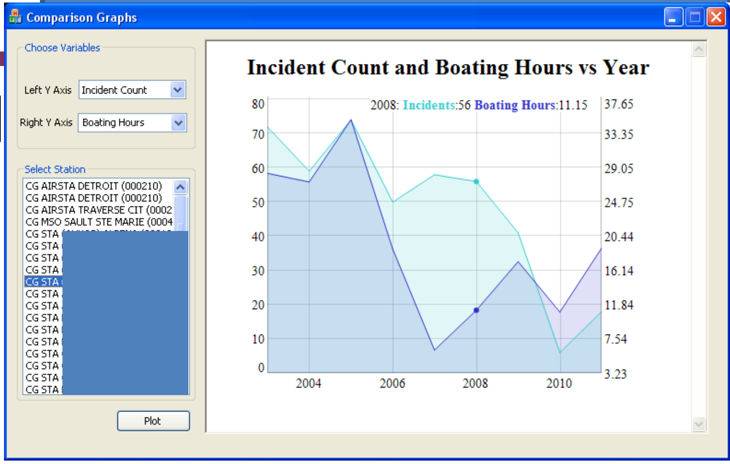


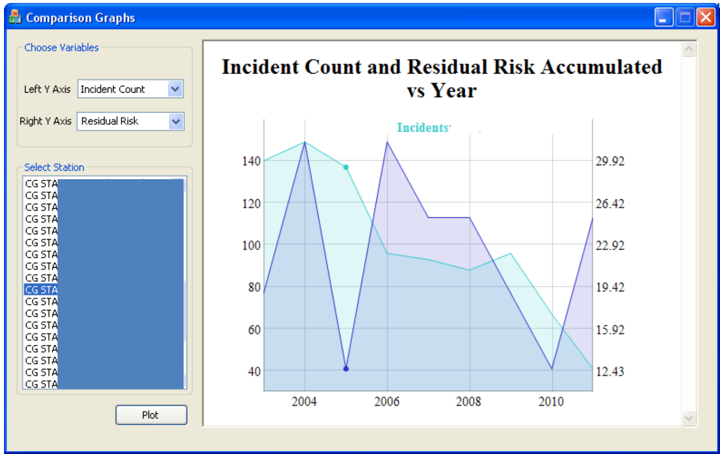


Resource Allocation and Risk-Based Decision Making

Leave Blank

- Explore risk-based decision making and utilize historical data for analysis and prediction
  - Total Risk, Mitigated Risk, Residual Risk
  - Explore 11 different USCG missions
  - Explore allocation of assets with different capabilities
  - Explore staffing, utilization, assets vs. risk measures
  - Perform What-If scenarios

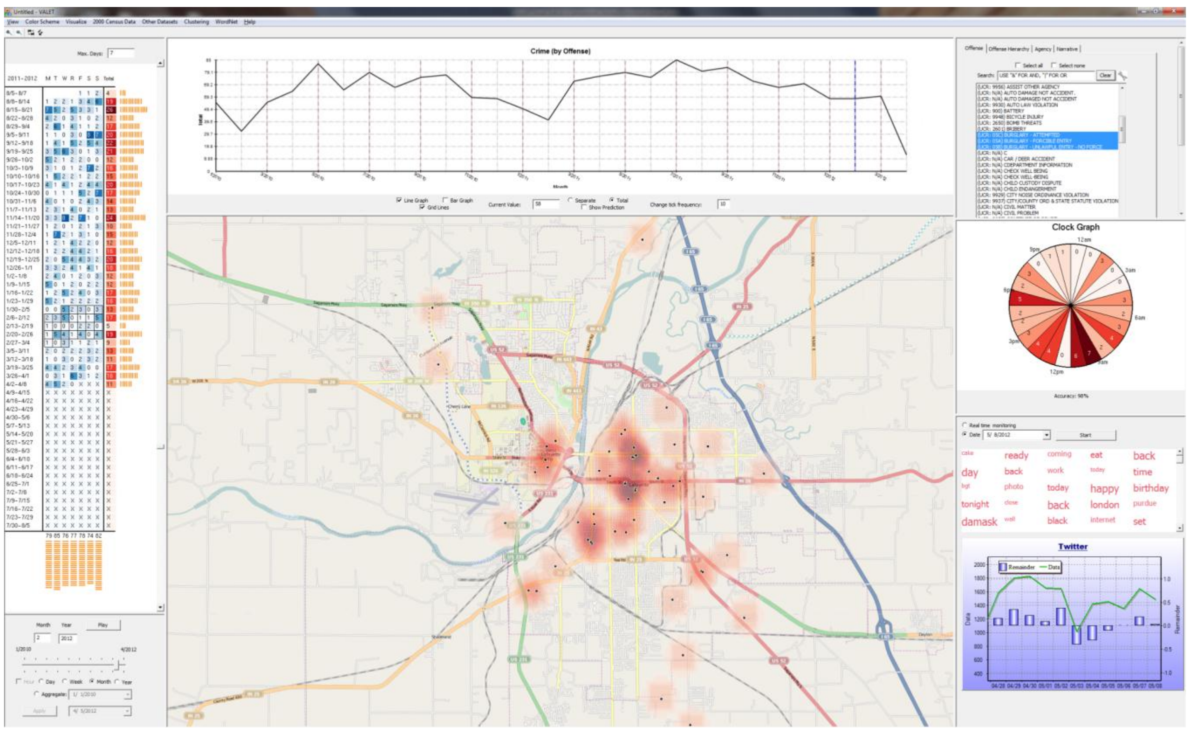


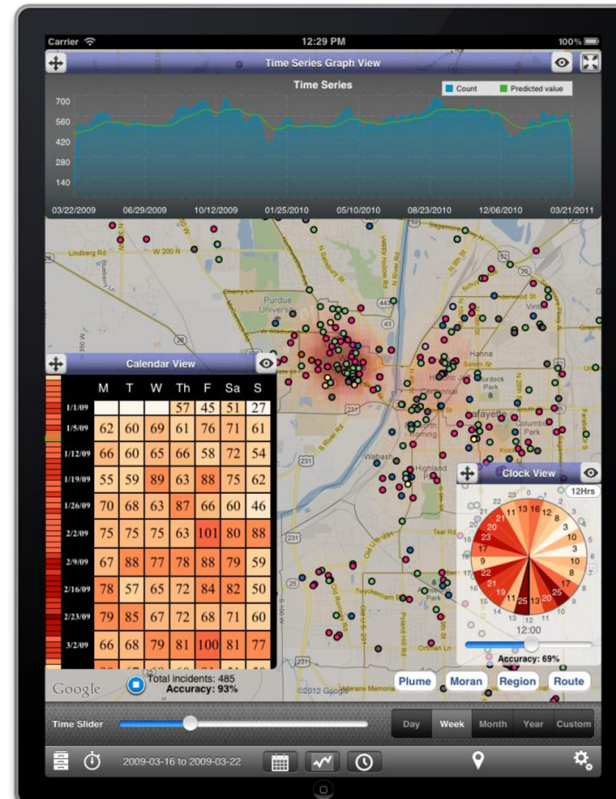


VACCINE

Visual Analytics Law Enforcement Toolkit (i)VALET

Leave Blank

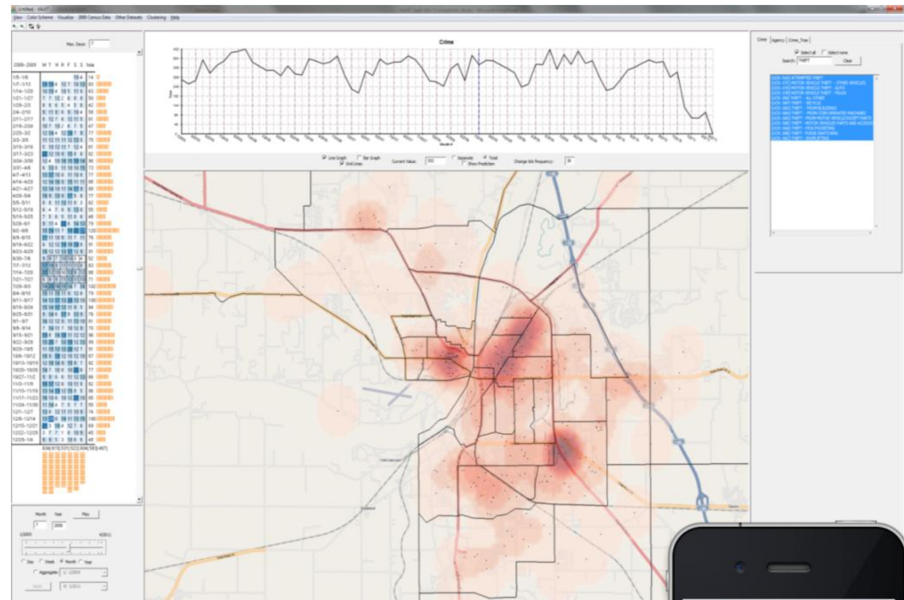




Visual Analytics Law Enforcement Toolkit
(VALET, iVALET)

Leave Blank

Interesting VA features:
- Time a first class citizen
- Temporal exploration and comparisons
- Integrated predictive models
- Interactive multivariate correlation and visual analytics
- Novel statistical predictive model incorporated for planning
- Novel interactive anomaly detection from social media



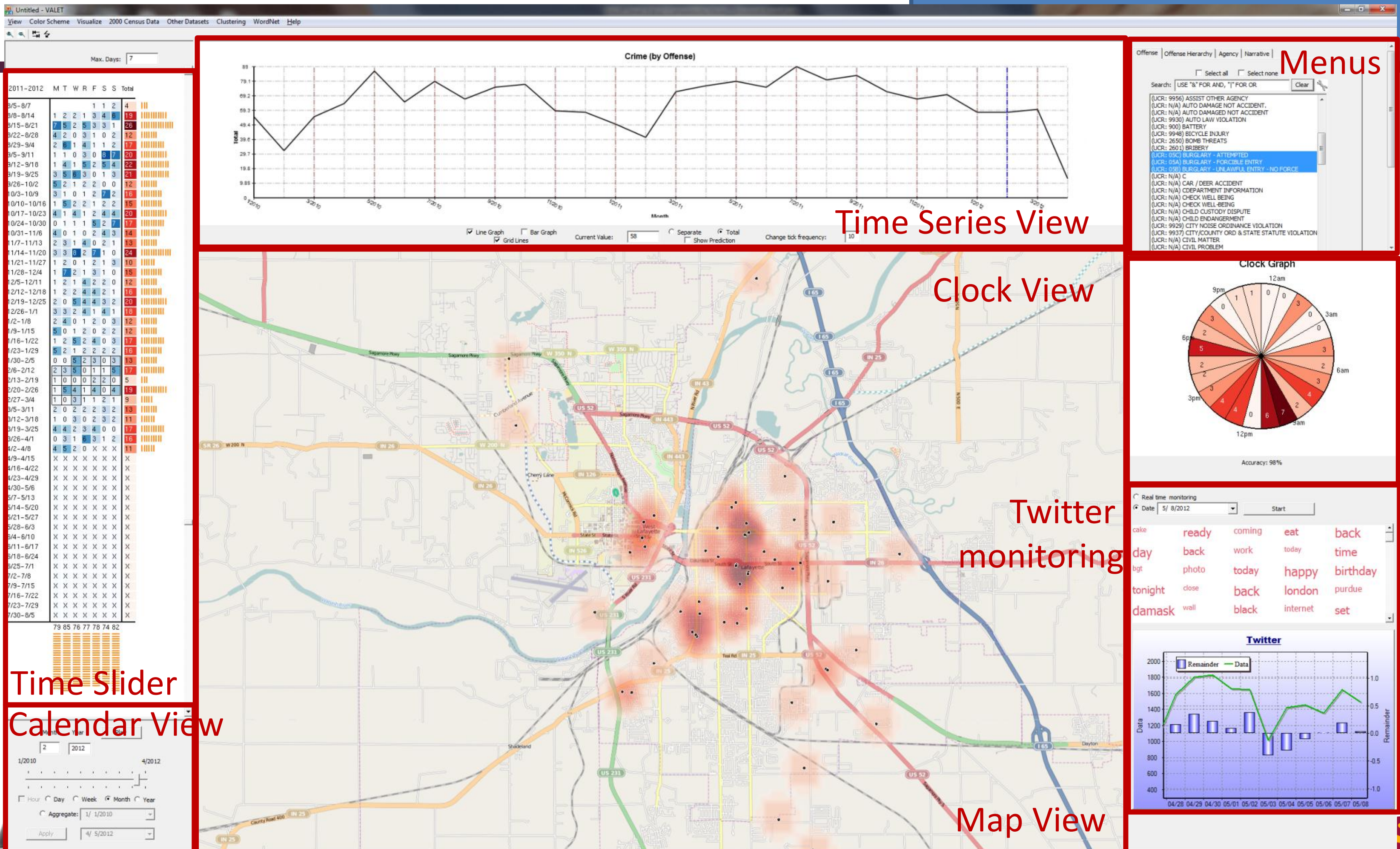
VALET delivered:
- Spring 2011: WL, Lafayette Police

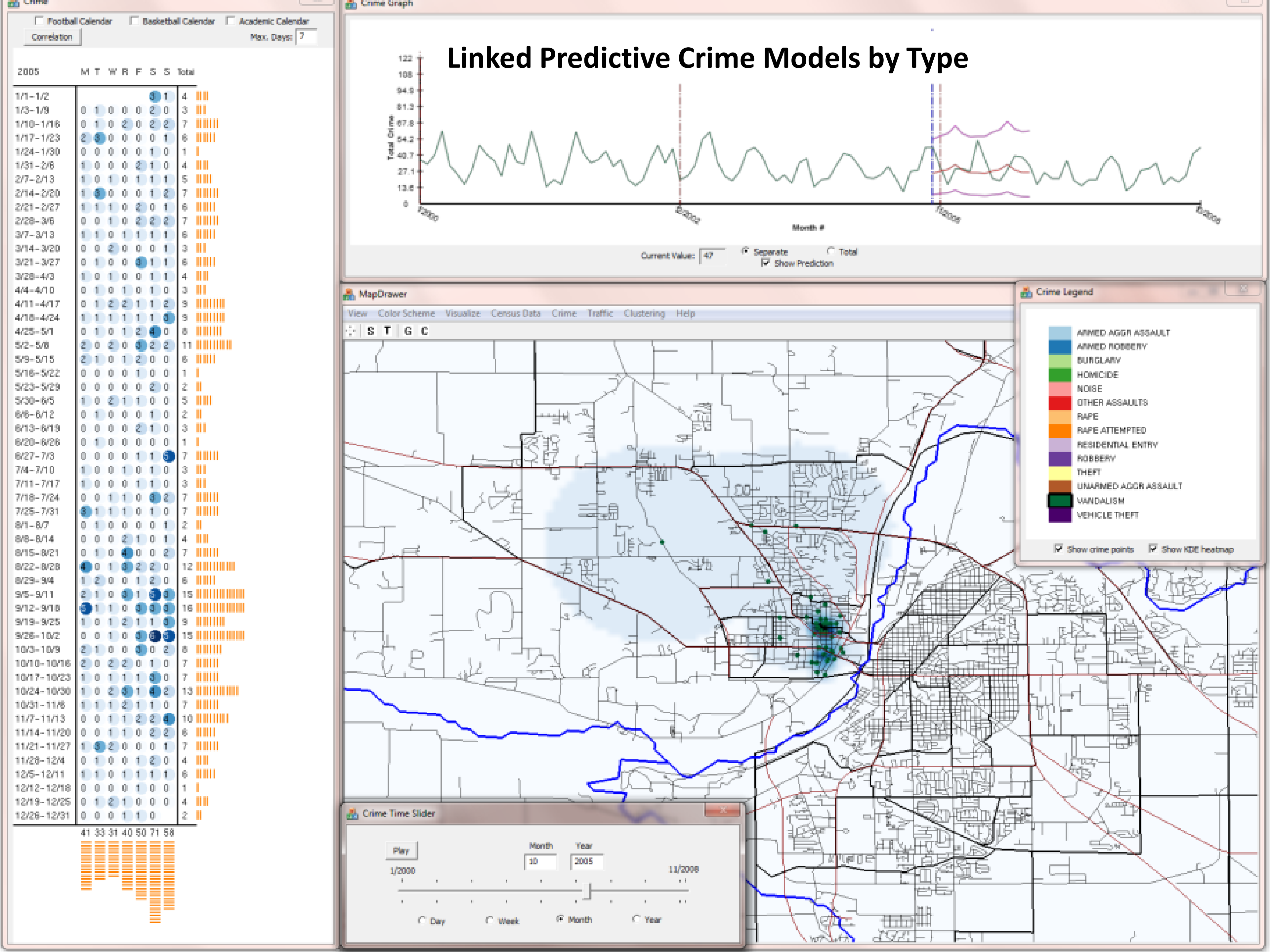
iVALET delivered:
- October 2011: Purdue, WL Police

VACCINE

VALET Overview

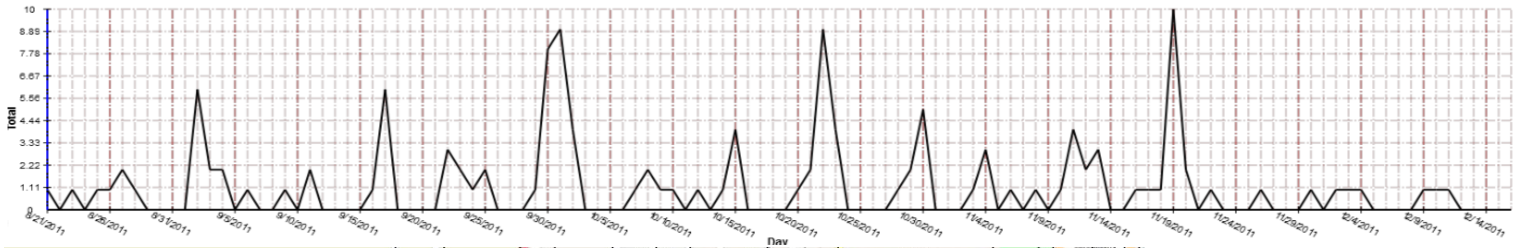
Leave Blank

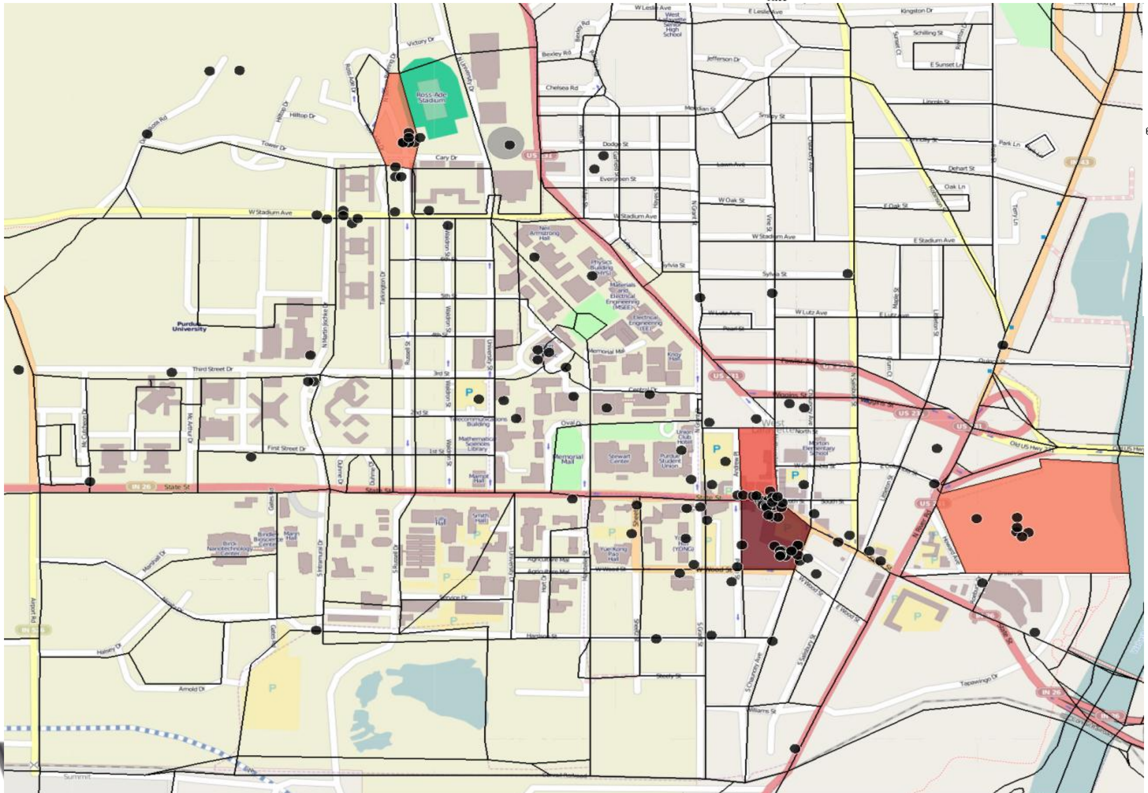


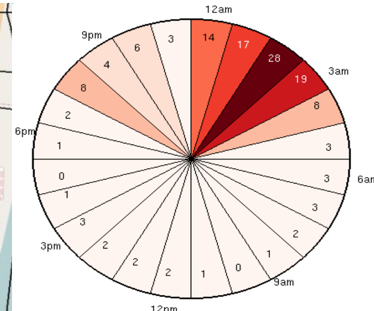

Linked Predictive Crime Models by Type
Crime
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
Max. Days: 7
2005
M T W R F S S Total
Crime Graph
Total Crime
Month #
Current Value: 47
Separate
Total
Show Prediction
MapDrawer
View Color Scheme Visualize Census Data Crime Traffic Clustering Help
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Show crime points
Show KDE heatmap
Crime Time Slider
Play
Month
Year
10
2005
1/2000
11/2008
Day
Week
Month
Year

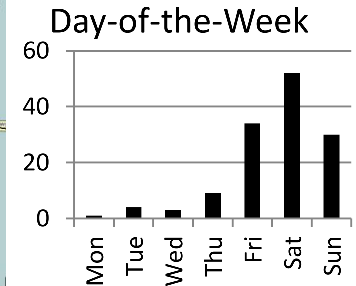
Example:
Drunkenness / Public Intoxication

Leave Blank



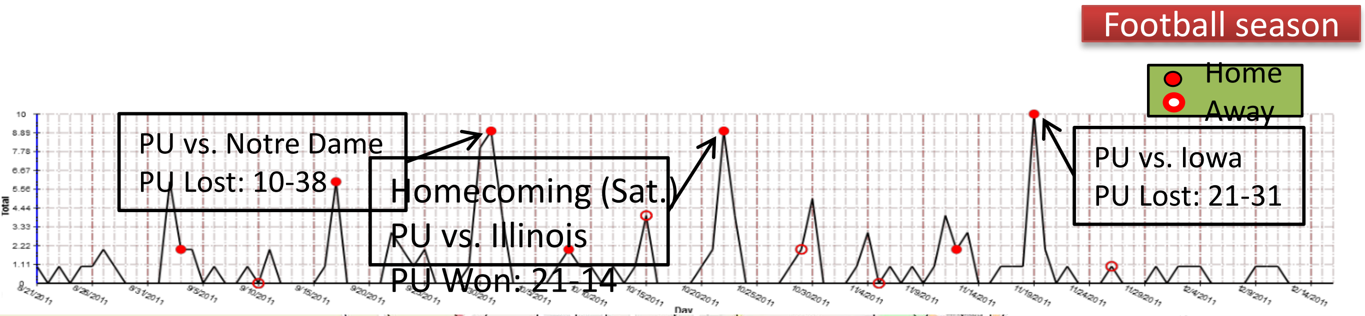


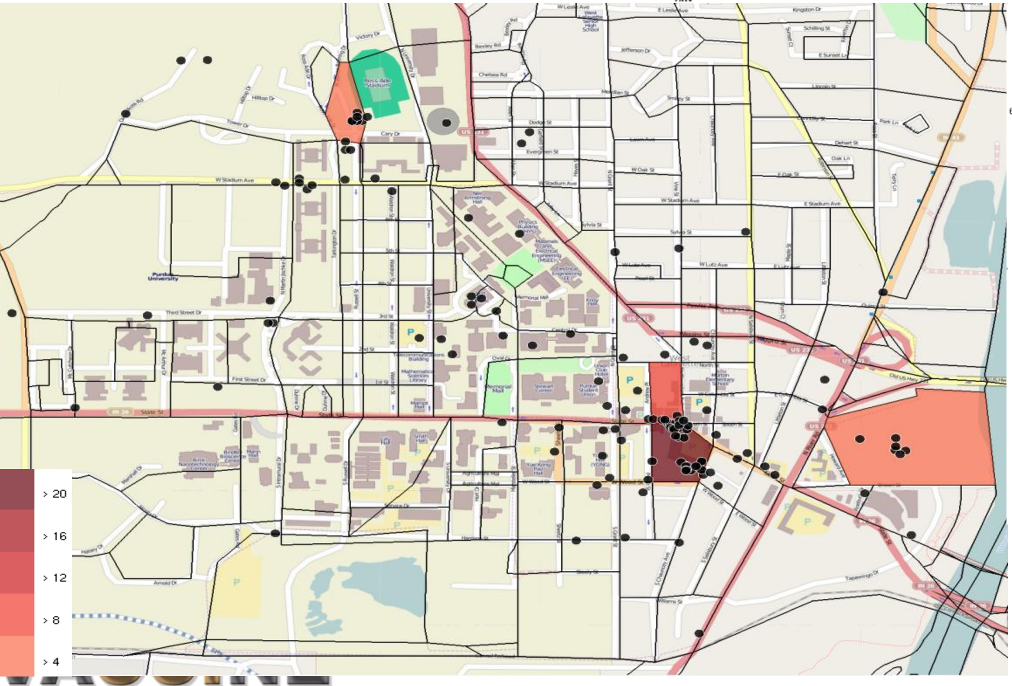


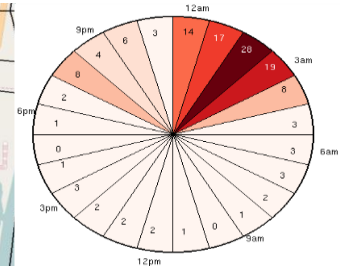


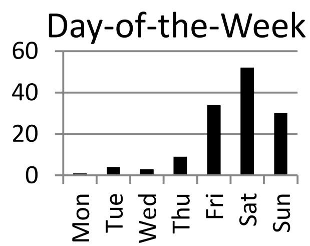
Example: Drunkenness / Public Intoxication

Leave Blank









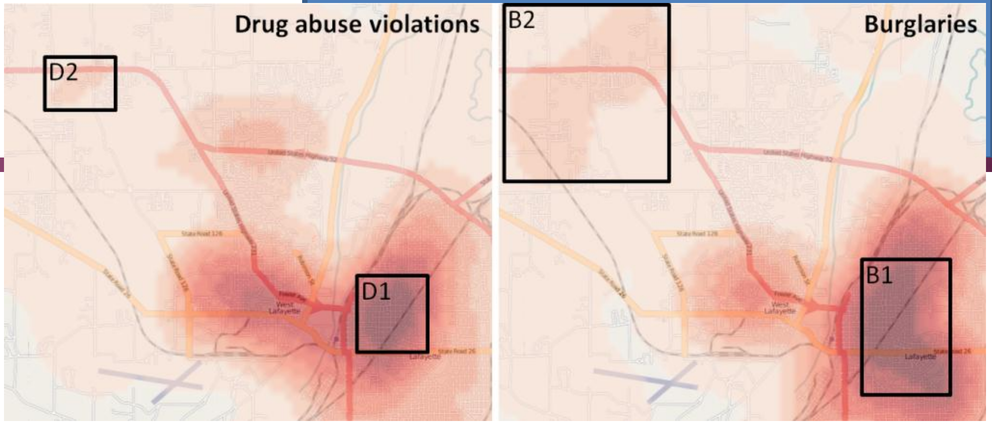
VALET Feature: Correlating Datasets at Multiple Granularities

Leave Blank

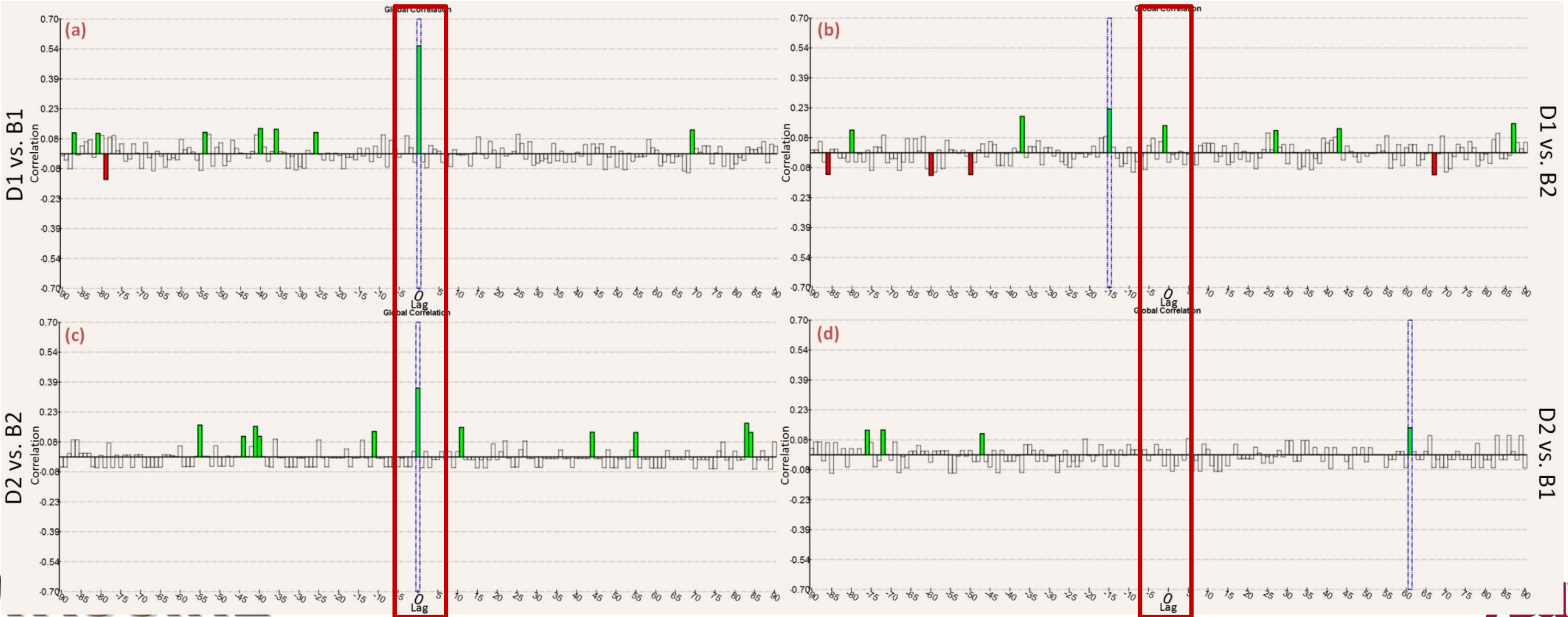
- Correlative analysis framework for exploring correlations among multi-variate/-source data
- Determining forces that determine nature of relationships among datasets
- Detection of periodic properties among different data variables
- Allows correlations at multiple **spatial** and **temporal** granularities

Case Study: Drug Abuse Violations vs. Burglaries

Leave Blank



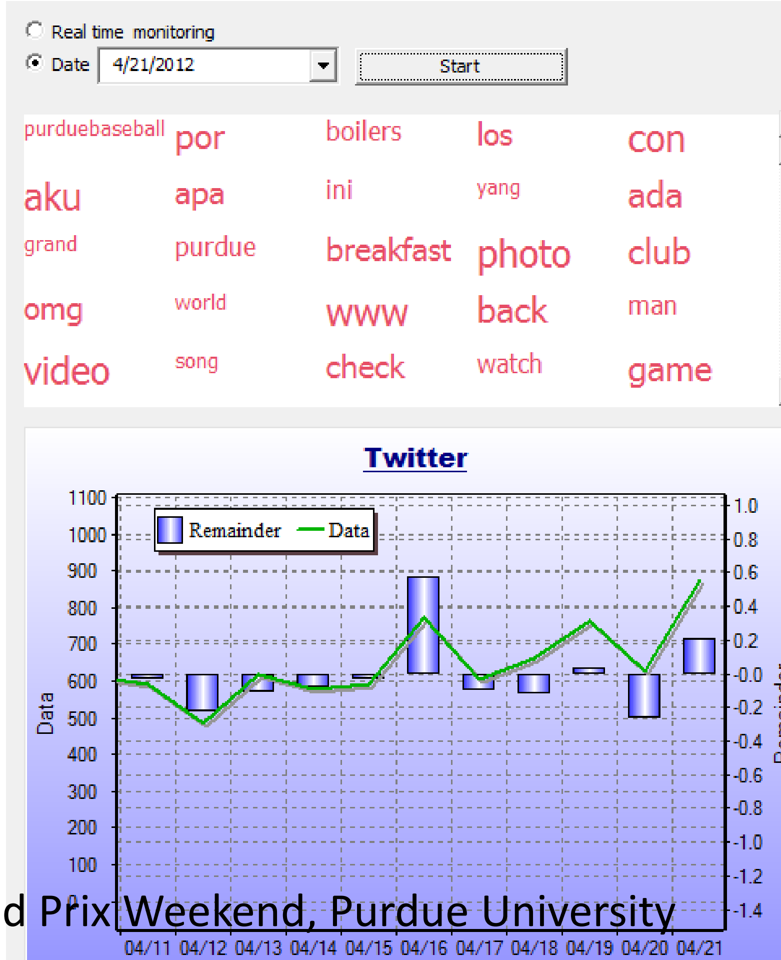
- Over geospatial hotspots



VALET Feature: Real-time Twitter Monitoring and Integration

Leave Blank

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud

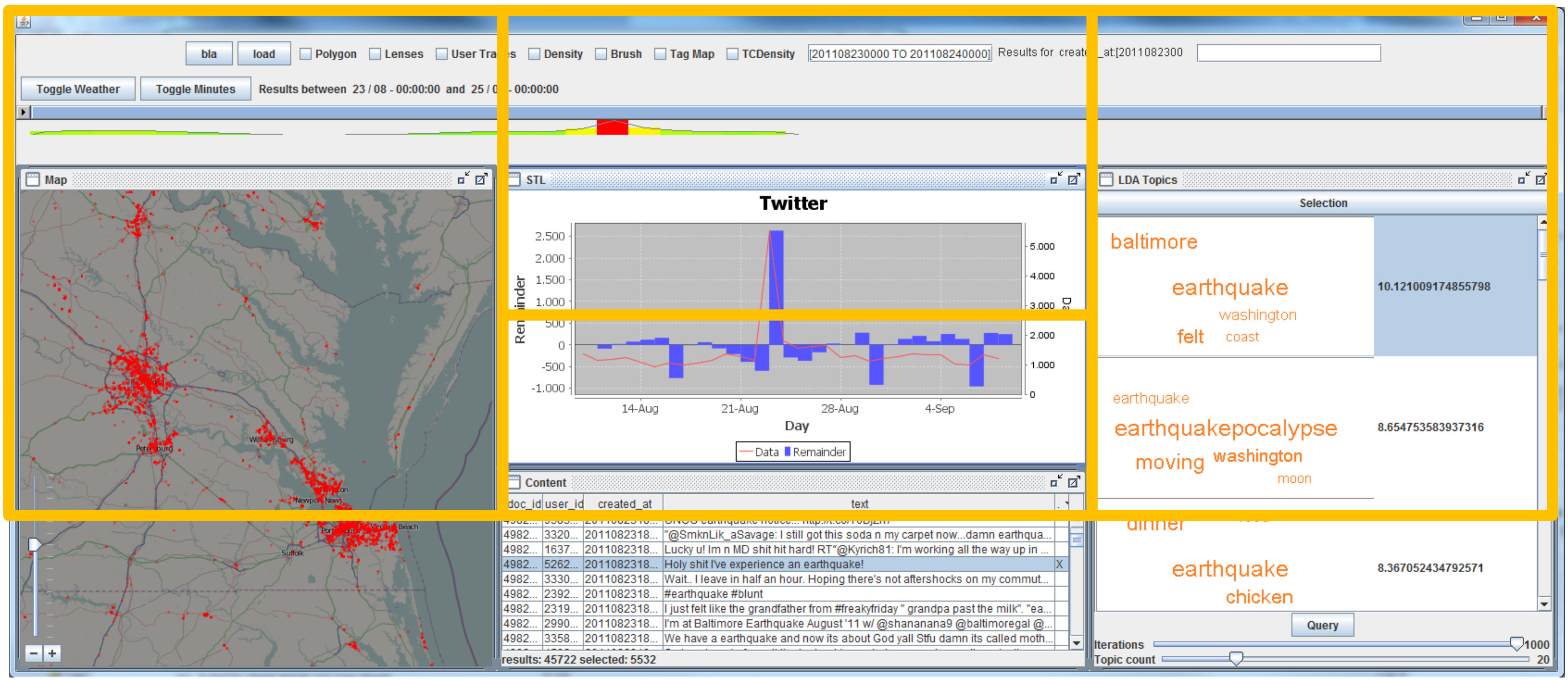


Grand Prix Weekend, Purdue University

VACCINE

Spatiotemporal Social Media Analytics for Abnormal Event Detection

Leave Blank

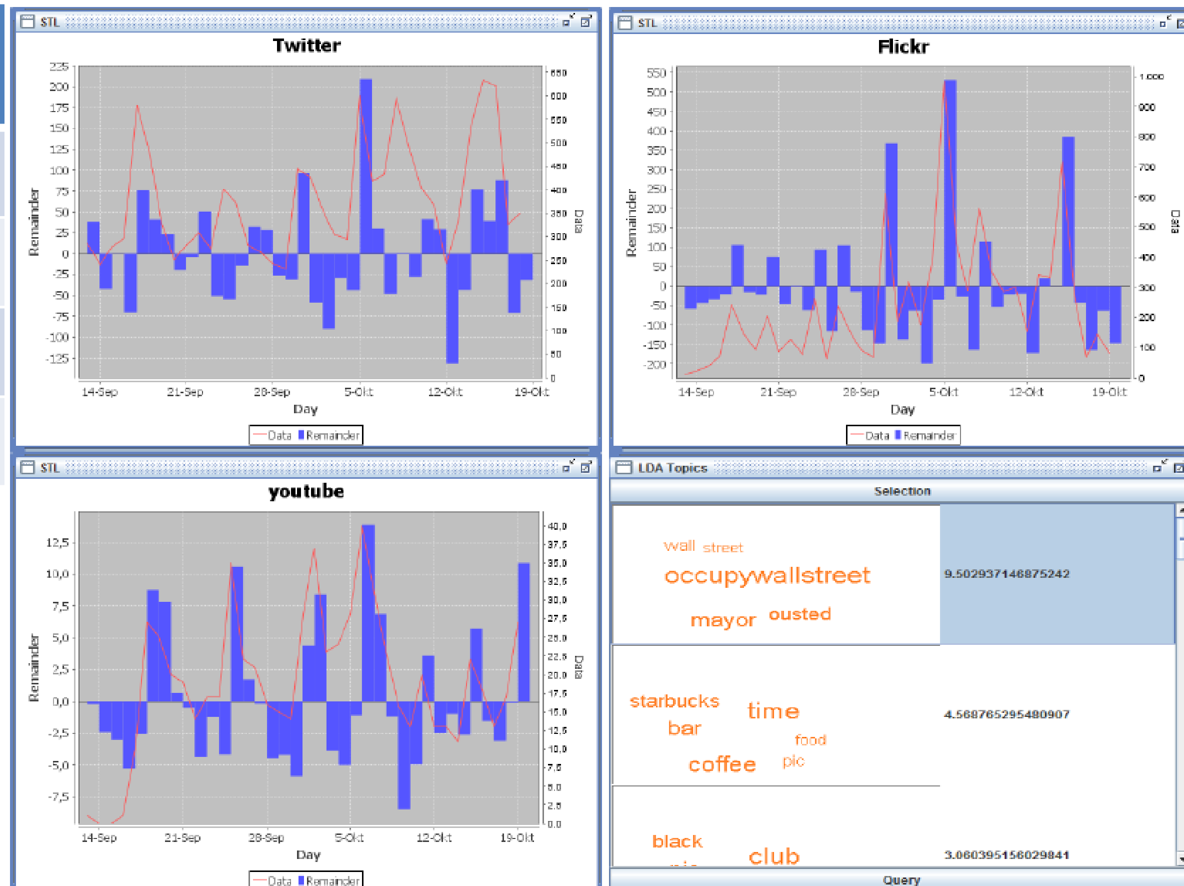


Correlation Between Multiple Social Media

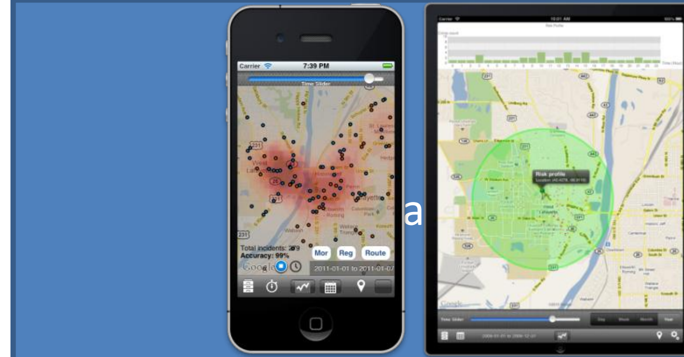
Leave Blank

- Occupy Wall Street event

Date	Twitter (z-score)	Flickr (z-score)	Date	YouTube (z-score)
9/17, 2011	2.53	0.97	9/18, 2011	5.98
9/30, 2011	3.20	3.50	10/2, 2011	1.38
10/5, 2011	3.17	3.36	10/6, 2011	2.75
10/15, 2011	4.90	4.53	10/15, 2011	2.19



iVALET

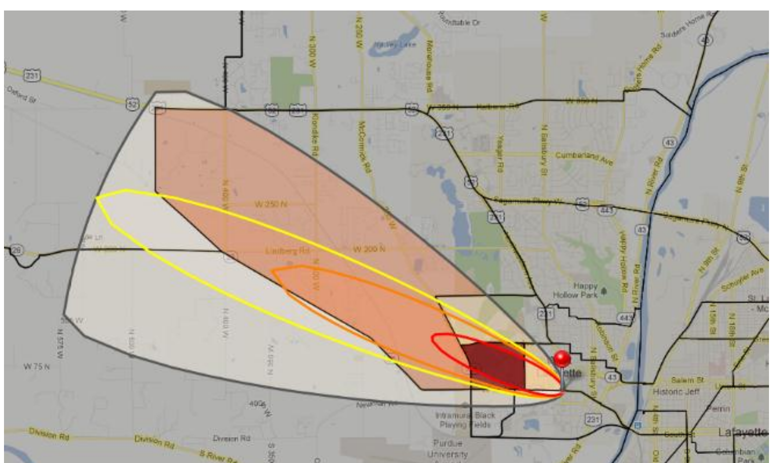


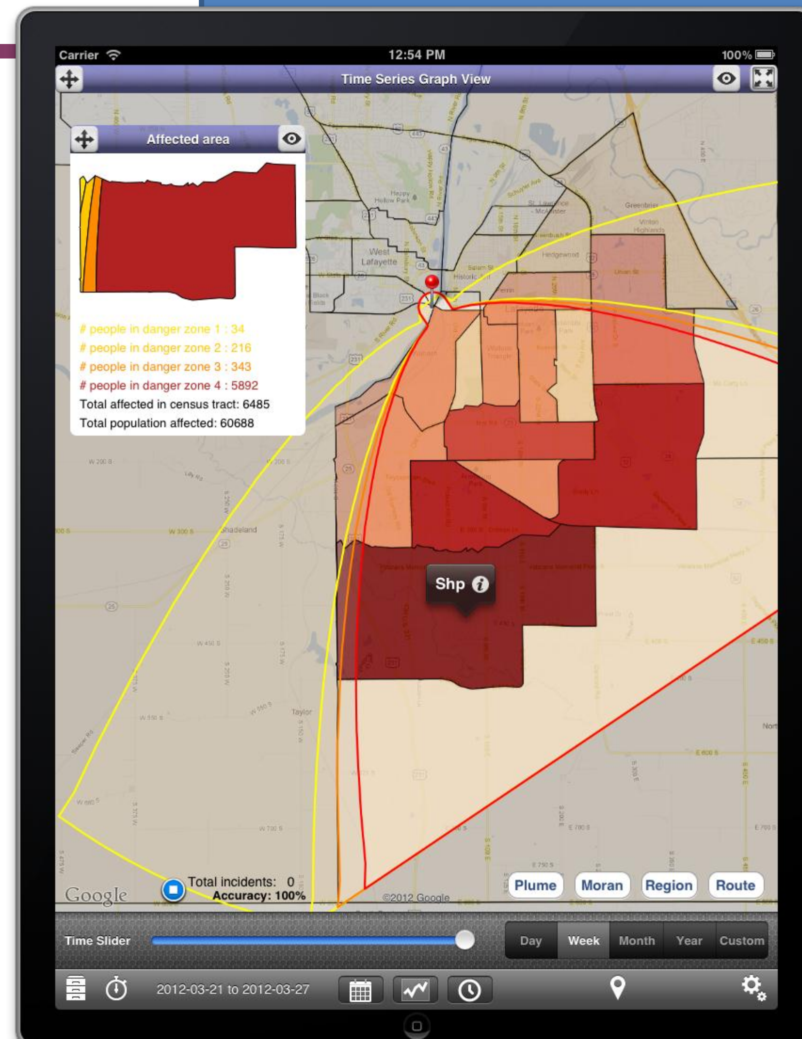
- Explore criminal, traffic and civil data on-the-go
- Risk assessment
- Use current spatial + temporal context into analysis

VACCINE

MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

- Chemical release plume modeling identifies census tracts with the highest number of expected people affected





Jigsaw:
Visual Analytics for Investigative Analysis
and Exploration of Document Collections

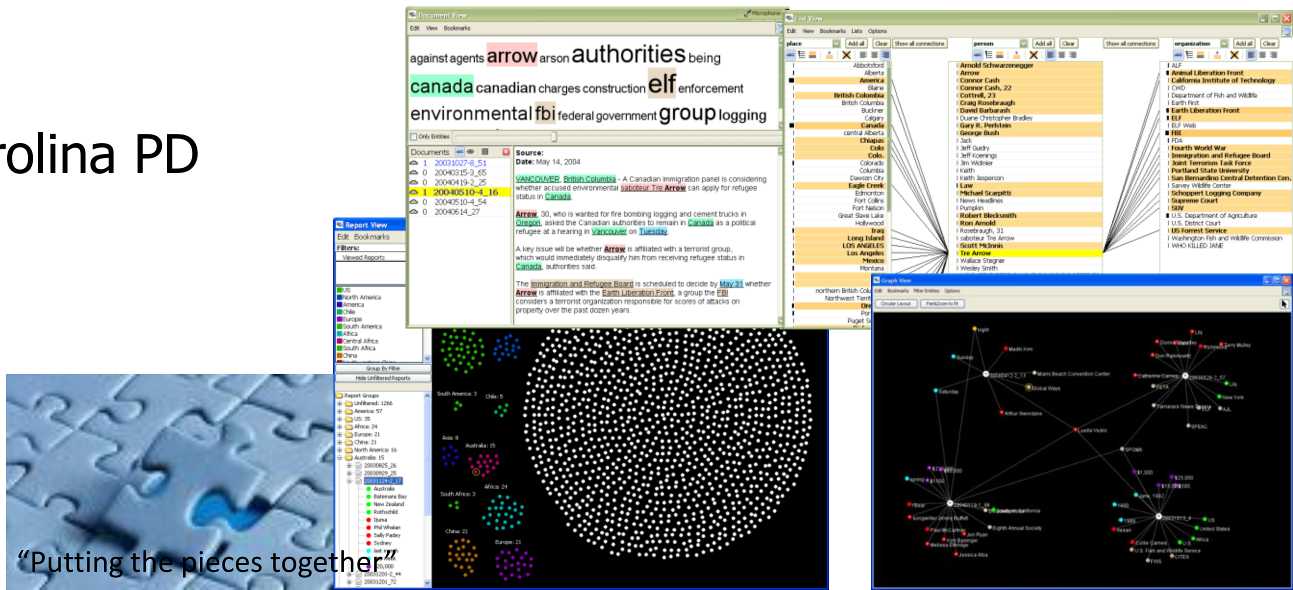
Leave Blank

<u>Goal</u>: Assist investigators with understanding, sense-making, and analysis of large, unstructured and structured document collections

<u>Approach</u>: Provide multiple visual perspectives on the documents and entities within them, highlighting connections between entities

Users: Over 150 downloads – active users include the following:

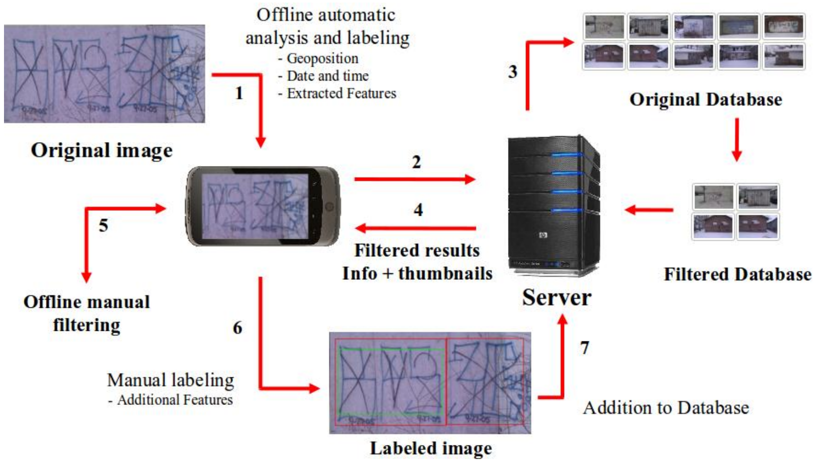
- Seattle PD
- Rock Hill, South Carolina PD
- Indianapolis PD
- West Lafayette, PD (student embedded)
- TSA



“Putting the pieces together”

Gang Graffiti Recognition and Analysis Using a Mobile Telephone (GARI)

Leave Blank



GARI delivered:

- Summer 2011:
  - IMPD gang detectives
- August 2011:
  - IMPD at large
  - Ind Fusion Gang Task Force
- October 2011:
  - Gang detectives across Indiana

IMPACT:

- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be “interpreted”
- More than 40 users and 450 graffiti images acquired

VACCINE

ASU

Crime Analytics & SA Issues and Techniques

Leave Blank

- Fuse data from a variety of sources
  - Law enforcement records management
  - Weather and phases of the moon
  - Street light locations, bus routes
  - Tracking release data of offenders
  - Civil court data
  - Social Media
  - Local event calendar
- Reliable predictive models
- Understandability and trust of predictions
- Main Question: What helps officers, detectives, chief do their variety of jobs?

Data, Models, Reductions, and Visual Representations

Leave Blank

Data Models

Leave Blank

- Data models are structured forms suitable for computer-based transformations
- These structures exist in the original data or are derivable from the original data
- Structures retain the information and knowledge content and the related context within the original data
- These structures are transformable into lower-dimensional representations for visualization and analysis

J. J. Thomas and K. A. Cook (Ed.). *Illuminating the Path: The R&D Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.

Data Models vs. Conceptual Models

Leave Blank

- Data models are mathematical abstractions
  - We can perform numerical operations, addition, subtraction, etc.
- Conceptual models are our mental constructs
  - These contain semantic structure and support reasoning
  - Think of giving directions to someone using landmarks

Data Types

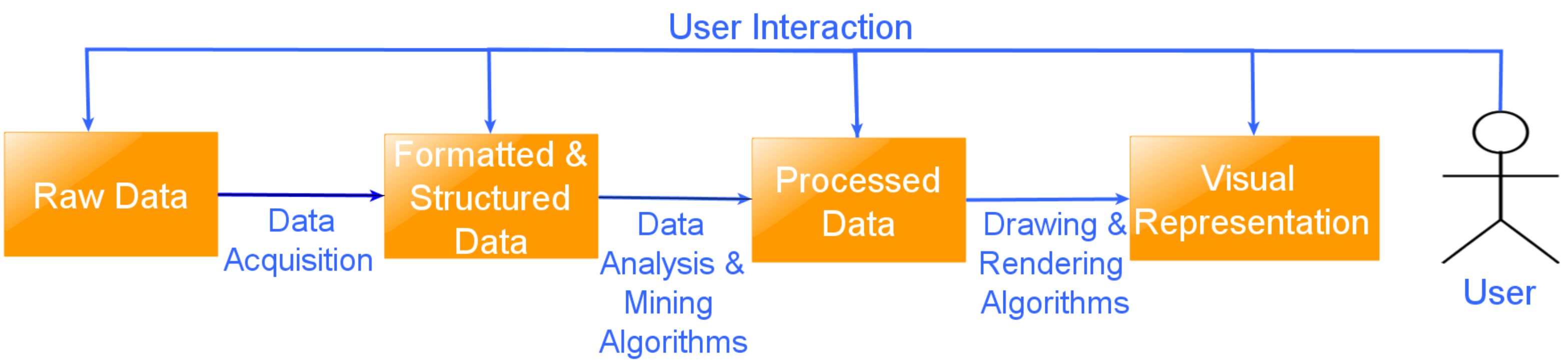
Leave Blank

- **Nominal**
  - Data whose categories have no implied ordering
  - Examples include political affiliations of a population
- **Ordinal**
  - Data that has a specified order, but no specified distance metric
  - Examples include beverage sizes at McDonalds (Small, medium, large)
- **Interval**
  - Data that has measurable distances
  - Examples include periods of time (second, minute, etc.) – the zero point is arbitrary
- **Ratio**
  - Same as interval, but include a zero point
  - Example include Celsius scale, height above sea level

1- SS Stevens, "On the Theory of Scales of Measurement," *Science*, 103(2684):677-680, 1946.

The Visual Analytics Pipeline

Leave Blank



- We want to take these different data types and map them to an appropriate visual representation
- **Data Analysis** – data are prepared for visualization (smooth, interpolate, transform)
- **Filtering** – A subset of the data (usually user defined) is selected for visualization
- **Mapping** – Data are mapped to geometric primitives and their attributes
- **Rendering** – Geometric data are transformed to image data

Image taken from http://www.pafkiet.edu.pk/Default.aspx?tabid=693

Haber, R. & McNabb, D. A. (1990). Visualization idioms: A conceptual model for scientific visualization systems. In Visualization n Scientific Computing.

VACCINE

Visual Representations

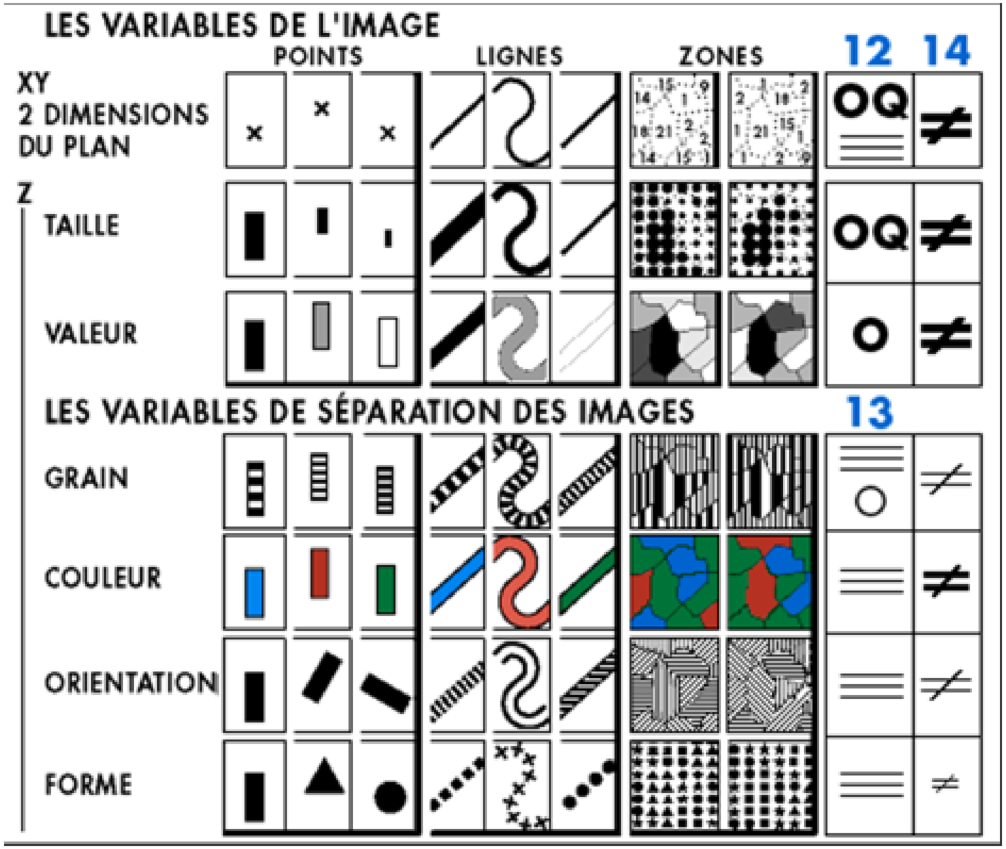
Leave Blank

- Visual representations translate data into a visible form, highlighting important features, such as commonalities and anomalies
- Visual representations make it easy for users to perceive salient aspects of their data quickly
- These visual representations augment the cognitive reasoning process with perceptual reasoning which enhances the analytical reasoning process

Bertin's Visual Variables

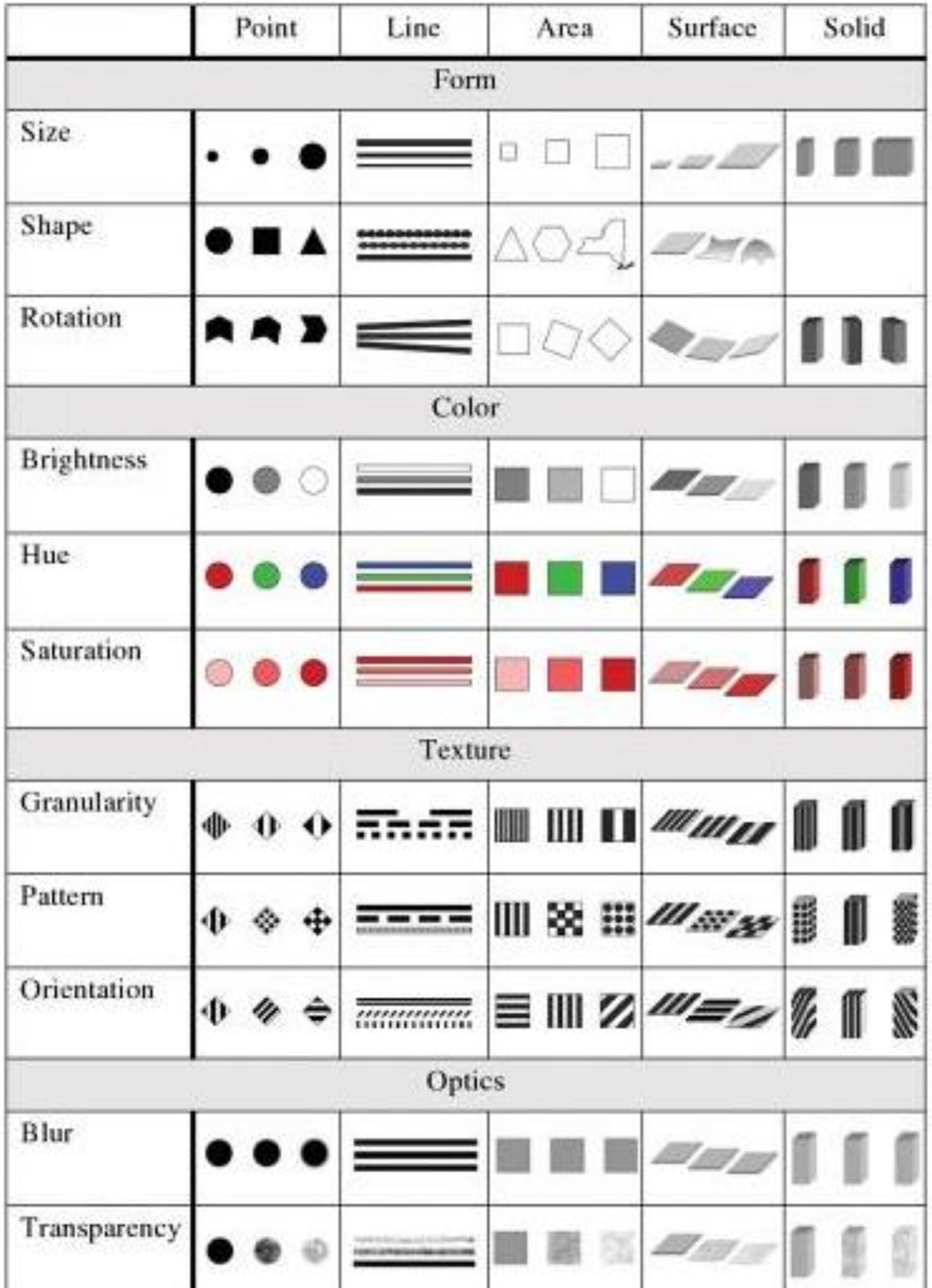
Leave Blank

- Visualization is concerned primarily with a mapping to visual form
- [x,y]
  - Position
- [z]
  - Size (Taille)
  - Value (Valeur)
  - Color (Couleur)
  - Texture (Grain)
  - Orientation
  - Shape (Forme)



J Bertin (1967), *The Semiology of Graphics*

VACCINE



Leave Blank

L Wilkinson (2005) *The Grammar of Graphics*, p. 317

How Many Variables to Use?

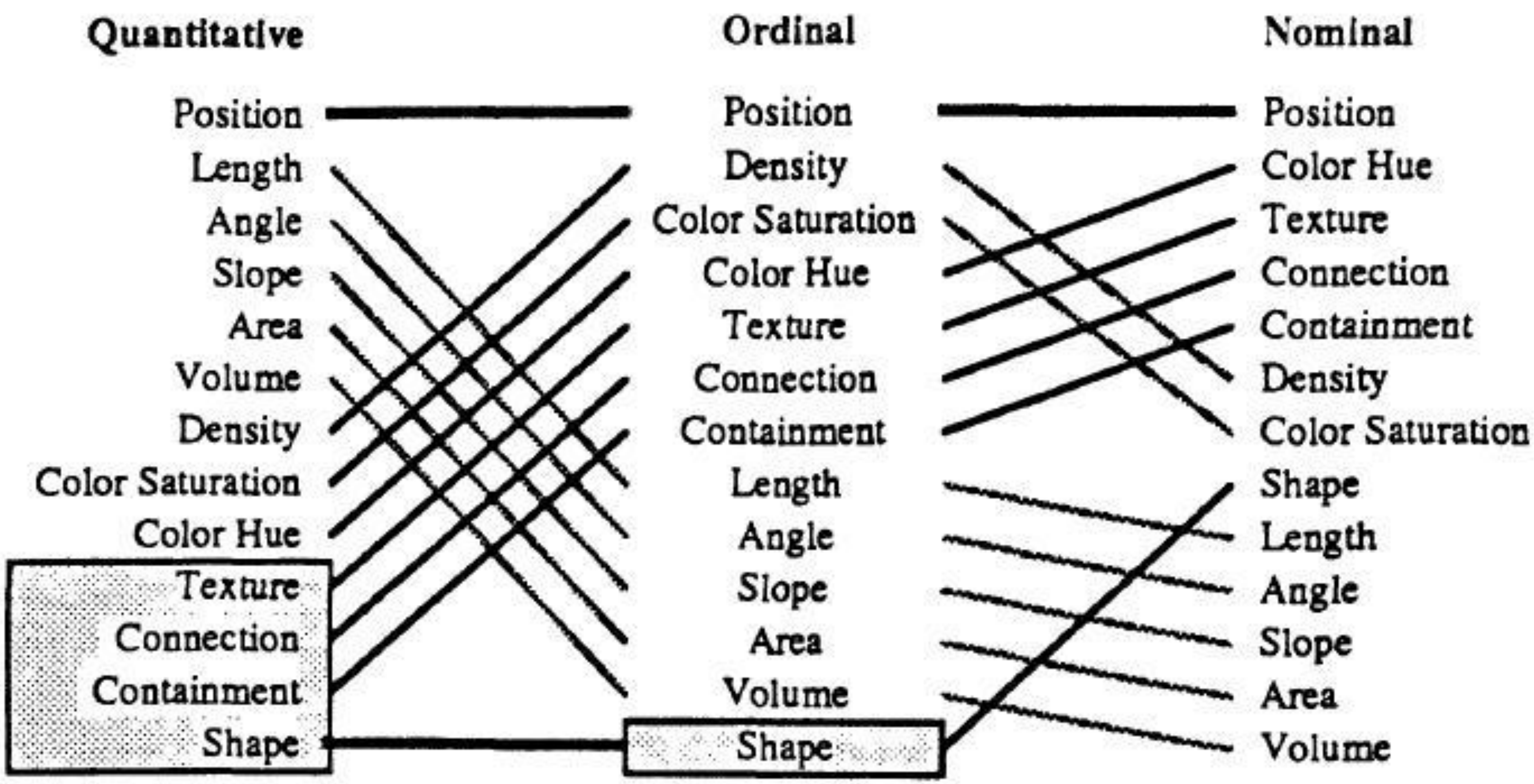
Leave Blank

- The number of visual variables necessary for the representation is at least equal to the number of components in the information
- With three components, the information can be perceived as a single image
- Otherwise, we need several images (this is often termed *small multiples* and will be discussed in a future lecture)
- The number of components is the best basis for a classification of graphic constructions

J Bertin (1967), *The Semiology of Graphics*

How Do I Choose the Encoding?

Leave Blank



J Mackinlay. Automating the Design of Graphical Presentations of Relational Information, ACM TOG 5:2, 1986.

VACCINE

Expressing Data With Color

Leave Blank

- Results from research on visual attention can be used to assign visual features to data values
- One of the key components of visually representing data is choosing the appropriate color scale
- There is no "best" color scale
- Choice depends on[1,2]
  - data type
  - problem domain
  - visual representation
  - Questions the analyst is asking of the data
- While there is no "best" choice, there are design principles

1 - H. Levkowitz and G. T. Herman, "Color scales for image data," IEEE Computer Graphics and Applications, vol. 12, pp. 72-80, 1992
2 - P. Rheingans, "Task-based color scale design," Proceedings of Applied Image and Pattern Recognition, pp. 35-43, 1999.

Design Principles for Color Schemes

Leave Blank

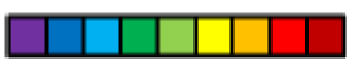
- *Order*[1] – Given a univariate data type, the color scale that is chosen to map the data must represent a perceived ordering
- *Separation*[2] – Important differences between ranges of the variable should be represented by colors that can be perceived as being different
  - Not only should they be perceived as different, but also equal
- *Aesthetics*[3] – color map should be aesthetically pleasing, contain a maximum perceptual resolution, and ordering should be intuitive

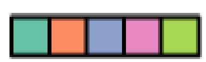
1 - B. E. Trumbo, “Theory for Coloring Bivariate Statistical Maps," *The American Statistician*, vol. 35, no. 4, pp. 220-226, 1981.
2 - H. Levkowitz and G. T. Herman, "Color scales for image data," IEEE Computer Graphics and Applications, vol. 12, pp. 72-80, 1992
3 - K. Moreland, “Diverging Color Maps for Scientific Visualization," *Proceedings of the 5th International Symposium on Visual Computing*, December 2009.

Univariate Color Schemes

Leave Blank

- **Qualitative scheme**
  - Rainbow color scale is one of the most commonly used, but it is a poor color map in a large variety of domain problems
  - Ordering of the hues is unintuitive
  - Nominal data types can use this scale as no ordering is implied


Rainbow

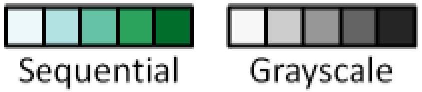

Qualitative

1 - D. Borland and R. M. Taylor, “Rainbow Color Map (Still) Considered Harmful,“ *Computer Graphics & Applications*, vol. 27, no. 2, pp. 14-17, 2007.
2 -P. Rheingans, “Task-based color scale design," *Proceedings of Applied Image and Pattern Recognition*, pp. 35-43, 1999.
3 - M. A. Harrower and C. A. Brewer, “ColorBrewer.org: An online tool for selecting color schemes for maps," *The Catrtographic Journal*, vol. 40, no. 1, pp. 27-37, 2003.

Univariate Color Schemes

Leave Blank

- **Sequential color scheme**
  - Simplest is the gray scale map where variable is mapped to brightness
  - Sequential maps represent ordered data
  - Dark colors typically represent high ranges, bright, low
  - Benefits are that the scale is intuitive
  - Weakness is that limited number of distinguishable colors can be represented

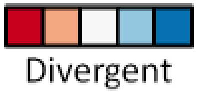


1 - D. Borland and R. M. Taylor, “Rainbow Color Map (Still) Considered Harmful,“ *Computer Graphics & Applications*, vol. 27, no. 2, pp. 14-17, 2007.
2 -P. Rheingans, “Task-based color scale design," *Proceedings of Applied Image and Pattern Recognition*, pp. 35-43, 1999.
3 - M. A. Harrower and C. A. Brewer, “ColorBrewer.org: An online tool for selecting color schemes for maps," *The Catrtographic Journal*, vol. 40, no. 1, pp. 27-37, 2003.

Univariate Color Schemes

Leave Blank

- **Divergent color scheme**



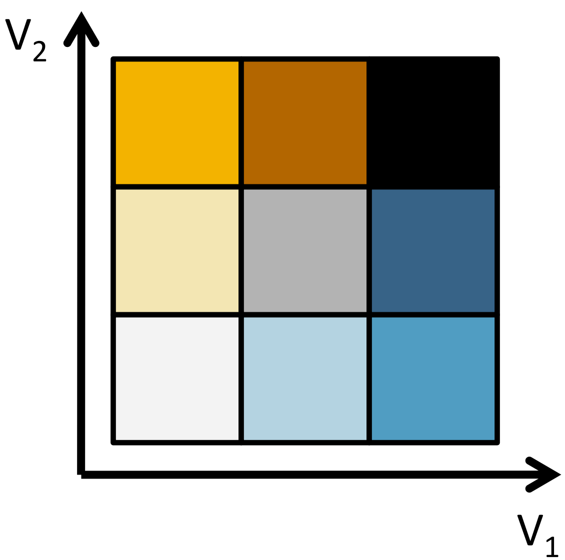
  - Provides means for variable comparisons
  - Best suited for ratio data where there is some meaningful zero point
  - Scale lacks a natural ordering of colors
  - Careful choices must be made in choosing high and low ends
  - Can use concept of cool (blues) and warm (reds and yellow) colors

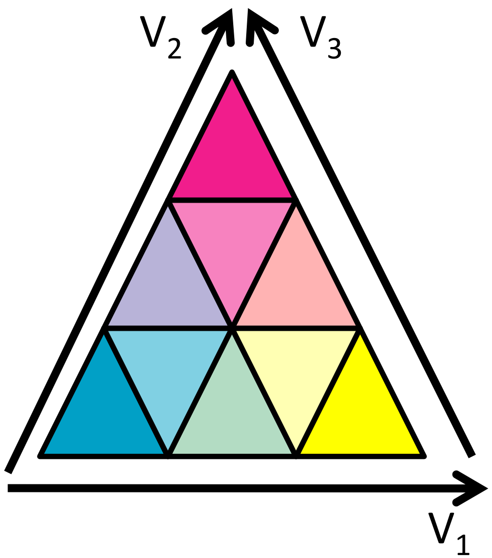
1 - C. Hardin and L. Maffi, *Color Categories in Thought and Language*, Cambridge University Press, 1997.
2 -P. Rheingans, "Task-based color scale design," *Proceedings of Applied Image and Pattern Recognition*, pp. 35-43, 1999.
3 - M. A. Harrower and C. A. Brewer, "ColorBrewer.org: An online tool for selecting color schemes for maps," *The Catrtographic Journal*, vol. 40, no. 1, pp. 27-37, 2003.

VACCINE

Multivariate Color Schemes

Leave Blank





M. A. Harrower and C. A. Brewer, “ColorBrewer.org: An online tool for selecting color schemes for maps," *The Catrtographic Journal*, vol. 40, no. 1, pp. 27-37, 2003.

VACCINE

Leave Blank

Cleveland's Heirarchy[1]

- Cleveland evaluated elements when isolated
- Tasks were restricted to magnitude and ratio comparisons
- Research indicates this hierarchy may be best in pre-attentive stages or when focusing only on portions of a graphic[2,3,4]

Best

1 - Position along a common scale

2 – Position along nonaligned scales

3 – Length

4 – Angle/Slope

5 – Area

6 – Volume

7 - Color

Worst

1 – WS Cleveland, *The Elements of Graphing Data*, 1985
2 – Spence, I., and Lewandowsky, S. (1991). Displaying proportions and percentages. *Applied Cognitive Psychology* 5, 61-77
3 – Simkin, D., and Hastie, R. (1987). An information processing analysis of graph perception. *Journal of the American Statistical Association*, 82, 454-465
4 – Carswell, C. M. (1992). Choosing specifiers: An evaluation of the basic tasks model of graphical perception. *Human Factors*, 4, 535-554

What if I Combine Encodings?

Leave Blank

- So, what happens when we combine several scales in a single display?
- Can we represent one quantitative dimension with color and another with orientation and expect a perceiver to respond to both dimensions?
- Do these things make psychological sense?

L.Wilkinson (2005) *The Grammar of Graphics*

Integral Versus Separable Dimensions

Leave Blank

- "A configuration has properties that have to be expressed as some form of interaction or interrelation between the components, be they features or dimensions."[1]
- *Integral dimensions* are not as easily decomposable by perceivers as *separable dimensions*
- Separating hue from brightness in a color (integral dimensions) are harder to decompose than say size and texture (separable dimensions)

1 – Garner, W. R. (1981). The analysis of unanalyzed perceptions. In M. Kubovy and J. R. Pomerantz (Eds.), *Perceptual Organization* (pp. 119-139). Hillsdale, NJ: Lawrence Erlbaum.

Combinatorics of Encodings

Leave Blank

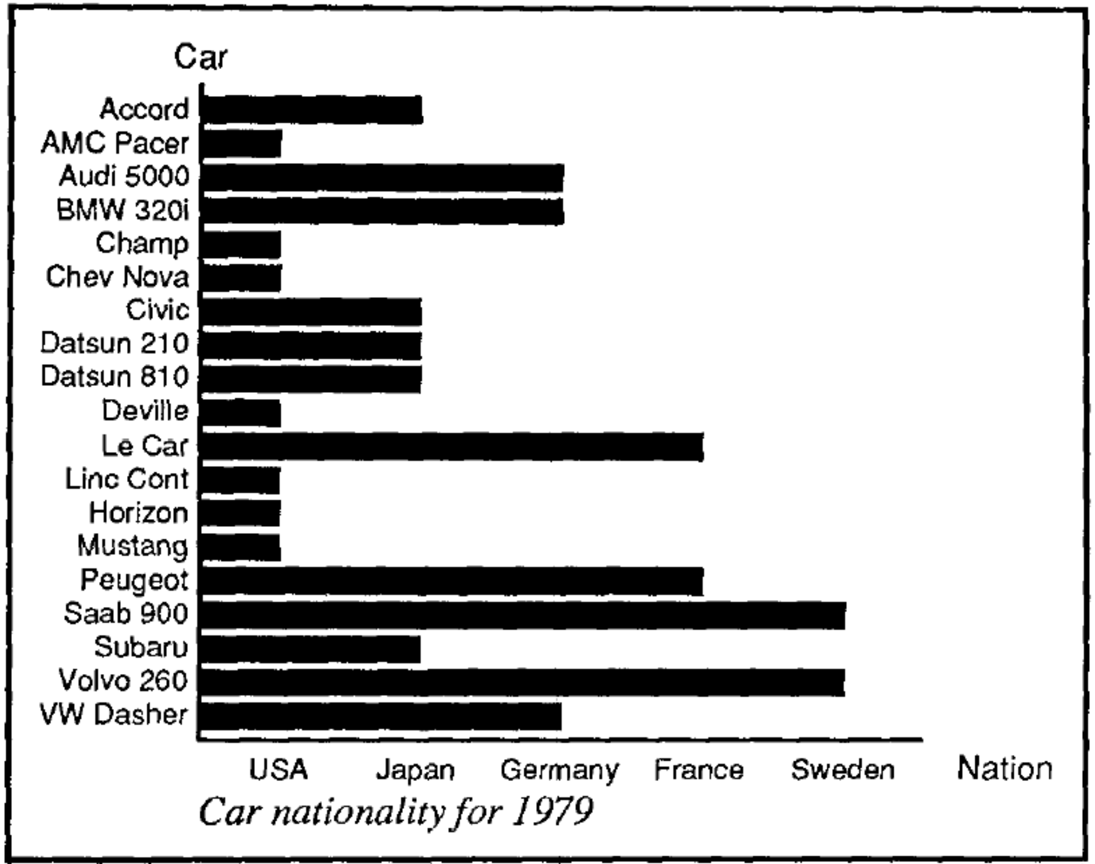
- Challenge:
  - Pick the best encodings from the exponential number of possibilities $(n+1)^8$

- Principle of Consistency: The properties of the image should match the properties of the data

- Principle of Importance Ordering: Encode the most important information in the most effective way

1 – This slide is borrowed from Pat Hanrahan's "From Data to Image" lecture: http://graphics.stanford.edu/courses/cs448b-04-winter/lectures/encoding/

Leave Blank

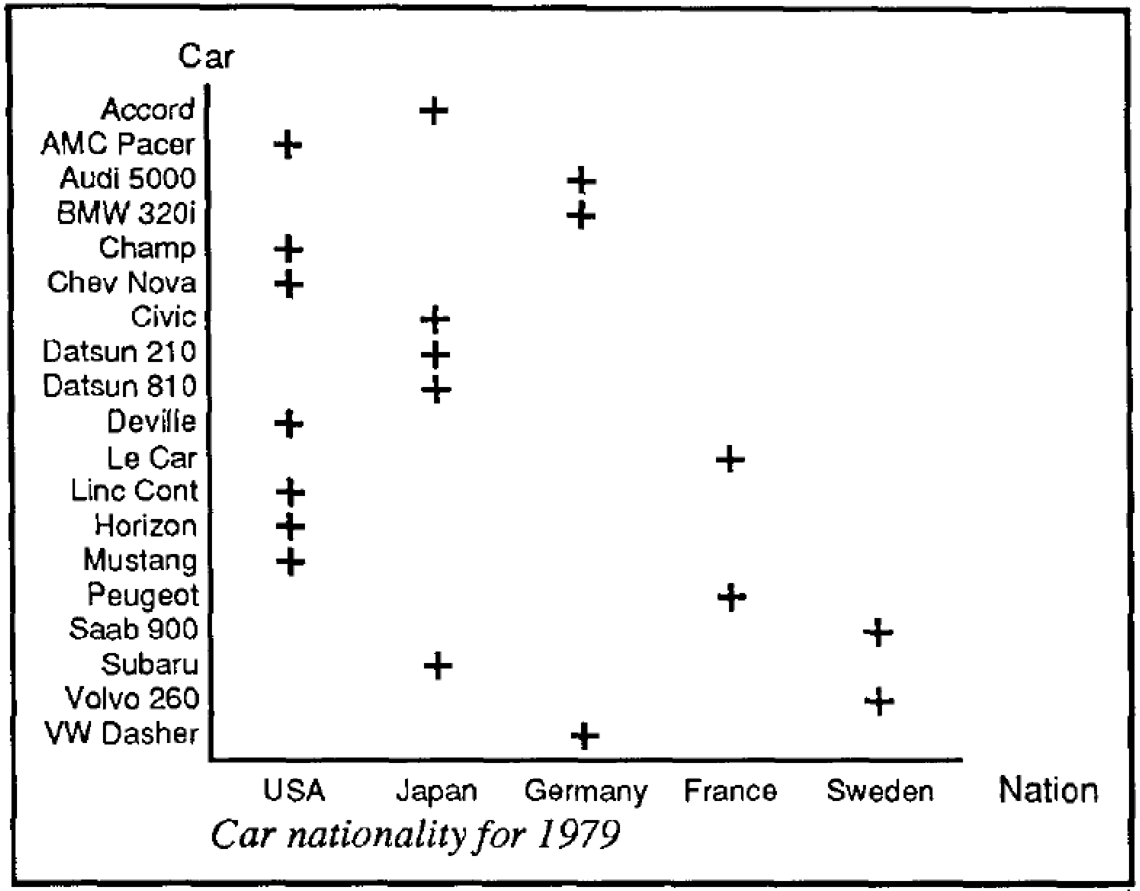


Car nationality for 1979

apt

1 – J. Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM Transactions on Graphics, 5(2): 110-141, 1986.

Leave Blank



Car nationality for 1979

apt

1 – J. Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM Transactions on Graphics, 5(2): 110-141, 1986.

The Curse of Dimensionality

Leave Blank

- A term coined by Bellman in 1961
- Refers to the problems associated with multivariate data analysis as the dimensionality increases the available data becomes sparse
- Sparsity is a problem for any method that requires statistical significance
- Sometimes data dimensions are redundant and can be reduced with minimal information loss
- In visualization we are also limited with screen space and the number of available visual variables, so choosing the most appropriate dimensions is key

Richard Ernest Bellman; Rand Corporation (1957). *Dynamic programming*. Princeton University Press.

The Curse of Dimensionality

Leave Blank

- So, what can we do?
  - We can incorporate prior knowledge of the data
  - We can smooth the target function
  - We can reduce the dimensionality

The Practicality of the Curse of Dimensionality

Leave Blank

- For a given sample size, there is a maximum number of features above which the performance of classifying samples will degrade rather than improve
- In most cases, the additional information that is lost when discarding some features is compensated by a more accurate mapping in the lower-dimensional space
- So, how do we know what features we can throw away?
- For visualization this implies that there are some features of a dataset that will be better to visualize (contain more information) than others!

Dimensional Reduction

Leave Blank

- Two approaches are available to reduce dimensionality
  - **Feature extraction:** creating a subset of new features by combinations of the existing features
  - **Feature Selection:** choosing a subset of all the features

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \rightarrow \begin{bmatrix} x_{i_1} \\ x_{i_2} \\ \vdots \\ x_{i_M} \end{bmatrix} \qquad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \rightarrow \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = f\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}\right)$$

  - Given a feature space $x_i \in R^N$ find a mapping $y = f(x): R^N \rightarrow R^M$ with M<N such that the transformed feature vector $y \in R^M$ preserves (most of) the information or structure in R^N
  - An optimal mapping is one that does not increase error

Principle Components Analysis

Leave Blank

- One of the most commonly applied dimension reduction techniques
- PCA is a deterministic analytical procedure that utilizes an orthogonal transformation to reduce a set of sample observations with potentially correlated variables into a set of uncorrelated variables called principal components
- The number of principal components will always be less than or equal to the original number of variables in the sample set

. T. Jolliffe, \Principal Component Analysis," *Springer Series in Statistics*, Springer, NY, 2002.

Principle Component Analysis

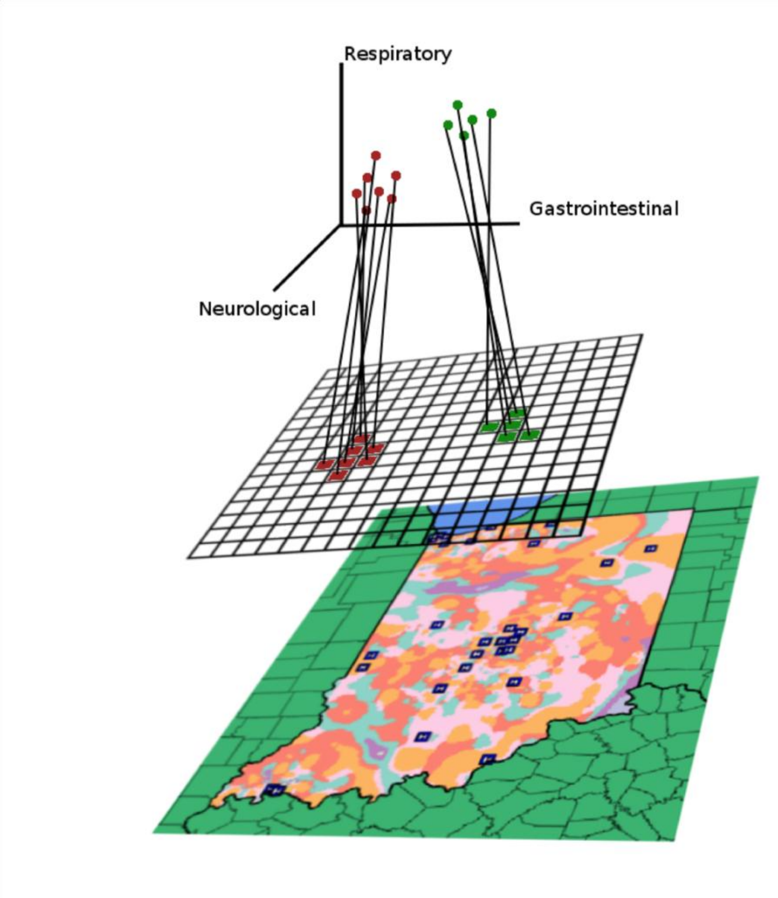
Leave Blank

- The main limitation of PCA is that it does not consider class separability since it does not take into account the class label of the feature vector
  - PCA simple performs a coordinate rotation that aligns the transformed axes with the directions of maximum variance
  - There is no guarantee that the directions of maximum variance will contain the most interesting/important features

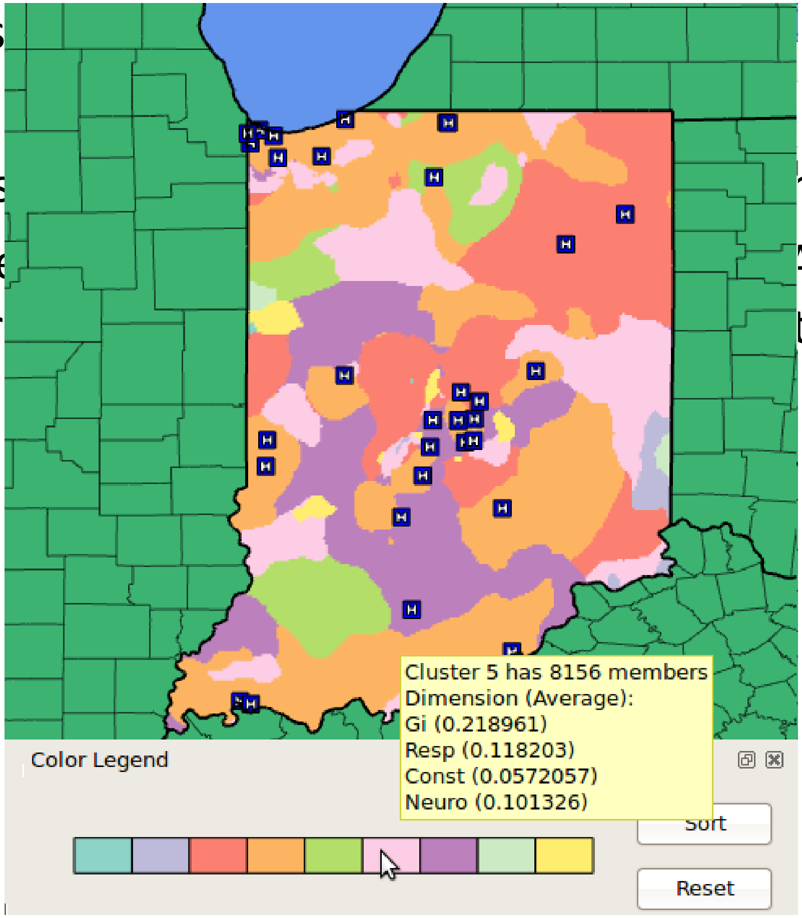
K-Means Clustering

Leave Blank

- K-means clustering
  – Goal is to partition n-observations into k-clusters where each observation



neares
w the
lian dis
s cluste
pal dir



hed
A
troid

VACCINE

ASU

Visual Analytics

Leave Blank

- We want to know how to best create images of data effectively in order to facilitate advanced data analysis
  - Spotting outliers
  - Discriminating clusters
  - Checking distributional and other assumptions
  - Examining relationships
  - Comparing mean differences
  - Observing time-based processing

Representation

Leave Blank

- What are the two main ways of presenting multivariate data sets?
  - Directly (textually) – Tables
  - Symbolically (pictures) – Graphs

- How do we decide which to use, and when?

1 – Descriptions on this slide are borrowed from John Stasko's "Multivariate Data & Tables and Graphs" lecture: http://www.cc.gatech.edu/~stasko/7450/Notes/data.pdf

Tables?

Leave Blank

- Use tables when
  - The document will be used to look individual value
  - The document will be used to compare individual values
  - Precise values are required
  - The quantitative info to be communicated involves more than one unit of measure
- Use graphs when
  - The message is contained in the shape of the values
  - The document will be used to reveal relationships among values

1 – Descriptions on this slide are borrowed from John Stasko's "Multivariate Data & Tables and Graphs" lecture: http://www.cc.gatech.edu/~stasko/7450/Notes/data.pdf

2 - S. Few, Show Me the Numbers: Designing Tables and Graphs to Enlighten,

Graphs?

Leave Blank

- Graph
  - Visual display that illustrates one or more relationships among entities
  - Shorthand way to present information
  - Allows a trend, pattern or comparison to be easily comprehended
- Critical to remain task-centric
  - Why do you need a graph?
  - What questions are being answered?
  - What data is needed to answer those questions?
  - Who is the audience?

1 – Descriptions on this slide are borrowed from John Stasko's "Multivariate Data & Tables and Graphs" lecture: http://www.cc.gatech.edu/~stasko/7450/Notes/data.pdf

Histograms

Leave Blank

- Typical first look visualization method
  - Shows the shape of the data distribution
- The choice of the histogram bin width greatly impacts the resultant visualization
- There is no “best” number of bins, instead, different bin sizes can reveal different features of the data
- Some methods work to determine an optimal number of bins, but these methods make assumptions on the underlying data distribution

Histogram Binning

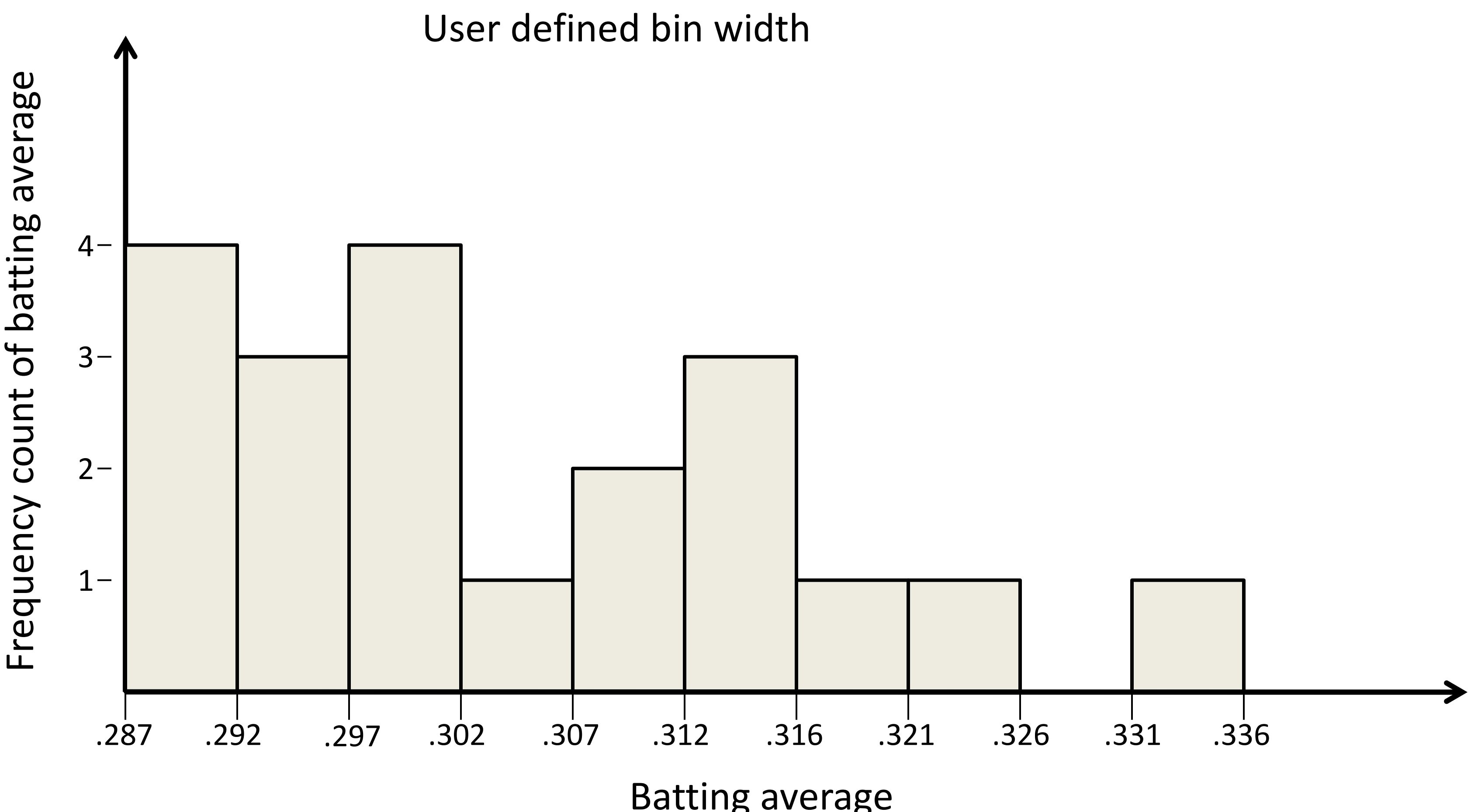
Leave Blank

- Number of bins (k) can be user specified or chosen from a suggested bin width (h) such that
- $k = \left\lceil \frac{\max x - \min x}{h} \right\rceil$
- Common choices for k include the square-root choice where $k = \sqrt{N}$
- Sturge's formula[1] where $k = \lceil \log N + 1 \rceil$
- Scott's choice[2] where $h = \frac{3.5\sigma}{N^{\frac{1}{3}}}$
- Freedman-Diaconis rule[3] $h = 2IQR(x)N^{-\frac{1}{3}}$

1 – H. A. Sturges. The choice of a class interval. *Journal of the American Statistical Association, p. 65-66, 1926*
2 - D. W. Scott. On optimal and data-based histograms. *Biometrika, 66(3):605-610, 1979.*
3 - Freedman, David; Diaconis, Persi, "On the histogram as a density estimator: L_2 theory". *Probability Theory and Related Fields* (Heidelberg: Springer Berlin) **57** (4): 453–476, 1981.
.

VACCINE ASU

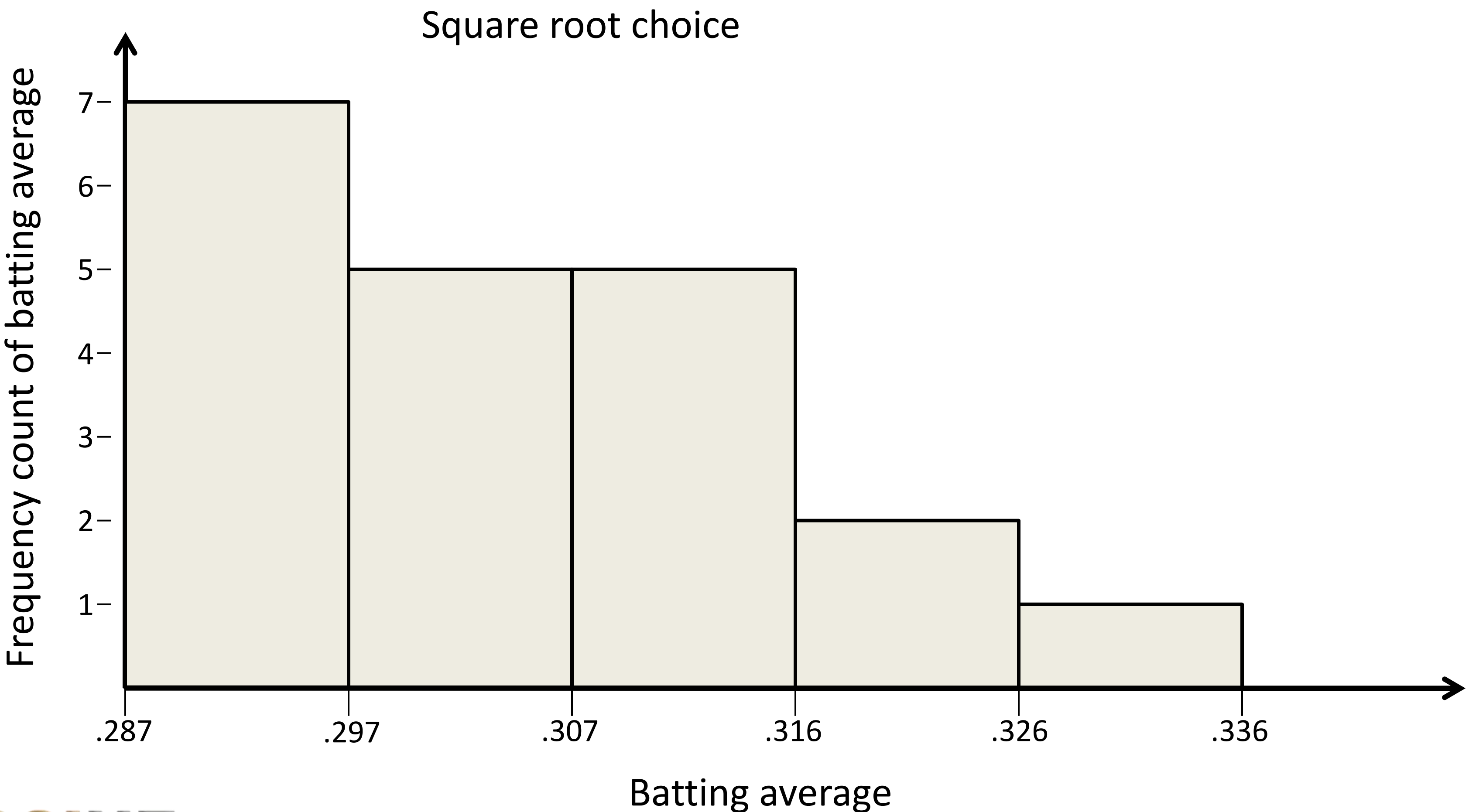
Leave Blank



VACCINE

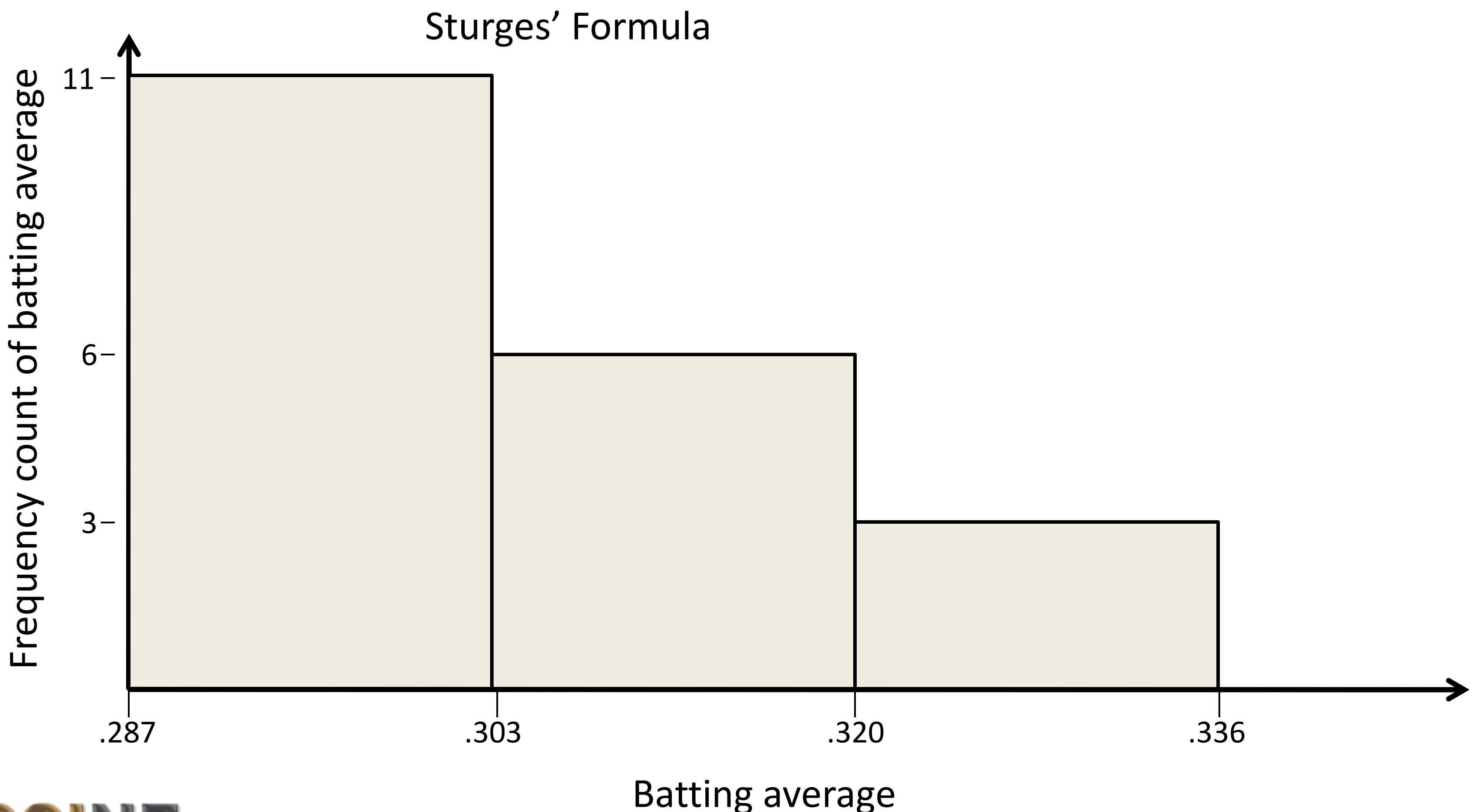
ASU

Leave Blank



Square root choice
Frequency count of batting average
7
6
5
4
3
2
1
.287
.297
.307
.316
.326
.336
Batting average

Leave Blank



VACCINE

ASU

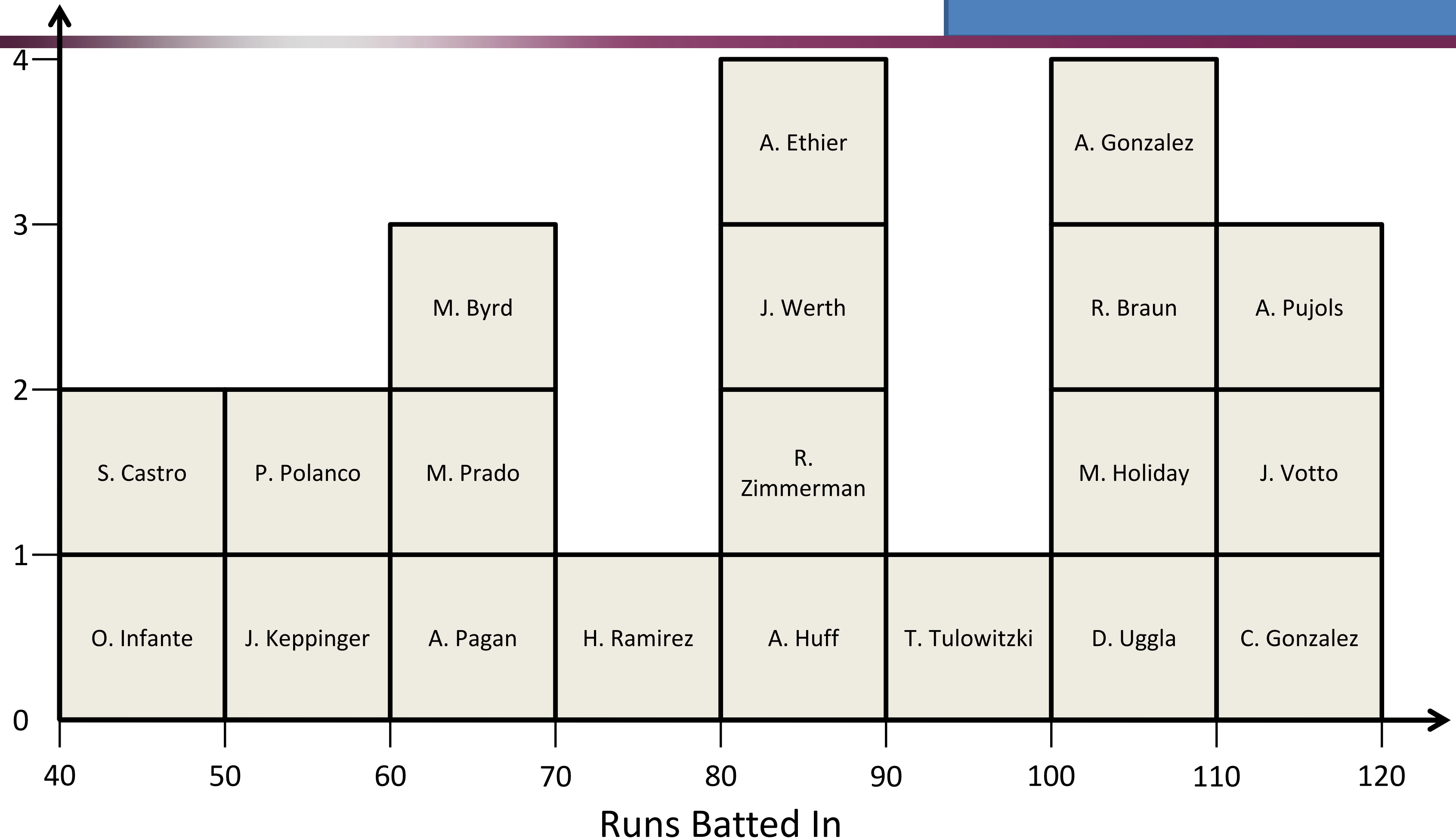
Histograms

Leave Blank

- Drawbacks
  - Density estimate depends on starting position of the bins
  - Discontinuities of the estimate are not due to the underlying density, only artifacts of the chosen bin location
  - As the number of dimensions grow, need many samples or else most bins would be empty
- Histograms are best suited for quick visualizations in 1 or 2 dimensions

Bivariate Case – Stacked Bar Graph

Leave Blank

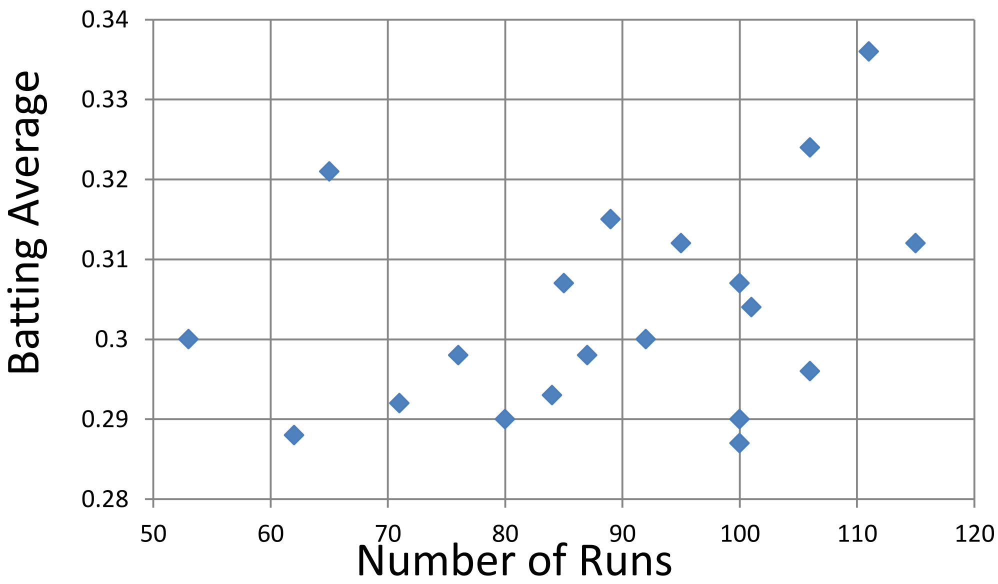


T. N. Dang, L. Wilkinson and A. Anand, "Stacking Graphic Elements to Avoid Over-Plotting," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1044-1052, 2010.

Bivariate Case - Scatterplot

Leave Blank

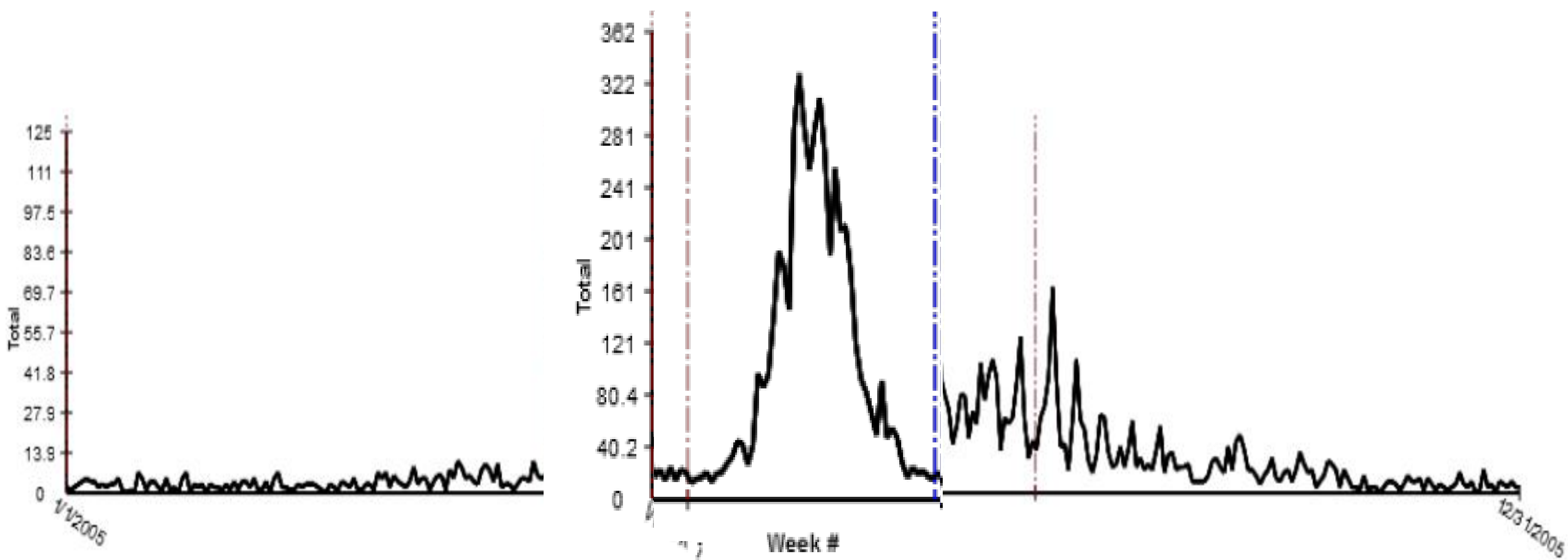
- Visualizes discrete data values along two axes
- Used as a means of analyzing bivariate relationships
- Quick means of assessing outliers, clusters and distributions
- Putting a line through the data can help assess trends, but can also mislead viewer



Time Series Visualization

Leave Blank

- We can think of a line-graph time series visualization as a form of a histogram where the x-axis bins the data by some increment of time
- Depending on the binning used, temporal data patterns can emerge



VACCINE Number of Coast Guard Search and Rescue Missions by Day Week

Non-Data Components of Graphs

Leave Blank

- Axes and legends can often be as important as the data themselves
- Poor axis choices and label choices can lead to confusing visualizations
- Axis tick labels provide cognitive context for most basic plot types
- They support estimation and contribute to the overall appearance of the graphic
- Cleveland suggests choosing the scales so that the data rectangle fills up as much of the scale-line rectangle as possible

W. S. Cleveland. The Elements of Graphing Data. Wadsworth Publ. Co., Belmont, CA, USA, 1985.

J. Talbot, S. Lin, P. Hanrahan, "**An Extension of Wilkinson's Algorithm for Positioning Tick Labels on Axes**," IEEE Transactions on Visualization and Computer Graphics 16(6): 1036-1043, 2010

Graph Aspect Ratios

Leave Blank

- Our ability to perceive trends and patterns in a given display is heavily influenced by the aspect ratio
- Aspect ratio affects densities, relative distances and orientations
- Several methods have been proposed for automatically selecting the aspect ratio
- Aspect ratio: a = width/height

J. Talbot, J. Gerth, P. Hanrahan, "Arc Length-based Aspect Ratio Selection," IEEE Transactions on Visualization and Computer Graphics 17(12): 2276-2282, 2011

Data Distributions

Leave Blank

- The type of data distribution affects not only the way the data should be analyzed, but also the way it should be visualized!
- Pre-conditioning data through transformations is a key step for appropriate analysis and visualization.

The Normal Distribution

Leave Blank

- The Normal (Gaussian) distribution has many features which make it popular

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

  – It can be fully characterized with two parameters
  – The probability of any value can be obtained by knowing how many standard deviations it is away from the mean
  – Many statistical measures and tests are well defined for this distribution
- Must be careful not to characterize non-Normal data as Normal

Skewness

Leave Blank

- Measure of the asymmetry of the probability distribution
- Can be positive, negative or undefined
- Negative skew indicates that the tail on the left side is longer than on the right
- Positive skew indicates that the tail on the right side is longer than on the left
- A zero value indicates the tails are relatively evenly distributed on both sides of the mean, but does not necessarily imply symmetry

Visualizing Skewed Data

Leave Blank

- Plotting skewed data compresses data values into smaller regions, resulting in lower visual fidelity[1]
- How to overcome skewness?
  - Remove data outliers
  - Interactive techniques (zoom, pan, brush)
- This can help for the visual analysis, but not always!
  - What if I want to compare data sets with different skews?
  - What if I want to perform statistical analyses?

1 - W. S. Cleveland. *Visualizing Data*. Hobart Press, 1993

The Power Transformation

Leave Blank

- We can transform data to a normal approximation by a power transformation[1]
- Such transformations reduce the effects of:
  - Skewness (transformations changes the range of data values and helps fit data onto the display)
  - Random noise (transformations can help show global trends)
  - Monotone spread

1 – J. W. Tukey. On the comparative anatomy of transformations. *Annals of Mathematical Statistics*, 28: 602-632, 1955.

The Power Transformation

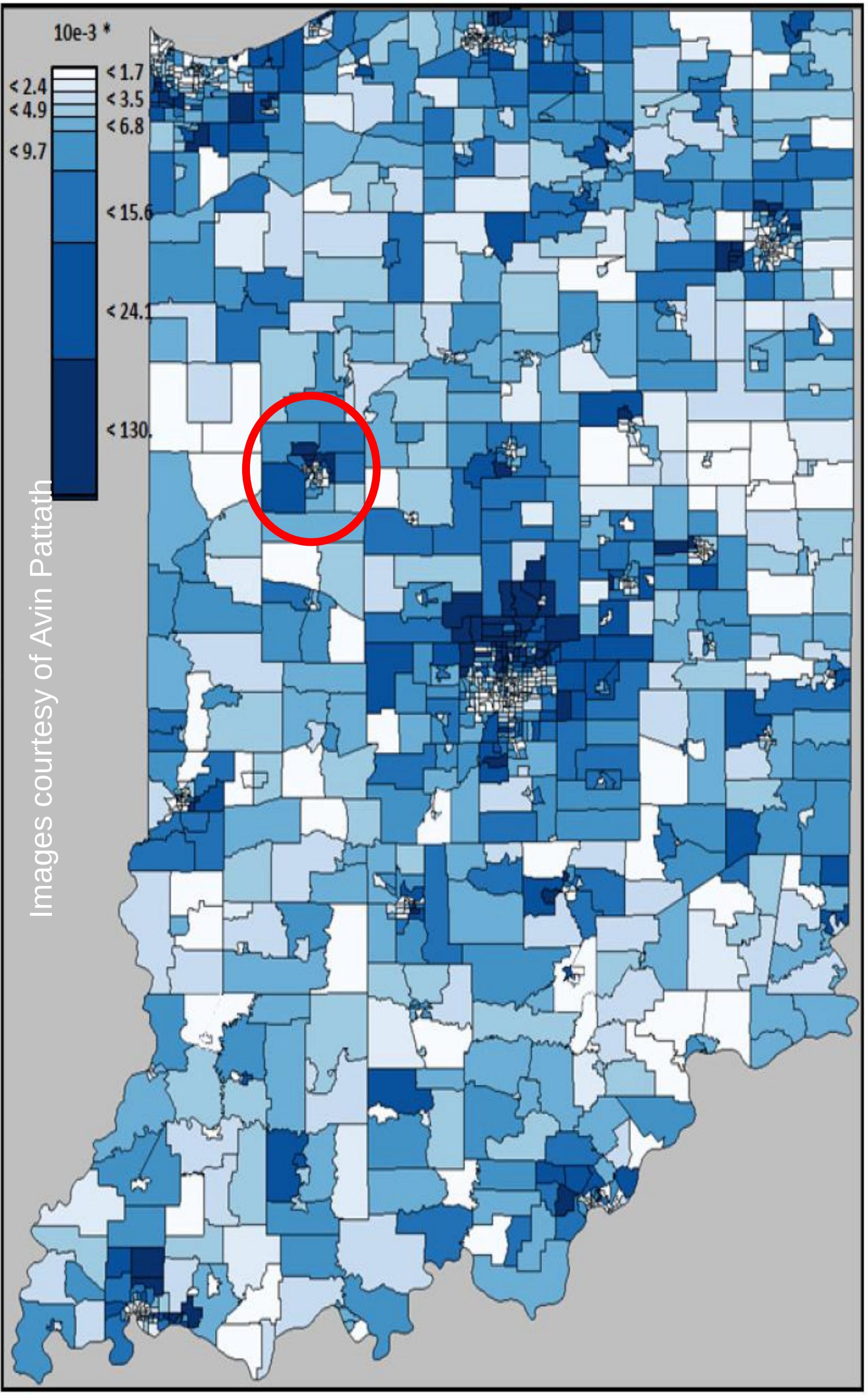
Leave Blank

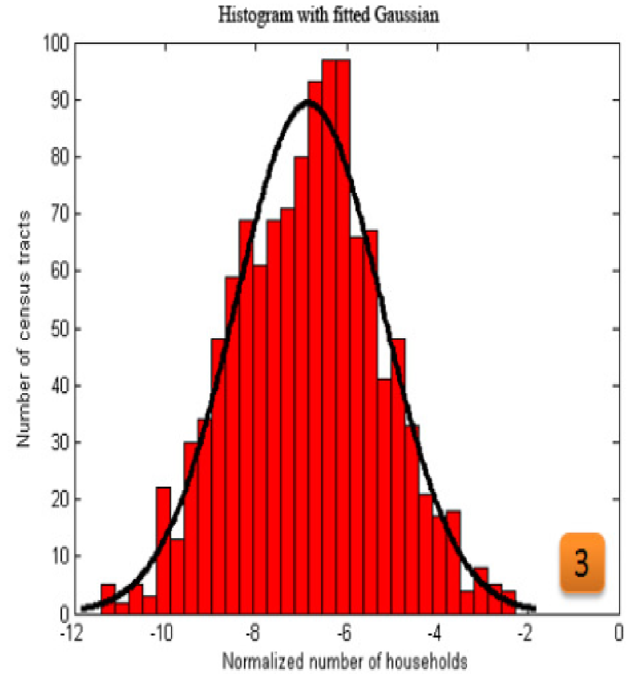
- The Box-Cox[1] Power Transformation
  - x is the observed (recorded) data and λ is the power

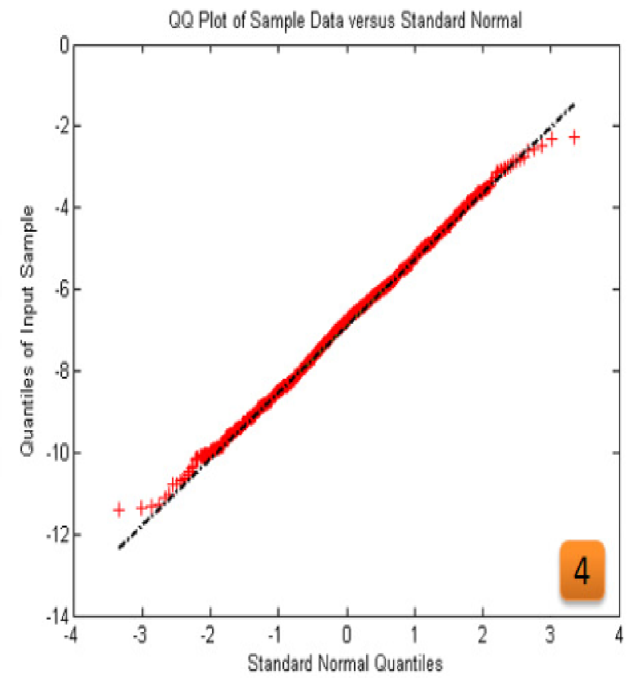
$$x(\lambda) = \frac{x^{\lambda} - 1}{\lambda}, (\lambda \neq 0)$$
$$= ln(x), (\lambda = 0)$$

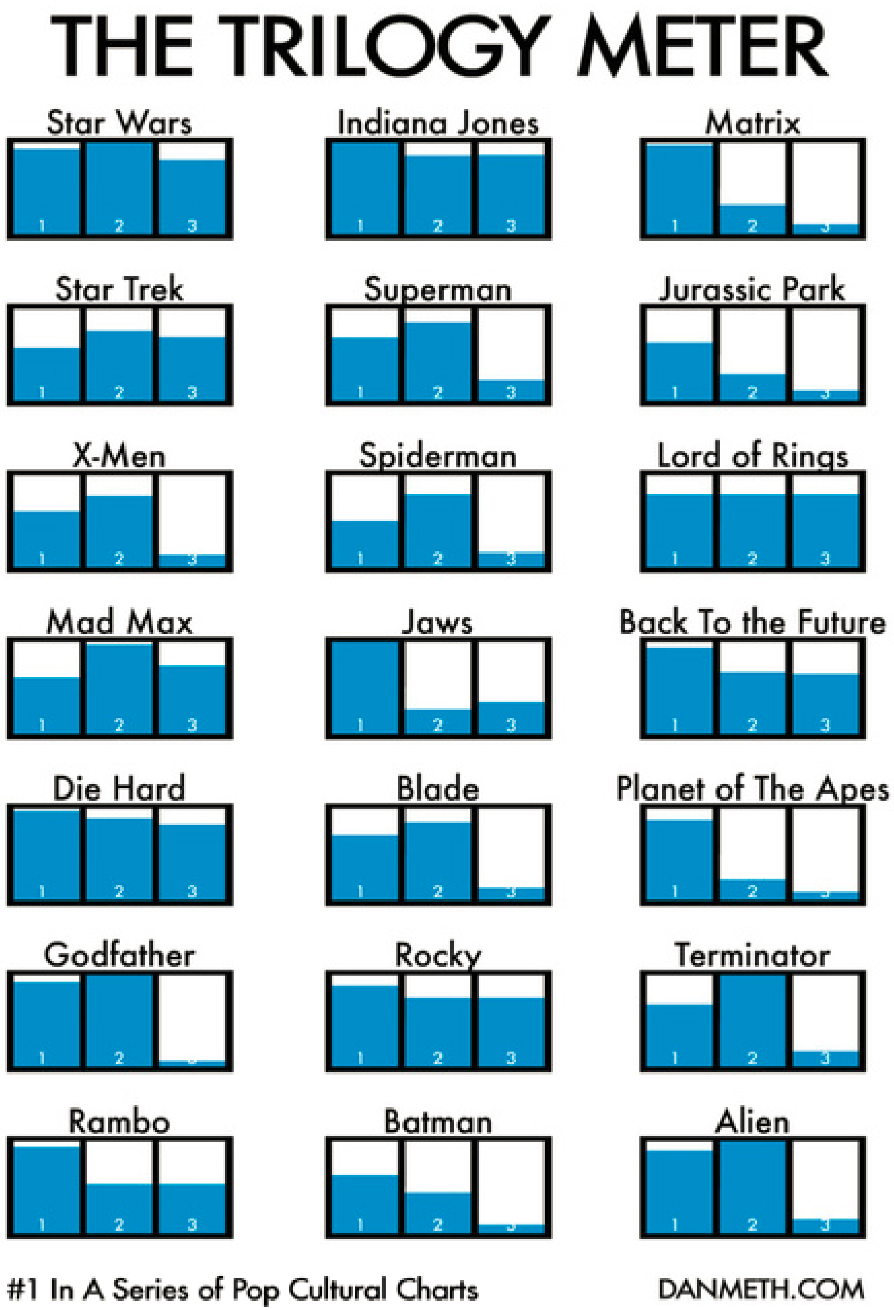
- The power (λ) is chosen by maximizing the logarithm of the likelihood function

1 – G. Box and D. Cox. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26(2):211-252, 1964.*


10e-3 *
< 1.7
< 2.4
< 3.5
< 4.9
< 6.8
< 9.7
< 15.6
< 24.1
< 130.
Images courtesy of Avin Pattath


Histogram with fitted Gaussian
Number of census tracts
Normalized number of households
3


QQ Plot of Sample Data versus Standard Normal
Quantiles of Input Sample
Standard Normal Quantiles
4



http://www.juiceanalytics.com/writing/better-know-visualization-small-multiples/

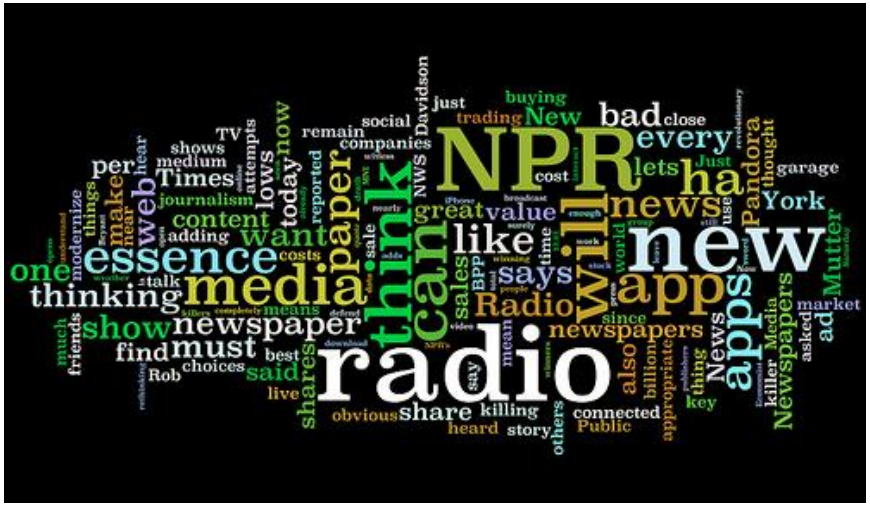
Leave Blank

Tag Clouds and Wordles

Leave Blank

- Visual representation for text data where words are placed and scaled based on some statistical measures
- Font size is typically determined by the number of instances a word is used

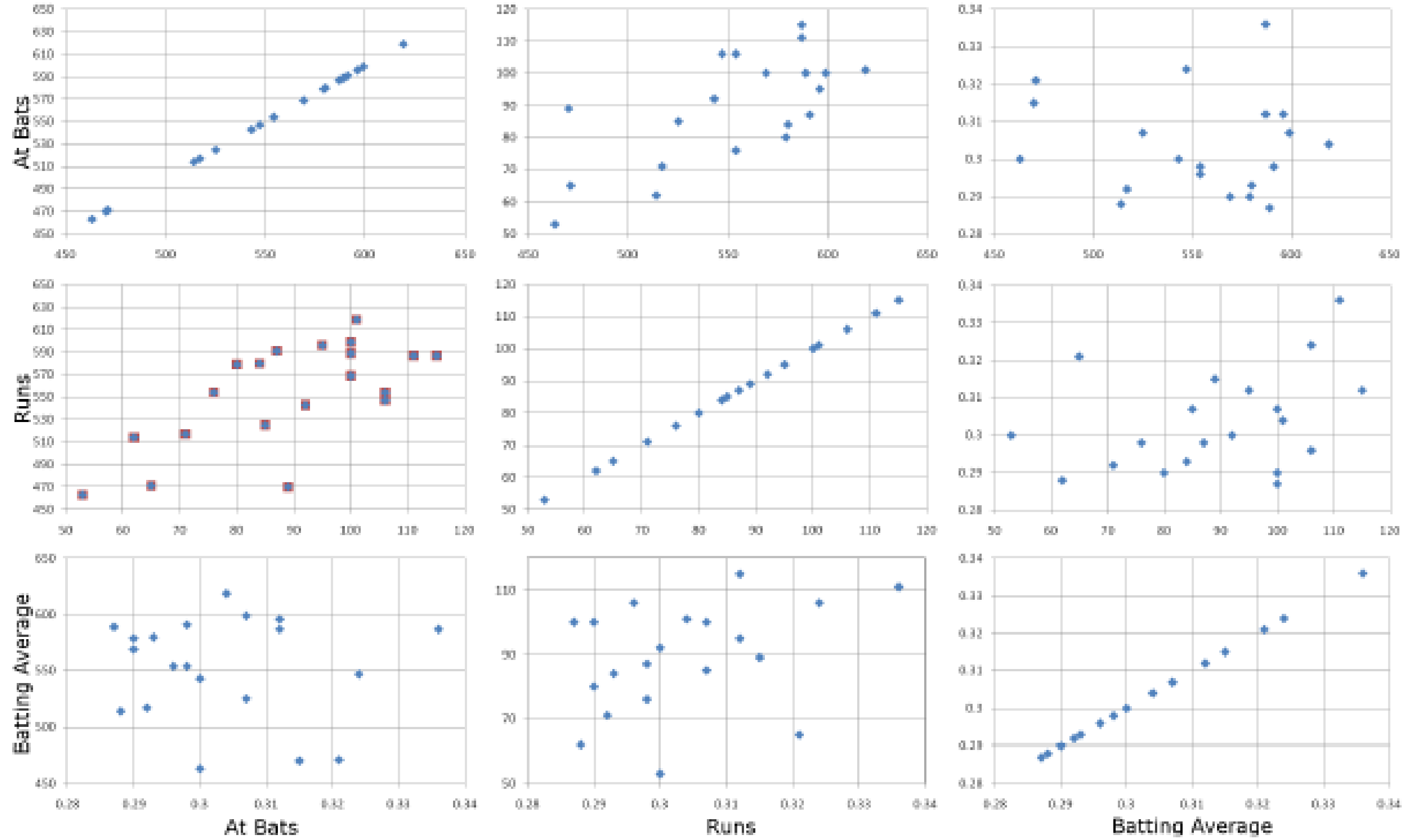
Text cloud comparing 2002 State of the Union Address by U.S. President Bush and 2011 State of the Union Address by President Obama. Created at TagCrowd.



Viegas, FB, Wattenberg, M, Feinberg, J, “Participatory Visualization with Wordle, IEEE Transactions on Visualization and Computer Graphics 15(6)1137-1144

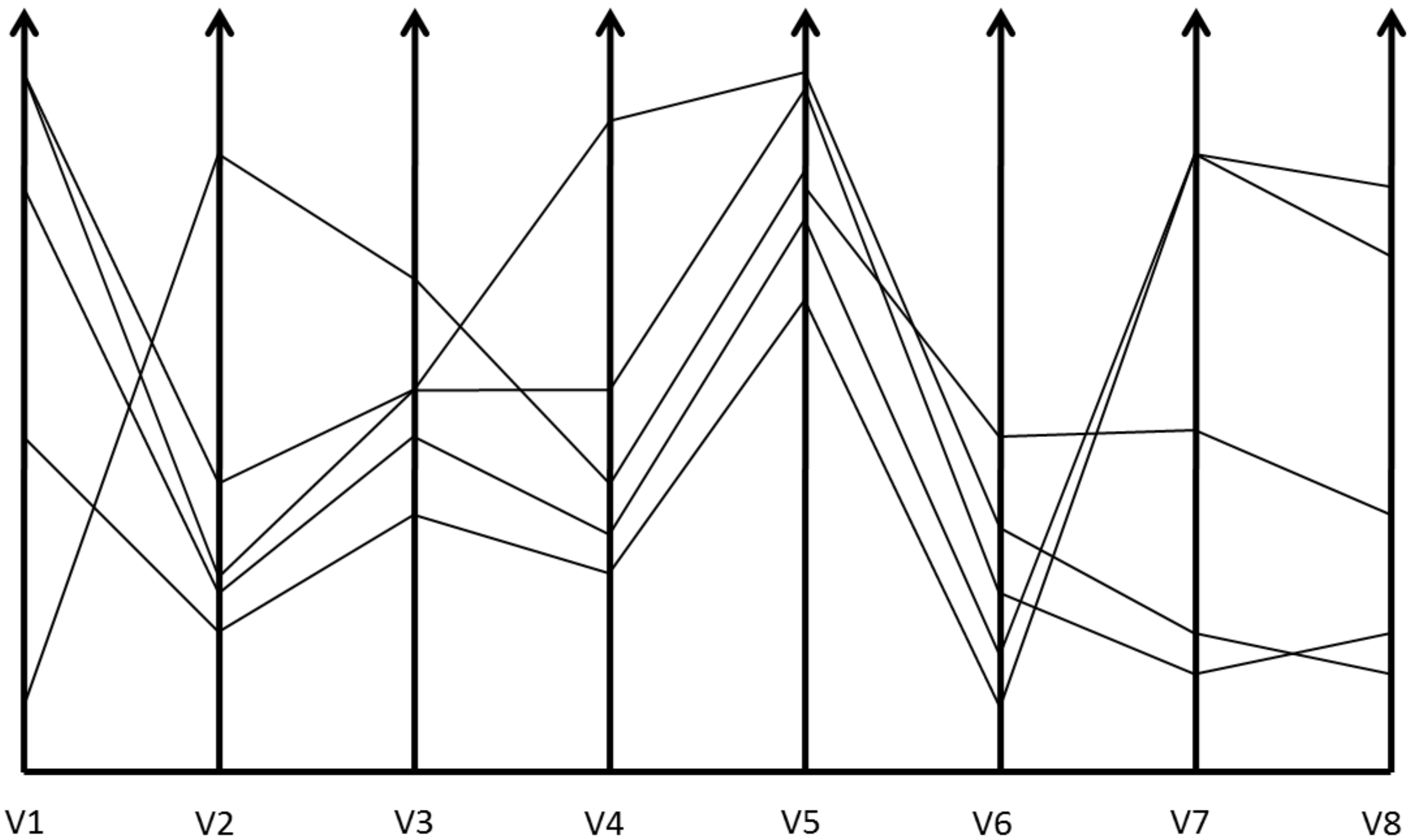
Multivariate Case - Scatterplot Matrix

Leave Blank



Parallel Coordinate Plots

Leave Blank



1 - Inselberg A (1985) The Plane with Parallel Coordinates. The Visual Computer1(4):69-91
2 - Ankerst M, Berchtold B, Keim DA (1998) Similarity clustering of dimensions for an enhanced visualization of multidimensional data. IEEE Symposium on Information Visualization pp 52-62

Issues With Parallel Coordinate Plots

Leave Blank

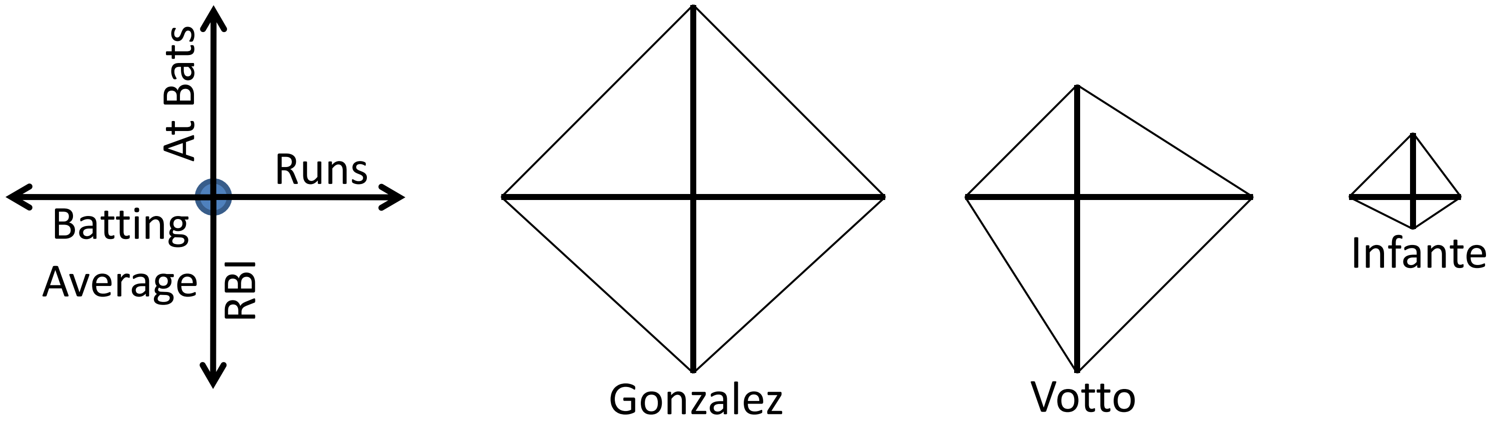
- Different variables can take different values with very different ranges
  - Need to normalize data ranges (maybe do a power transformation and then a normalization?)
- Order of the parallel coordinate plots has a major impact on the resultant visualization
- The more variables we plot, the more lines we get and the more clutter that we get

1 -Yang, J., Peng, W., Ward, M.O., Rundensteiner, E.A., Interactive hierarchical dimension ordering, spacing and filtering for exploration of high dimensional datasets. In *Proc. of IEEE Symposium on Information Visualization* (2003), pp. 105–112.
2 – Zhou, H., Yuan, X., Qu, Huamin, Cui, W., Chen, B., "Visual Clustering in Parallel Coordinates," Computer Graphics Forum 27(3) 1047-1054, 2008.

Star Plot

Leave Blank

- Lay out axes in a radial layout, length of a ray emanates from a central point
- Rays are then joined together by a polyline drawn around the outside

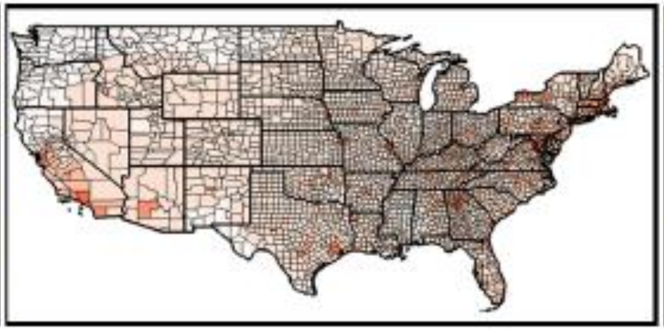


S. E. Fienberg,"Graphical Methods in Statistics," *The American Statistician*, vol. 33 pp. 156-178, 1979.

Geovisualization

Leave Blank

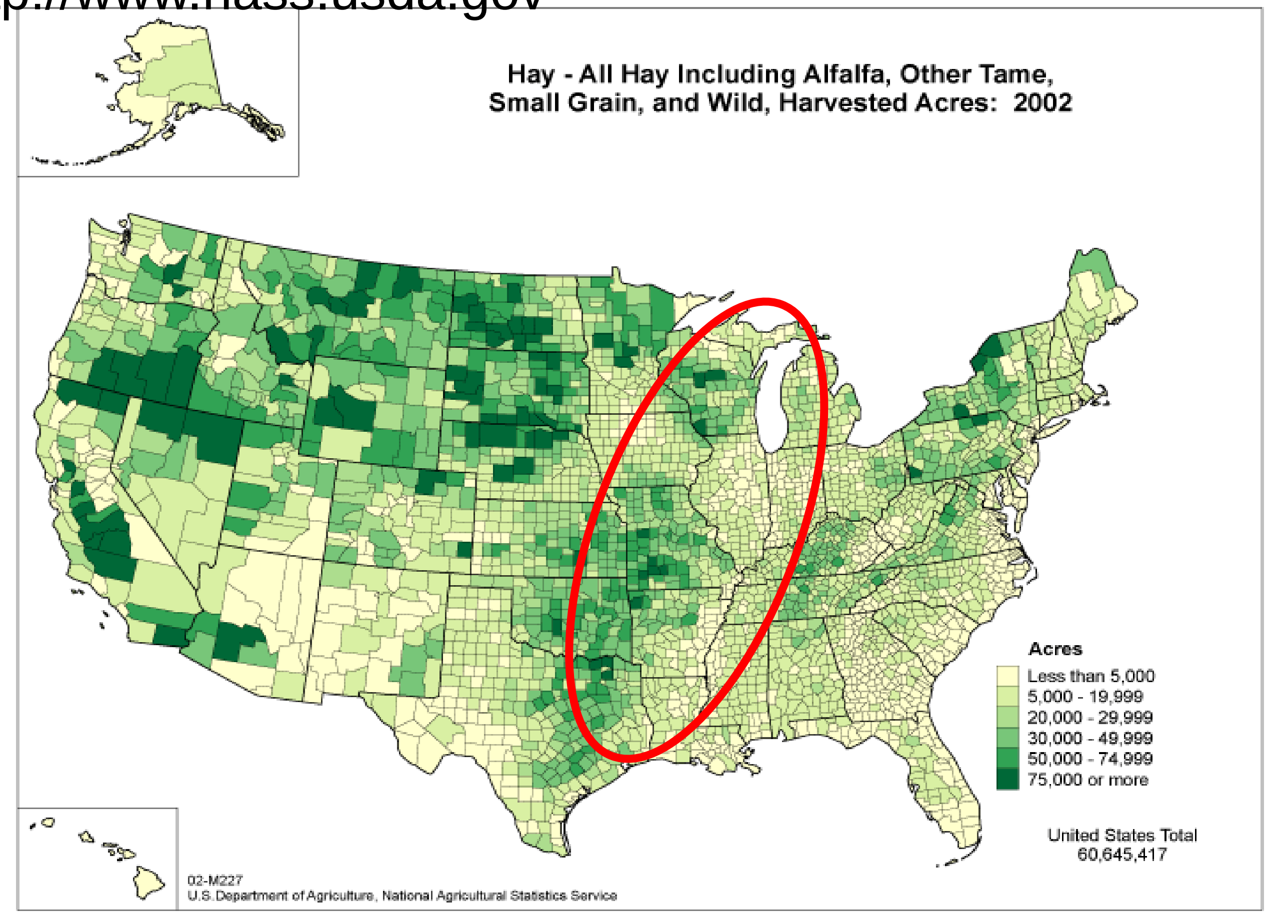
- Geovisualization primarily denotes tools and techniques that are designed to support analyses that focus on datasets with a geographic component
- Visual representations are designed and built utilizing cartographic principles
- Looking for trends over geographic regions




Choropleth Maps

Leave Blank

- http://www.nass.usda.gov



sured variable

ng) of the

1 nce to the Mapping of

P

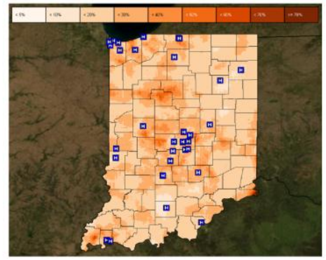
Coloring Choropleth Maps[1]

Leave Blank

- The number of colors depends on the number of classes (bins)
- Too many classes can overwhelm the user and distract them from seeing trends
- Too many classes can compromise legibility as colors become difficult to distinguish
- Typical cartographic rule of thumb is 5-7 classes
- Typical coloring schemes[8] include sequential, divergent and qualitative
- For more details on mapping as a visual representation, see [2]

1 – C. A. Brewer. http://colorbrewer.org
2 – A. MacEachren. *How Maps Work*. Guilford Press, 1995.

Isopleth Map



Leave Blank

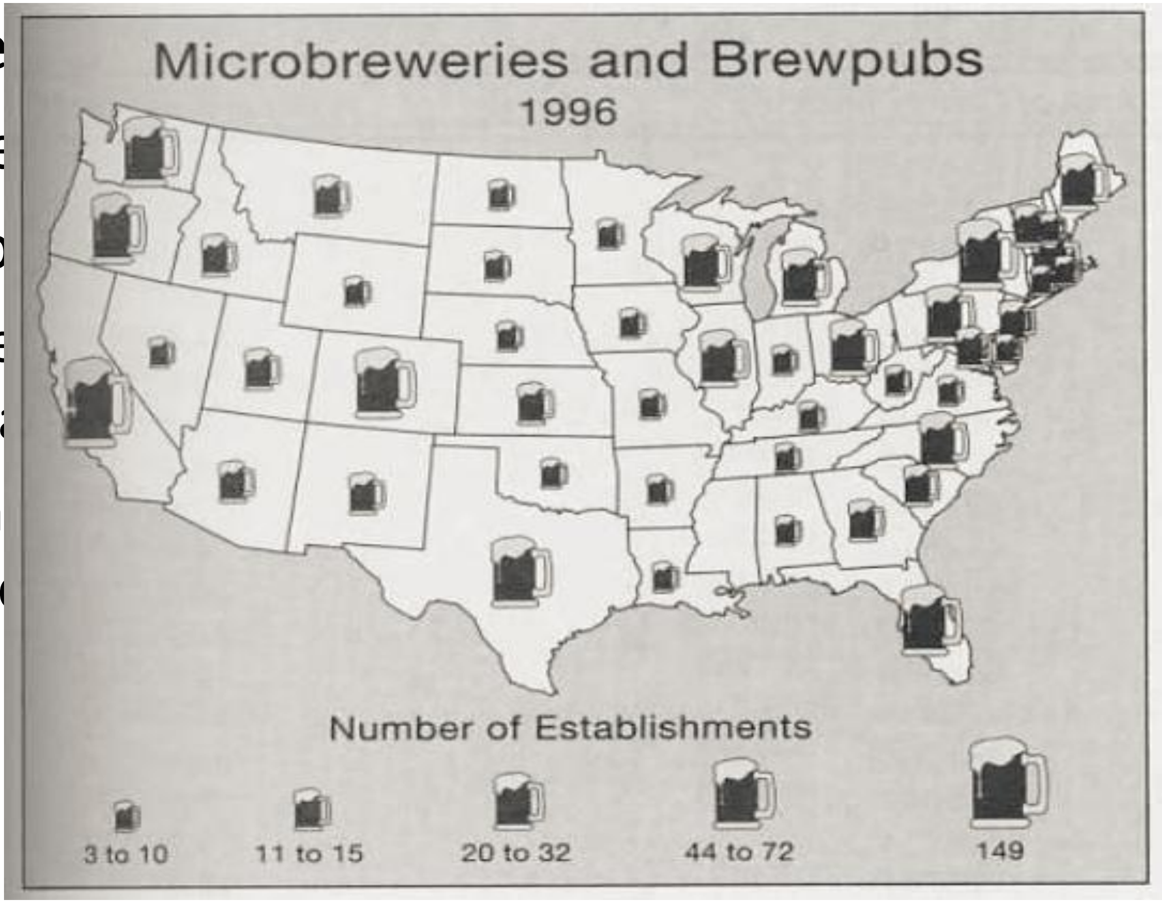
- Isarithmic map or contour map is created by interpolating a set of isolines between sample pints of known values
- Created from collection of point data and then interpolating unknown values between them
- Interpolation methods:
  - Kernel density estimation (see earlier lectures)
  - Inverse distance
  - Kriging

TA Slocum, RB McMaster, FC Kessler and HH Howard, “Thematic Cartography and Geovisualization,” Prentice Hall Series, 2009.

Proportional Symbol Mapping

Leave Blank

- Repre t locations
- Utilize ta
- Propo s of point data
  - True ocations (number of ca
  - Con rea, but data is conc e the centroid)



TA Slocum, RB McMaster, FC Kessler and HH Howard, "Thematic Cartography and Geovisualization," Prentice Hall Series, 2009.

Coordinated Multiple Views

Leave Blank

- Instead of trying to make the “best” visualization of all of our data, perhaps we can use multiple views
- Data can be expressed in a variety of ways
- Given the multivariate nature of data, a single statistical graphic may not be enough
- Interactive graphics systems can provide multiple representations of the data
- These representations can be *linked* or *coordinated*

North C, Shneiderman B (2000) Snap-together visualization Evaluating coordination usage and construction. International Journal of Human-Computer Studies 51:715-739

Interaction Types

Leave Blank

- Highlighting and focus
- Drill-down and hyperlinks
- Overview and context
- Changing parameters
- Changing representations
- Temporal fusion

A. Dix and G. Ellis. **Starting Simple - adding value to static visualisation through simple interaction.** *Proceedings of Advanced Visual Interfaces*, 1998

Interaction Types

Leave Blank

- How do we think of interacting with data?
  - Comparing and sorting
    - Provide a selection of data graphics to compare values and sort these values based on information
  - Adding/filtering variables
  - Highlighting
  - Aggregating/combining variables
  - Zoom and Pan
  - Rescale
  - Annotate
  - Bookmark

S. Few. *Now You See It*.

Interaction

Leave Blank

- Really, the key is user intent
  - Focusing on what a user wants to achieve through a specific interaction technique
- Interaction is done by a person for a purpose
  - Seeking more information
  - Solving a problem
  - Exploratory analysis
  - Analytic discourse

* - Material borrowed from John Stasko's lecture on Interaction: http://www.cc.gatech.edu/~stasko/7450/Notes/interaction.pdf

Interaction in Visualization

Leave Blank

- Broken down into 7 categories
  - Select
  - Explore
  - Reconfigure
  - Encode
  - Abstract/Elaborate
  - Filter
  - Connect

Ji Soo Yi; Youn ah Kang; Stasko, J.T.; Jacko, J.A.; , "Toward a Deeper Understanding of the Role of Interaction in Information Visualization," *IEEE Transactions on Visualization and Computer Graphics,* vol.13, no.6, pp.1224-1231, Nov.-Dec. 2007

Visual Analysis vs. Visual Analytics

Leave Blank

- The choice of visual representation will influence the user's analysis
- Massaging the data can help improve the visual representation
- It also makes some data analysis methods easier
- Adding interaction allows the user to form and explore their own hypotheses
- But, how do we find "important" features in the data?

Data Mining

Leave Blank

- Data mining domain has techniques for examining data, looking for patterns, looking for anomalies
- Can be used to enhance the visualizations, show us what is important
- Can also be used in exploratory analysis, “I think this looks interesting, show me similar trends.”

Temporal Analysis Methods

Leave Blank

- For temporal data, we can find anomalies using control chart methods
- Control charts consist of a statistic representing some measurement in time (number of patients in a hospital, value of a stock, etc.)
- The mean and standard deviation of the statistic is calculated given all the available samples
- If the current value is greater than some pre-set number of standard deviations from the mean, then an alert is generated (Shewart suggested three standard deviations)

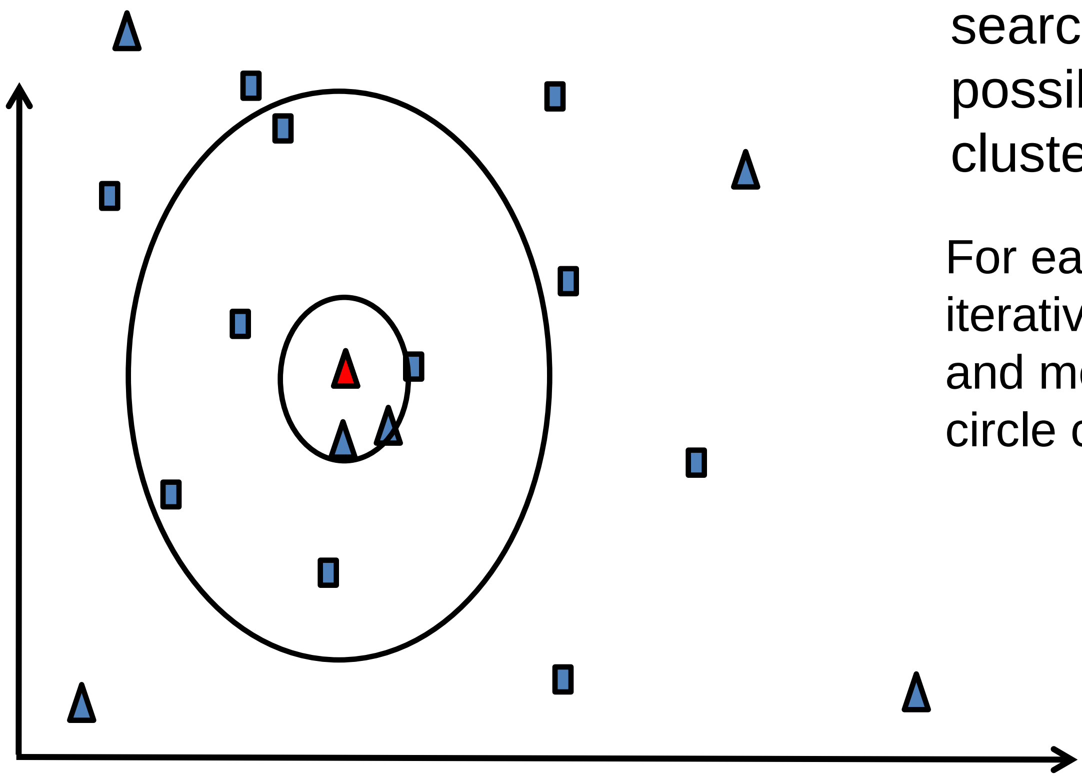
Finding Similarities in Time Series

Leave Blank

- Indexing problem
  - Find all stocks whose fluctuations are similar to stock X
- Subsequence Similarity
  - Find out other days in which stock x had similar movements as today
- Clustering
  - Group regions that have similar temporal patterns
- Rule Discovery
  - Find rules like if x goes up, y goes up, etc.

Scan Statistics[1]

Leave Blank



The idea is to exhaustively scan the search space to check for all possible statistically significant clusters

For each case and control, the algorithm iteratively draws circles containing more and more data samples until the largest circle contains N/2 samples.

- control

- case

1 – M. Kulldorff. http://www.satscan.org

Scan Statistics

Leave Blank

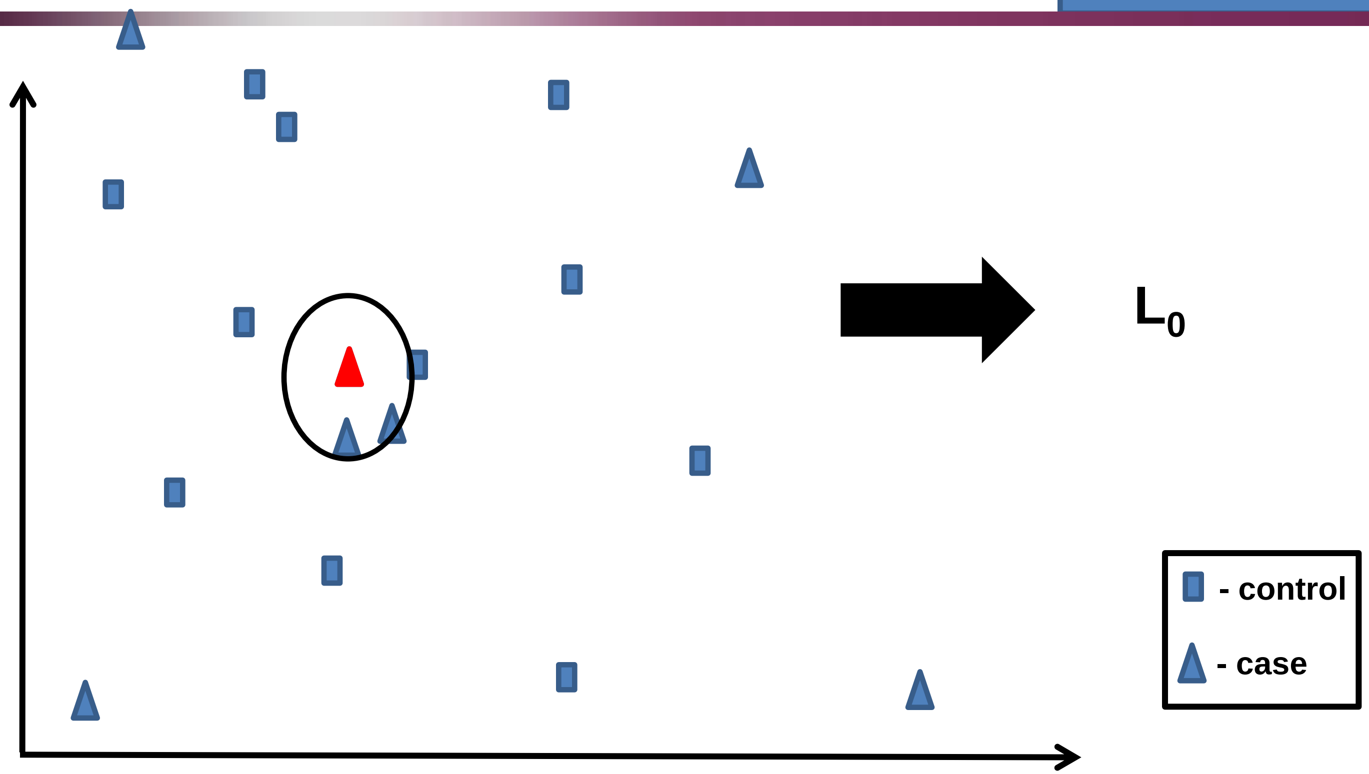
- For each circle (window) compute the likelihood function
- For Bernoulli distributions, the function is:

$$\left(\frac{c}{n}\right)^{c}\left(\frac{n-c}{n}\right)^{n-c}\left(\frac{C-c}{N-n}\right)^{C-c}\left(\frac{(N-n)-(C-c)}{N-n}\right)^{(N-n)-(C-c)} I()$$

- c is the number of cases in the window, C is the total number of cases in the data set, n is the number of cases + controls in the window, N is the total number of cases + controls, I() is an indicator function that is one if the number of cases is more then expected based on the null hypothesis and zero otherwise

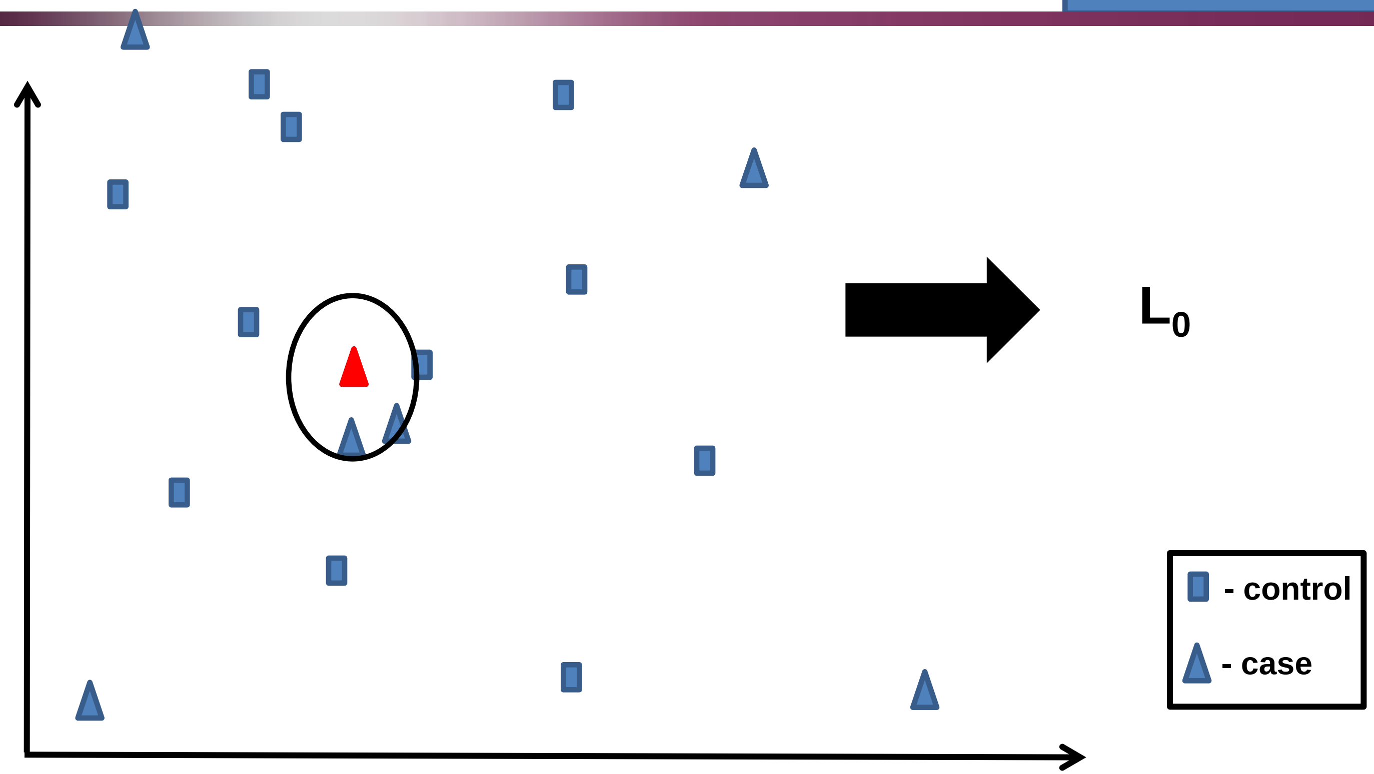
Scan Statistics

Leave Blank



Scan Statistics

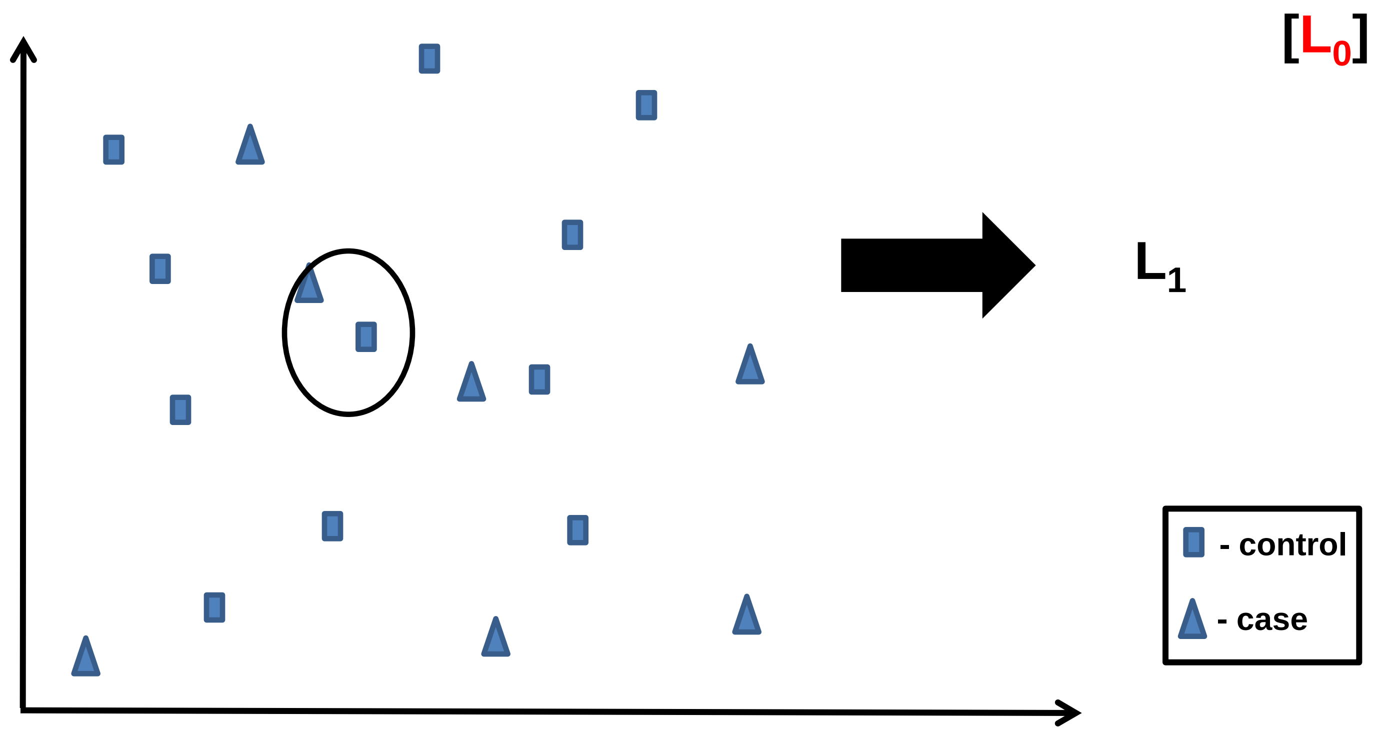
Leave Blank



VACCINE

Scan Statistics

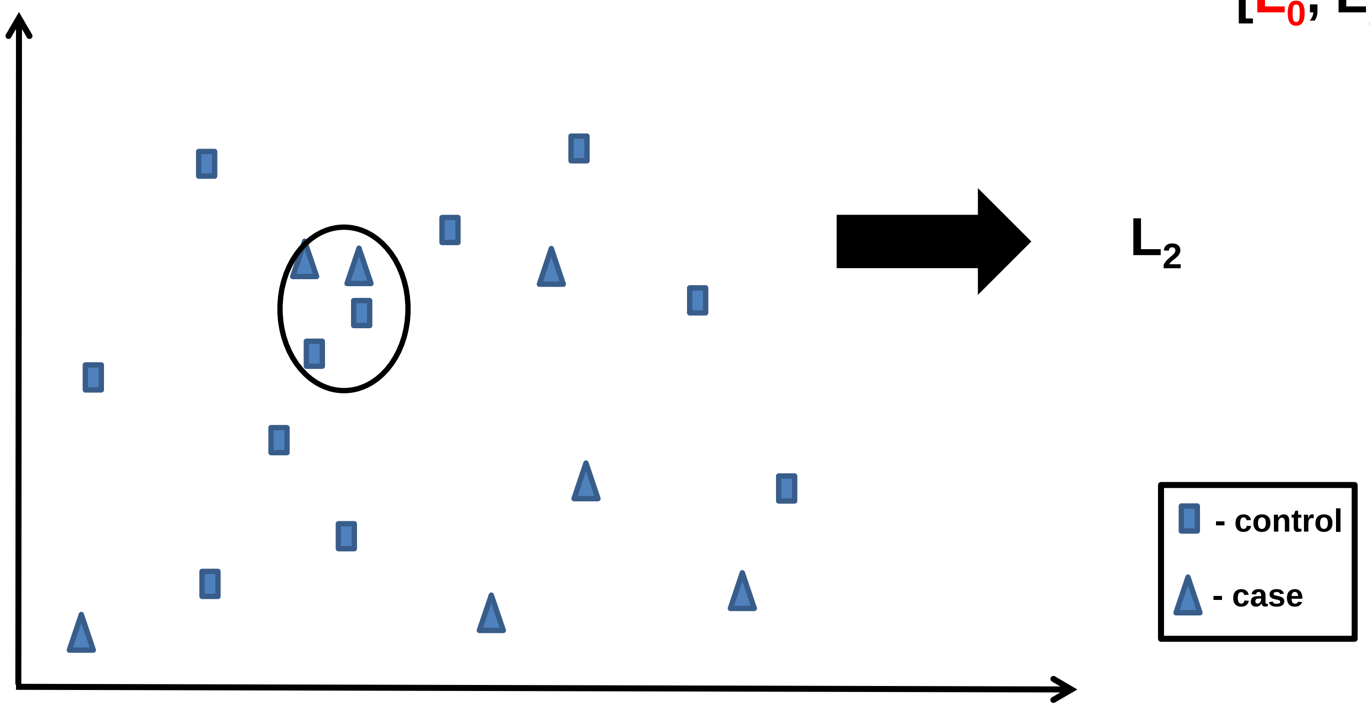
Leave Blank



Scan Statistics

Leave Blank

$[L_0, L_1]$



Scan Statistics

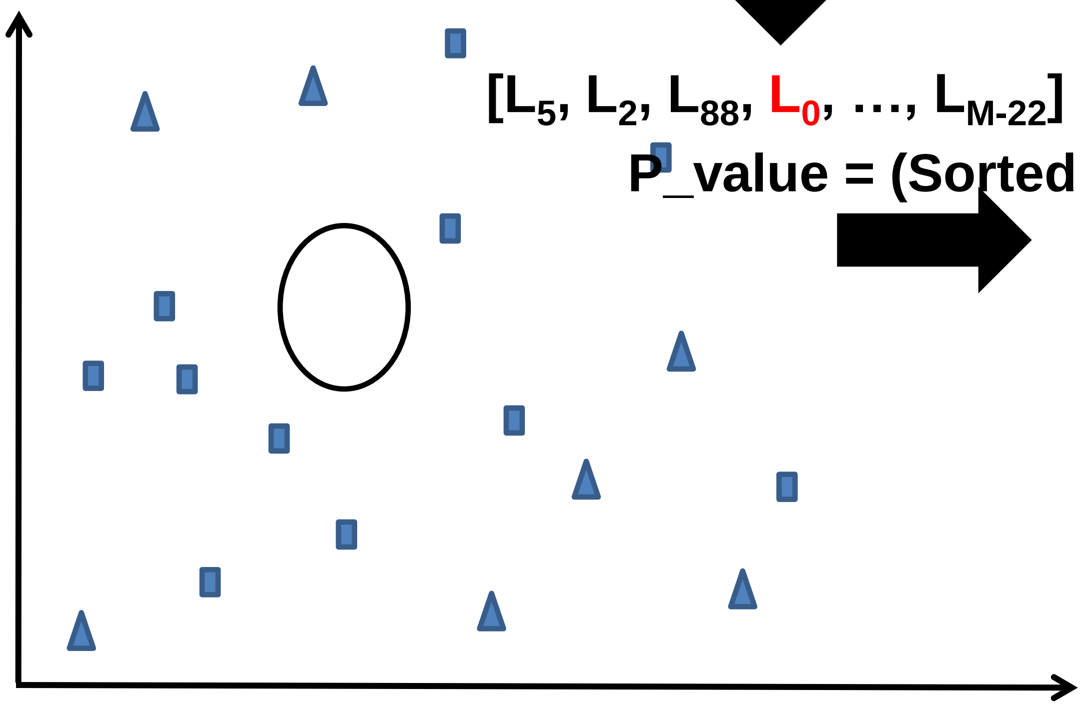
Leave Blank

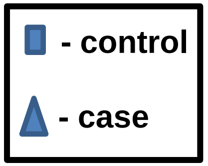
[L_0, L_1, L_2, ..., L_{M-1}]

[L_5, L_2, L_{88}, L_0, ..., L_{M-2}2]

P_value = (Sorted position of L0 in list)/M

L_M





Scan Statistics

Leave Blank

- Keeping the window in the same location, randomly redistribute the cases and controls
- Calculate the likelihood function for the new distribution
- Repeat 99, 999, xxx times and determine a p-value for the original distribution by sorting the likelihood functions

- Related work includes the Geographical Analysis Machine[1]

1 – S. Openshaw, M. Charlton, C. Wymer and A. Craft. A mark I geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographical Information Systems*, 1:335-358, 1987.
* - http://www.ccg.leeds.ac.uk/software/gam/

Summary

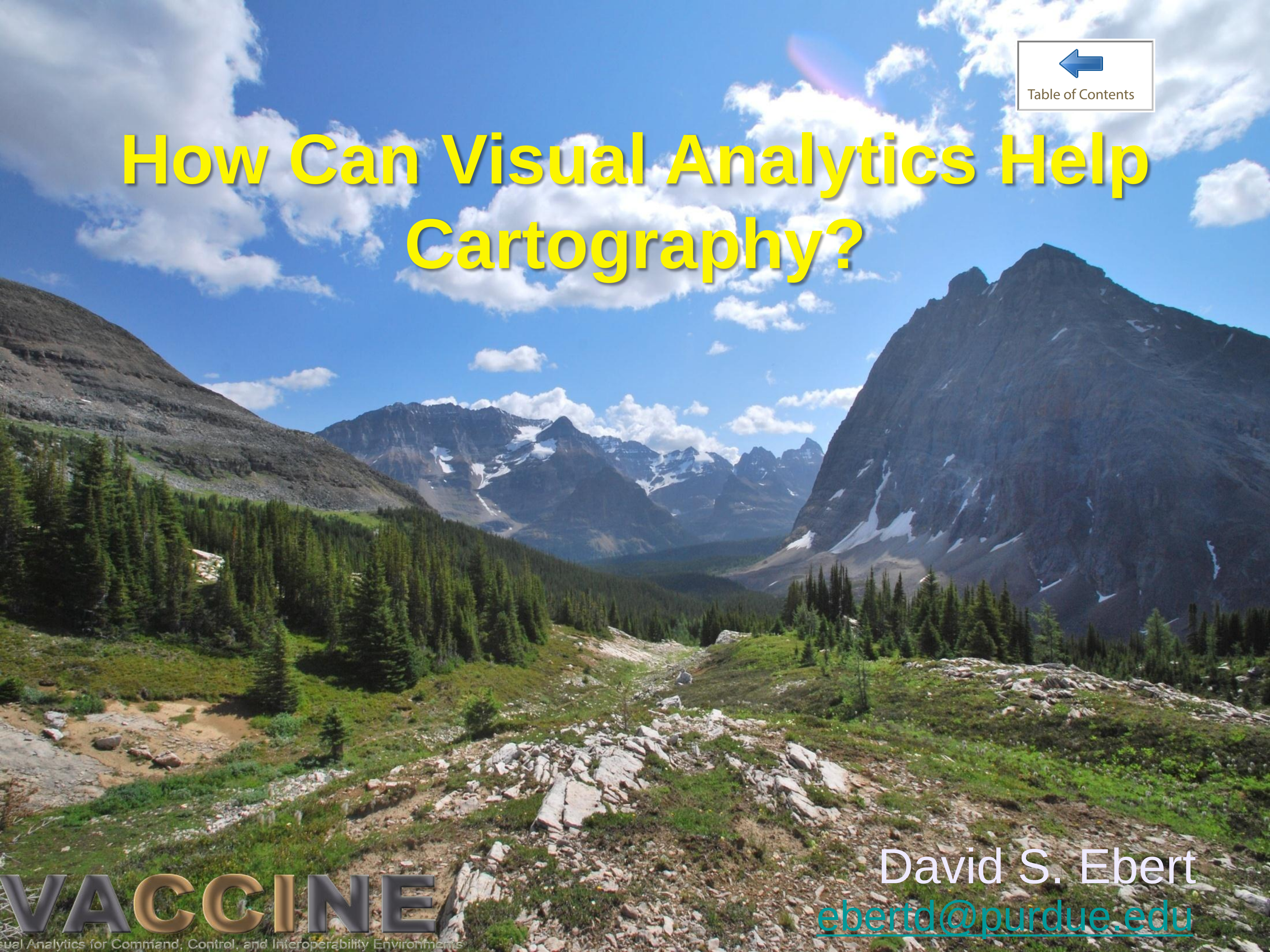
Table of Contents

Leave Blank

- Visual Analytics combines a wide breadth of topics
  - Visual representations
  - Perception
  - Statistics
  - Human Computer Interaction
  - Machine Learning and Data Mining
  - Cognition and Many More
- This tutorial only glossed over some of the more commonly used methods and functions as an introduction

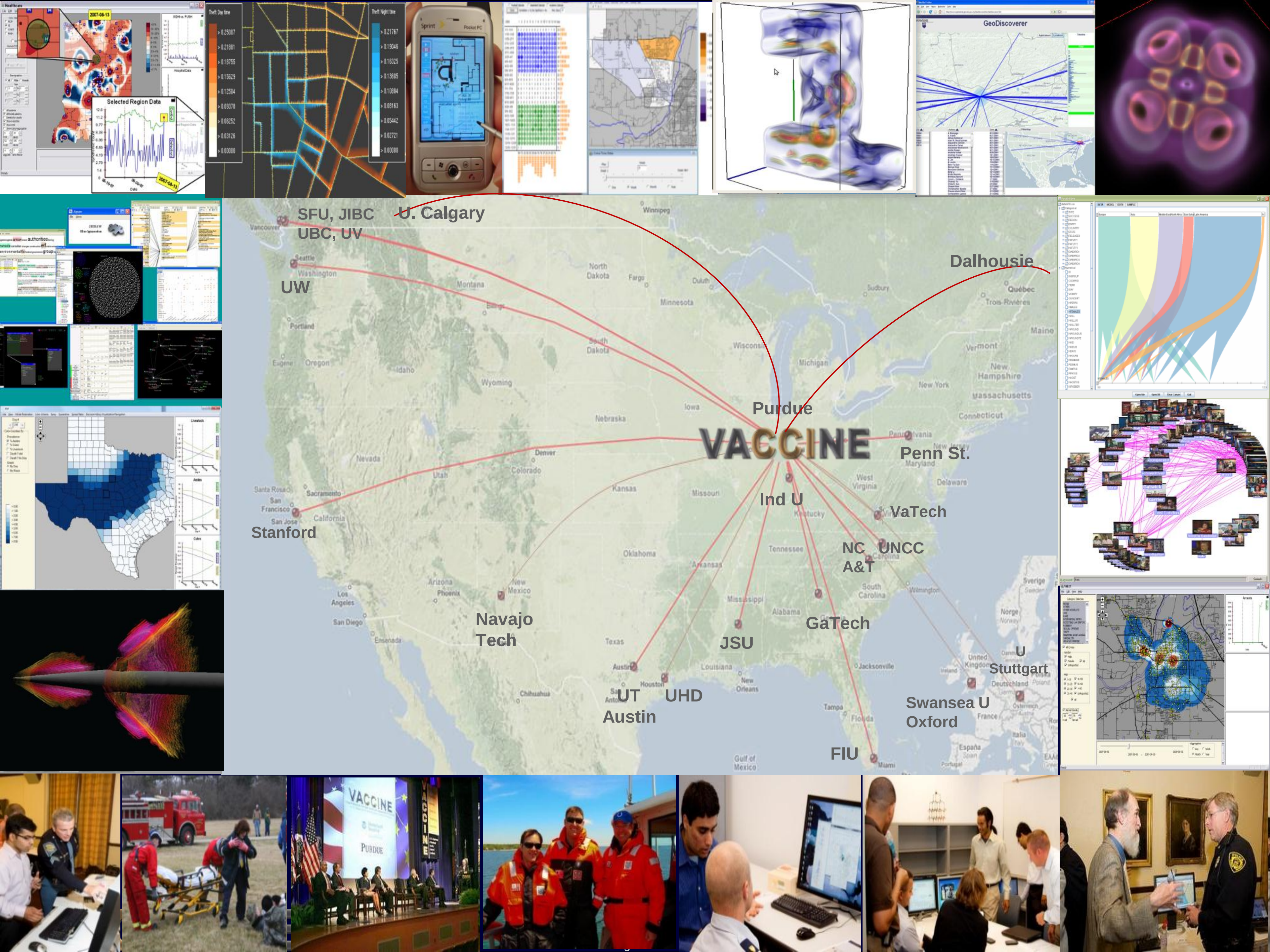
Table of Contents

How Can Visual Analytics Help Cartography?



David S. Ebert
ebertd@purdue.edu

VACCINE
ual Analytics for Command, Control, and Interoperability Environments


SFU, JIBC
UBC, UV
U. Calgary
UW
Dalhousie
Purdue
VACCINE
Penn St.
Ind U
VaTech
Stanford
NC A&T
UNCC
Navajo Tech
JSU
GaTech
U Stuttgart
UT Austin
UHD
Swansea U
Oxford
FIU
GeoDiscoverer
Selected Region Data
Theft Day time
Theft Night time

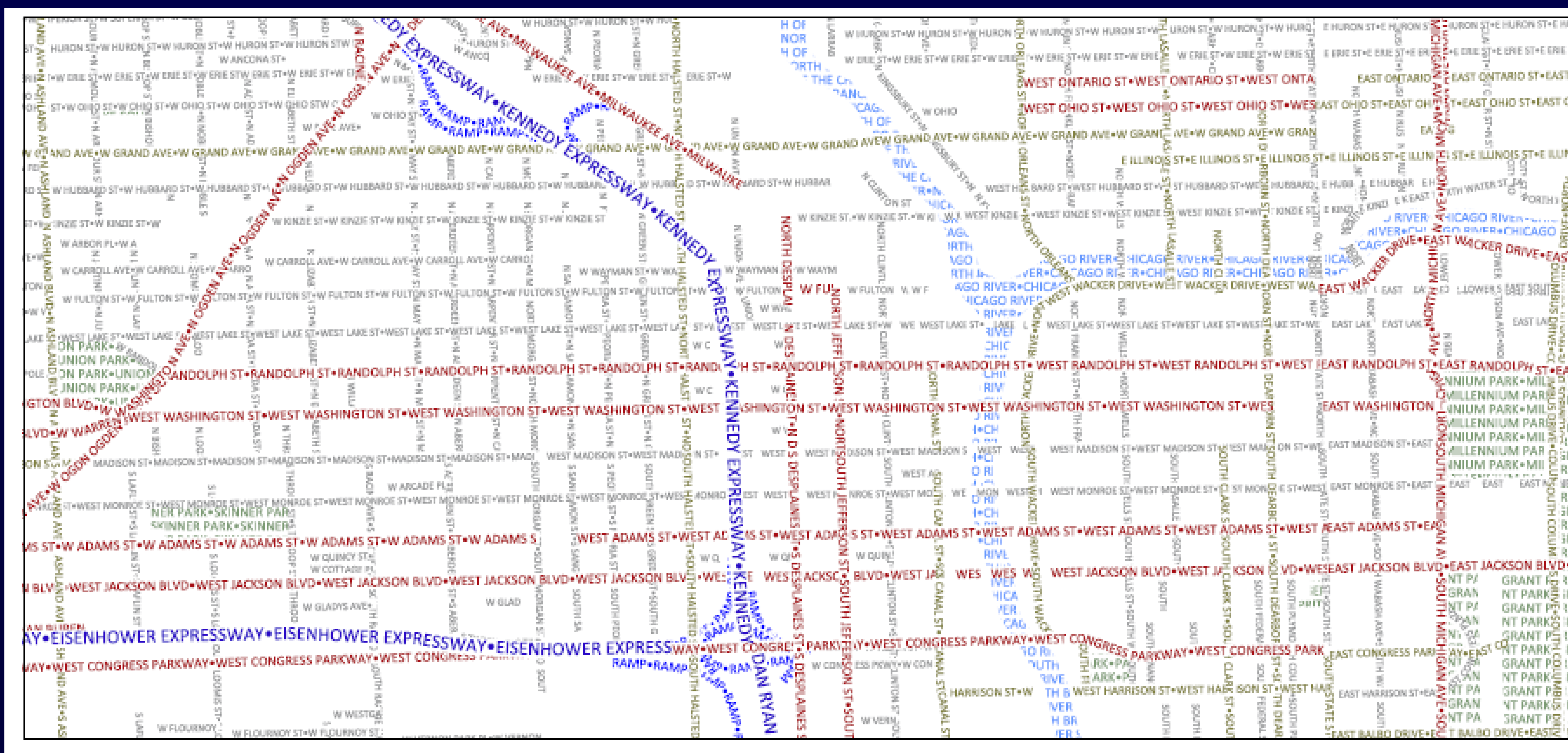
Path: Where Do We Want to Go?

- How can visual analytics help cartography?
- How can cartography help visual analytics?

VACCINE

How Can Cartography Help Visual Analytics?

- Inspiring new compact visual representations



How Can Cartography Help Visual Analytics?

- An amazing amount of information has a spatial context
- Spatialization is a natural mental model, but isn't always natural for visualization…

George Robertson

VACCI

Path: Where Do We Want to Go?

Answers and challenges

An example

Illustrative examples

- Science visual analytics
- Health visual analytics
- Public safety and law enforcement

Visual analytic challenges and future

How Can Visual Analytics Help Cartography: Answer

- Interactive, evolving, analytics and simulation driven analytical, correlative, predictive, decision making environments

- Of course!

- The interactive human analysis through interaction is key

VACCINE

The Science of Interaction

- Definition: The study of methods by which humans create knowledge through the manipulation of an interface.
  - …interaction and inquiry are inextricable. It is through the interactive manipulation of a visual interface – the analytic discourse – that knowledge is constructed, tested, refined and shared."
    Stasko et al., IV2009

5 Challenges:

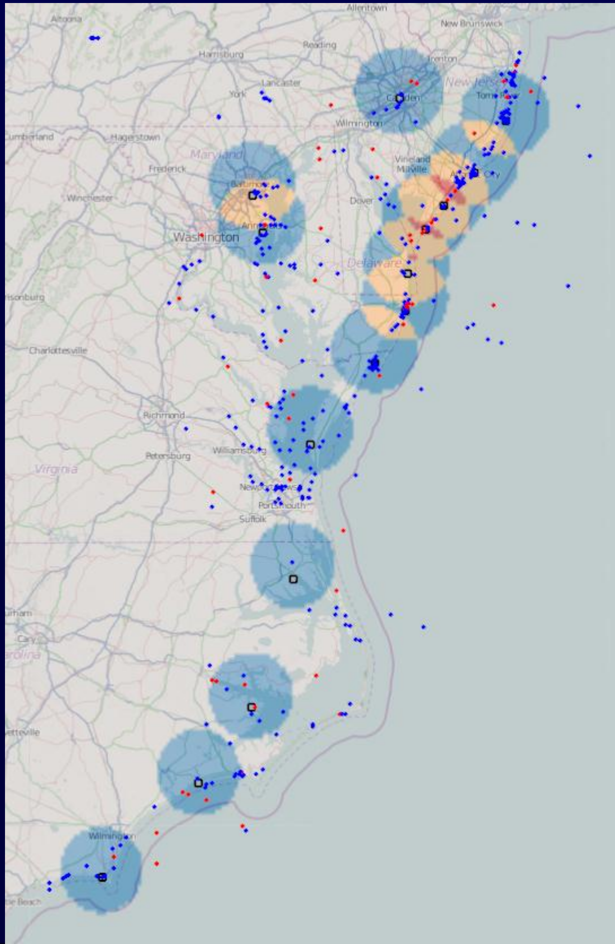
1. **Visual Discourse -** Interactive visual thinking tools for the exploration, understanding, collaboration, description, explanation, decision, dissemination, persuasion of concepts & data.
2. **Multi-modal Sensemaking In The Large -** Enable large-scale, distributed teams to interactively make sense of big multi-modal data and problems.
3. **Fluid Interaction -** Designing fluid, high-bandwidth, and powerful interaction models and paradigms for the purpose of individuals reaching their full potential in viewing, analyzing, and understanding large and complex data.
4. **Collaboration –** beyond space and time- synergizing technologies and human users
5. **Mixed Initiative Data Discovery and Manipulation -** Balancing active user input with systematic guidance to enable visual data manipulation and analysis.

Answers:

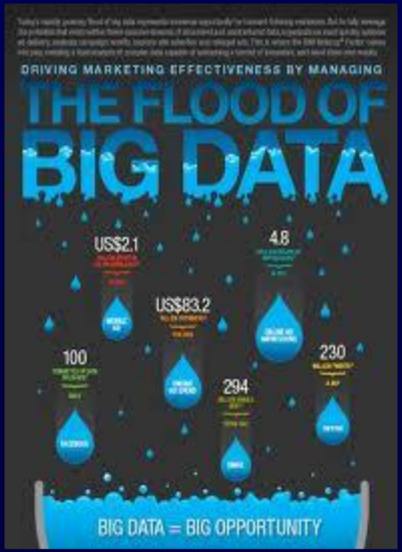
- Novel integrated, interactive visualization and analytics approaches for
  - Multisource, potentially conflicting data (more than layers)
  - Big data
  - Uncertain data
  - Multiscale data
  - Multimodal data
  - Temporal data
- Risk-based decision making
- Interactive hypothesis testing
- Integrated simulations and predictive models

Visualization, Data, and Decision Making

- At some level most decision making is driven by data
- What role does visual representation play?
- Good or bad – people believe charts and graphics – “seeing is believing”



Big Data Is The Rage



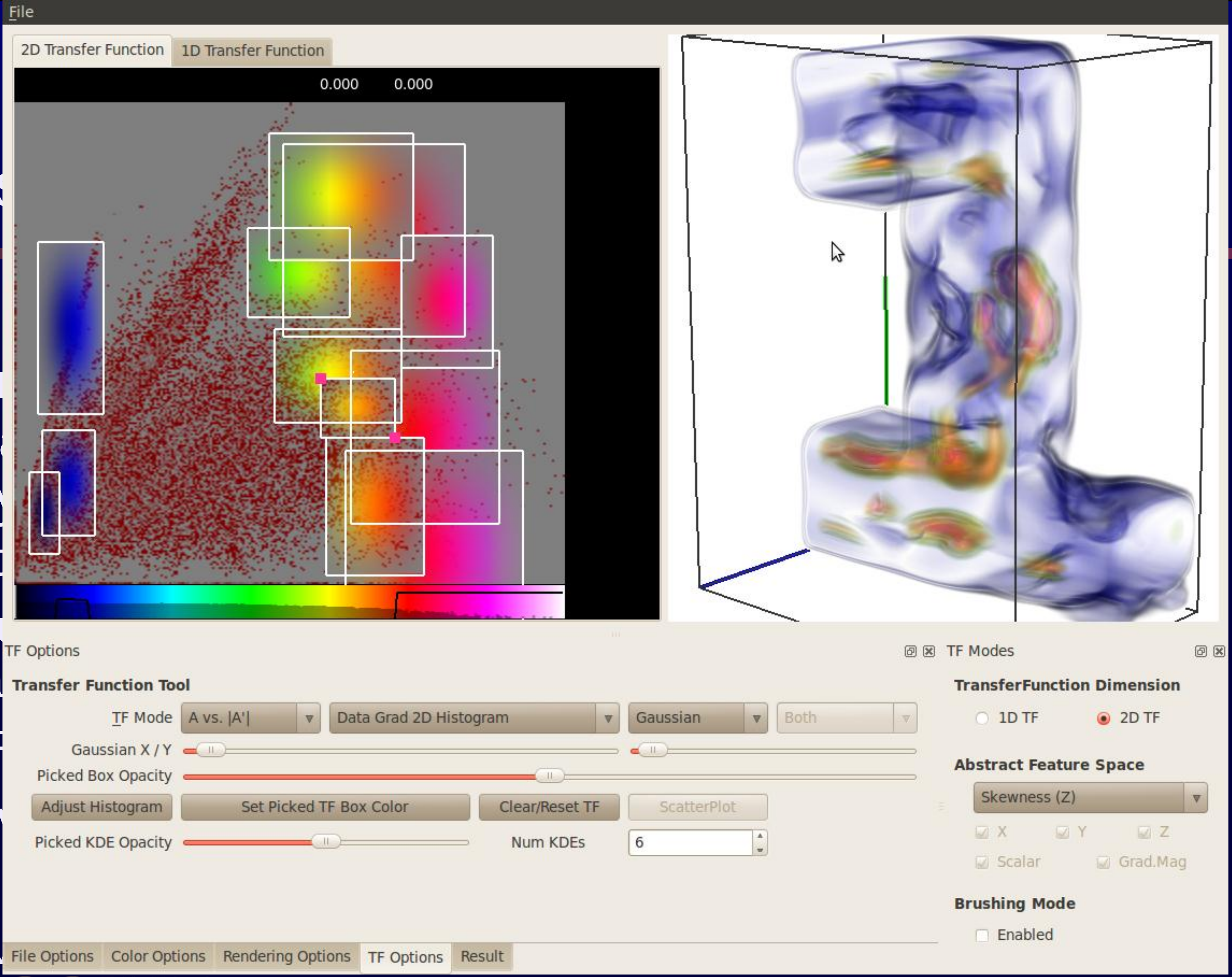
Visual Analytics Is More Than Big Data Analytics

VA Approaches to Big Data

- Don’t display all the data!!
- Extract relevant information at the appropriate natural scales using analytics and statistical models to reduce data space so that it is mappable to the visual space
- User interaction to guide above process
- Should we use big iron or cell phones ?

Vis

- Utili techn inter analy clust
  - Cl ba
  - GP
- Dev HP env



VACCINE

Multisource, Multiscale, Evolving, Uncertain Data

- Challenges
  - Signature extraction for visual cognition, analysis, correlation
  - Common scales and alignment (e.g., temporal, spatial, aggregation level, …)
  - Temporal reasoning
  - Natural representations of temporal aspects
  - Natural representations of confidence, certainty, risk
  - Natural integration of spatial and nonspatially (scaled) data

Path: Where Do We Want to Go?

Answers and challenges

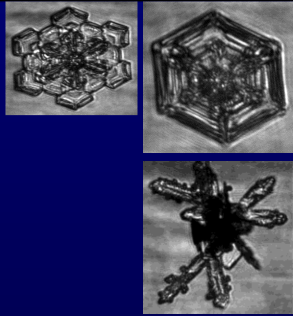
An example

Illustrative examples

- Science visual analytics
- Health visual analytics
- Public safety and law enforcement

Visual analytic challenges and future

Example Challenging Problem - Atmospheric Science: Multi-scale Interactions (in the words of a cloud physicist)



1 mm
1 kHz

No observing platform can measure the quantities of interest over all needed spatial and temporal scales needed

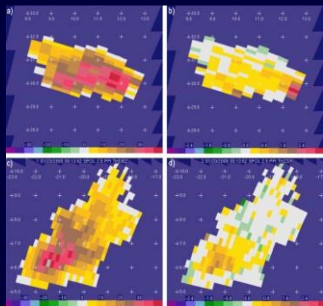
No numerical model can simulate the quantities of interest over all needed spatial and temporal scales

→We observe/simulate over a subset of the pertinent scales, using different instruments/models, and must assimilate these results to understand the "big picture"

Visual analytics is crucial for this task

1km
5min

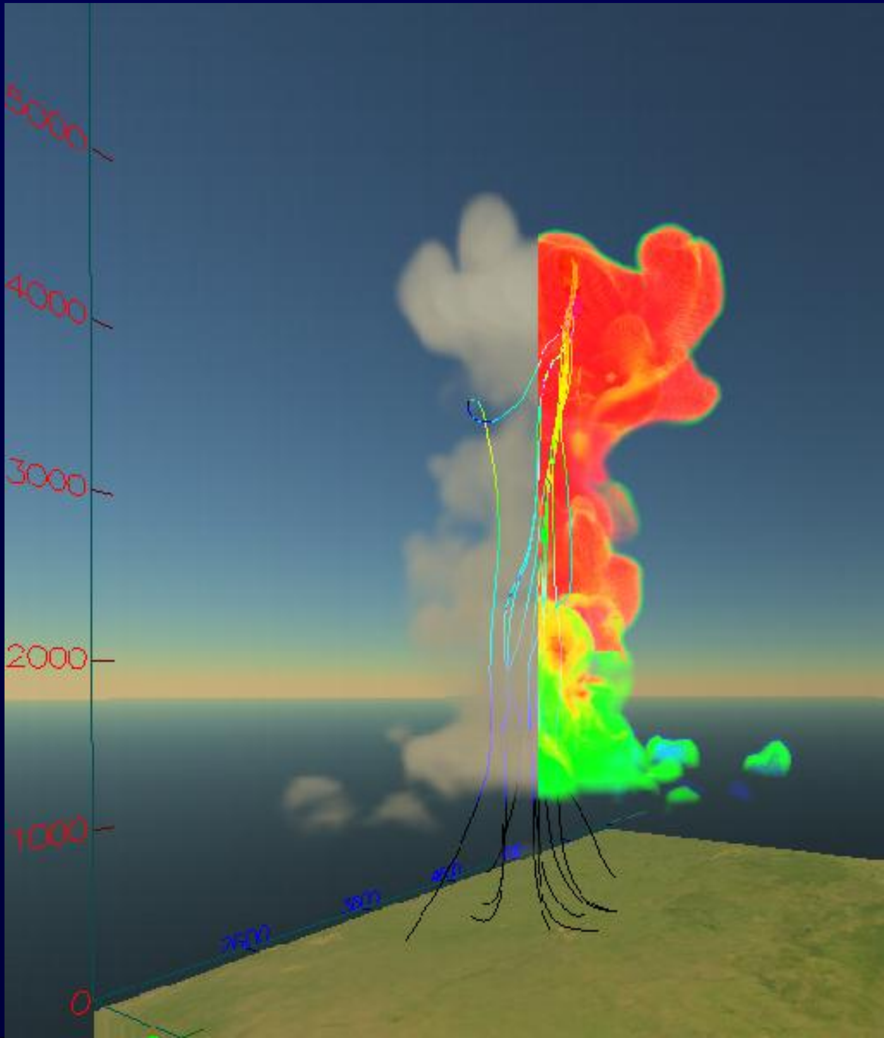
Issues: Multi-scale, multi-system, multisource, massive, data & simulations

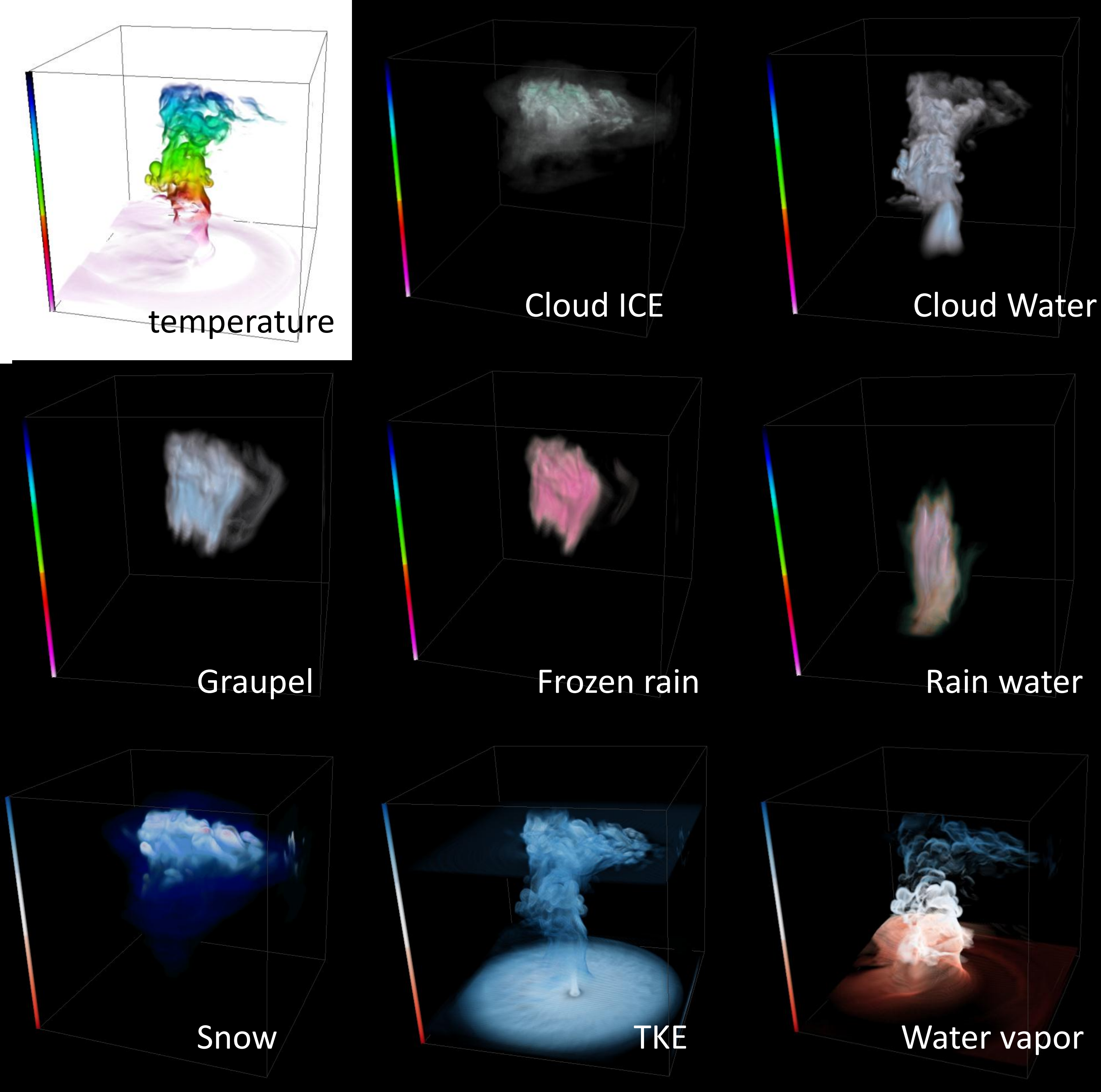


VACCINE

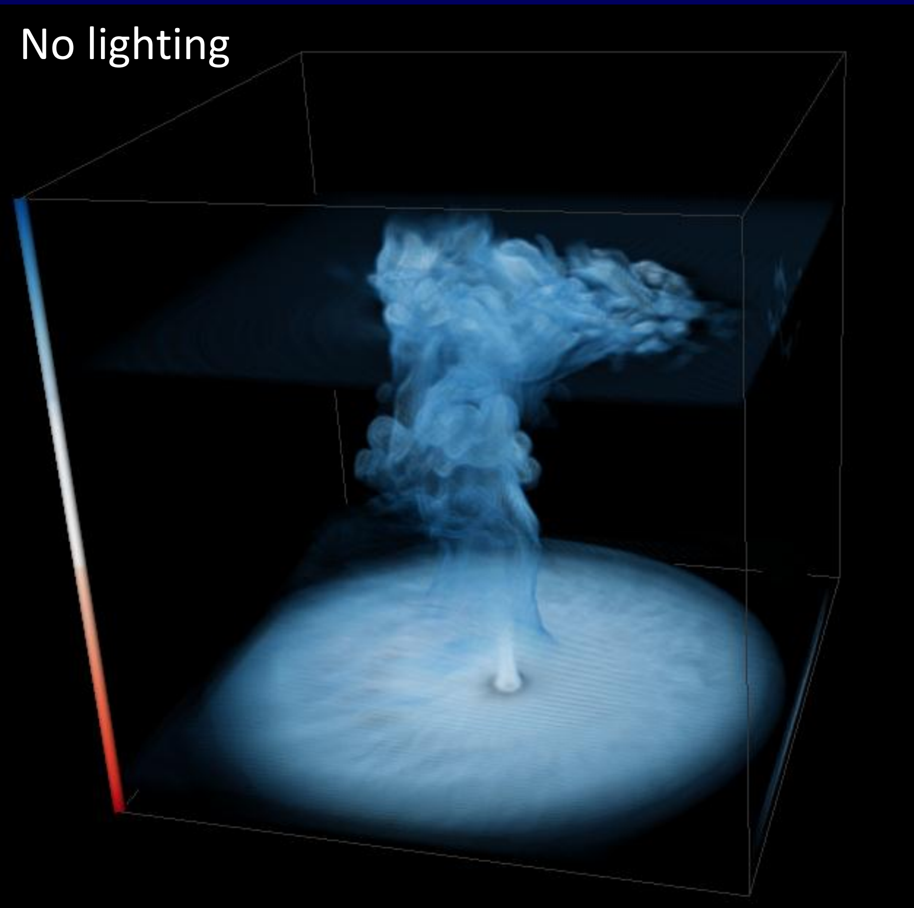
One Solution in Use: Our Atmospheric Visual Analytic Environment

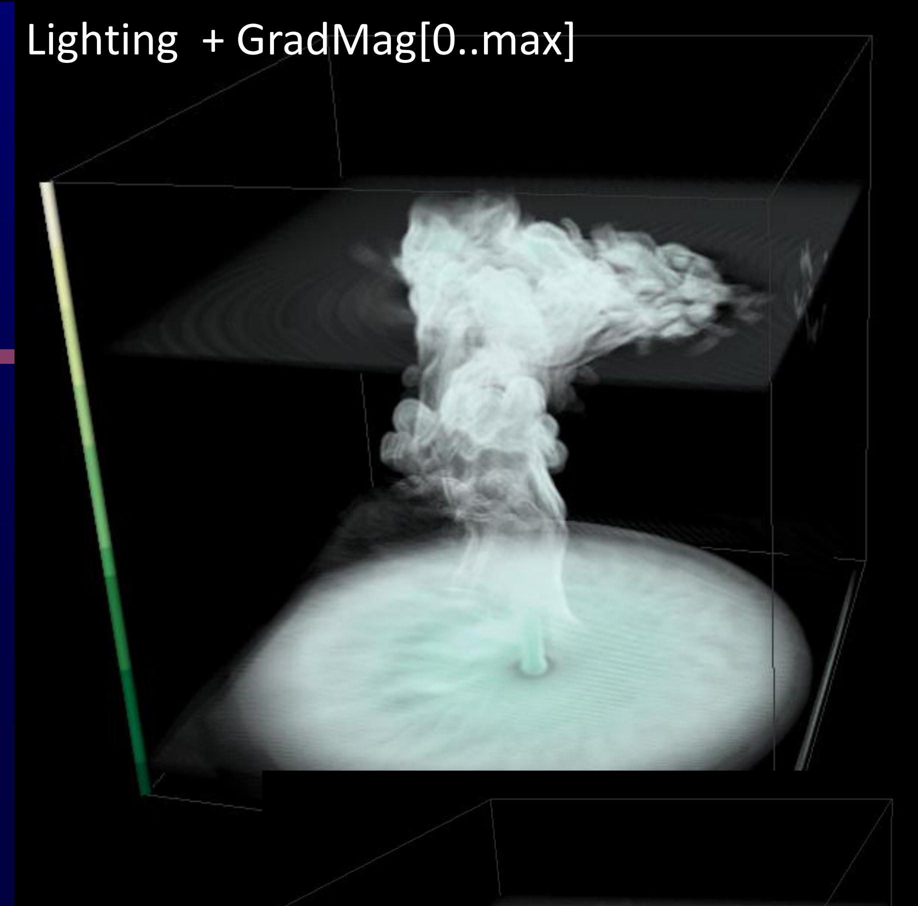
- Utilize multiple rendering styles
- Provide interactive data exploration and user directed analysis
- Allow user specified analysis and queries on the fly
- Allow interactive correlative analysis of multisource data

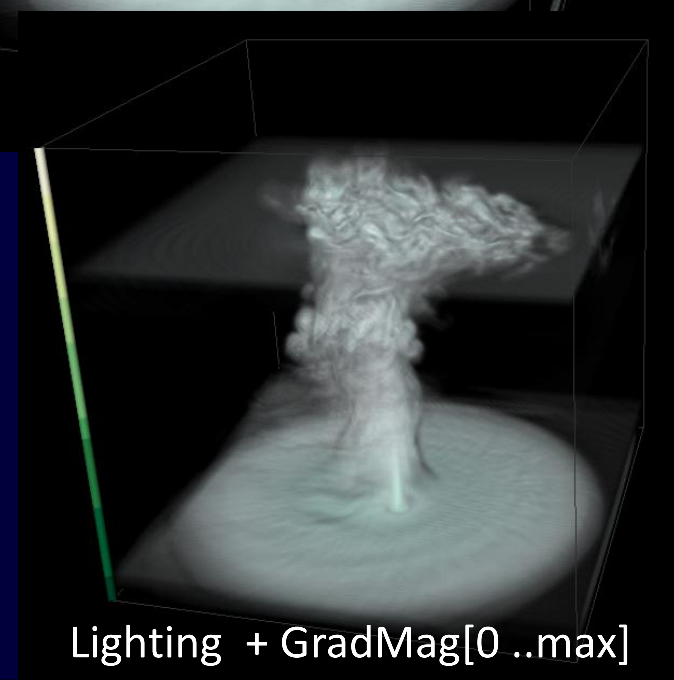



temperature
Cloud ICE
Cloud Water
Graupel
Frozen rain
Rain water
Snow
TKE
Water vapor

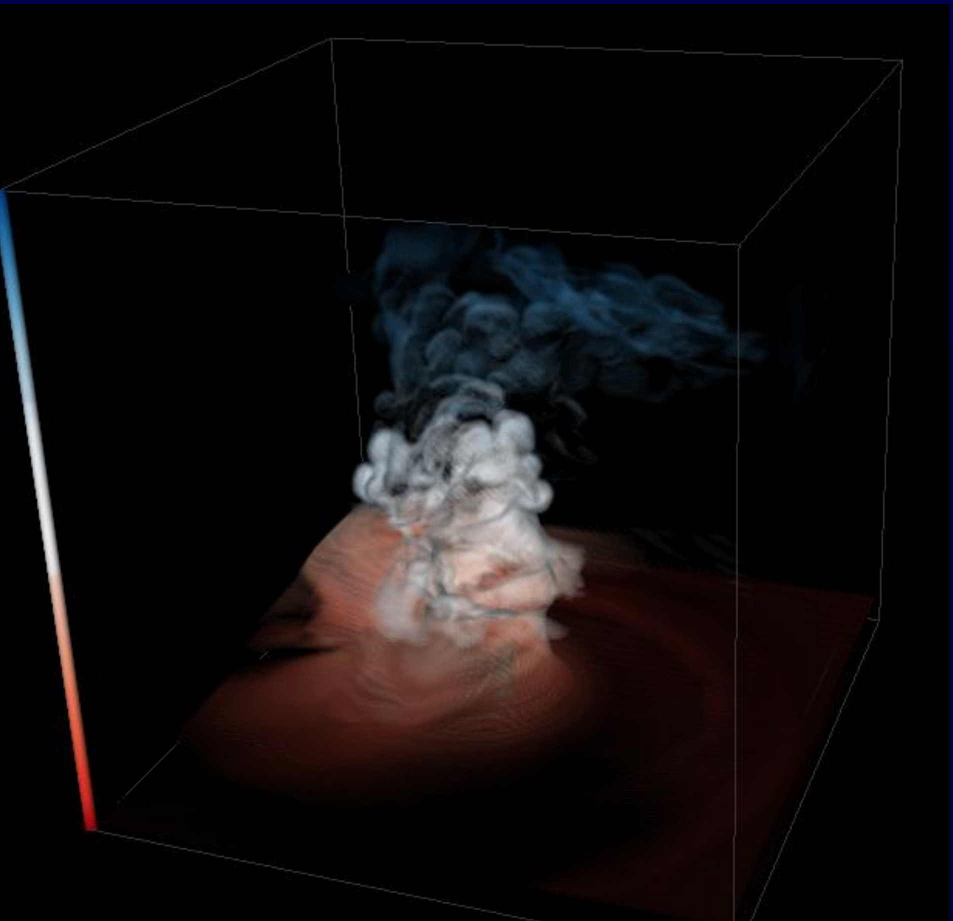
TKE



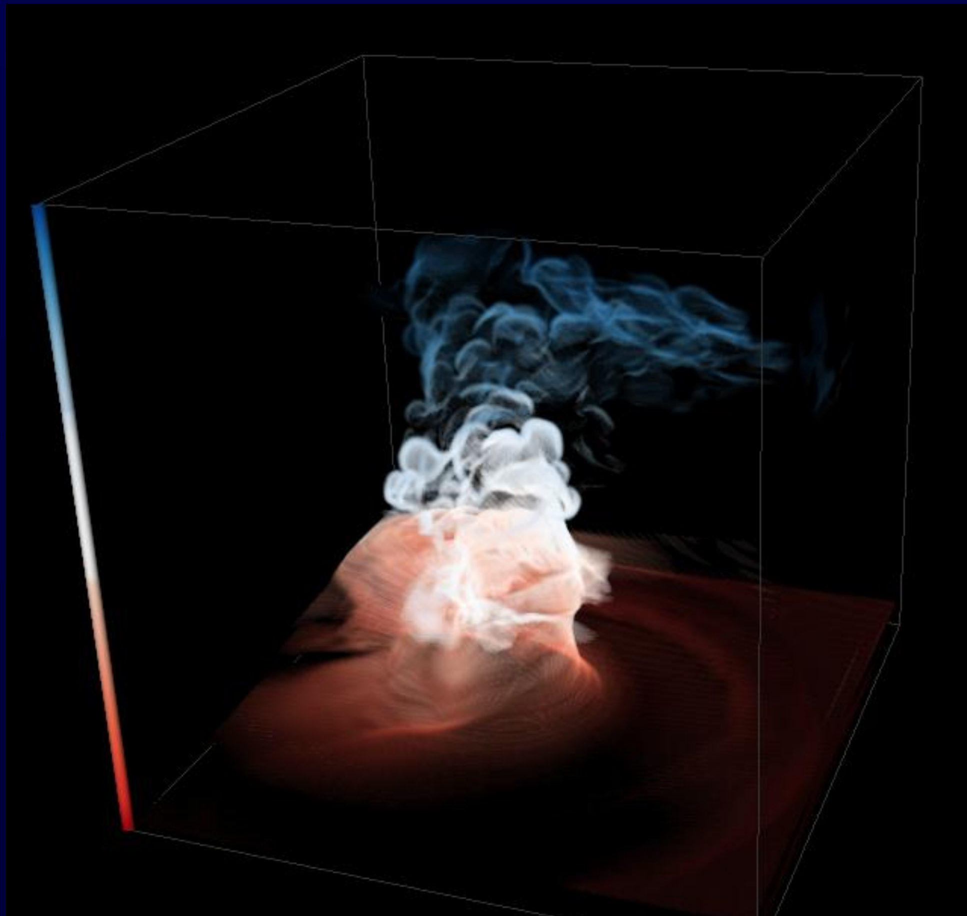




Water vapor



Lighting + GradMag[0.x ..max]



No Lighting + GradMag[0.x ..max]

Path: Where Do We Want to Go?

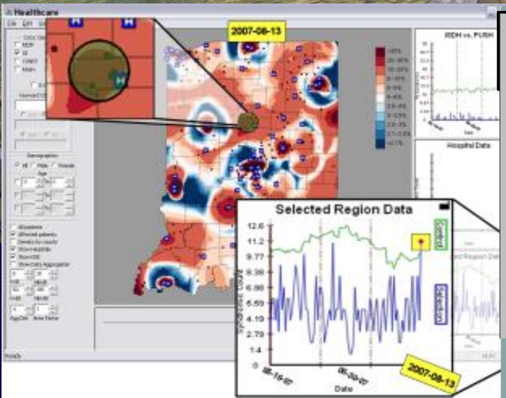
- Answers and challenges
- An example
- Illustrative examples
  - Science visual analytics
  - Health visual analytics
  - Public safety and law enforcement
- Visual analytic challenges and future

Example: Public Health Visual Analytics

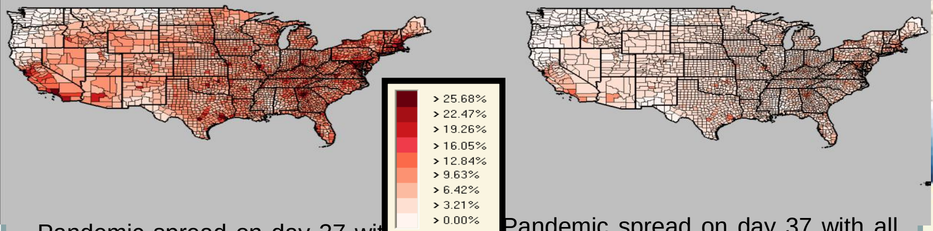
Health Surveillance

Pandemic Simulation and Planning

Zoonoses Planning and Response

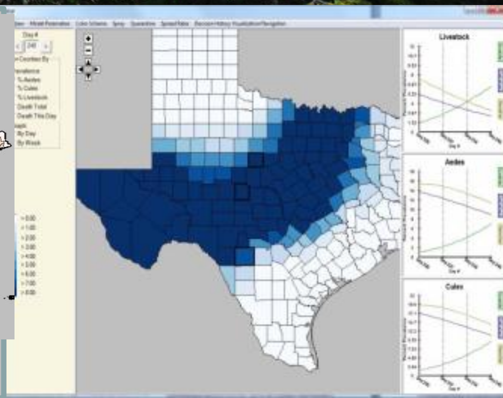


Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.



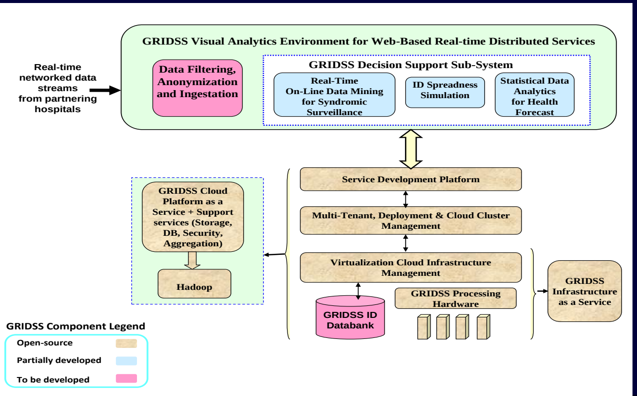
Pandemic spread on day 37 with no decision measures implemented

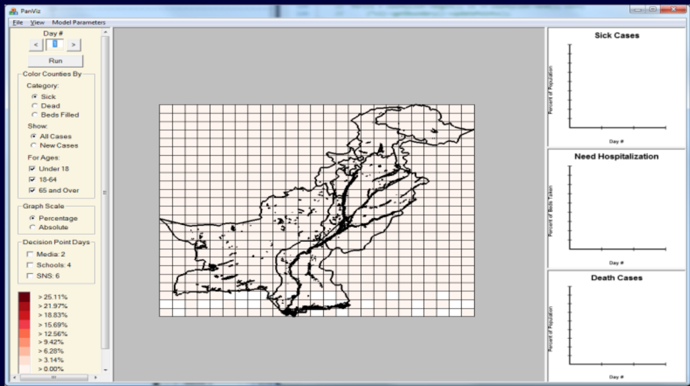
Pandemic spread on day 37 with all decision measures implemented



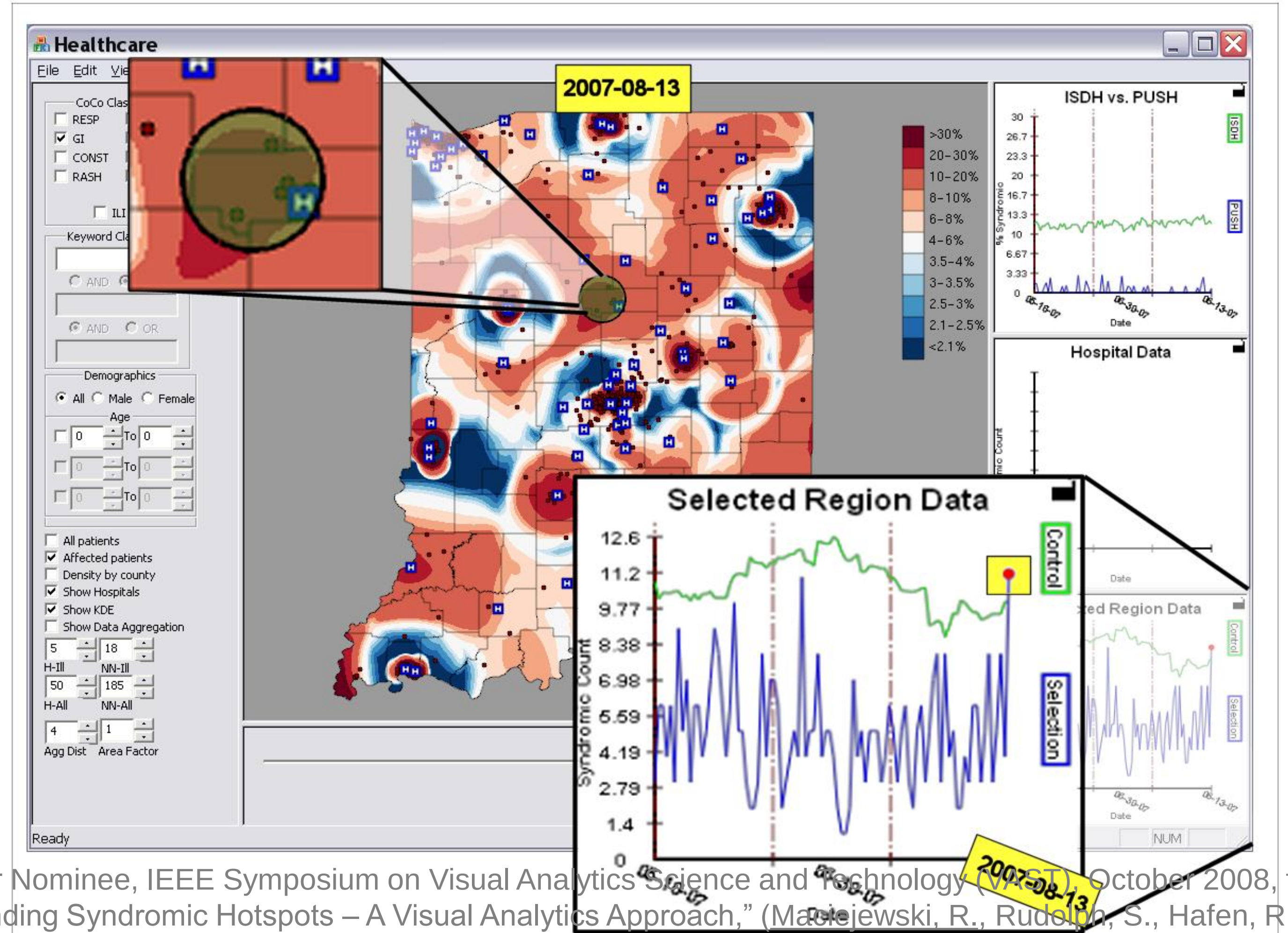
Visual Analytics For Syndromic Surveillance

- Syndromic surveillance: detection of adverse health events focusing on pre-diagnosis information to improve response time
- Pre-diagnosis information can consist of multiple data sources:
  - Over the counter medicine sales
  - News reports on emerging diseases
  - Pro-med news feeds
  - Emergency department chief complaints





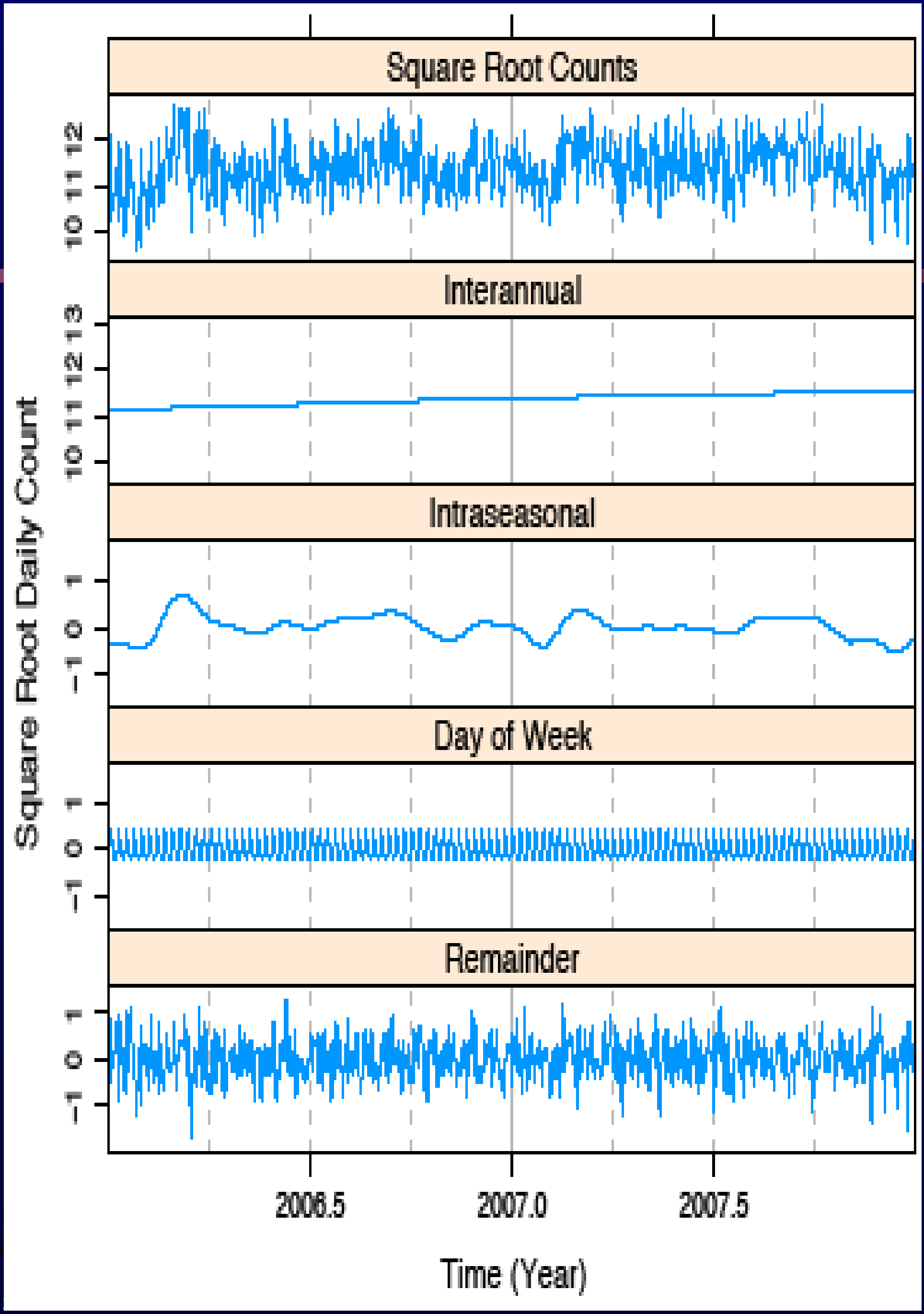
Hypothesis Generation and Exploration



Best Paper Nominee, IEEE Symposium on Visual Analytics Science and Technology (VAST), October 2008, for "Understanding Syndromic Hotspots – A Visual Analytics Approach," (Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W., Grannis, S., Wade, M., Ebert, D.).

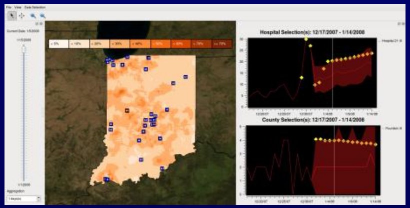
Time Series Modeling

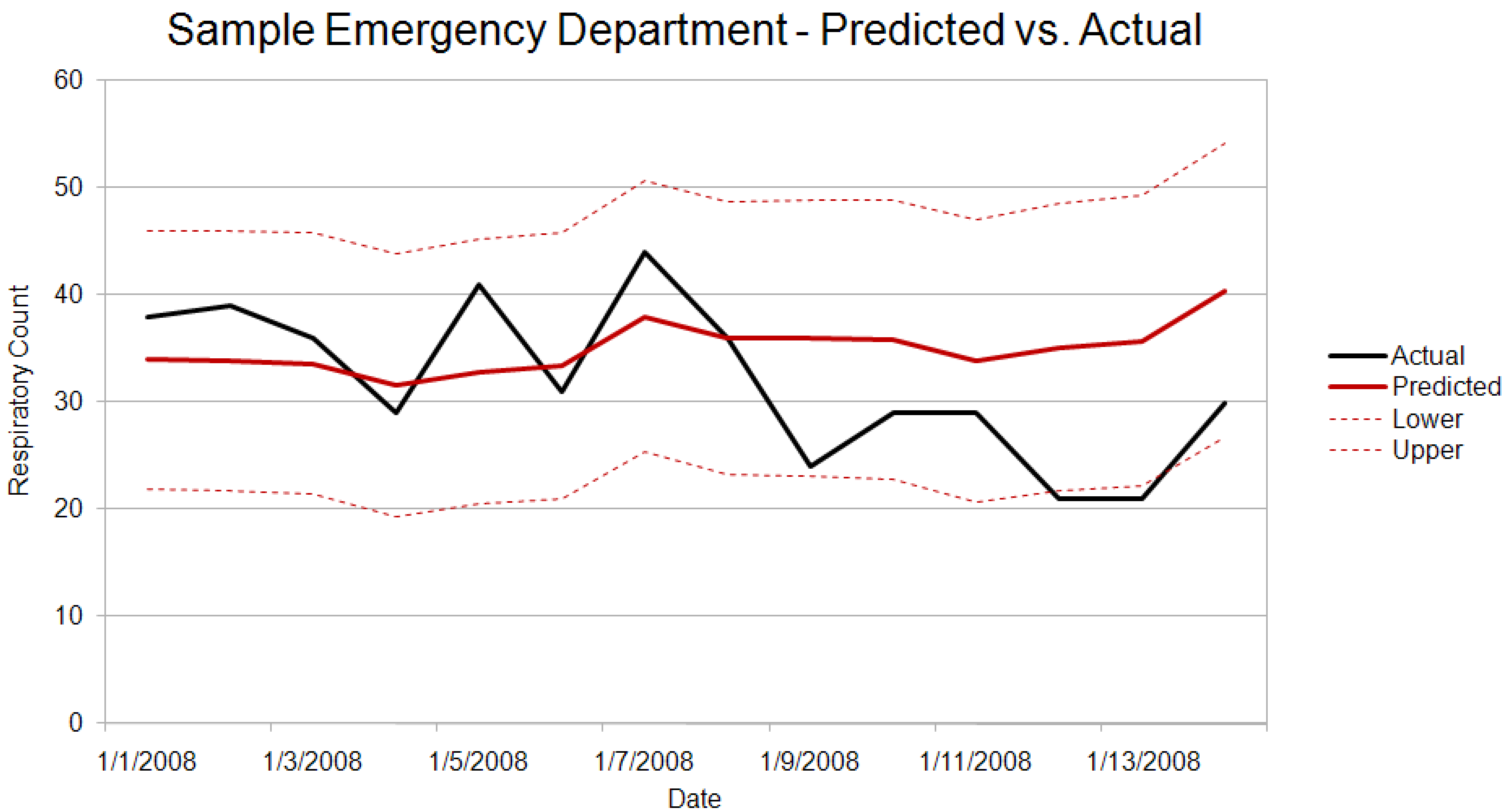
- Seasonal-Trend Decomposition Based on Loess
  - Time series can be viewed as the sum of multiple trend components
  - For each data signal, components are extracted
  - Can then analyze correlation between components



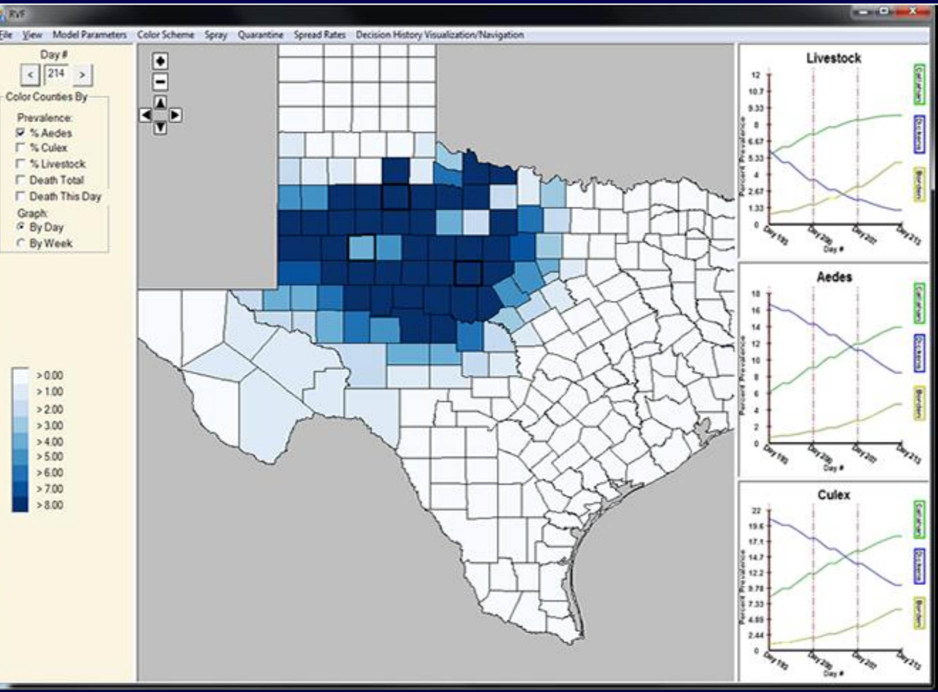
Hafen, R., Anderson, D., Cleveland, W., Maciejewski, R., Ebert, D., Abu
Visualizing, and Monitoring Disease Counts," *BMC Medical Informatics a*

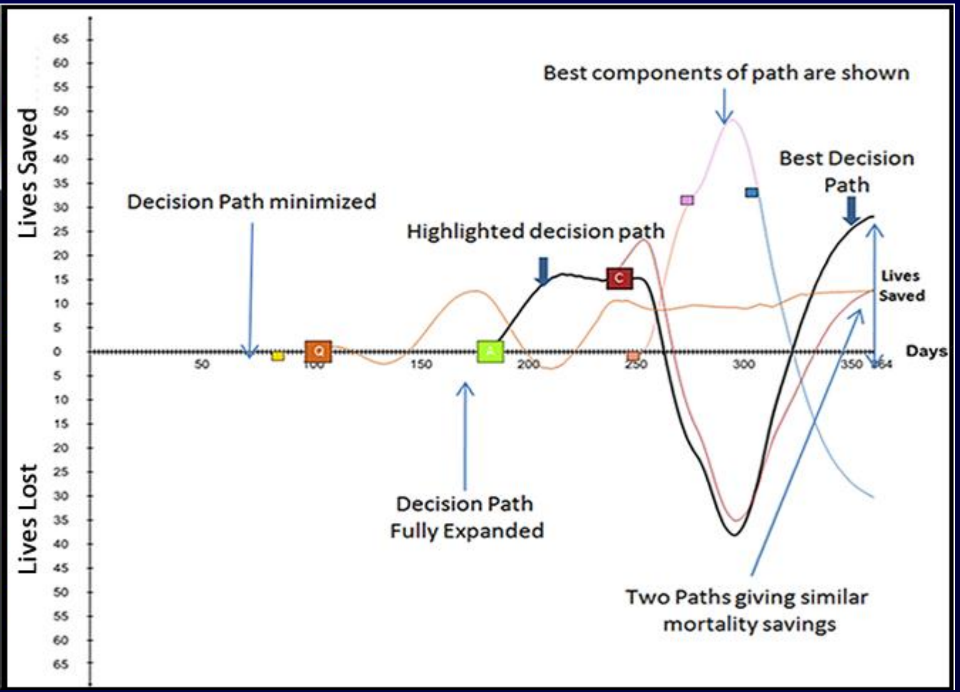
Predictive Visual Analytics





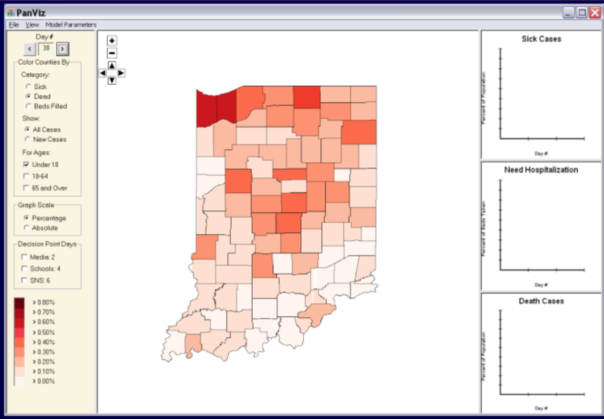
Advanced Decision Support Tools: Rift Valley Fever



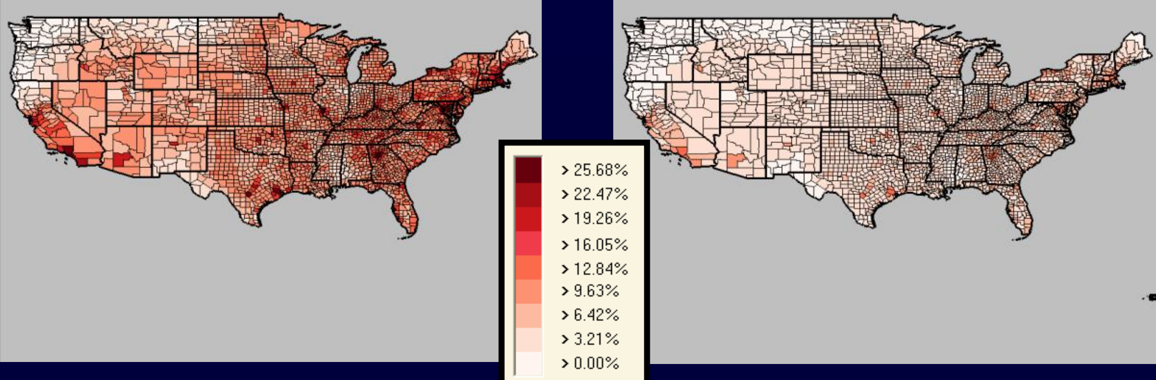


Modeling a Pandemic

- Pandemic Influenza Planning Tool
- Models user specified:
  - Pandemic influenza characteristics
  - County population, demographics, hospital beds
- Decision measures
  - Strategic National Stockpile deployment
  - School Closures
  - Media Alerts



Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.



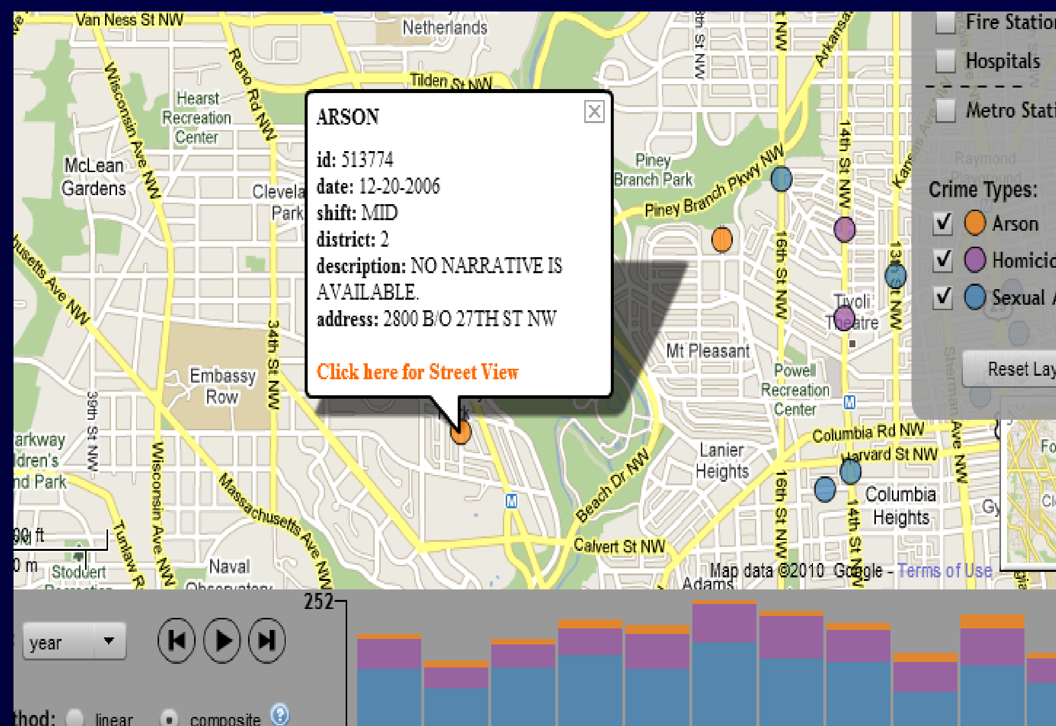
Pandemic spread on day 37 with no decision measures implemented

Pandemic spread on day 37 with all decision measures implemented

VACCINE

Visual Analytics Uses for Crisis Management and Response

- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis
- Purpose:
  - Planning for resiliency
  - Long-term analysis
  - Predictive analytics
  - Training
  - Detection
  - Investigation
  - Response
  - Recovery, remediation



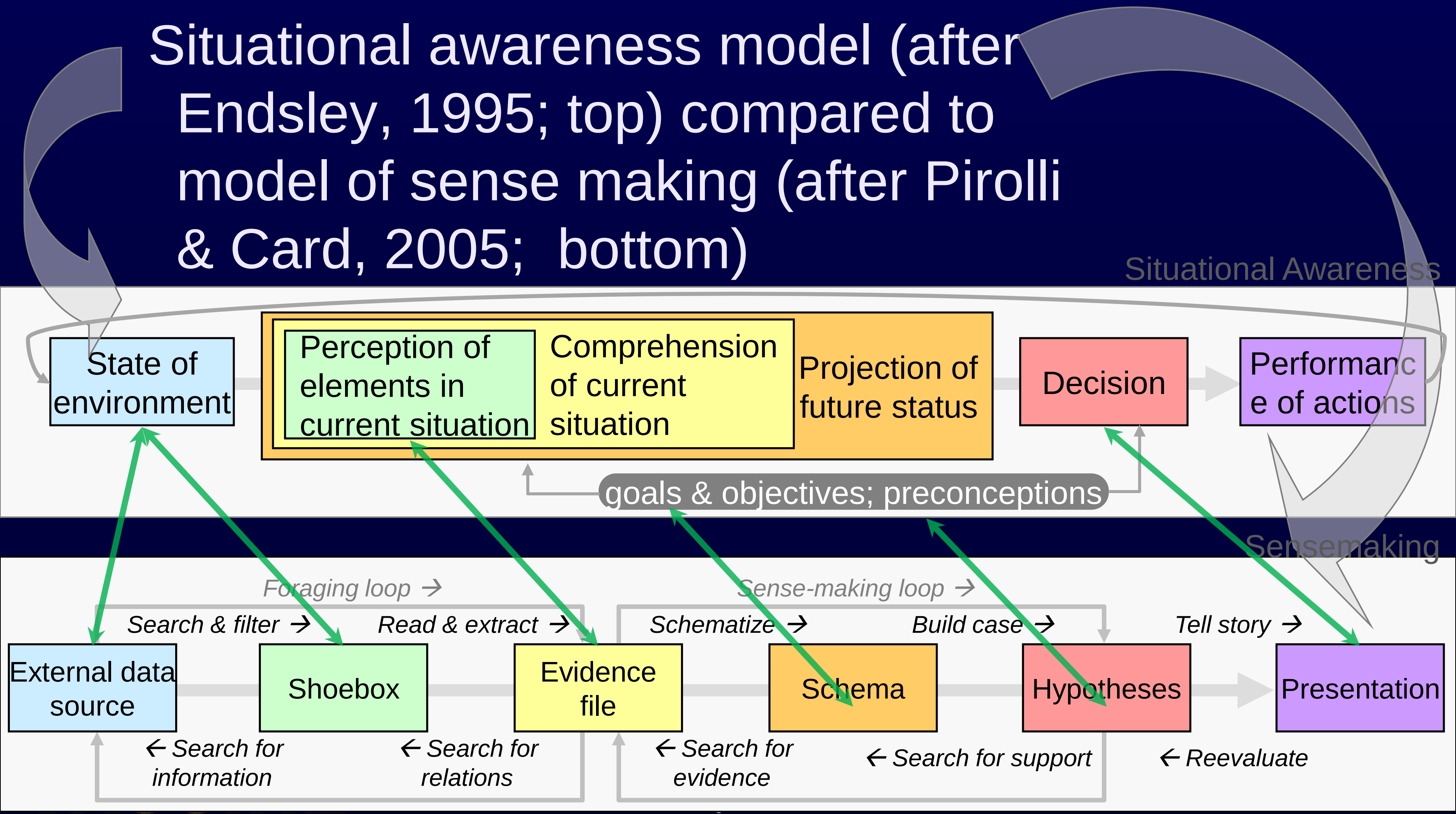
VACCINE

Visual Analytics vs. Situational Awareness

Situational Awareness ←→ Sensemaking

situational awareness as a sensemaking process (courtesy Alan MacEachren)

Situational awareness model (after Endsley, 1995; top) compared to model of sense making (after Pirolli & Card, 2005; bottom)



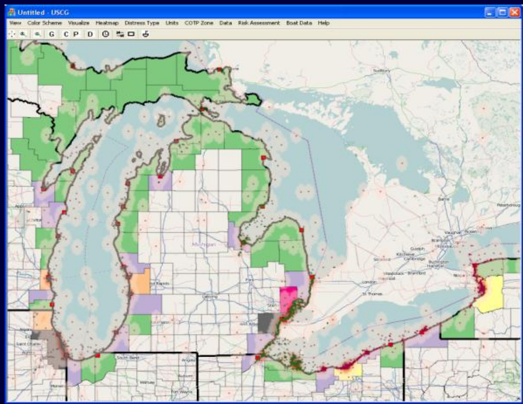
Situational Awareness, Planning, Investigation, Response Visual Analytics

- Situational Awareness and Assessment from massive data
  - In-field and desktop historical and current incident/crime information
  - Innovative spatiotemporal visualization, analysis, prediction
  - Image analysis tools
  - Integrated statistical forecasting and simulation models
  - Advanced social media analytics
- Tools for Intelligence analysis
- Planning and response support
  - Integrated simulation, resource allocation, census, weather data

VACCINE

Visual Analytics Uses for Public Safety

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia massive data integration & analysis



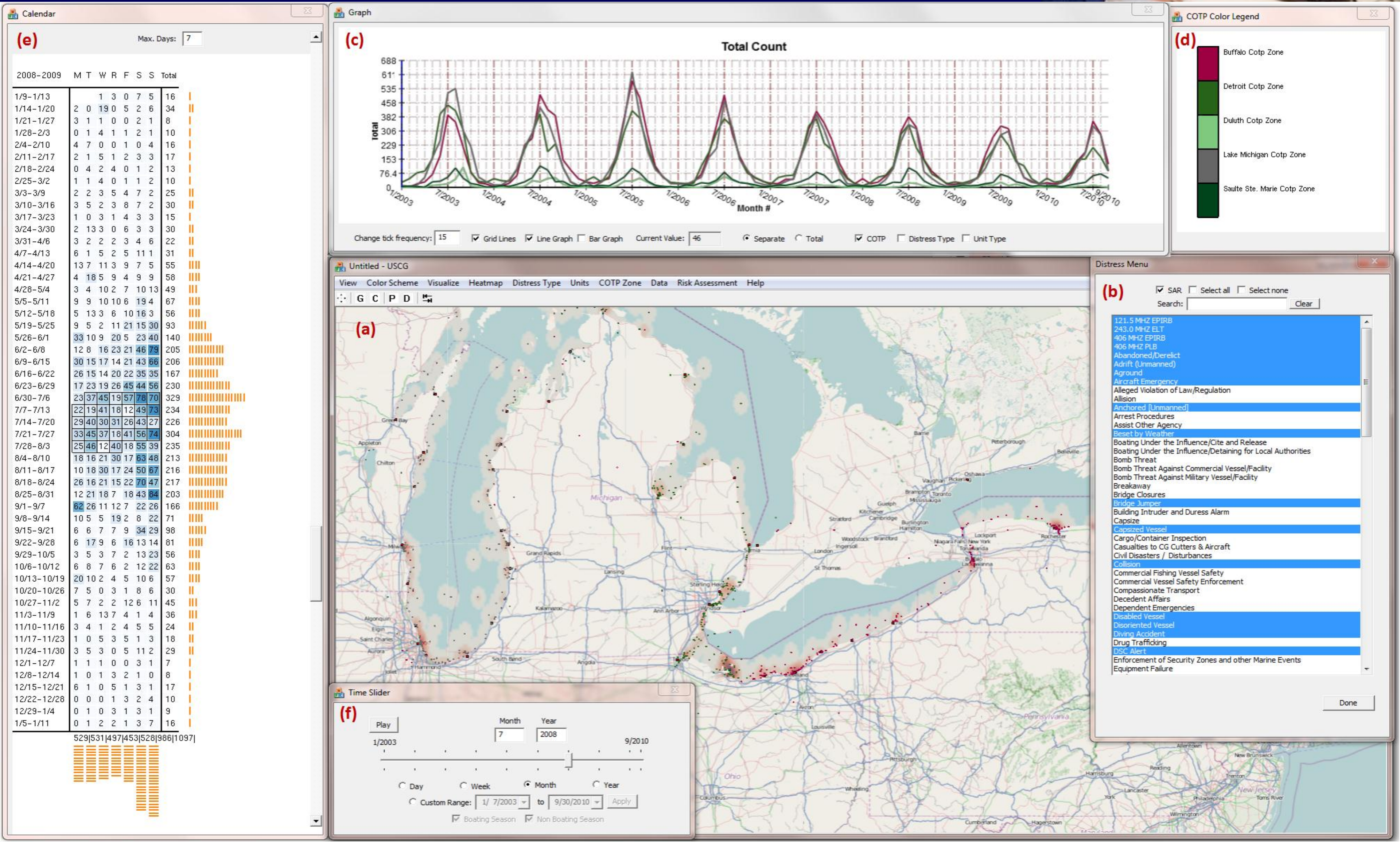
VACCINE

Example Maritime Safety Visual Analytics Environment

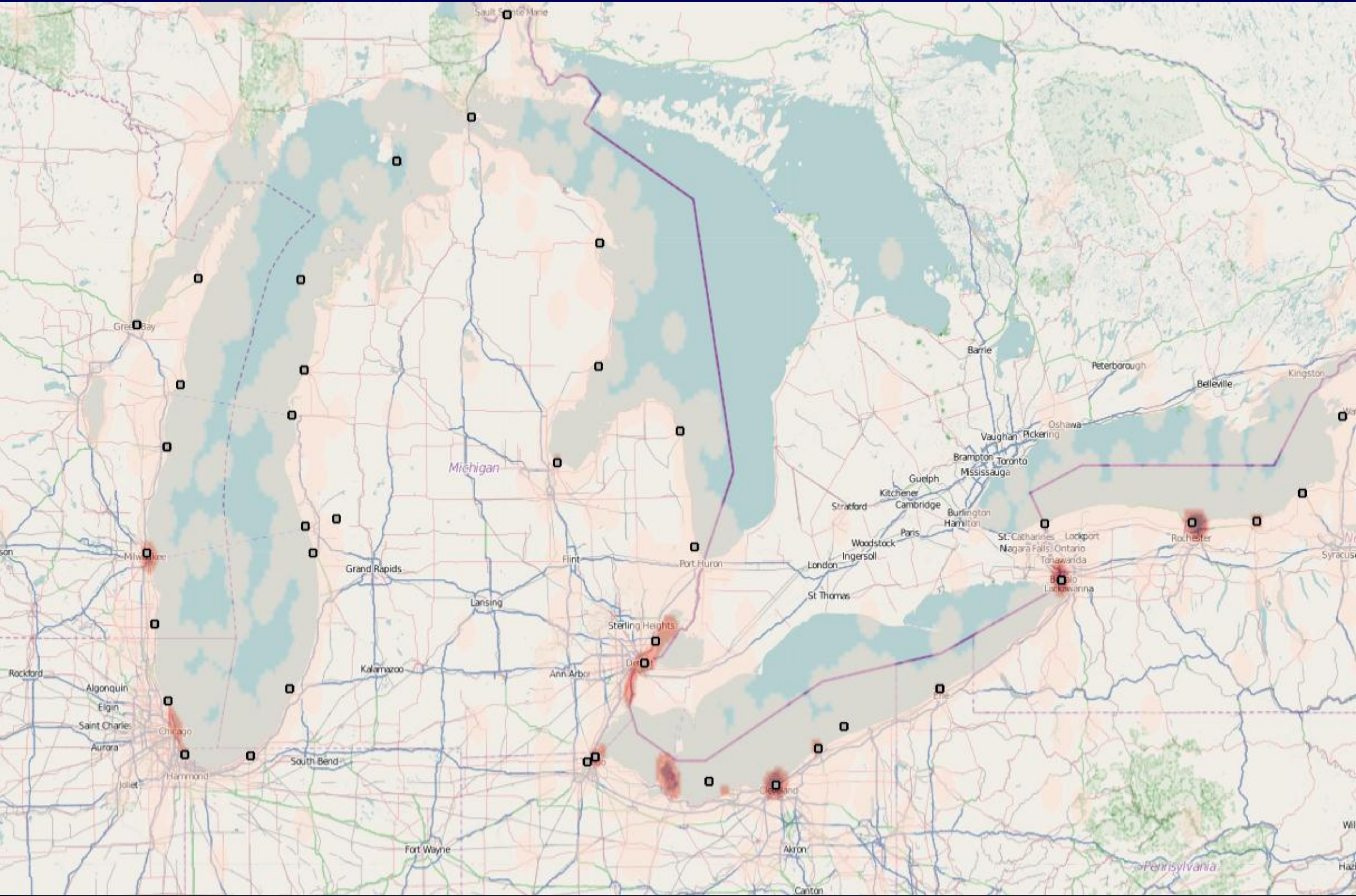
- Supports decision making and risk assessment
- Interactive exploration and analysis of trends, patterns and anomalies
- Allows analysis of risks associated with
  - Closing one or more Coast Guard stations
    - Find optimal stations that absorb work load of the closing station
  - Allocating new resources
    - Impact on safety and efficiency of operation
- Currently being used by analysts at the U.S. Ninth District, Fifth District, HQ, and Atlantic Commands

VACCINE

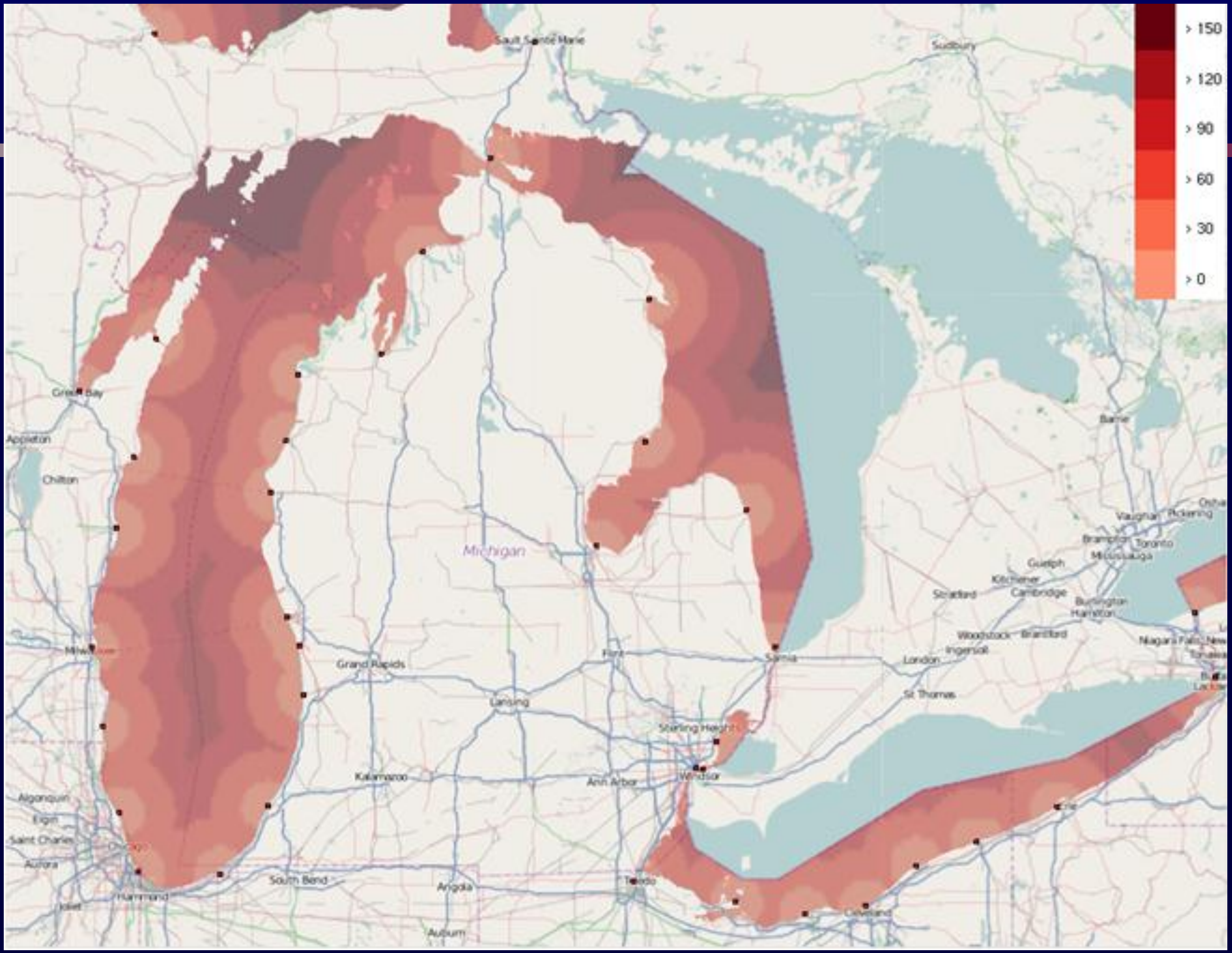
Example: USCG D9 Search And Rescue Operational Analysis



Total Risk Based on SAR Cases

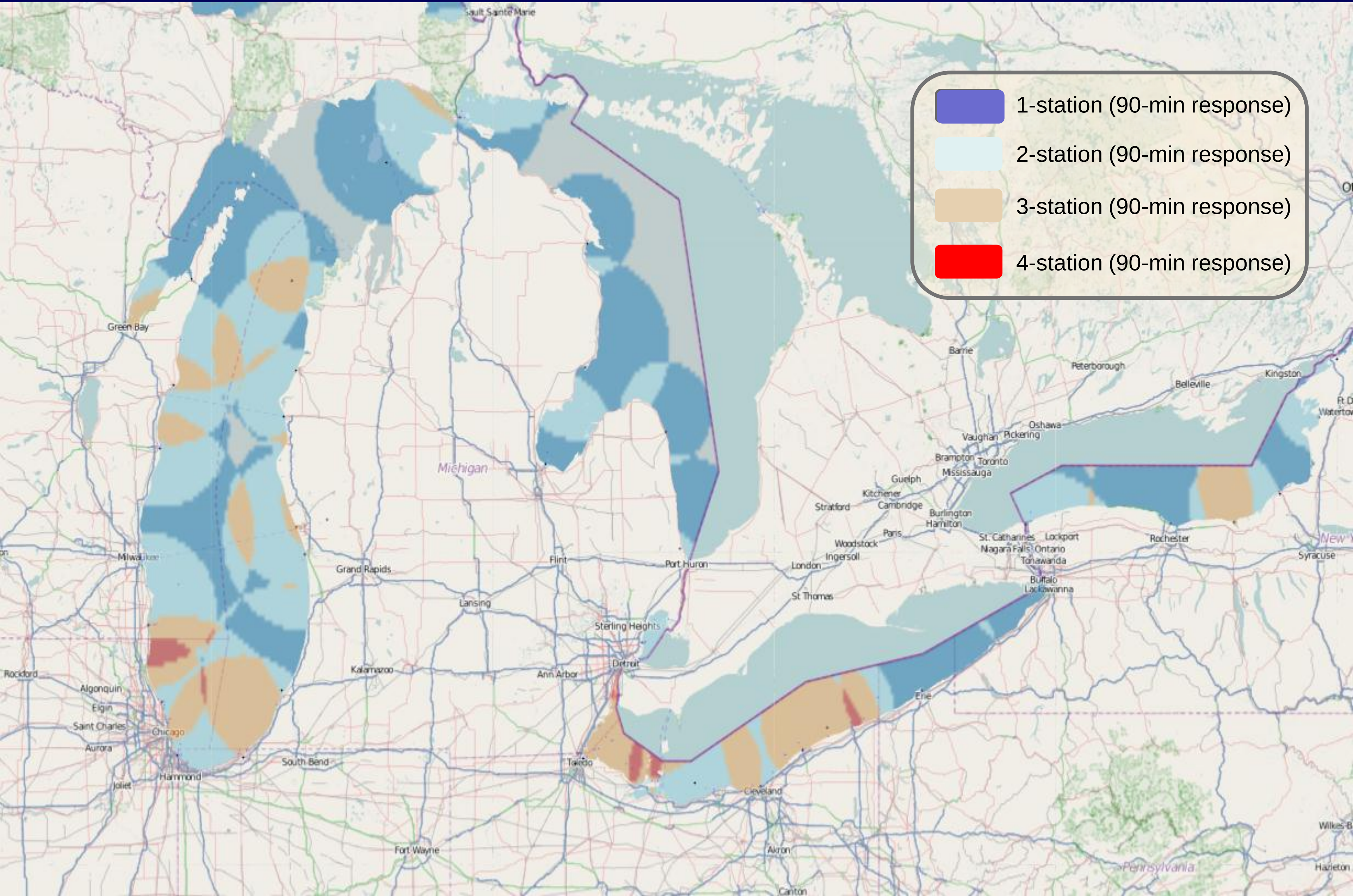


System Features: SAR Risk Profile

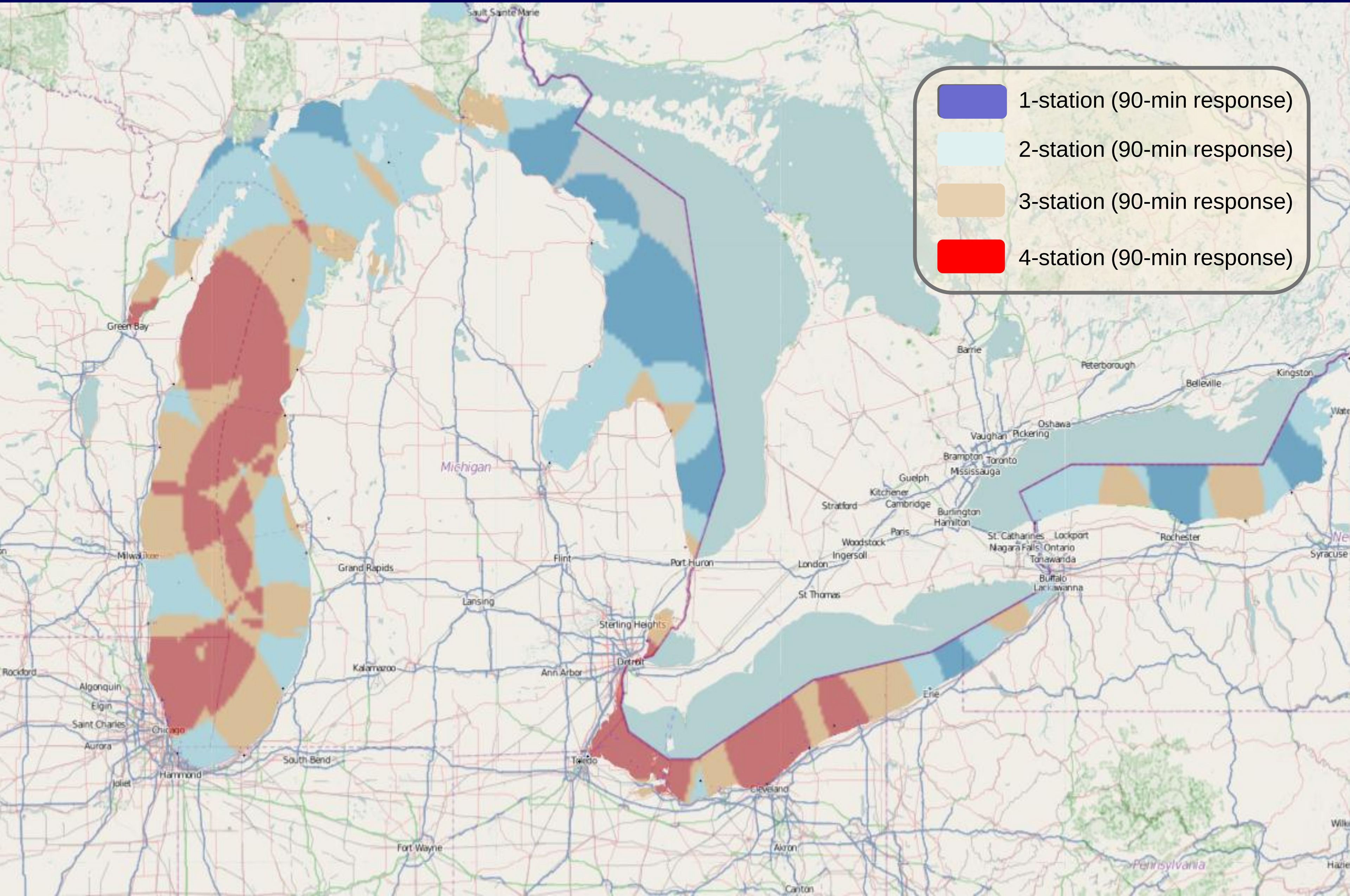


Time taken by CG stations to deploy an asset to the Great Lakes to respond to a SAR incident.

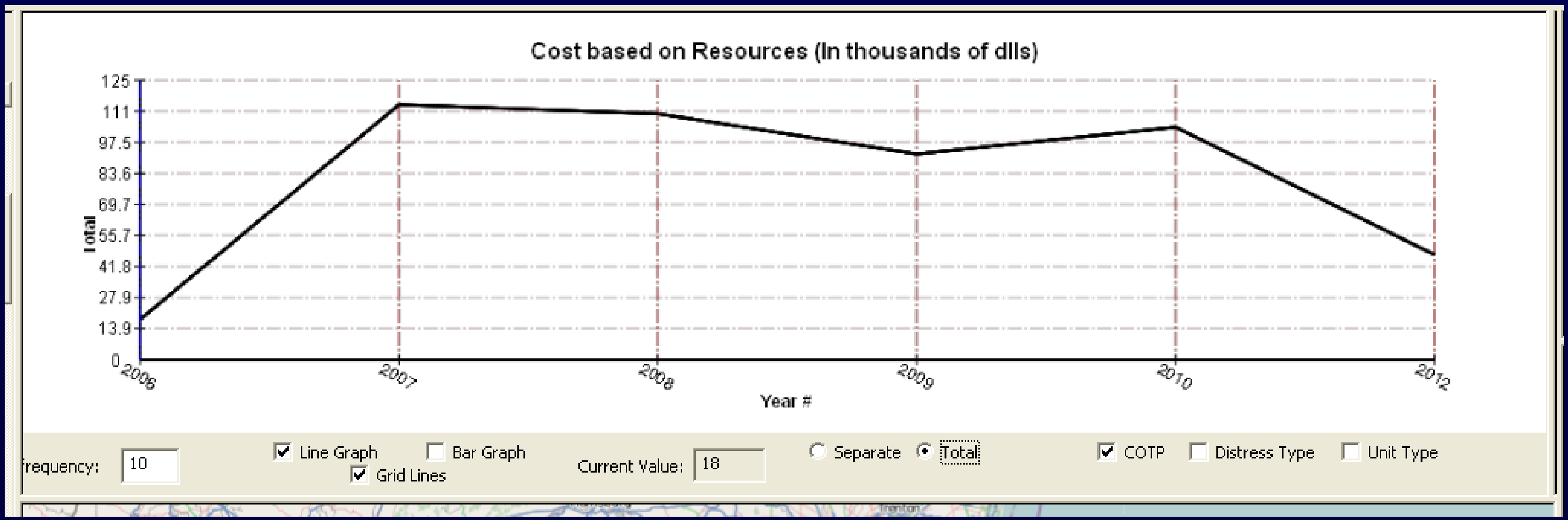
Response Efficiency – Current Assets

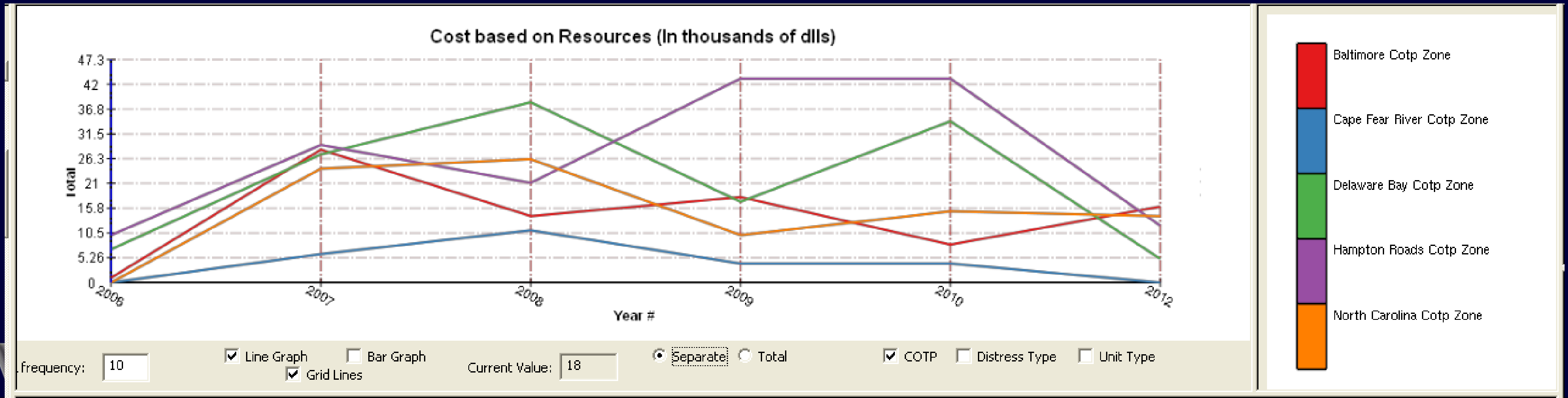


Response Efficiency – Potential Future Assets



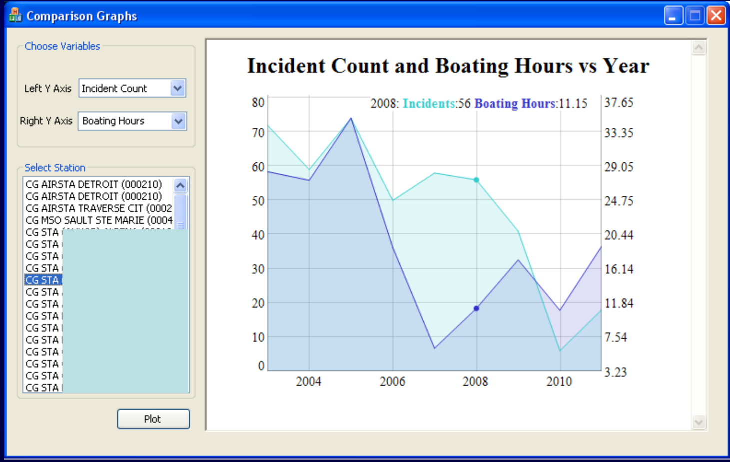
D5: Cost based on Resources used for the Open-Suspended Sorties over 5 years (2006-2011)

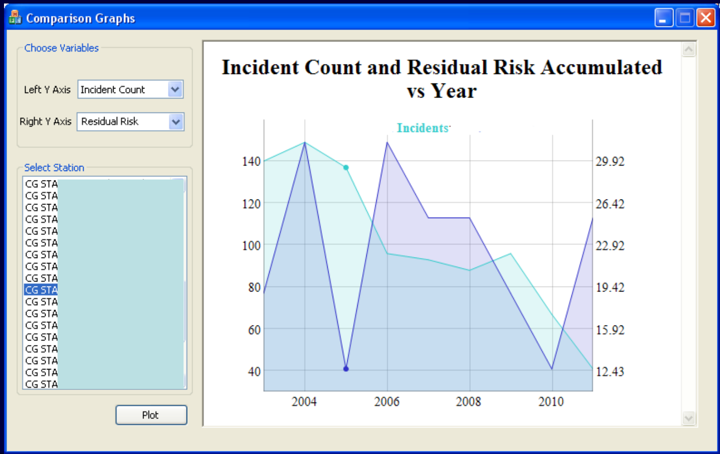




Resource Allocation and Risk-Based Decision Making

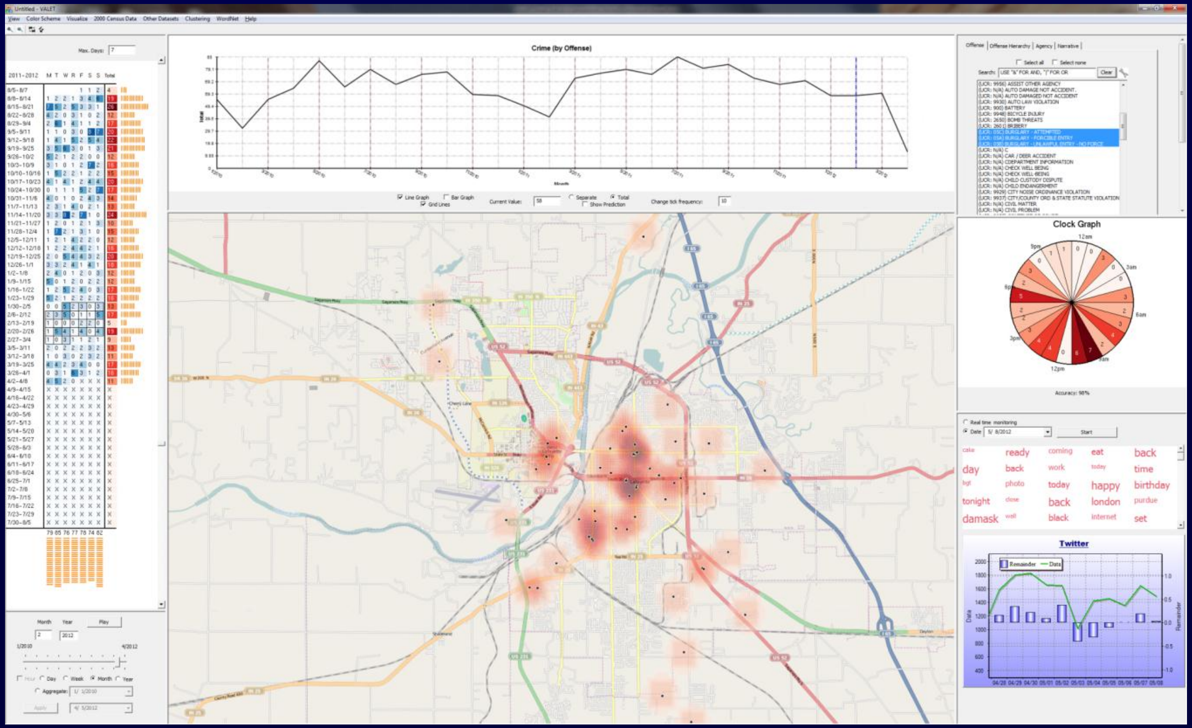
- Explore risk-based decision making and utilize historical data for analysis and prediction
  - Total Risk, Mitigated Risk, Residual Risk
  - Explore 11 different USCG missions
  - Explore allocation of assets with different capabilities
  - Explore staffing, utilization, assets vs. risk measures
  - Perform What-If scenarios

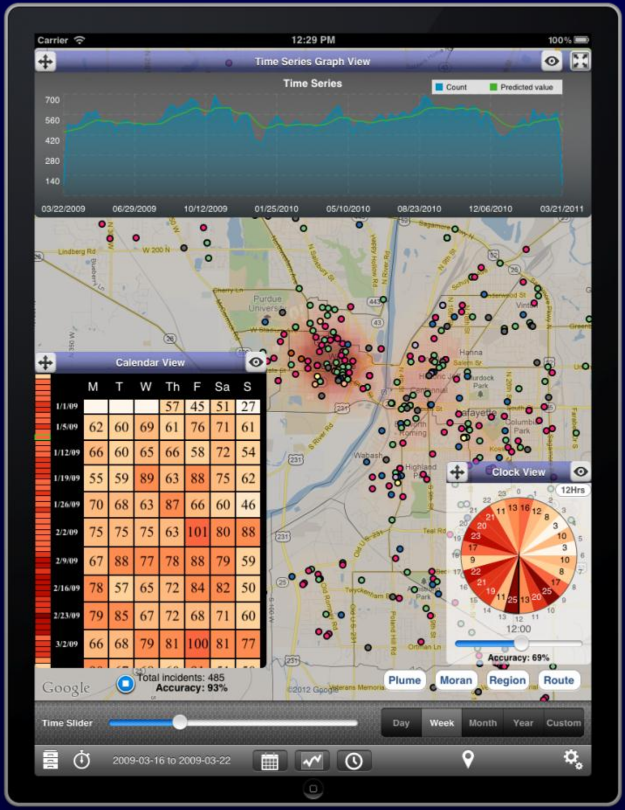




VACCINE

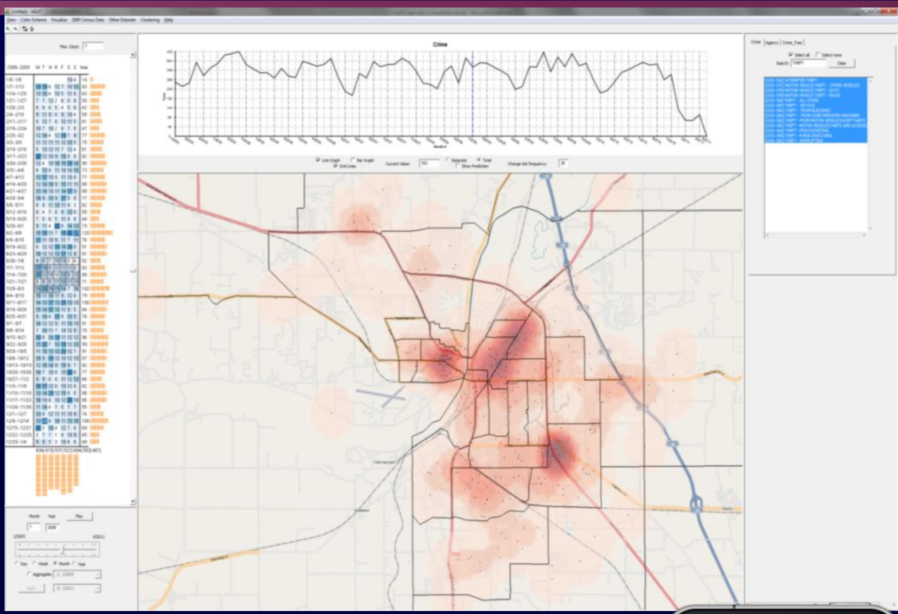
Visual Analytics Law Enforcement Toolkit
(i)VALET





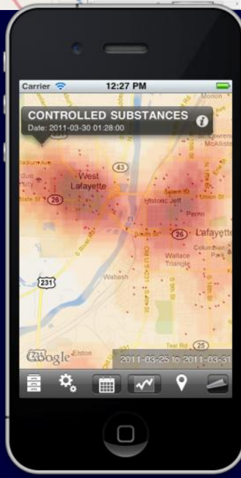
VACCINE

Visual Analytics Law Enforcement Toolkit (VALET, iVALET)



Interesting VA features:

- Time a first class citizen
- Temporal exploration and comparisons
- Integrated predictive models
- Interactive multivariate correlation and visual analytics
- Novel statistical predictive model incorporated for planning
- Novel interactive anomaly detection from social media



VALET delivered:

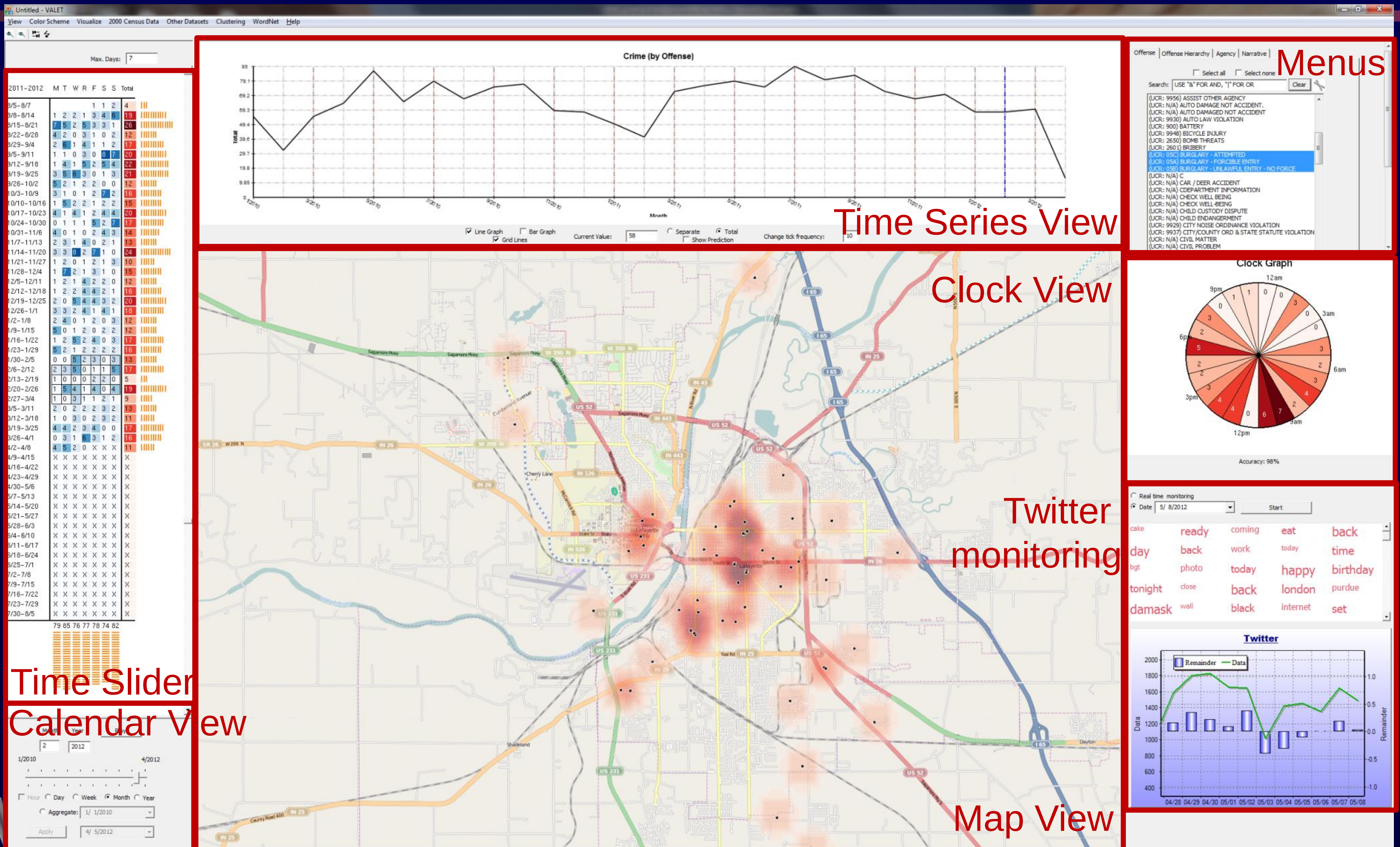
- Spring 2011: WL, Lafayette Police

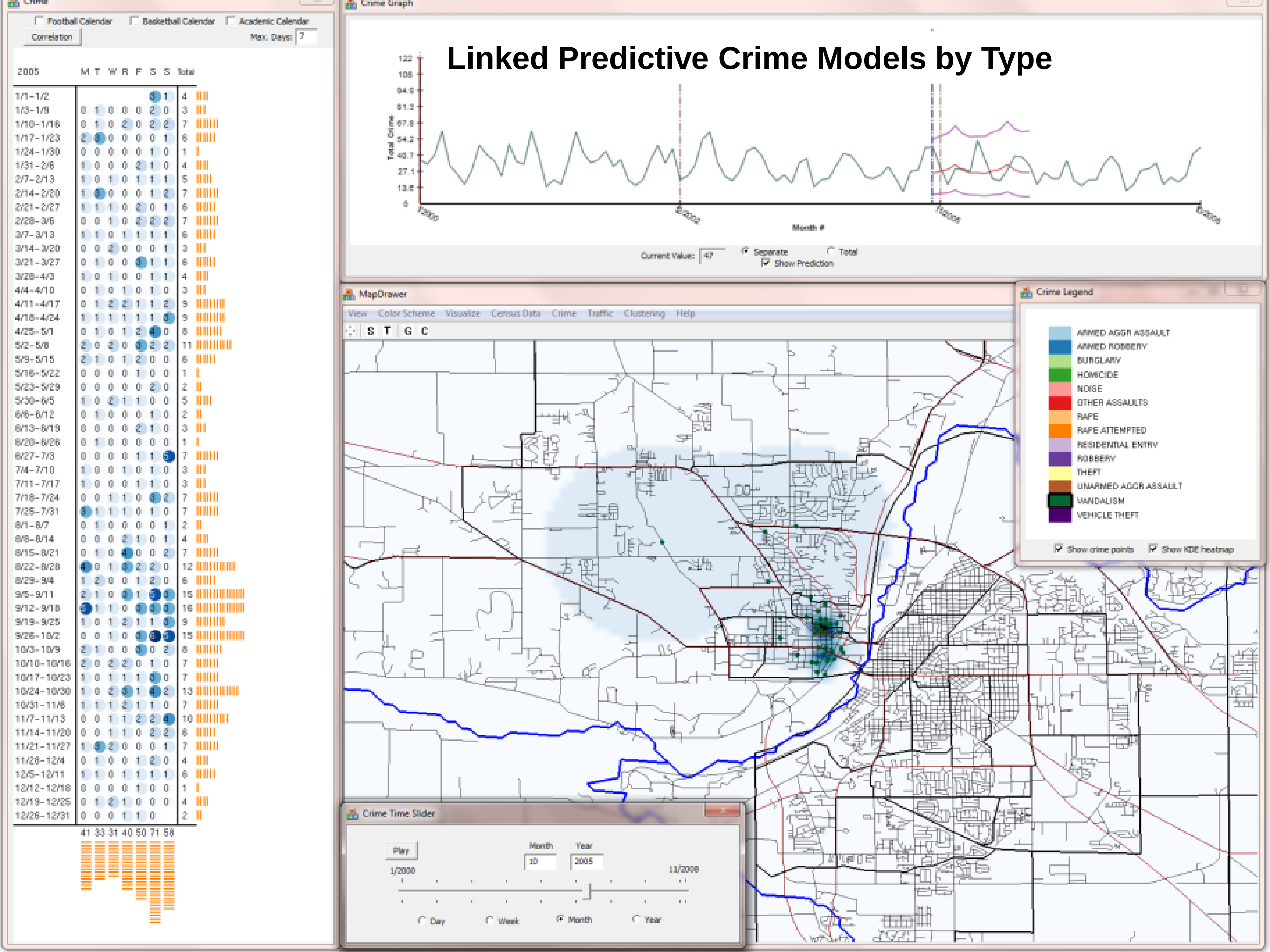
iVALET delivered:

- October 2011: Purdue, WL Police

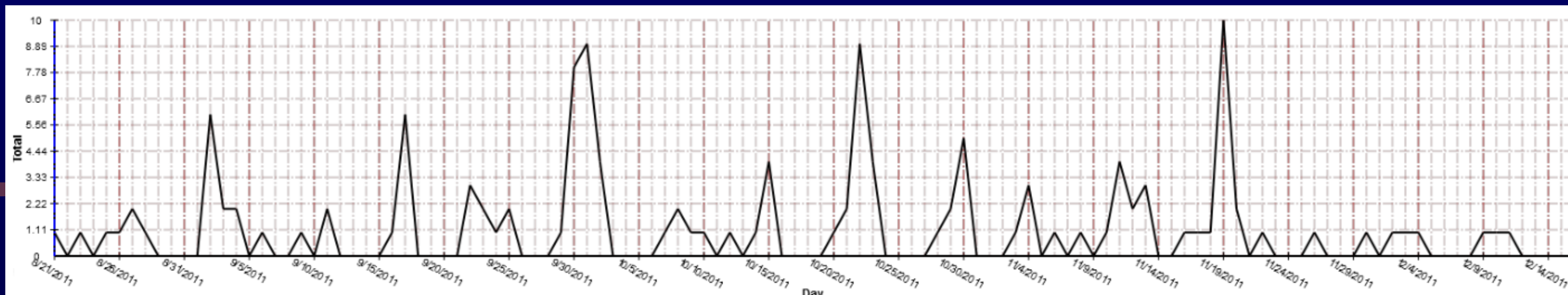
VACCINE

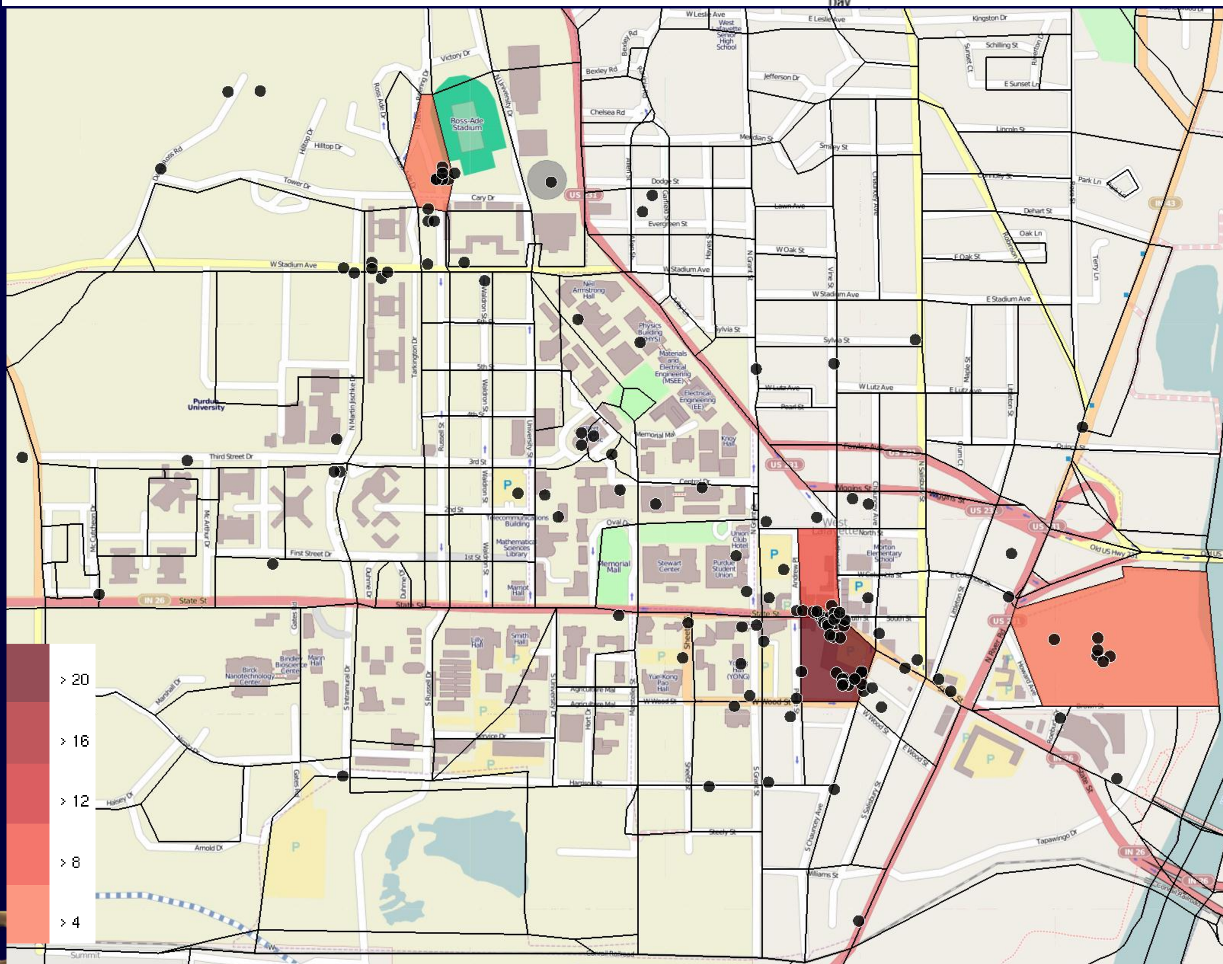
VALET Overview

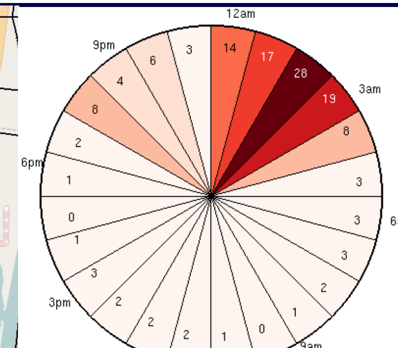



Linked Predictive Crime Models by Type
Crime
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
Max. Days: 7
2005
M T W R F S S Total
Crime Graph
Total Crime
Month #
Current Value: 47
Separate
Total
Show Prediction
MapDrawer
View Color Scheme Visualize Census Data Crime Traffic Clustering Help
S T G C
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Show crime points
Show KDE heatmap
Crime Time Slider
Play
Month
Year
10
2005
1/2000
11/2008
Day
Week
Month
Year

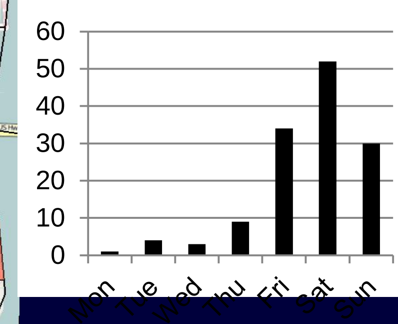
Example: Drunkenness / Public Intoxication



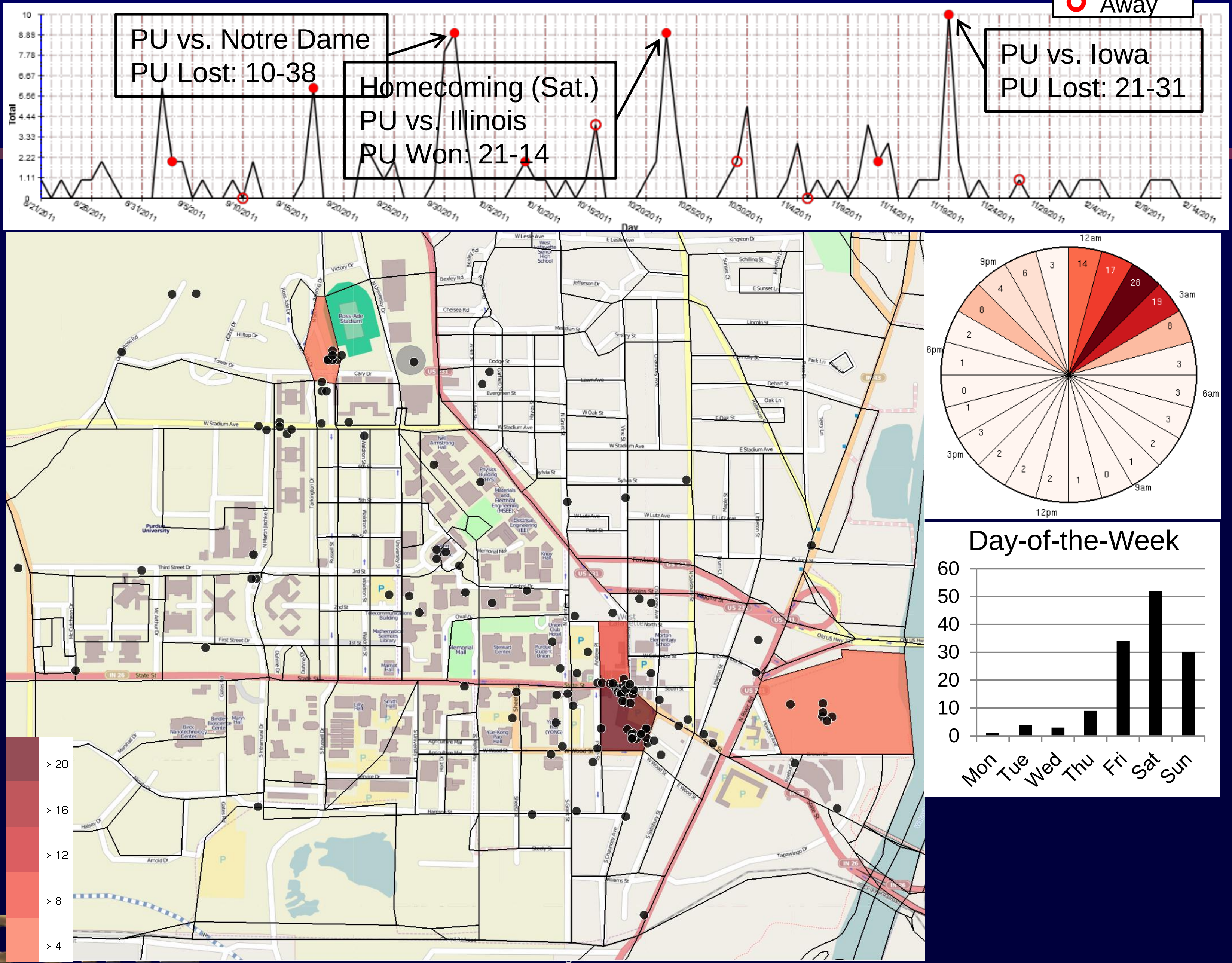




Day-of-the-Week



Example: Drunkenness / Public Intoxication

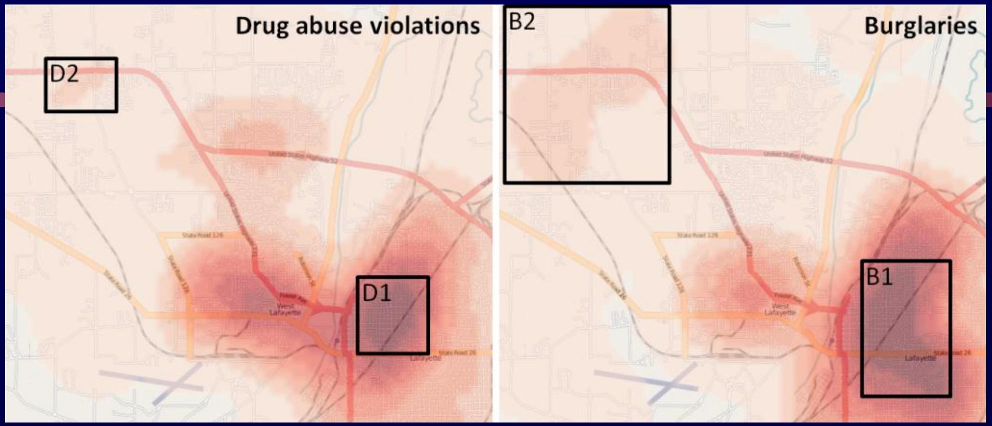


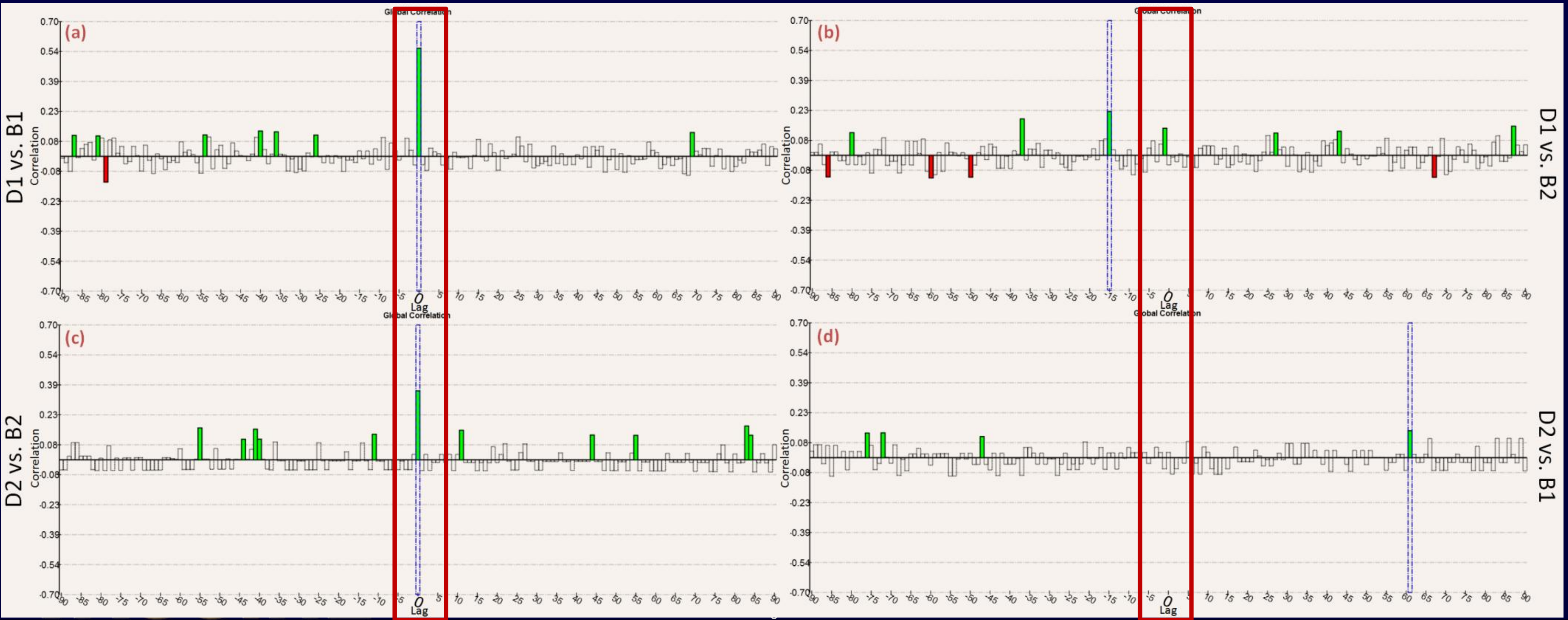
VALET Feature:
Correlating Datasets at Multiple Granularities

- Correlative analysis framework for exploring correlations among multi-variate/-source data
- Determining forces that determine nature of relationships among datasets
- Detection of periodic properties among different data variables
- Allows correlations at multiple spatial and temporal granularities

Case Study: Drug Abuse Violations vs. Burglaries

- Over geospatial hotspots



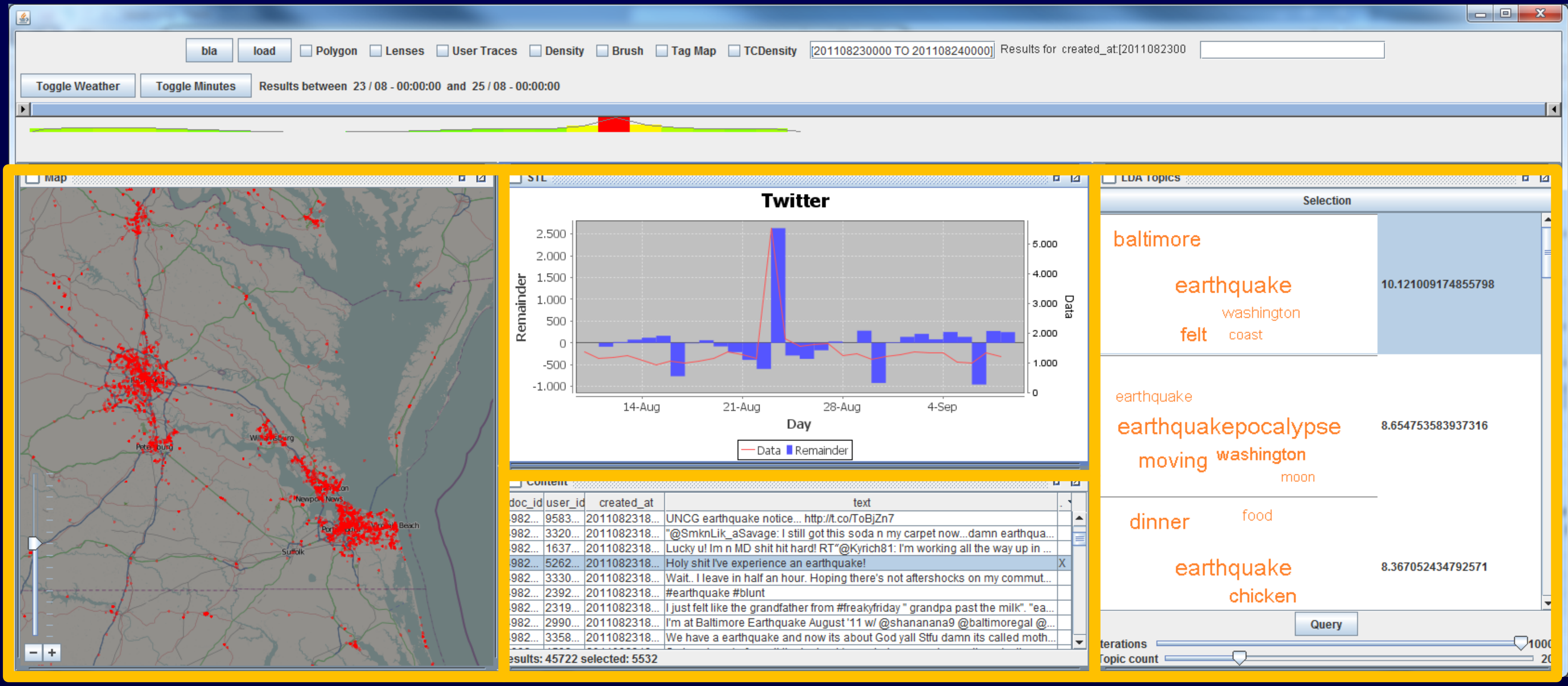


VALET Feature:
Real-time Twitter Monitoring and Integration

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud

Grand Prix Weekend, Purdue University

Spatiotemporal Social Media Analytics for Abnormal Event Detection

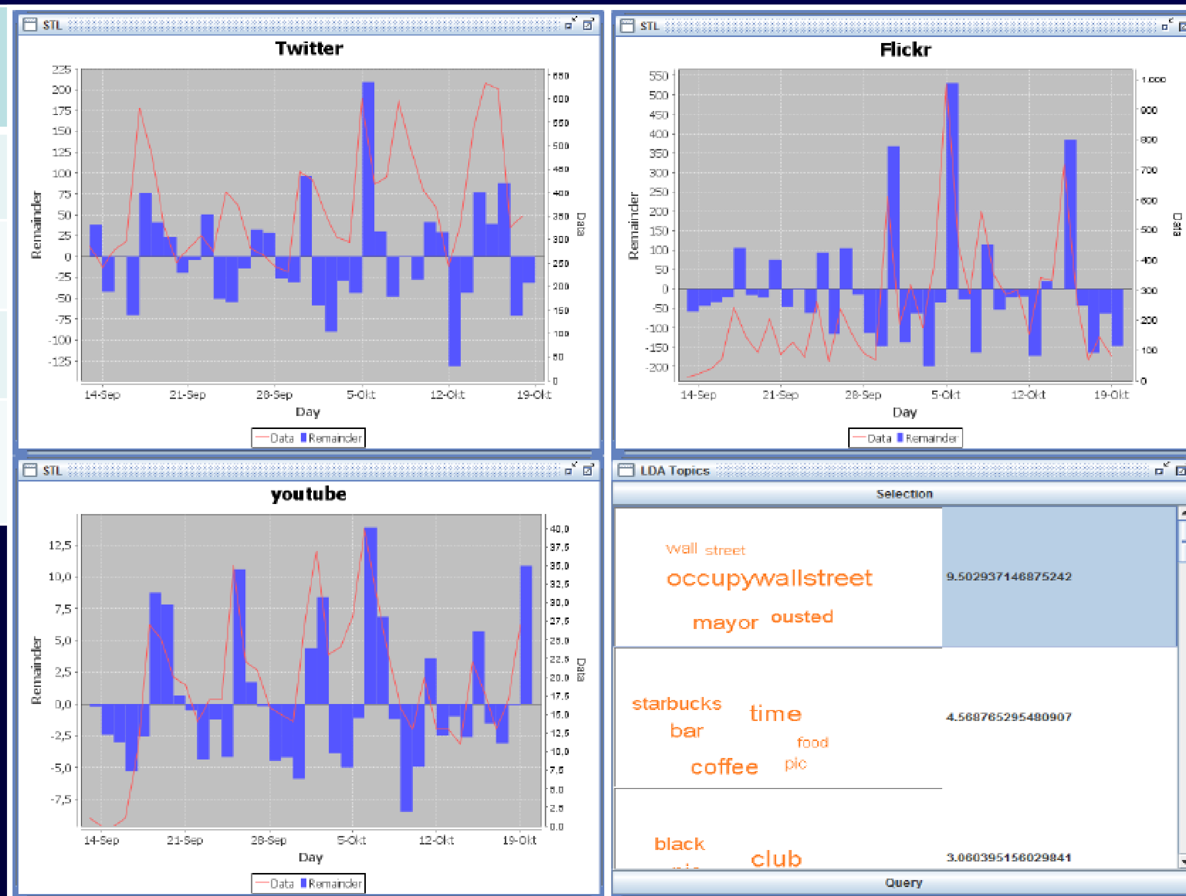


VACCINE

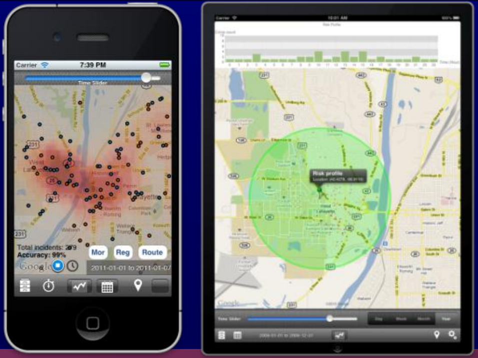
Correlation Between Multiple Social Media

- Occupy Wall Street event

Date	Twitter (z-score)	Flickr (z-score)	Date	YouTube (z-score)
9/17, 2011	2.53	0.97	9/18, 2011	5.98
9/30, 2011	3.20	3.50	10/2, 2011	1.38
10/5, 2011	3.17	3.36	10/6, 2011	2.75
10/15, 2011	4.90	4.53	10/15, 2011	2.19



iVALET

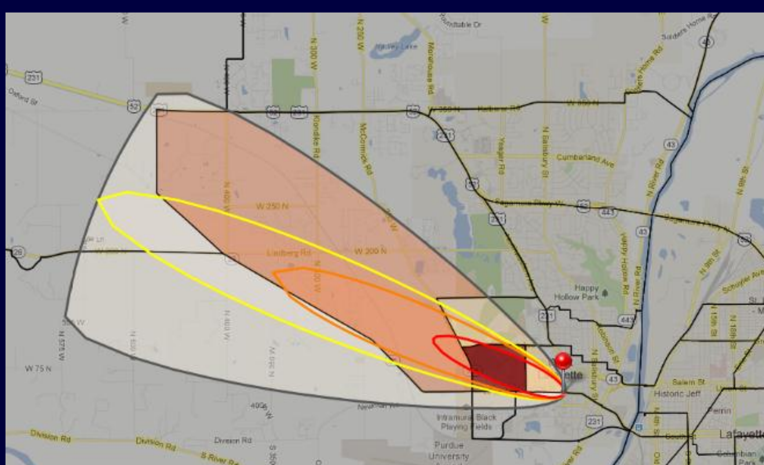


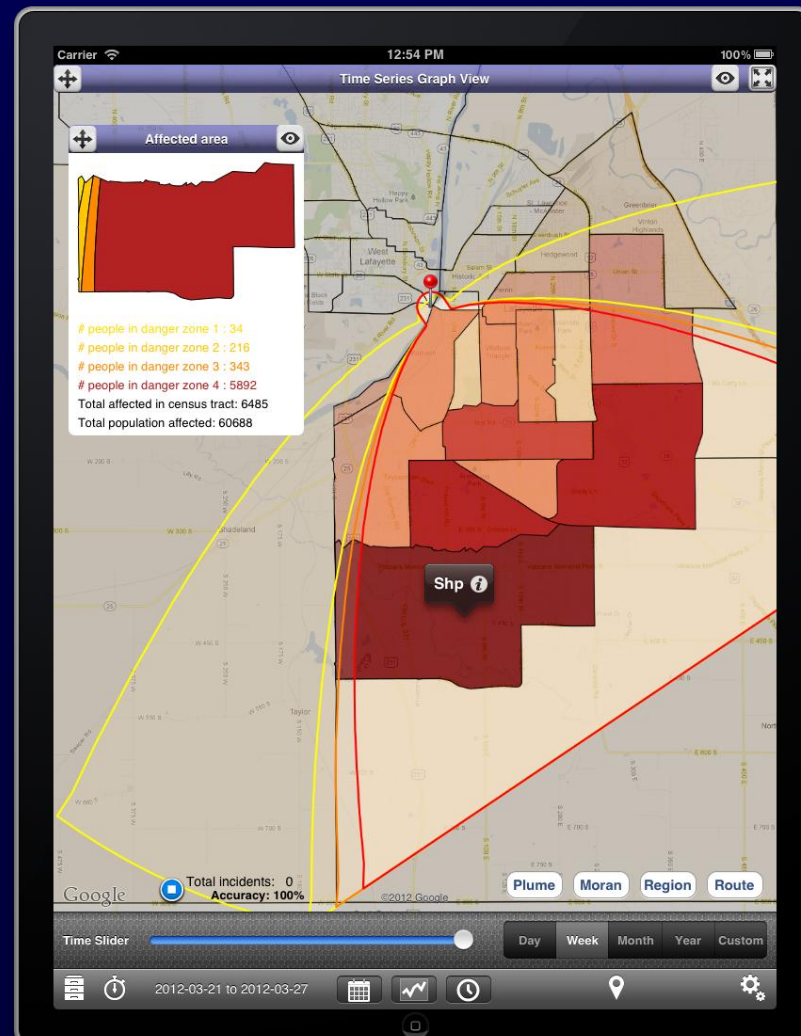
- Explore criminal, traffic and civil data on-the-go
- Risk assessment
- Use current spatial + temporal context into analysis

VACCINE

MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

- Chemical release plume modeling identifies census tracts with the highest number of expected people affected





VACCINE

Jigsaw:
Visual Analytics for Investigative Analysis and Exploration of Document Collections

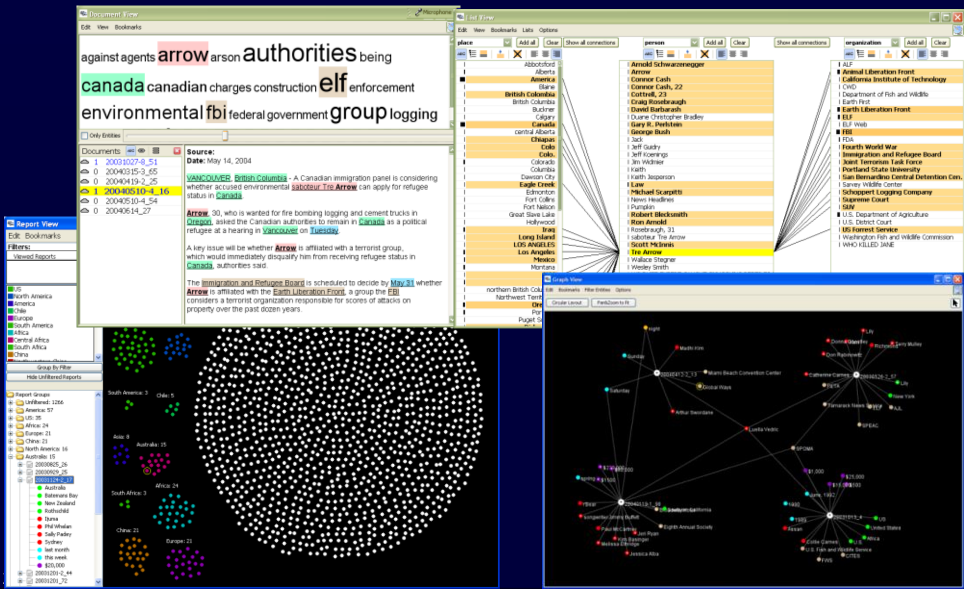
"Putting the pieces together"

Goal: Assist investigators with understanding, sense-making, and analysis of large, unstructured and structured document collections

Approach: Provide multiple visual perspectives on the documents and entities within them, highlighting connections between entities

Users: Over 150 downloads – active users include the following:

- Seattle PD
- Rock Hill, South Carolina PD
- Indianapolis PD
- West Lafayette, PD (student embedded)
- TSA



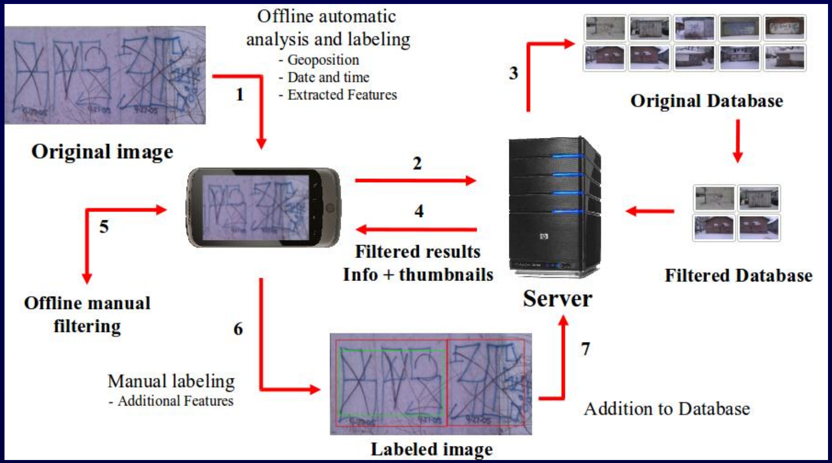
VACCINE

Gang Graffiti Recognition and Analysis Using a Mobile Telephone (GARI)

Edward J. Delp, Mireille Boutin

Collaborating Institution(s): Purdue University

End-User(s): Gary Coons, Chief/Indianapolis Department of Public Safety Division of Homeland Security and IMPD, Indiana Fusion Center Gang Task Force



IMPACT:

- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be "interpreted"
- More than 40 users and 450 graffiti images acquired

GARI delivered:

- Summer 2011:
  - IMPD gang detectives
- August 2011:
  - IMPD at large
  - Ind Fusion Gang Task Force
- October 2011:
  - Gang detectives across Indiana

VACCINE

Crime Analytics & SA Issues and Techniques

- Fuse data from a variety of sources
  - Law enforcement records management
  - Weather and phases of the moon
  - Street light locations, bus routes
  - Tracking release data of offenders
  - Civil court data
  - Social Media
  - Local event calendar
- Reliable predictive models
- Understandability and trust of predictions
- Main Question: What helps officers, detectives, chief do their variety of jobs?

Directions Forward

-

Challenges

Six Challenges for Proactive & Predictive Visual Analytics

1. Creating computer-human visual cognition environments
2. Integrating interactive simulation and analytics
3. Solving specific scale issues and cross-scale issues
4. Managing uncertainty and time
5. Enabling risk-based decision making environments
6. Developing the Science of Interaction for Visual Analytics

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu

ebertd@purdue.edu

VACCINE

What We Do With Visual Analytics

- Enable effective decision making through interactive visual analytic environments
- Enable effective communication of information
- Provide quantitative, reliable, reproducible evidence
- Enable user to be more effective from planning to detection to response to recovery
- Enable proactive and predictive visual analytics
- Enable effective situational awareness (perception, comprehension, projection)

What's Needed for Proactive and Predictive Visual Analytics?

Table of Contents

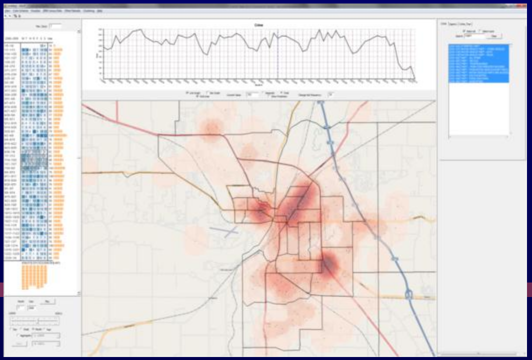
- Reliable and reproducible models and simulation
- Understanding of the data
  - Distribution and skewness, errors, appropriate analysis techniques
- Understanding of the sources and types of data
- Comparable or Correlative sources of data
  - Appropriate transformations applied to enable meaningful comparison and correlation
- Understanding of the use and problem to be solved!

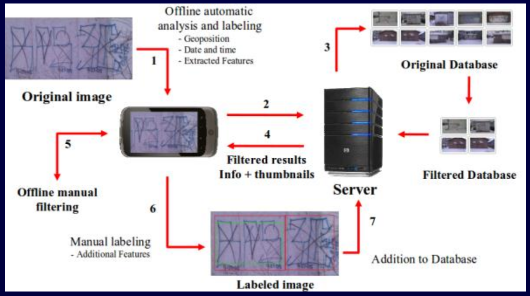
Innovation with Impact
Table of Contents
VACCINE
SFU, JIBC
UBC
UW
Stanford
Navajo Tech
UT Austin
UHD
JSU
Ind U
Purdue
Penn St.
VaTech
NC A&T
UNCC
GaTech
FIU
U Stuttgart
Swansea U
GeoDiscoverer
Selected Region Data
Theft Day time
Theft Night time

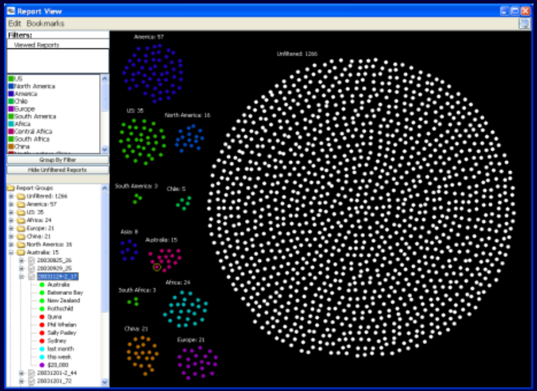
About VACCINE

- What we do
  - Help people make effective decisions from massive data
  - **Big Data Visual Analytics**
- Who we can help
  - Law enforcement
  - Public safety and health
  - Scientific discovery
  - Business production, resource allocation, risk-based decision making
- Types of technologies
  - Novel visualization techniques and analytics
  - Mobile graphics, volume graphics, procedural abstraction of complex and massive data

Law Enforcement



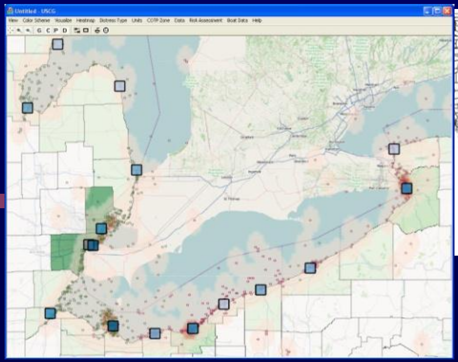


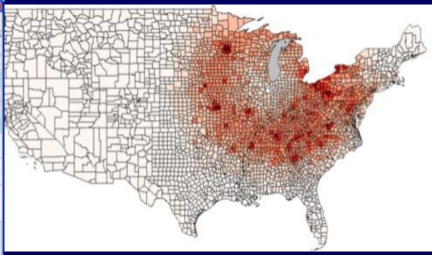


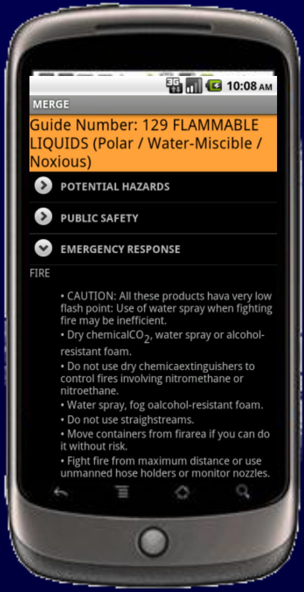
- Law enforcement visual analytics
  - VALET & iVALET (iPhone/iPad)
    - Visual Analytics Law Enforcement Toolkit
    - Analyze crime patterns and time of day problems
- Gang activity analytics
  - GARI
    - Gang Graffiti Recognition and Interpretation using a mobile telephone
    - Allowing police to catalog and analyze gang graffiti images, better track and determine gang activities
- Document visual analytics
  - JIGSAW
    - Visualization for investigative analysis
    - Discovery of hidden relationship and threats across documents

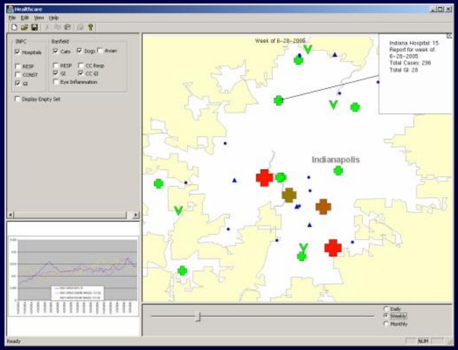
Public Safety & Health

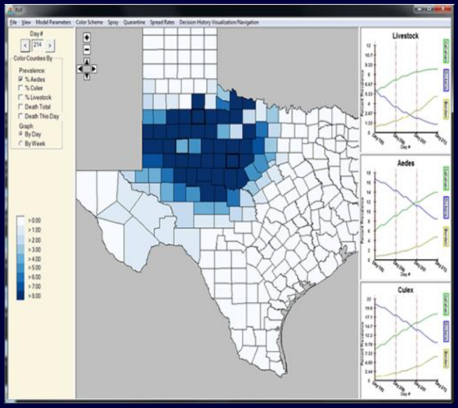
- Public safety visual analytics
  - cgSARVA
    - Coast Guard Search And Rescue Visual Analytics
  - MERGE
    - Mobile Emergency Response Guide
- Public health visual analytics
  - LAHVA
    - Linked Animal-Human Health Visual Analytics
  - RVF
    - Rift Valley Fever
    - Decision support environment for epidemic modeling and responses
  - PanVis
    - Pandemic influenza modeling and visualization tool
  - Cancer Care Engineering





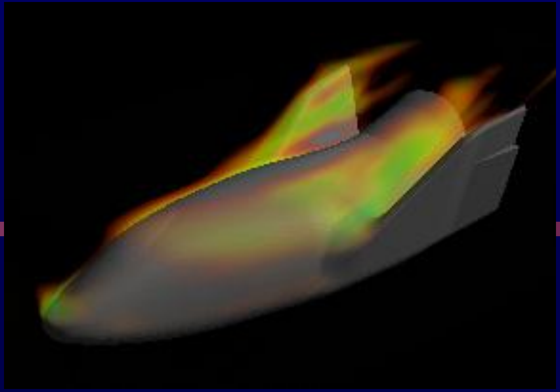


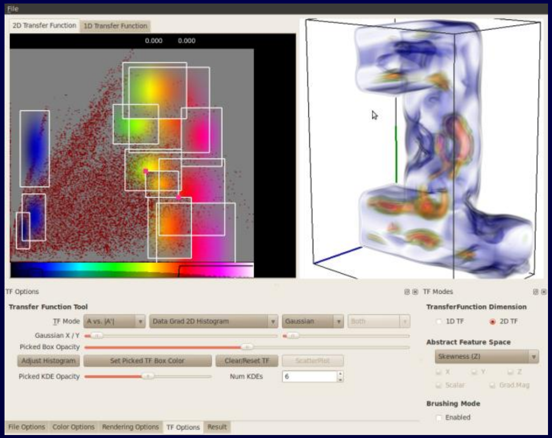


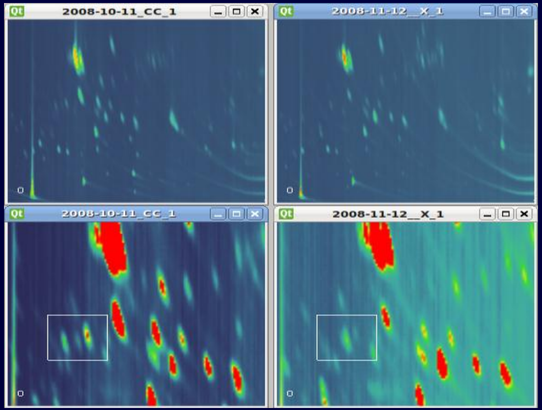


Scientific Discovery

- Flow dynamics visualization
  - Providing insights on large flow data
  - Visualization linked with simulations
- Nanohub
  - Information-assisted data analysis and visualization of nanoelectronic models
- Cancer biomarker
  - Visual analysis suite for exploring multiple samples of data from cancer care engineering



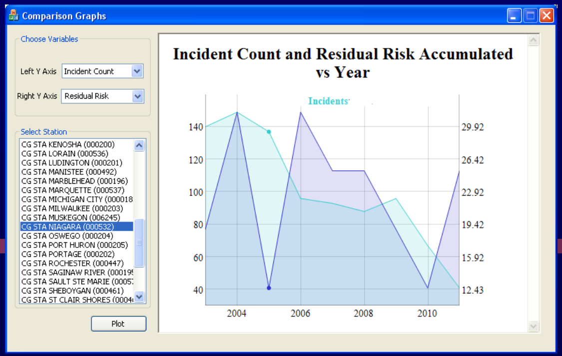


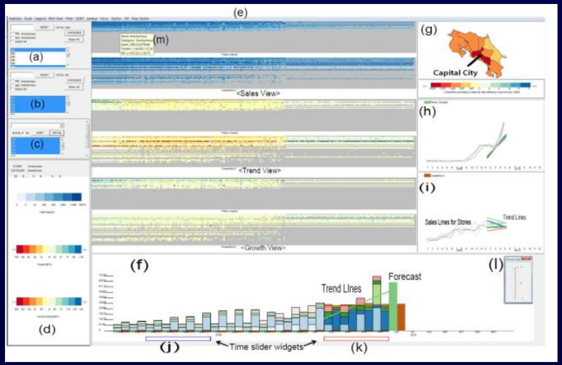


VACCINE

Business Visual Analytics

- Risk-based decision making and resource allocation
  - Coast guard operational risk assessment model
  - Helping to prioritize efforts to minimize risk
- Competitive Intelligence
  - Visual analytics system for business intelligence
- EconVIS
  - Visual analytics in various economic problems
  - Improving decision making and identifying key motivations in knowledge creation





VACCINE:

Who We Are

Who We Are:
International Team of Experts
- 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- **Simon Fraser University**
- **University of British Columbia**
- **Justice Institute of British Columbia**
- **Ontario Institute of Technology**
- **Dalhousie University**
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- **University of Stuttgart**
- **University of Swansea, UK**
- **Oxford**
- **University of Calgary**
- **University of Manitoba**
- **Carleton University**
- **University of Victoria**

VACCINE

VACCINE's Value

Our Value / Solution: Enable users to be more effective through innovative interactive visualization, analysis, and decision making tools

- Provide the right information, in the right format within the right time to solve the problem
- Turn data deluge into a pool of relevant, actionable knowledge
- Enable user to be more effective from planning to detection to response to recovery
- Enable effective communication of information

Approach: Partner-driven solutions and research

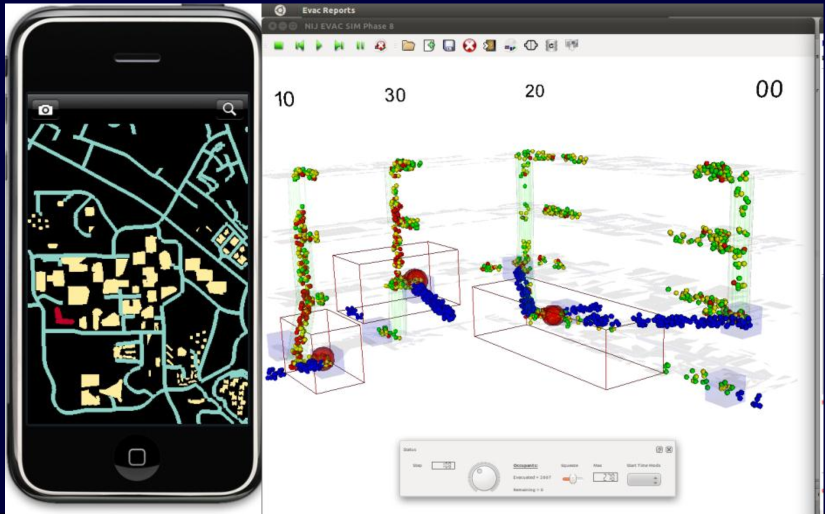
VACCINE:

Some of Our Projects

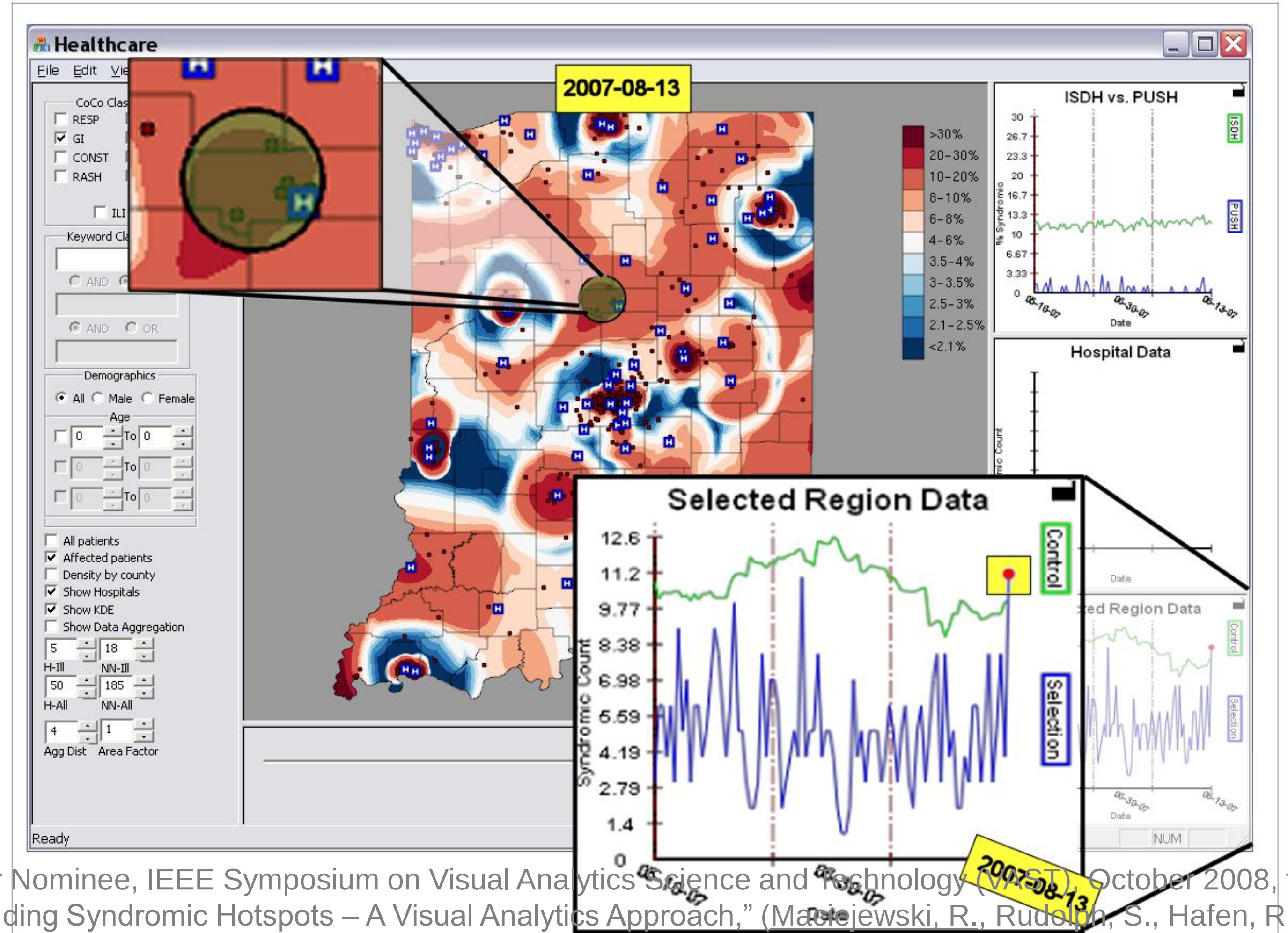
Example Projects

- Health Applications
- TSA
- Public Safety and Law Enforcement
- USCG
- Cybersecurity

UNCC Emergency Evac



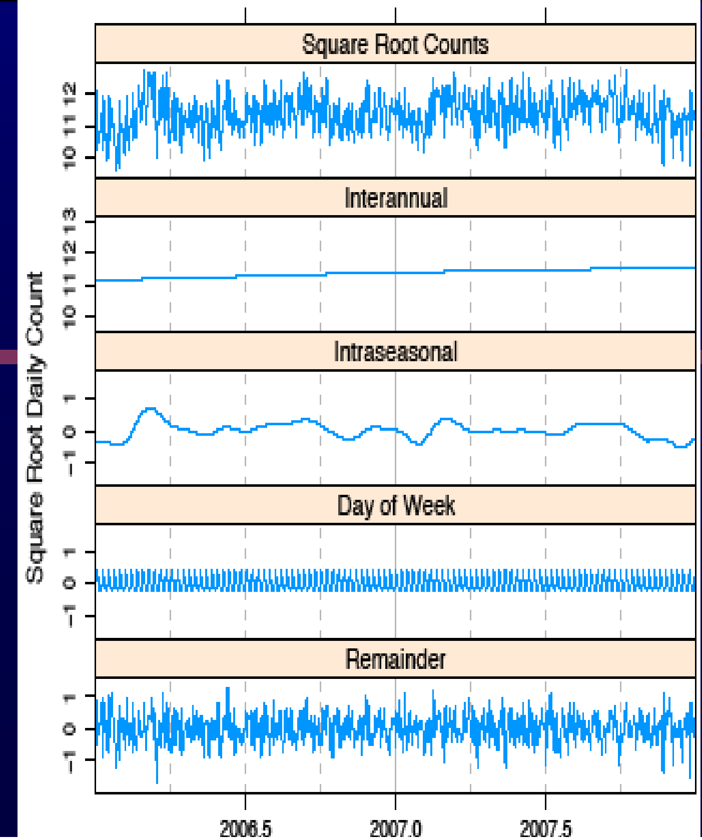
Visual Analytics for Syndromic Surveillance: Hypothesis Generation and Exploration

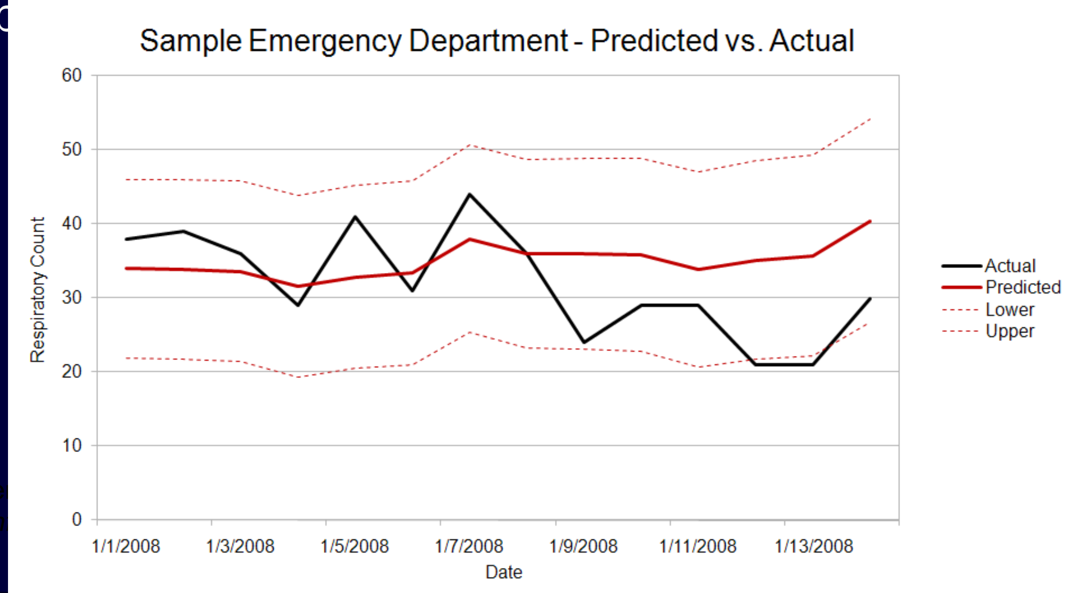


Best Paper Nominee, IEEE Symposium on Visual Analytics Science and Technology (VAST), October 2008, for “Understanding Syndromic Hotspots – A Visual Analytics Approach,” (Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W., Grannis, S., Wade, M., Ebert, D.).

Predictive Visual Analytics

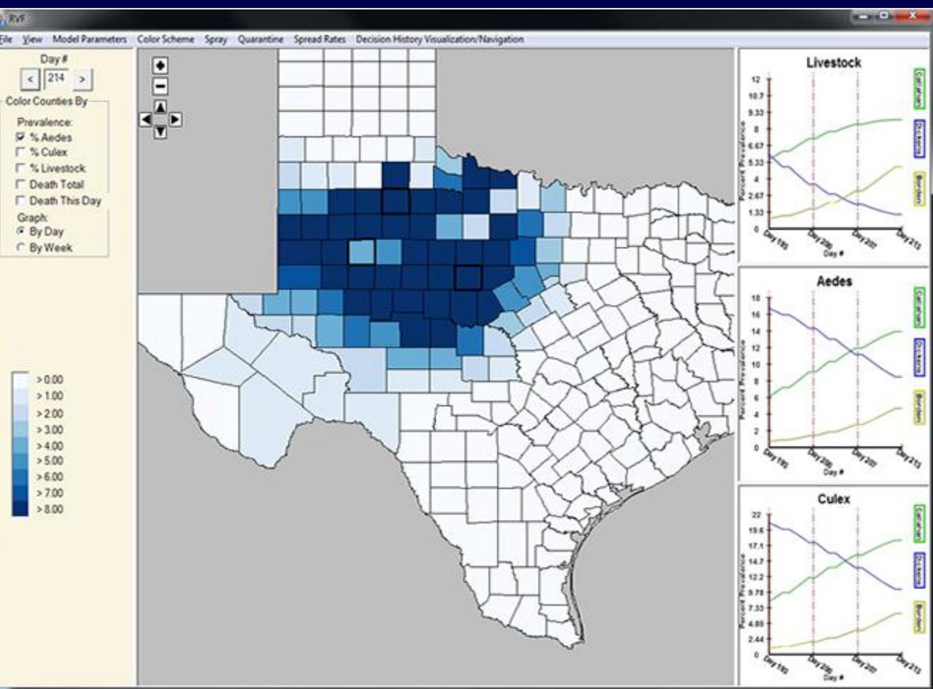
- Time Series Modeling:
  - Seasonal-Trend Decomposition Based on Loess
    - Time series ~ sum of multiple trend components
    - For each data signal, components are extracted
    - Can then analyze correlation b components
- **Predictive Visual Analytics using STL**

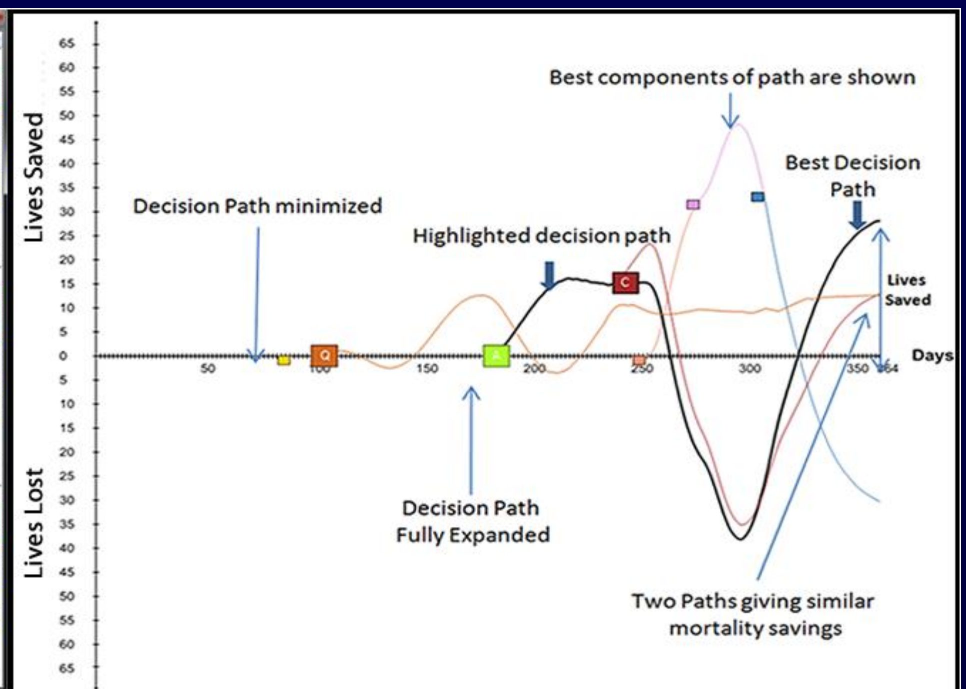




Hafen, R., Anderson, D., Cleveland, W., Maciejewski, R., Ebe... Visualizing, and Monitoring Disease Counts," *BMC Medical In...*

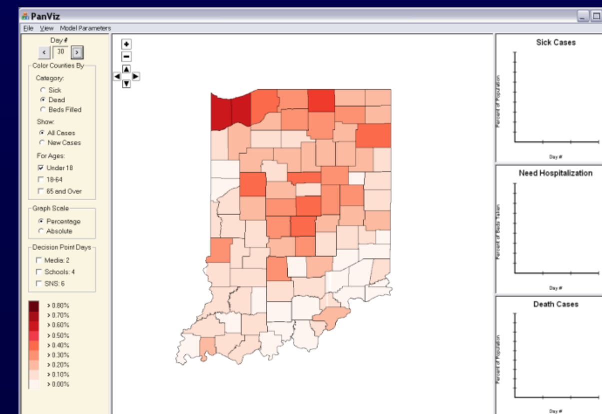
Advanced Decision Support Tools: Rift Valley Fever



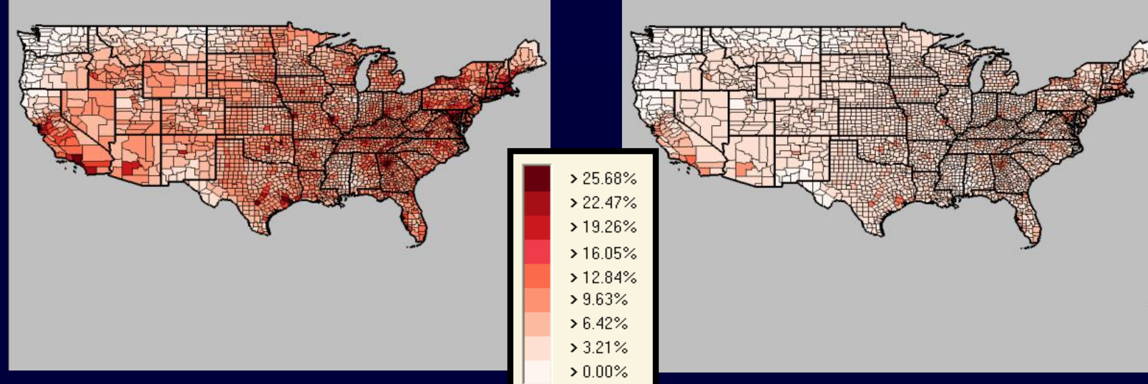


Modeling a Pandemic

- Pandemic Influenza Planning Tool
- Models user specified:
  - Pandemic influenza characteristics
  - County population, demographics, hospital beds
- Decision measures
  - Strategic National Stockpile deployment
  - School Closures
  - Media Alerts



Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.



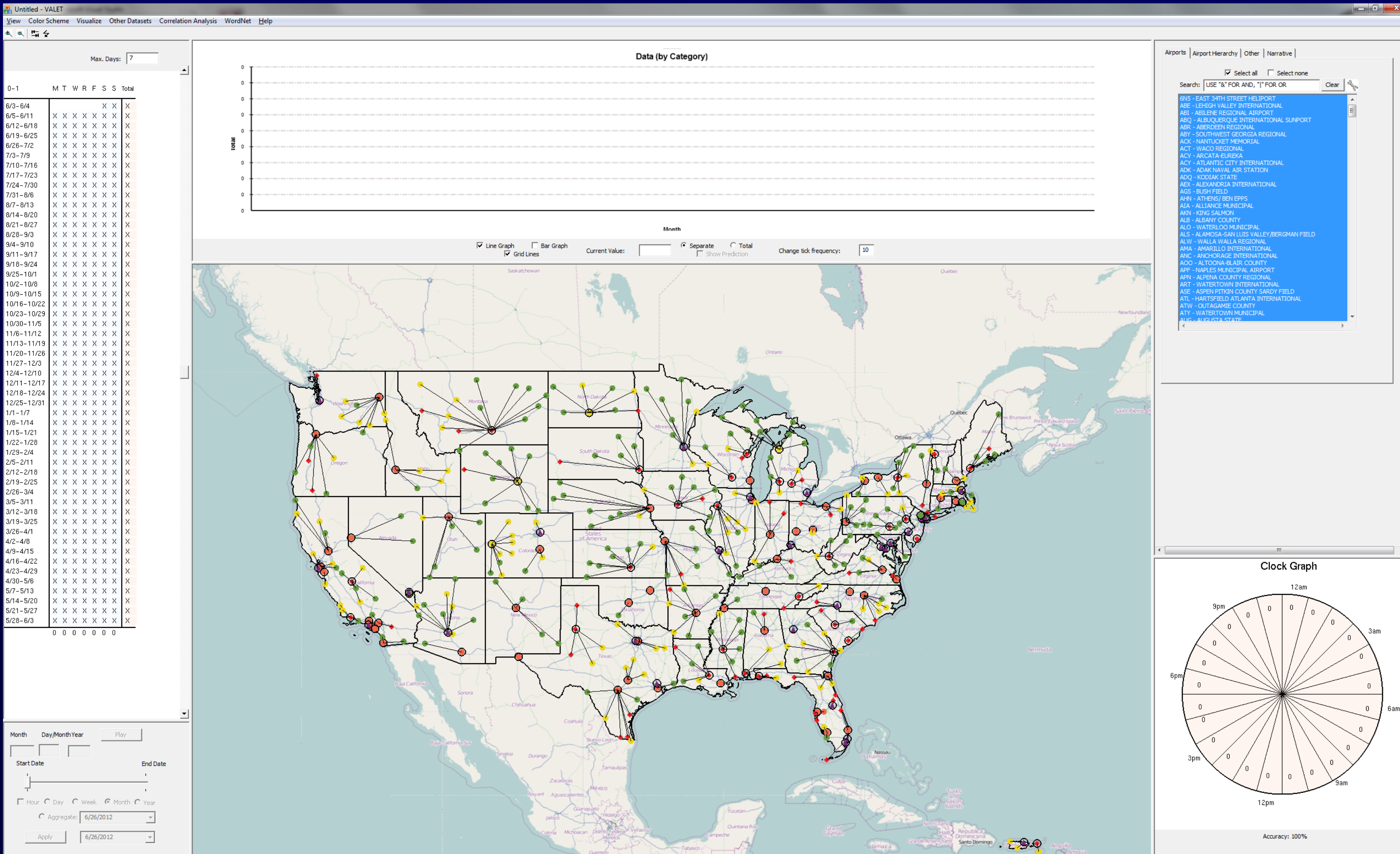
Pandemic spread on day 37 with no decision measures implemented

Pandemic spread on day 37 with all decision measures implemented

TSA Projects

- Visual Analytics for multifaceted operational analysis
  - Analysis and correlative analysis of
    - Operational performance (throughput per lane, use of AIT, targets)
    - Security performance (incidents)
    - Financial performance (expenses vs. performance metrics)
  - Spatial, temporal, spoke, hub, and regional analysis
- Customer satisfaction visual analytics
  - Jigsaw use to understand and find patterns in compliment/complaint database
  - Potential linkage to above data
- MERGE for rail inspectors

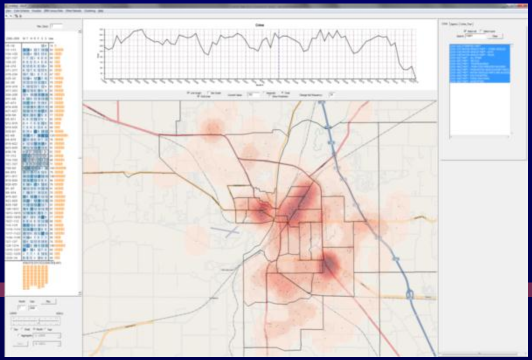
TSA Visual Analysis Prototype

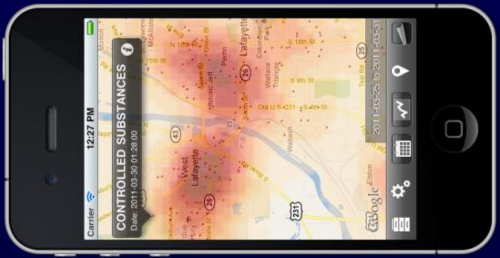


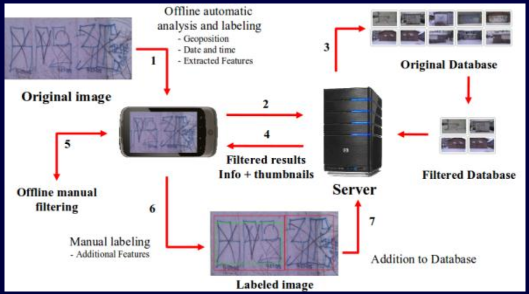
▲ Category X (28) ● Category I (57) ◆ Category II (77) ● Category III (127) ● Category IV (157) ○ Hub (90) Infrequent Public Charter ● Coordination Centers (121) Total (446)

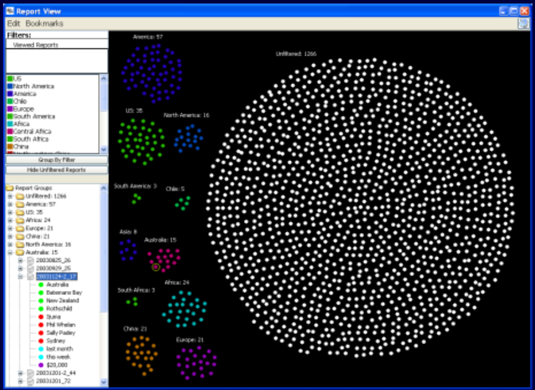
?
?

Public Safety & Law Enforcement



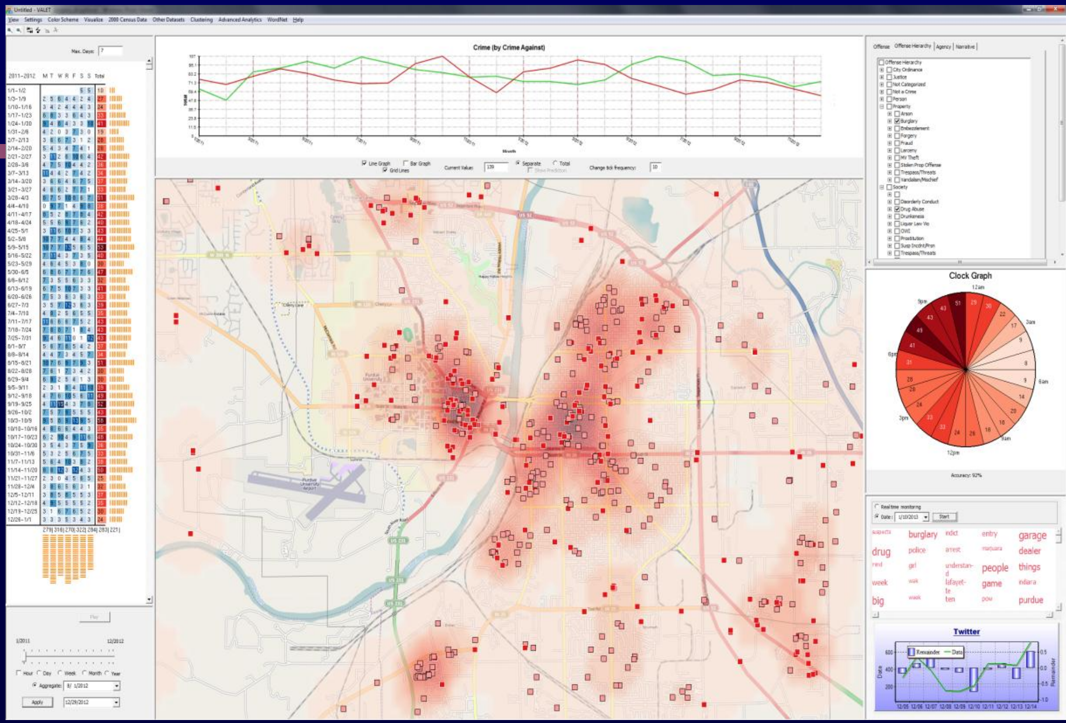






- Law enforcement visual analytics
  - VALET, iVALET (iPhone/iPad), CrimeViz
    - Visual Analytics Law Enforcement Toolkit
    - Analyzing crime patterns and time of day problems
- Gang activity analytics
  - GARI
    - Gang Graffiti Recognition and Interpretation using a mobile telephone
    - Allowing police to catalog and analyze gang graffiti images, better track and determine gang activities
- Document visual analytics
  - JIGSAW
    - Visualization for investigative analysis
    - Discovery of hidden relationship and threats across documents

Visual Analytics Law Enforcement Toolkit (VALET, iVALET)



Impacts:

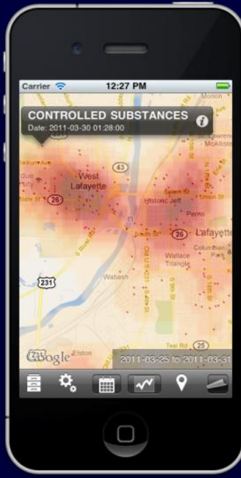
- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (February 2013) for community-based policing
- Investigating correlation factors
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning
- Incorporating predictive alerts

VALET delivered:

- Spring 2011: WL, Lafayette Police

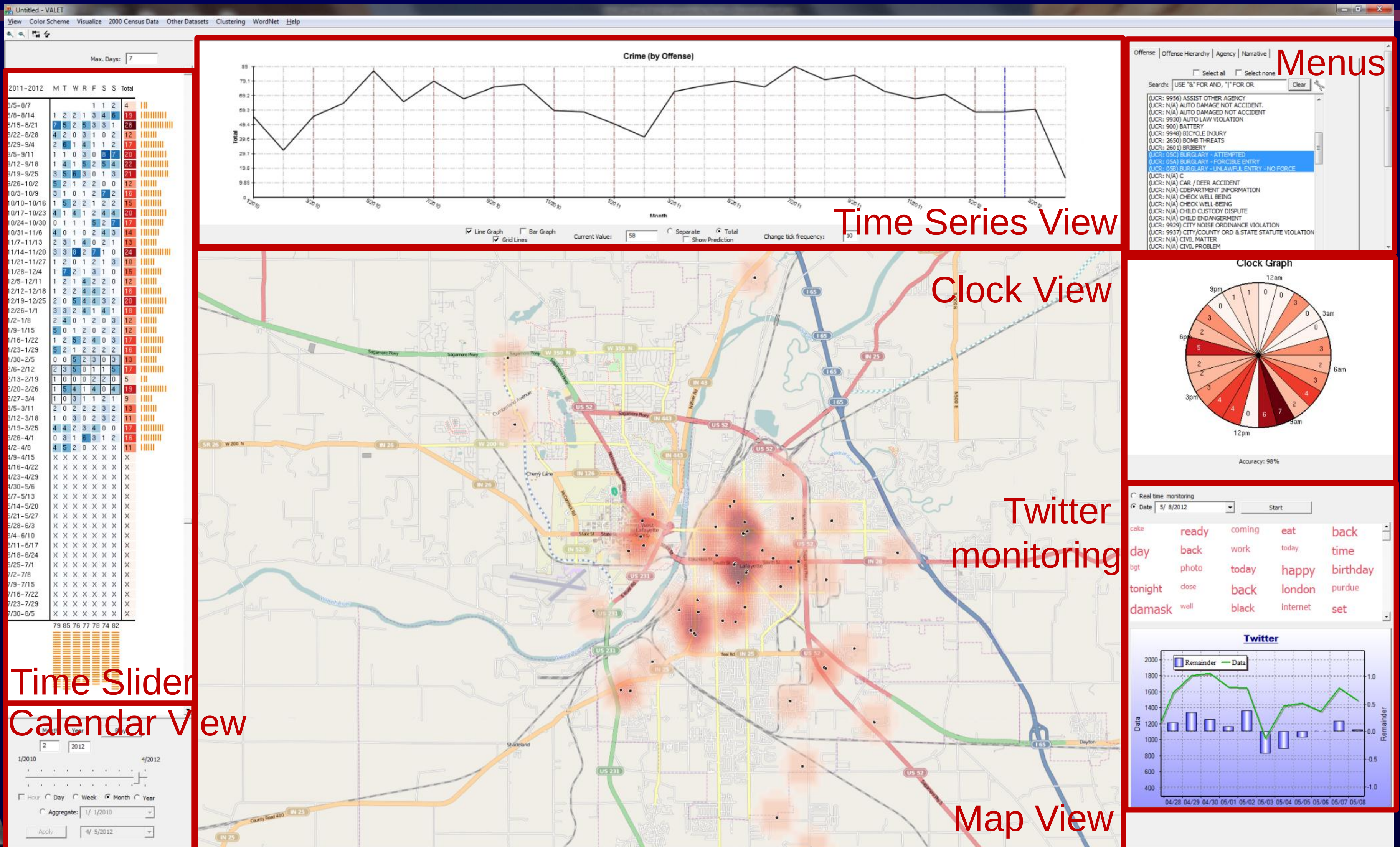
iVALET delivered:

- October 2011: Purdue, WL Police

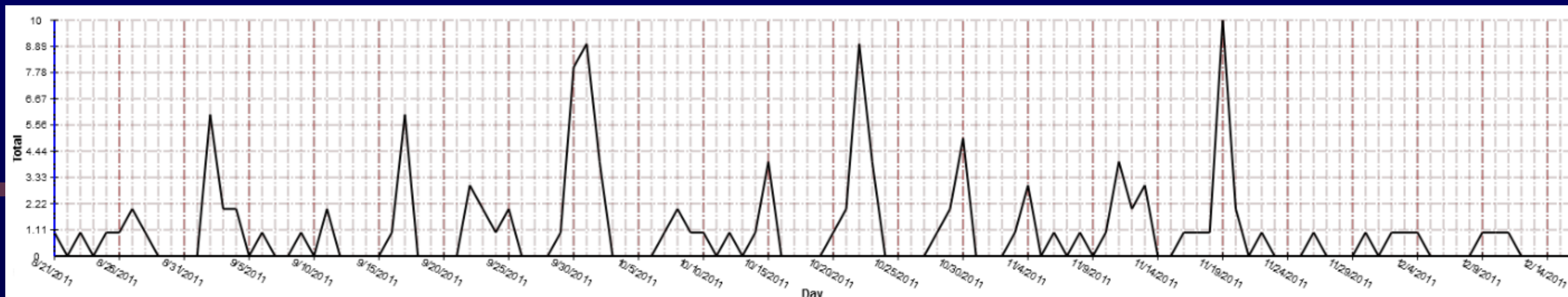


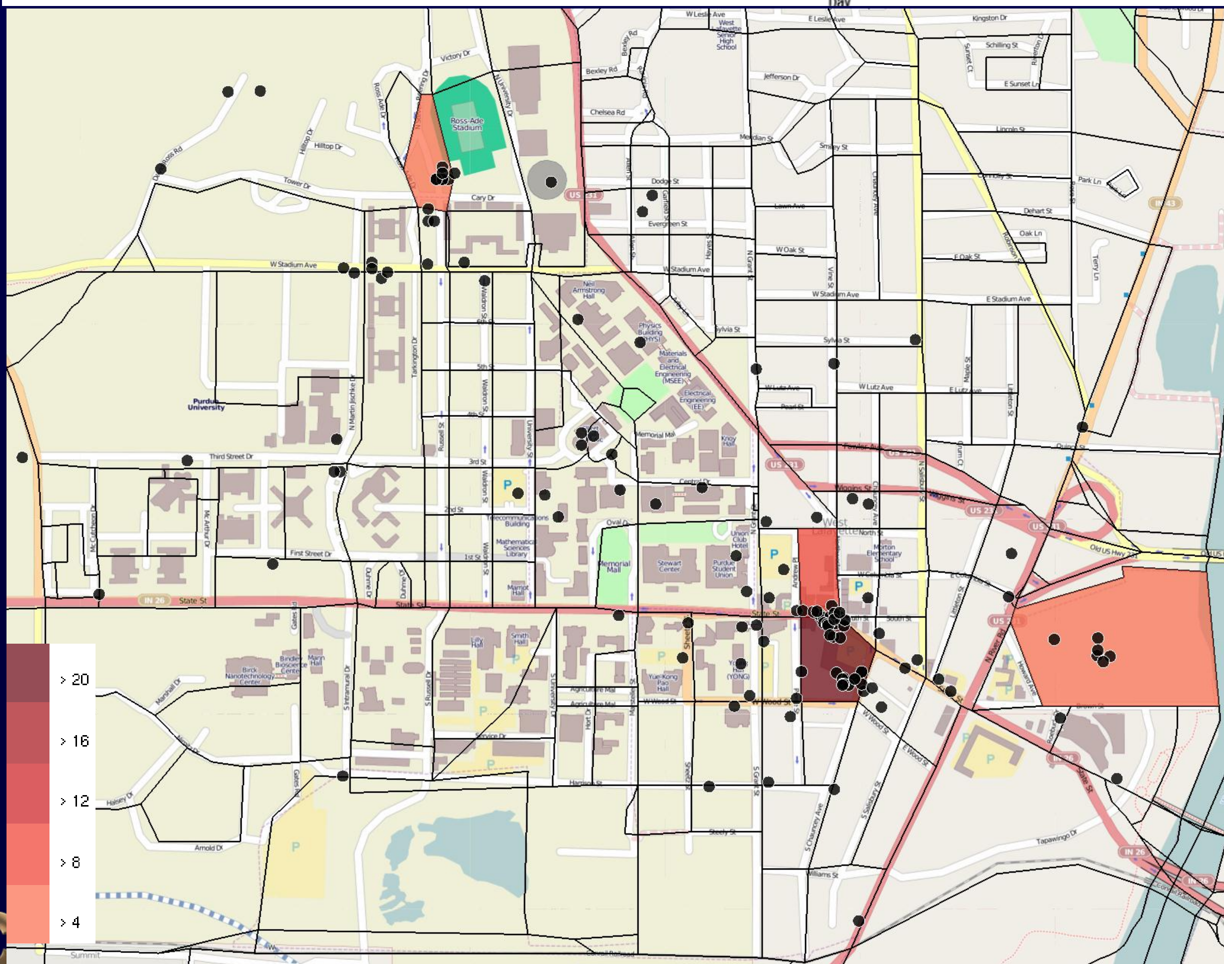
VACCINE

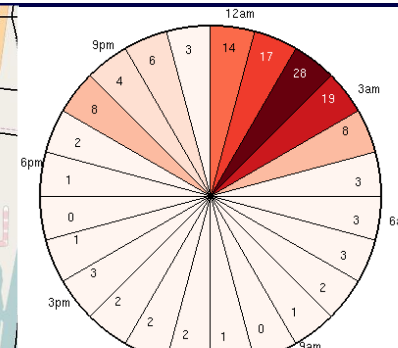
VALET Overview



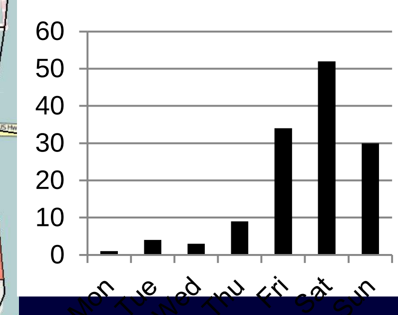
Example: Drunkenness / Public Intoxication



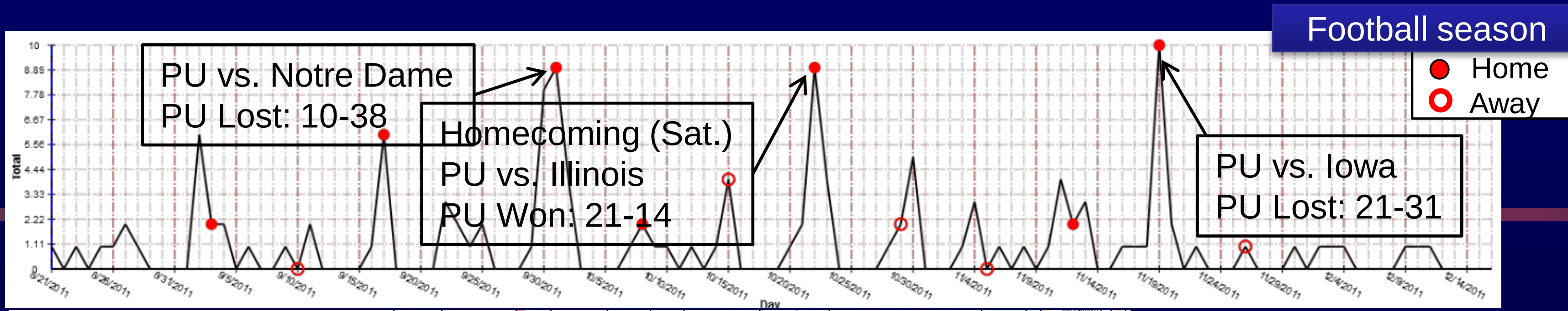


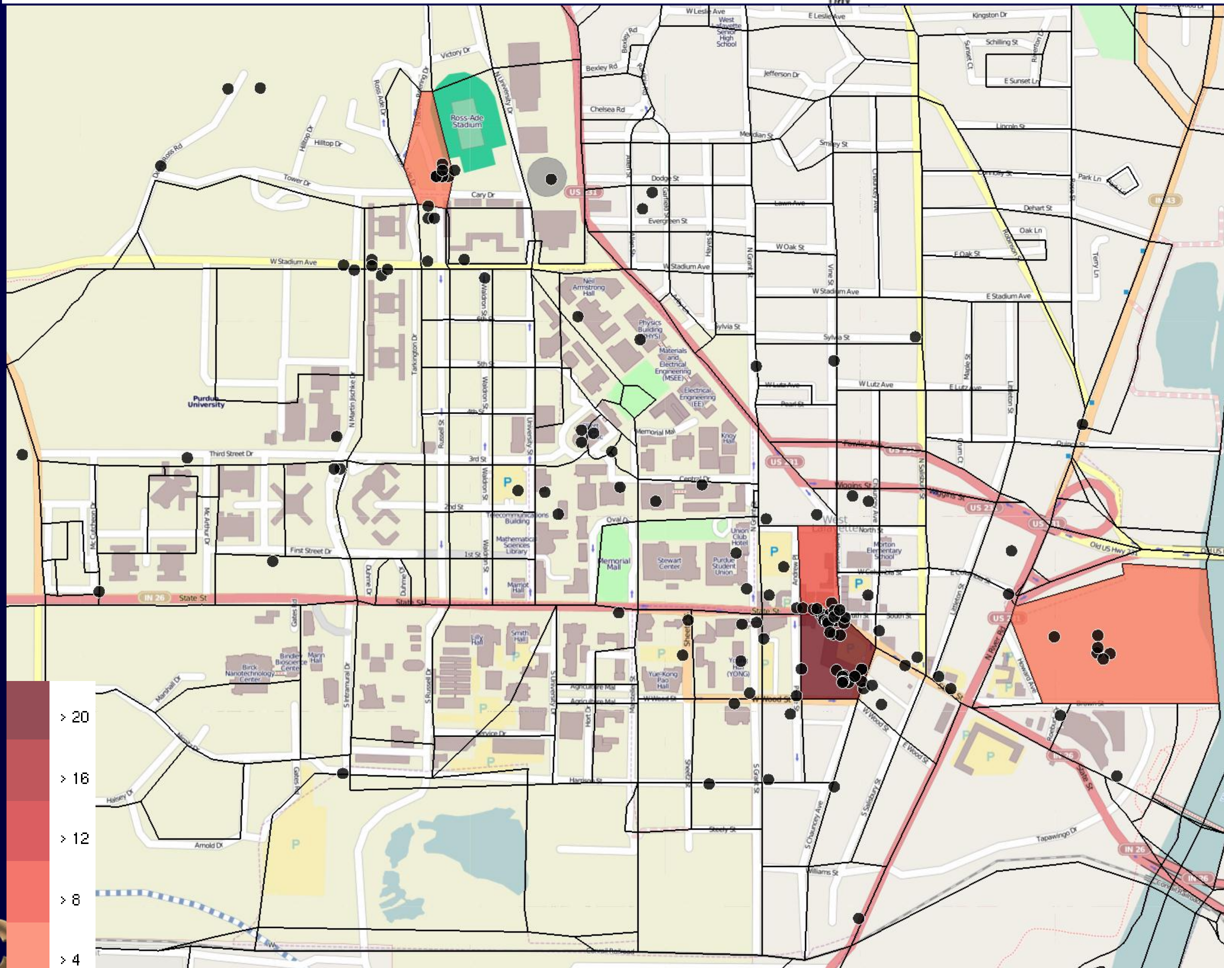


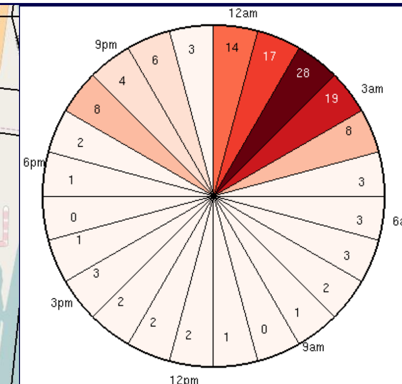
Day-of-the-Week

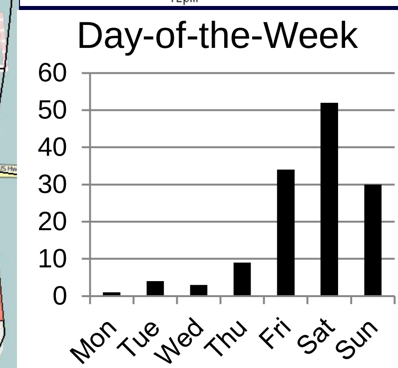


Example: Drunkenness / Public Intoxication





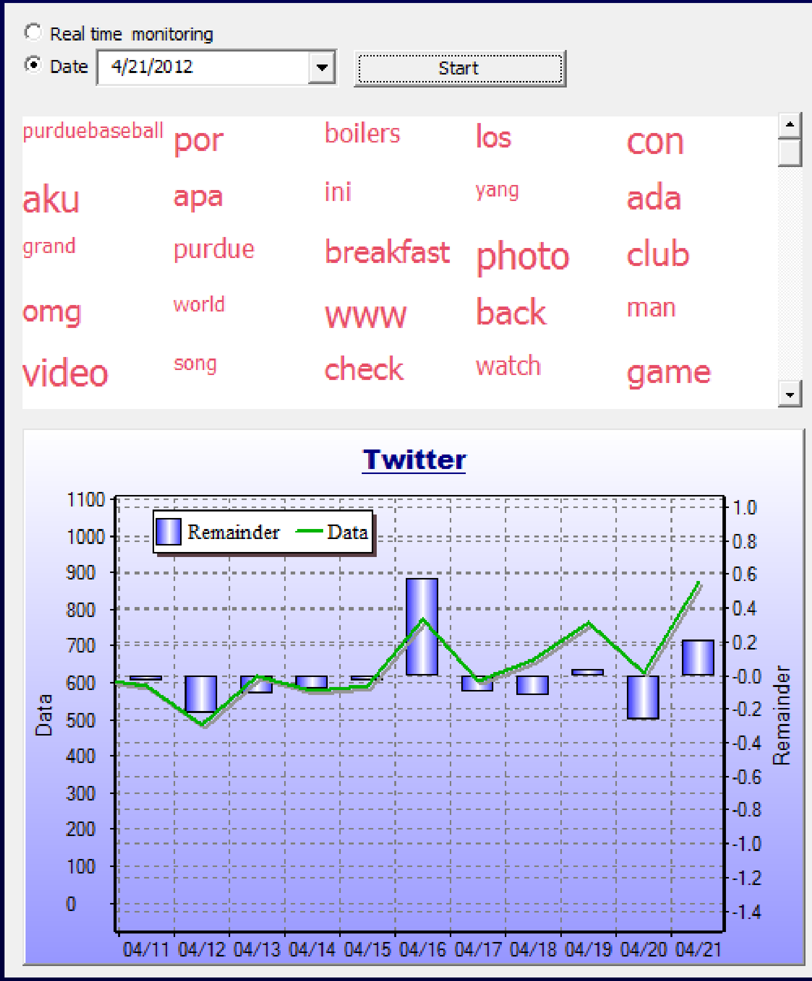




Social Media:Real-time Twitter Monitoring and Integration into Tools

(Purdue, Stuttgart, Penn St.)

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud



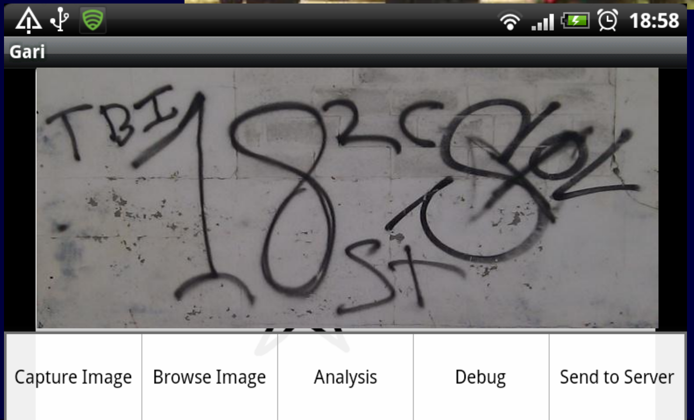
Grand Prix Weekend, Purdue University


Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
11/14/2010
01/01/2011
02/13/2011
03/27/2011
Calendar View
M T W Th F Sa S
1/1/10 36 46 29
1/4/10 38 48 44 43 42 32 30
1/11/10 62 53 53 56 59 45 34
1/18/10 62 49 53 72 61 52 49
1/25/10 98 66 57 66 64 55 49
2/1/10 58 63 50 51 73 51 42
2/8/10 50 46 48 63 52 39 35
2/15/10 48 65 48 54 55 67 51
2/22/10 51 66 48 56 57 46 47
3/1/10 50 59 42 54 82 45 41
3/8/10 50 57 64 78 63 36 30
Clock View
12Hrs
10:00
Accuracy: 64%
Choro
Plume
Moran
Region
Route
©2012 Google
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20

We Have an App For That!

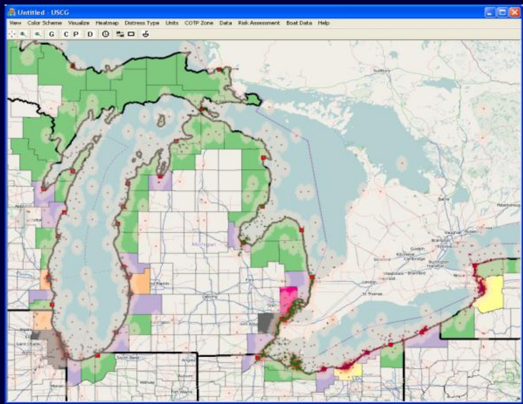
Our Mobile Tools

- VALET
- Evacuation Planning
- Rosetta Phone
- Hazmat app
- Gang graffiti app
- Tatoo app



Visual Analytics Uses for Risk-Based Decision Making

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



USCG: Effective Risk-based Decision Making and Resource Allocation Visual Analytics

- Evaluate current and historical mission area:
  - Demands
  - Risks (total, mitigated, residual)
  - Resource allocation
  - Return on investment
- Evaluate courses of action
- Evaluate above at both Strategic and Tactical/Operational level

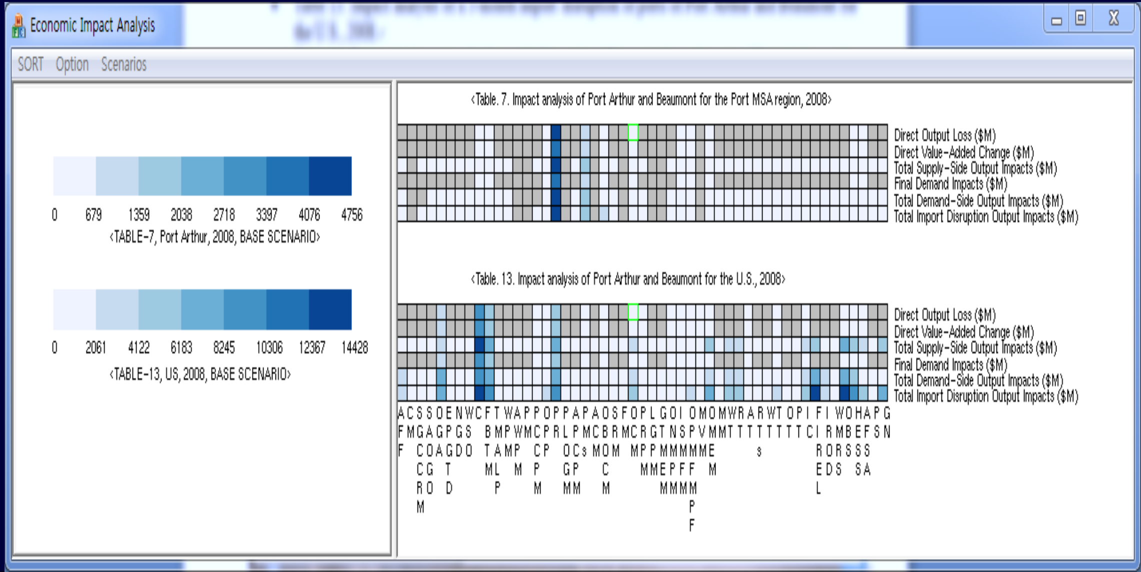
VACCINE

USCG Port Closure Economic Impact VA

Partners: USC CREATE, USCG RDC, USCG D7, USCG LANT

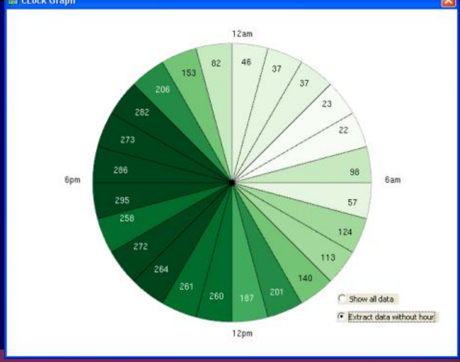
IMPACT:

- Provided tool for use analysis and planning for impact of port closure in Port Arthur, Tx
- Economic sector impact, local and national impact
- Impact and effectiveness of alternative mitigation strategies



VACCINE

Risk-Based Allocations



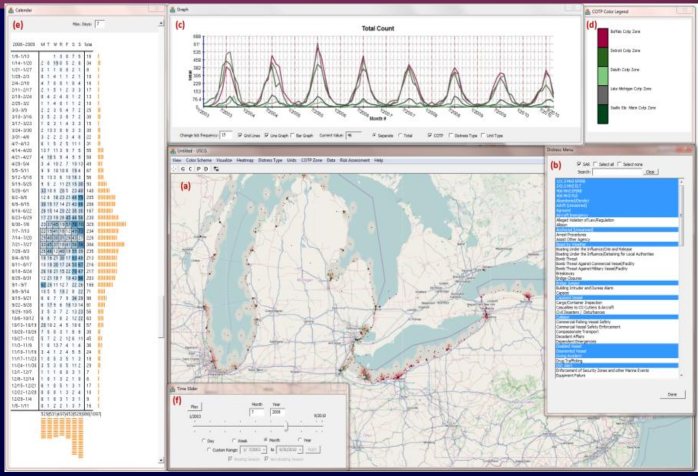
- Comparative visual analysis of mission cases/hours vs. staffing hours
- Comparative visualization of resources vs. risk
- Trend visual analytics
  - Increase/decrease in resource allocation
  - Increase/decrease in risk (total, mitigated, residual)
  - Increase/decrease in incidents
- Exploration of alternatives and effect on risk
- Predictive analytics based on historical data (STL and EWMA)

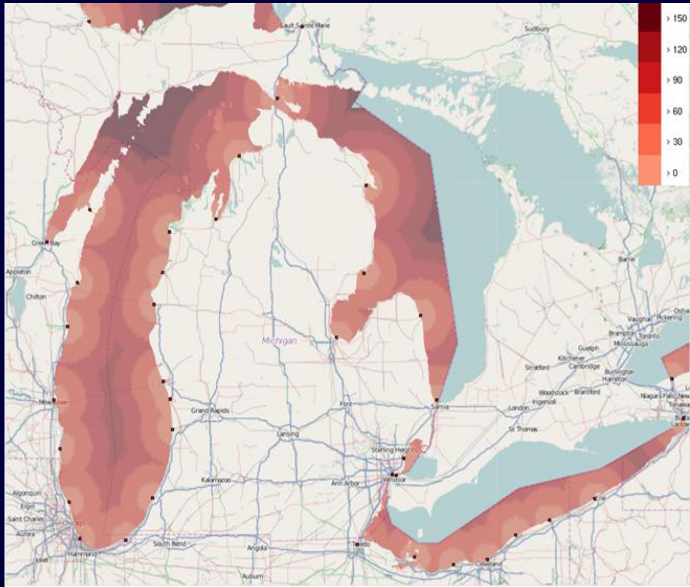
U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7, USCG D9, USCG D5, USCG HQ 771

IMPACTS:

- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Hurricane Irene resource allocation decision based on cgSARVA analysis and visualization
- Informed Commandant's budget testimony and recommendations to Congress
- Key component of USCG D9 reallocation plan for 2011-2012 based on decreased budget
- Key component of Coastal Operations Allocation Suite of Tools (COAST) – USCG HQ



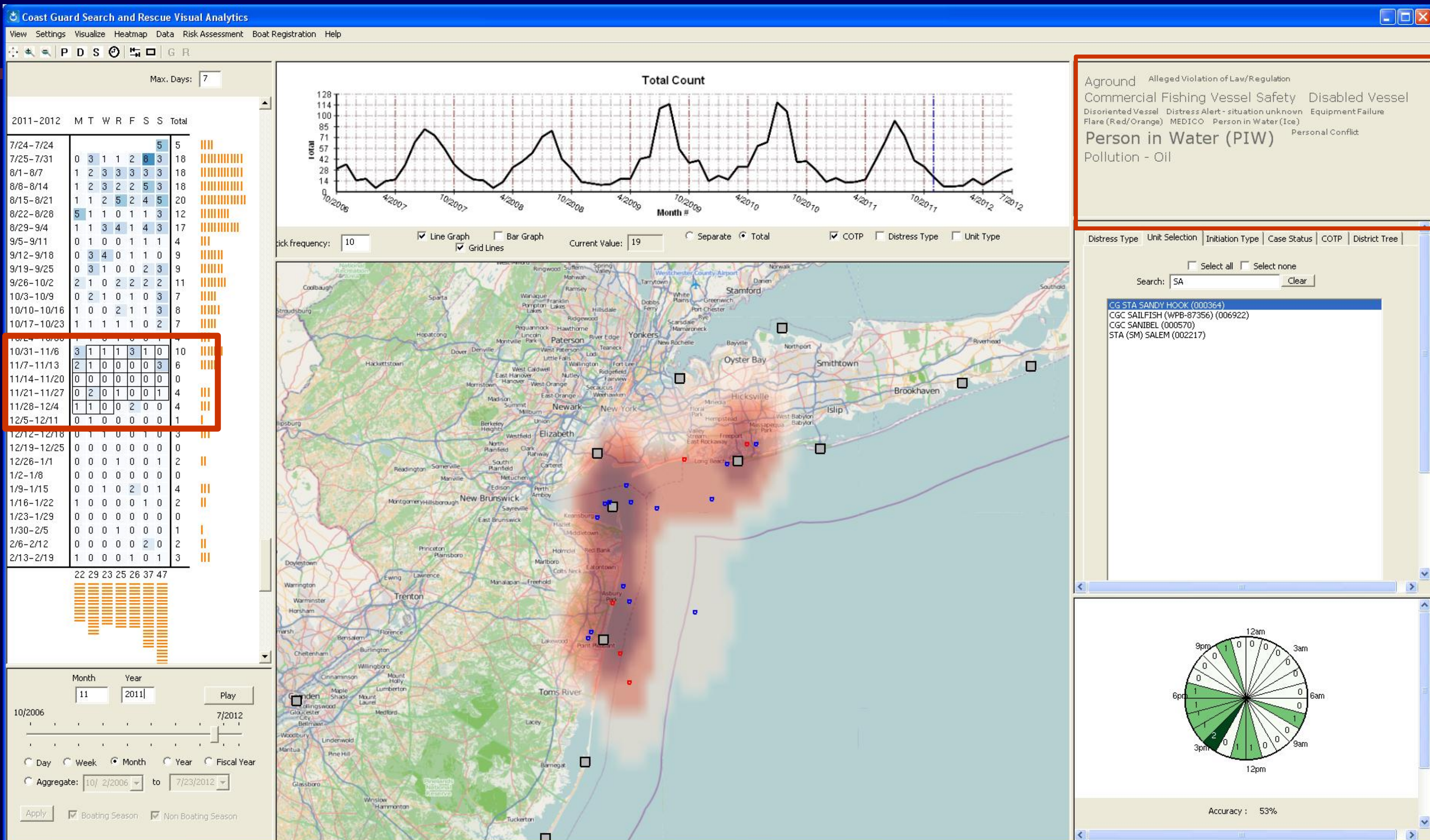


VACCINE

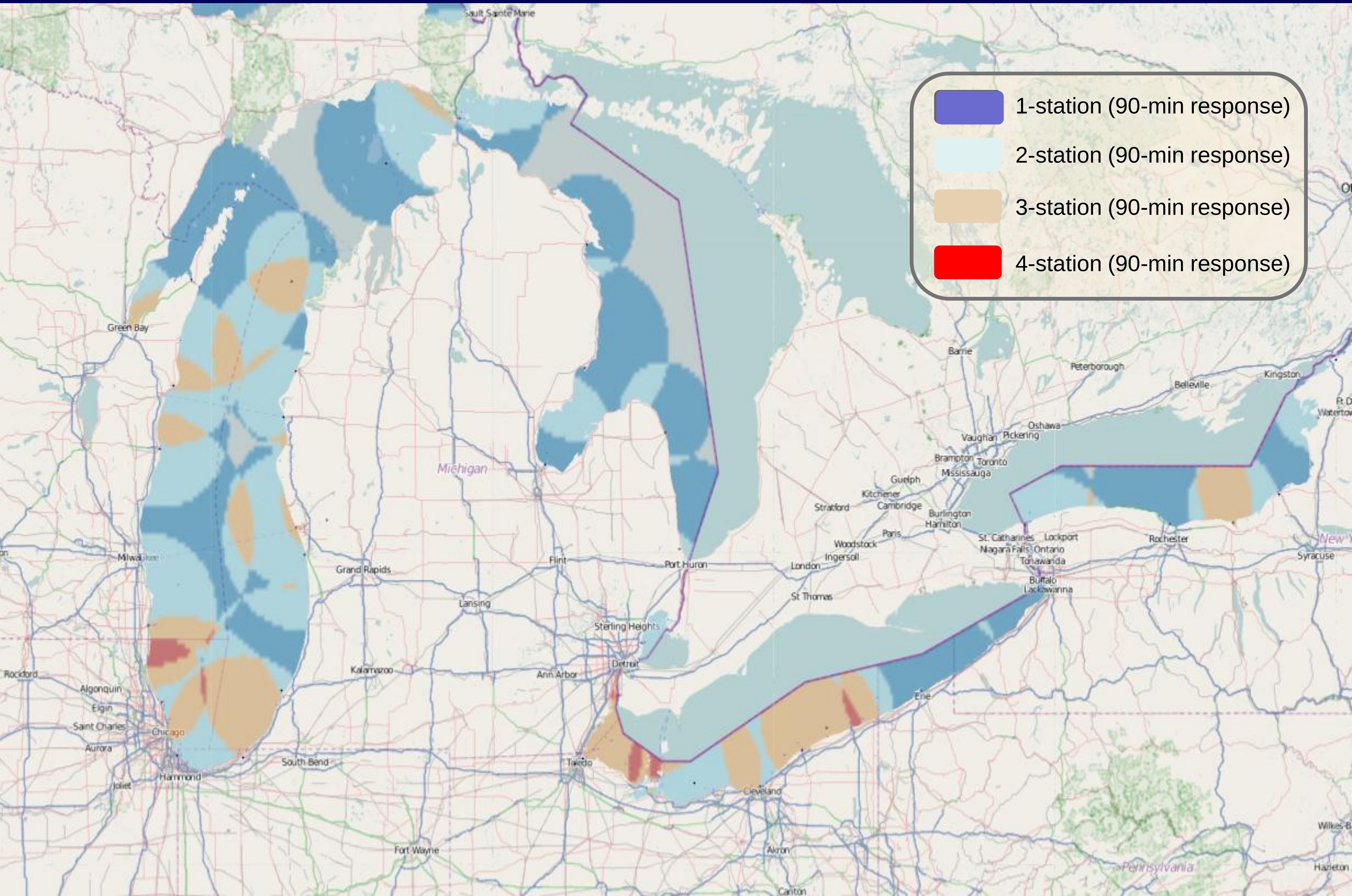
Jones Beach, Sandy Hook and Manasquan Inlet

Closest Cases (not necessarily reported under them)

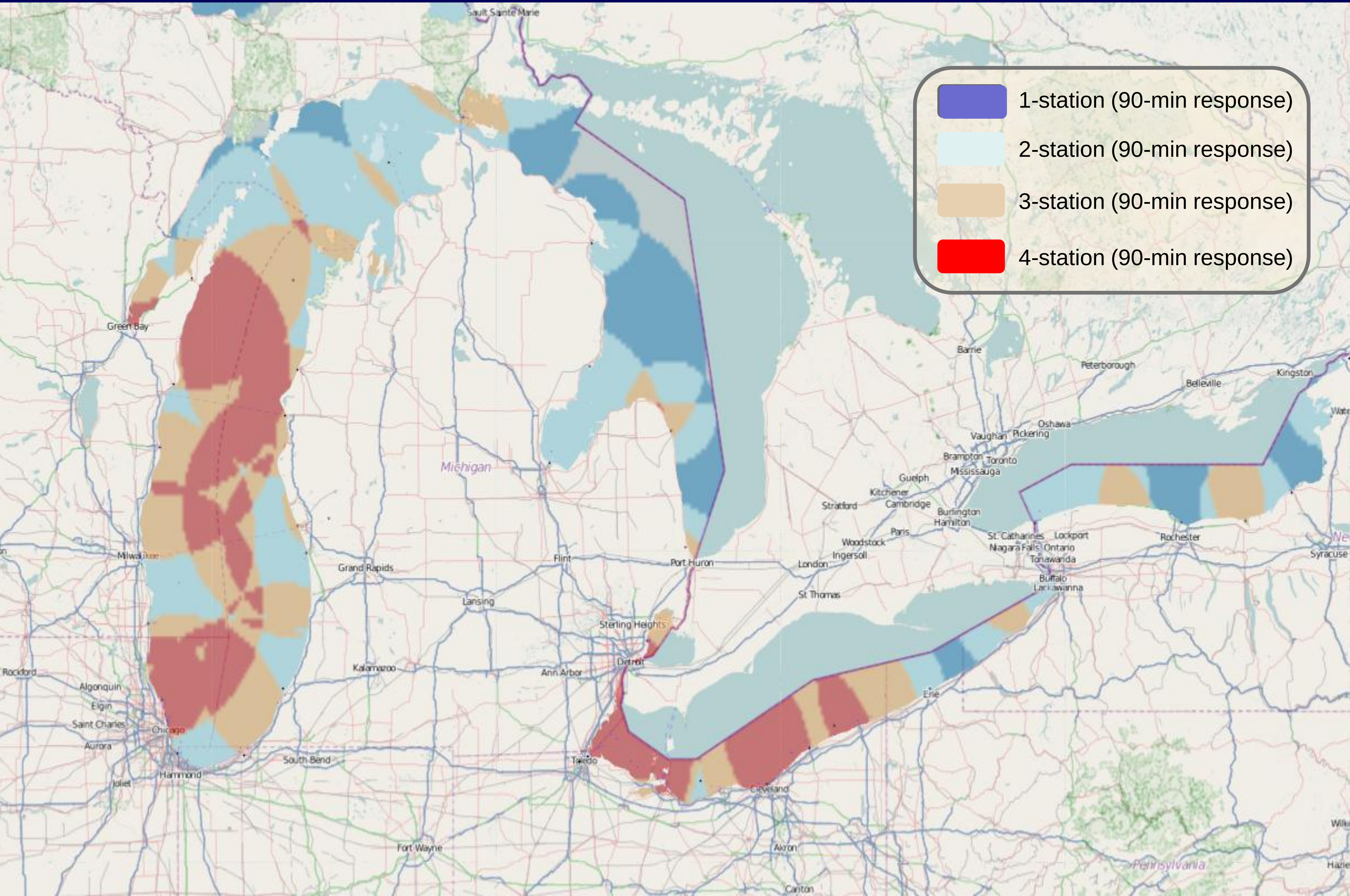
All Cases November 2011



Response Efficiency – Current Assets

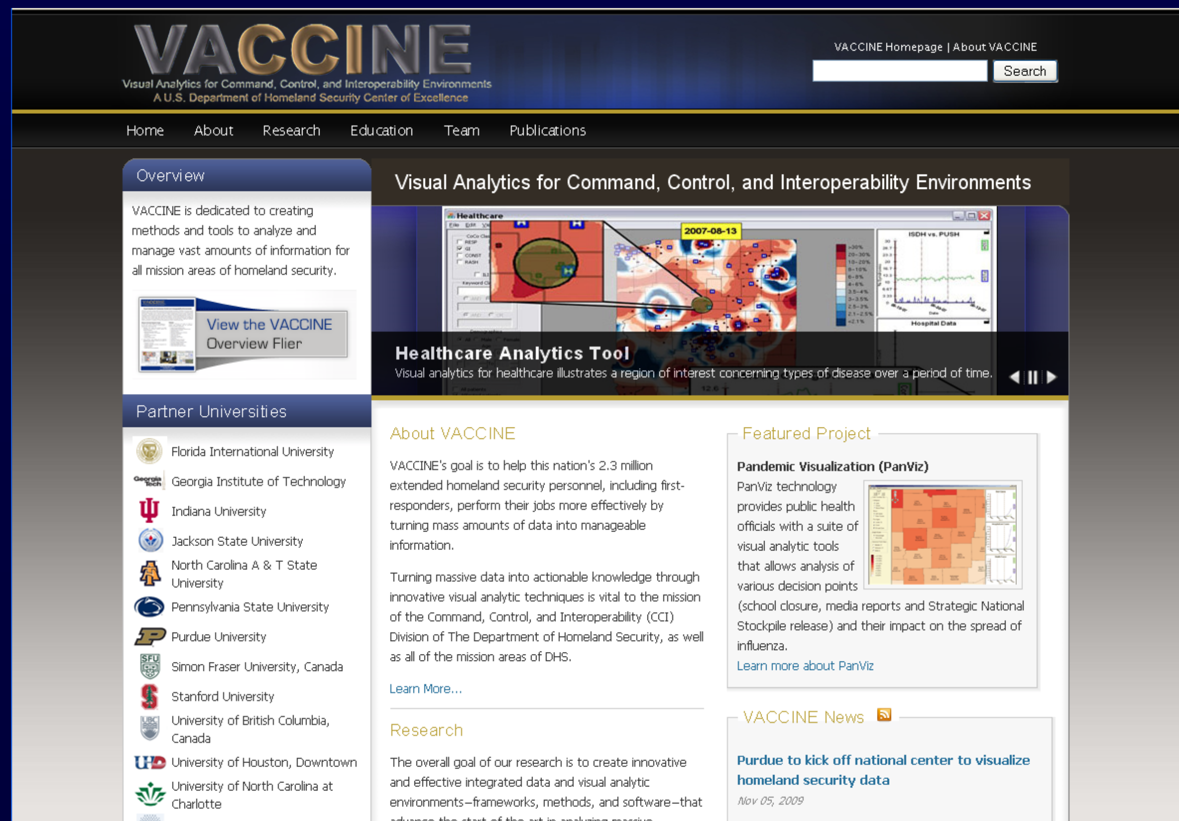


Response Efficiency – Potential Future Assets



For Further Information

www.VisualAnalytics-CCI.org



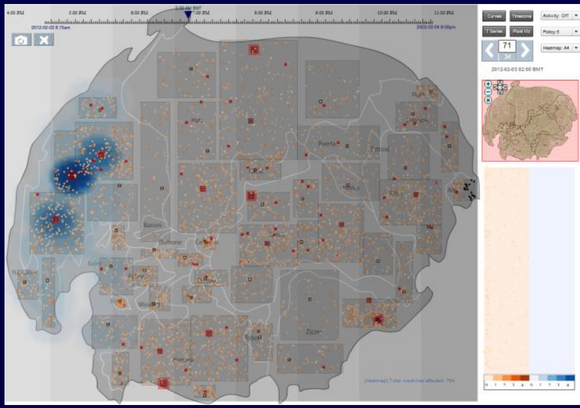
vaccine@purdue.edu

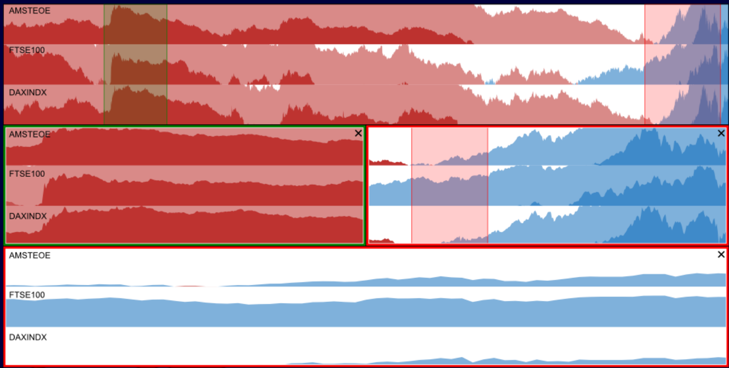
ebertd@purdue.edu

VACCINE

Example Work in Cybersecurity

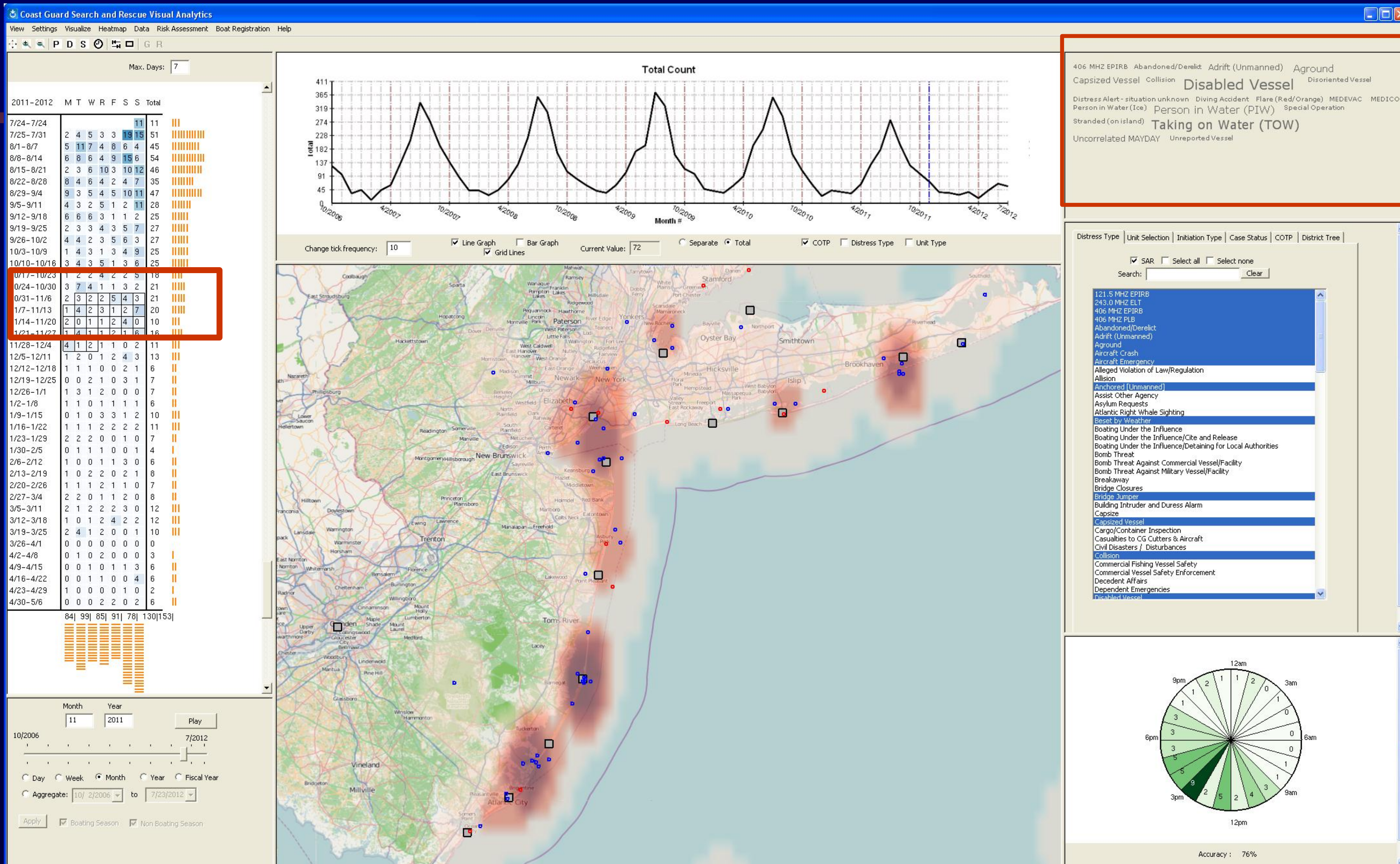
- Corporate Insider Threat Detection (Oxford, Leicester, Cardiff)
- Sensor Forensics (Purdue)
- SemanticPrism (Purdue)
- Multiscreen, Multiview Interactive Cyber Investigation (VaTech, PNNL)
- Log Visualization (Purdue)





Example: Risks and Consequences From Sandy:

SAR Cases November 2011 NJ/NYC Area



Overview of VACCINE Projects

Table of Contents

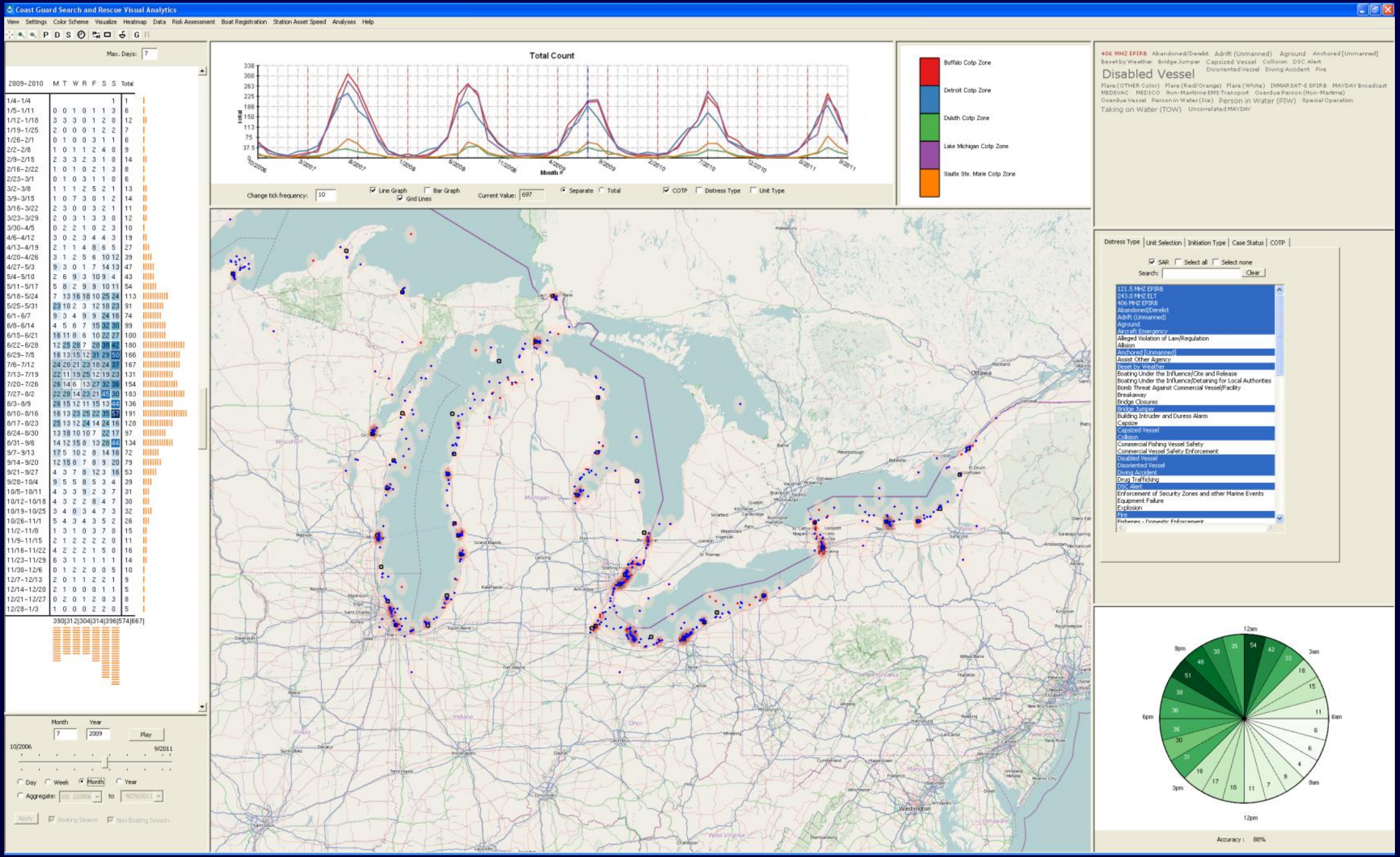
COAST
USCG Risk-Based Resource Allocation for Coastal Mission
Built on Coast Guard Search and Rescue Visual Analytics

VACCINE

May, 2012

VACCINE

System Overview



VACCINE

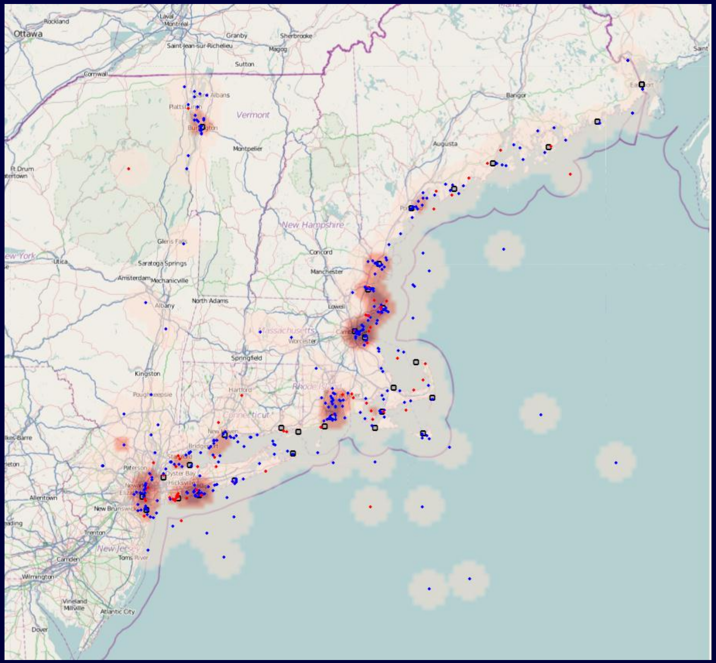
Capabilities and advantages

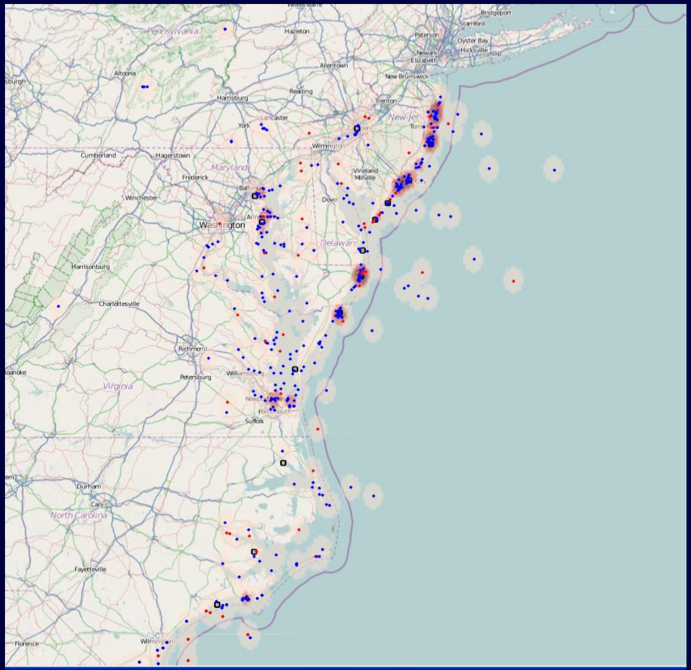
- Geo-temporal visualization of incidents on a map
- Calendar view of incidents
- Word Cloud of Incident Types
- Filter data by distress type, unit (stations), initiation types, cases status, and COTP owner.
- Heatmaps for: incidents, risk numbers, weighted incidents and cost.
- **Risk assessment capabilities for resource allocation and station closures**
- **Boat and air assets**

VACCINE

Recent Updates

- Addition of detailed filters for case initiation and case status, along with their respective heat maps.
- Expansion to D1 and D5 districts.

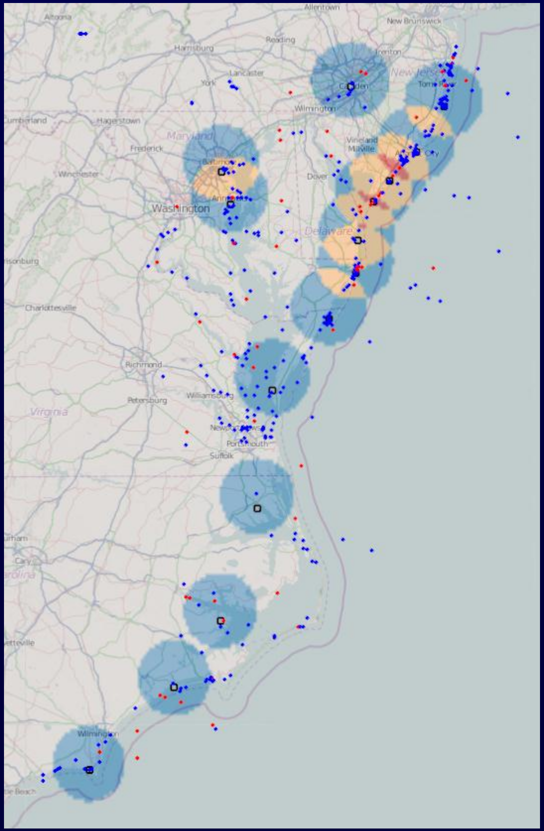
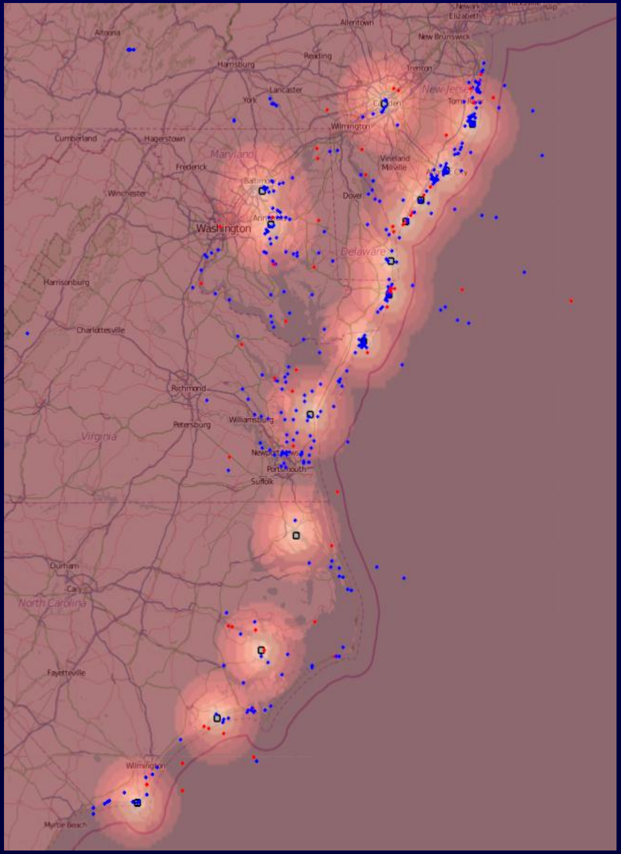




VACCINE

Recent Updates

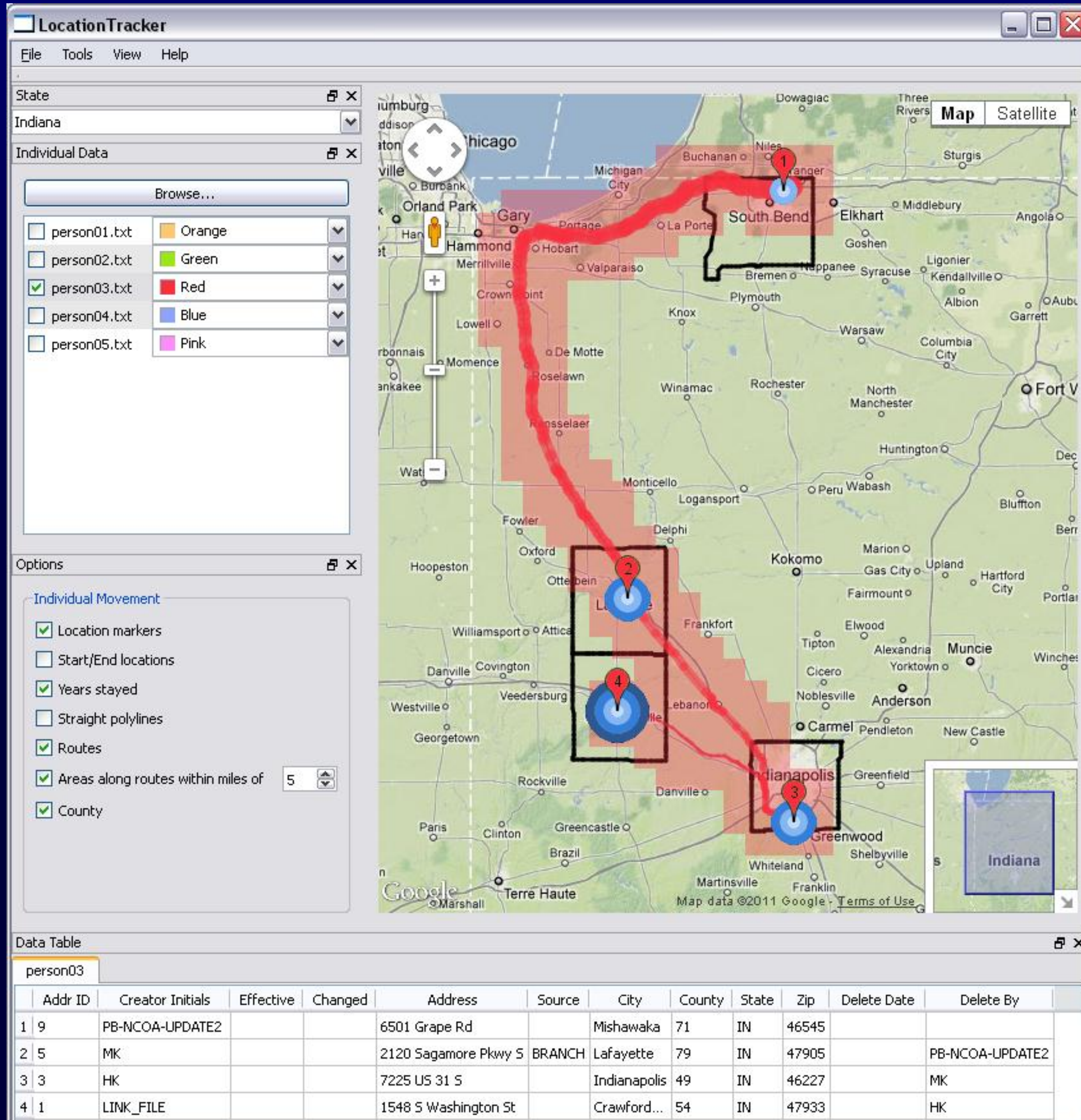
- Addition of custom response rings
- Addition of **redundancy areas in coverage**
- **New station utility metrics**

VACCINE

TRIP

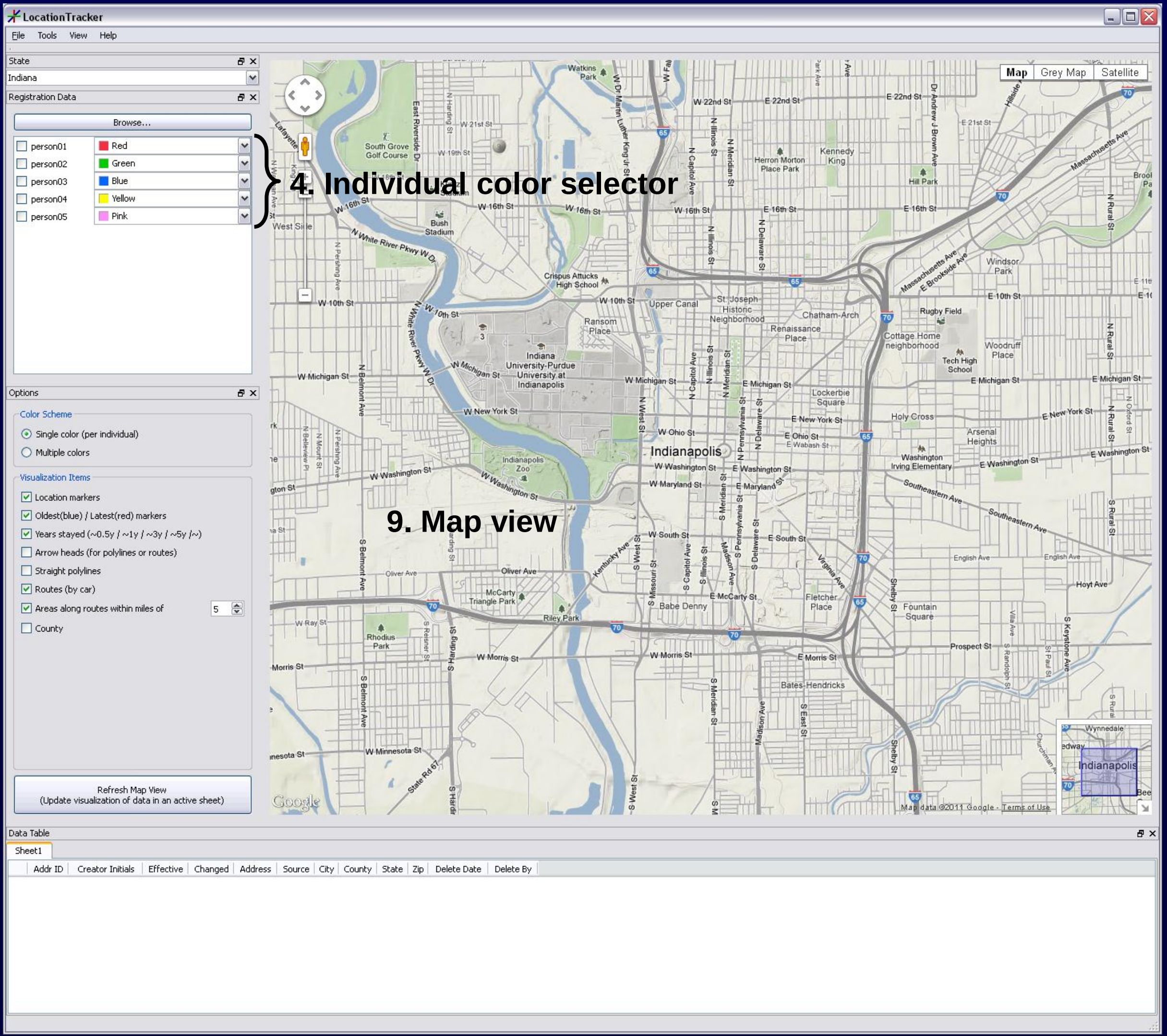
Indiana Fusion Center: Historical Movement Visual Analytics



VACCINE

System Overview

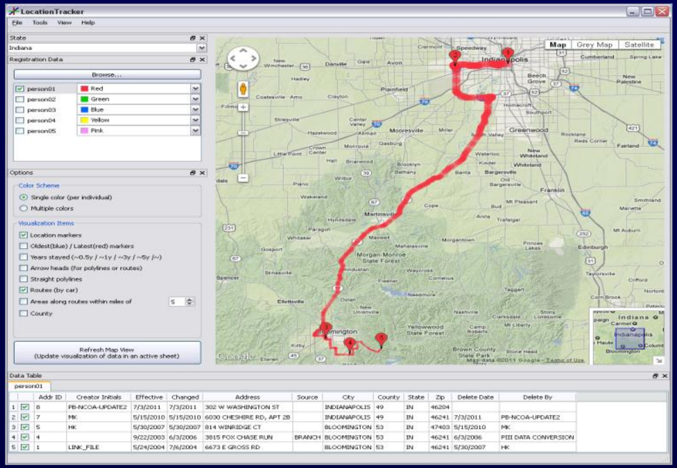
1. State selector
2. Data browse button
3. Data list view
5. Color scheme selector
6. Visualization options
7. Refresh button
8. Table view

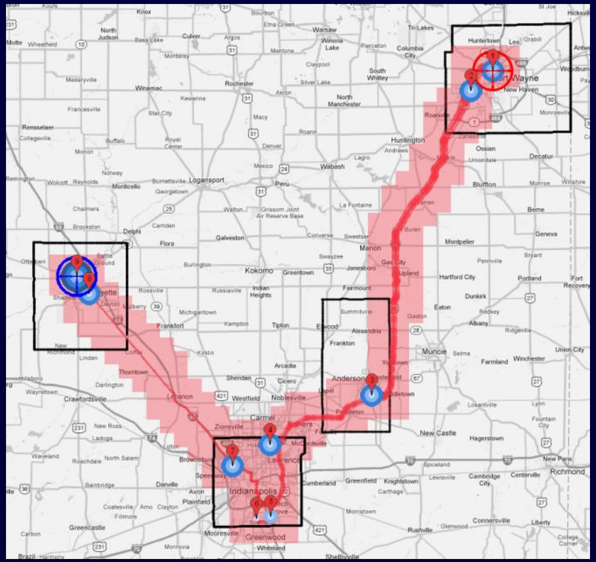


VACCINE

Capabilities and Advantages

- Geocoding
- Routes
- Reachable areas
- Linked table and map views
- Temporal cues
- Various map types
- Various visualization items
- Multiple data comparison
- Selective visualization





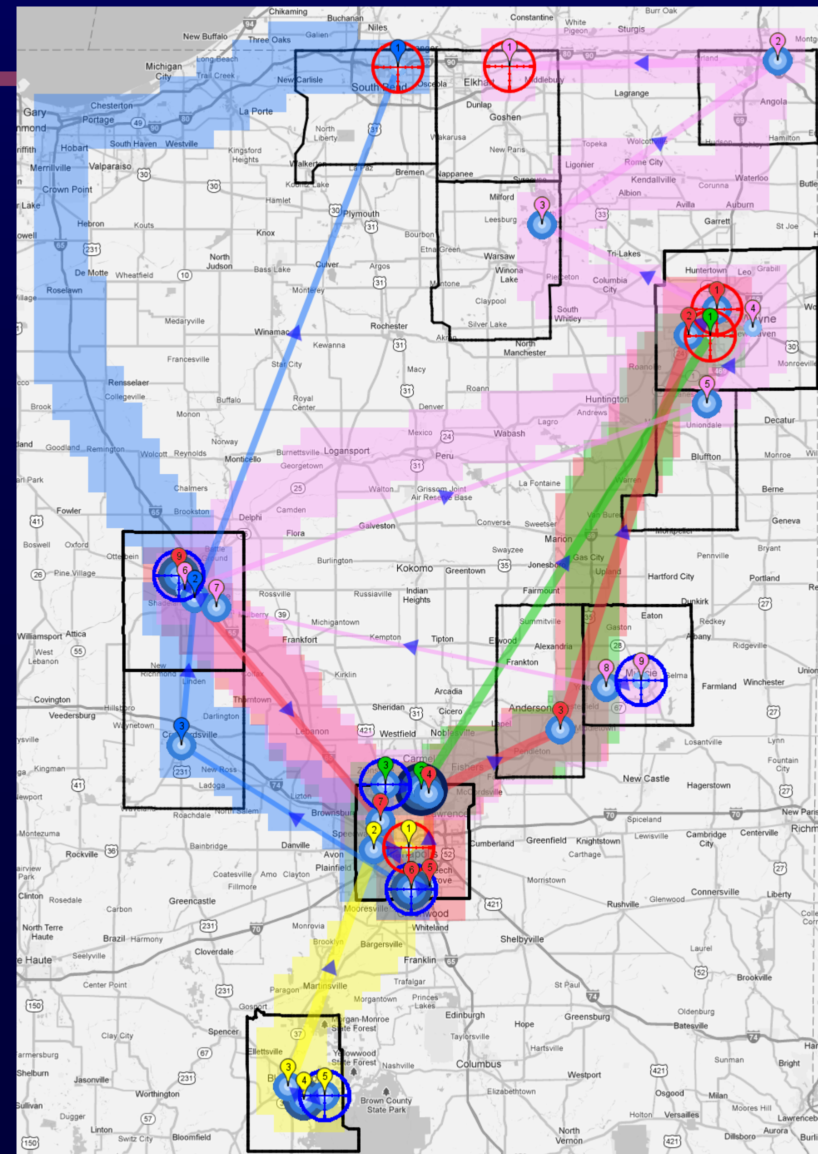
VACCINE

Multiple Data Comparison

- Example of 5 individual data

person01	Yellow
person02	Green
person03	Blue
person04	Red
person05	Pink

  - Living in same or close county
    - Person02 (Green) and Person04 (Red)
    - Person03 (Blue) and Person05 (Pink)
  - Shared reachable areas
    - Person02 (Green) and Person04 (Red)
  - Recent movement into Indianapolis
    - Person01 (Yellow)



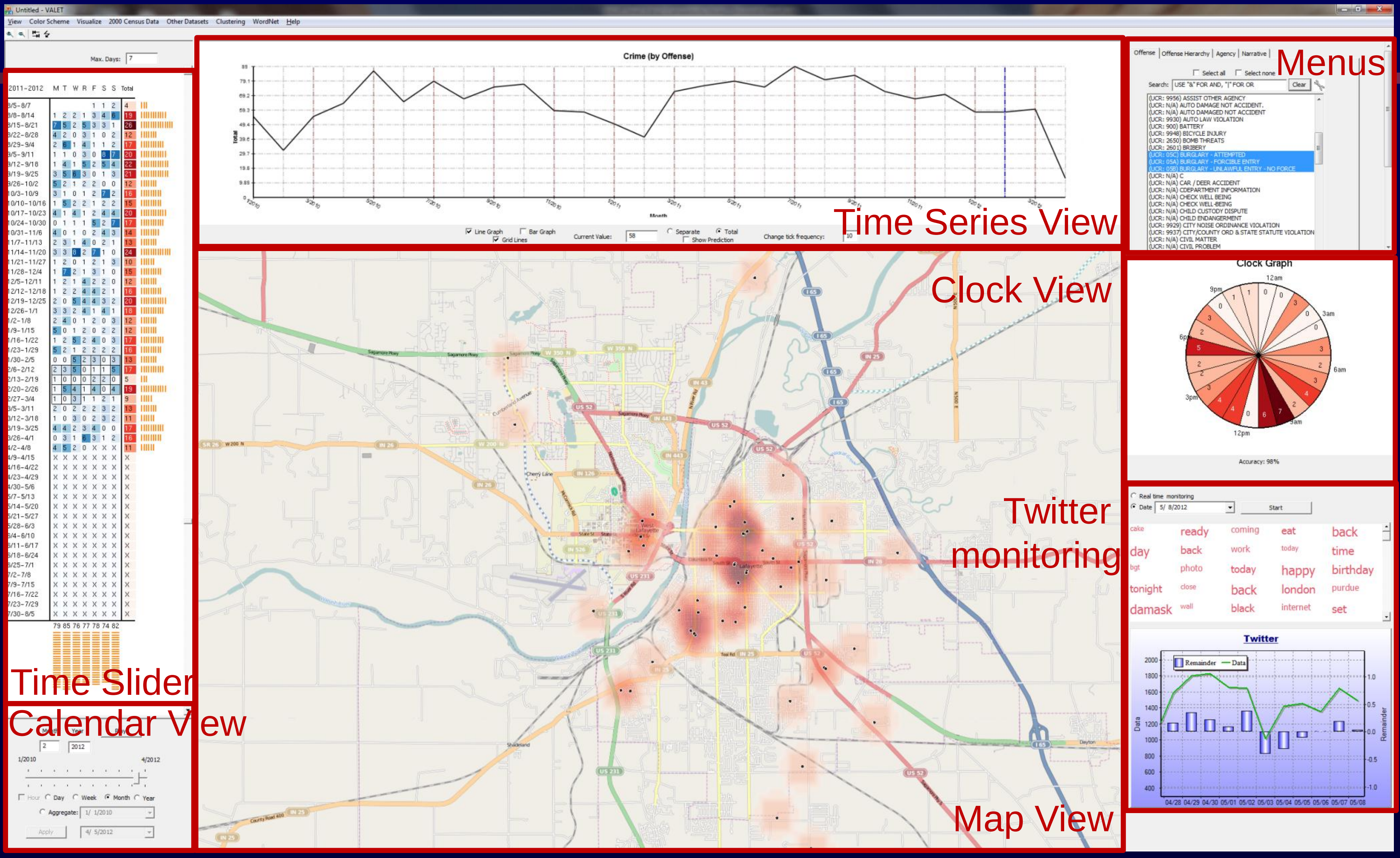
VACCINE

Crime Analytics: VALET – Features and Benefits

- Analysis of spatiotemporal law enforcement data
- Criminal data exploration and analysis
- Linked statistical and geographic visualizations
- Comparison of datasets for potential correlations
- Detection of patterns at different granularity levels
- Predictive analysis
- Resource allocation strategy development
- Twitter integration for real-time data monitoring

VACCINE

System – Overview



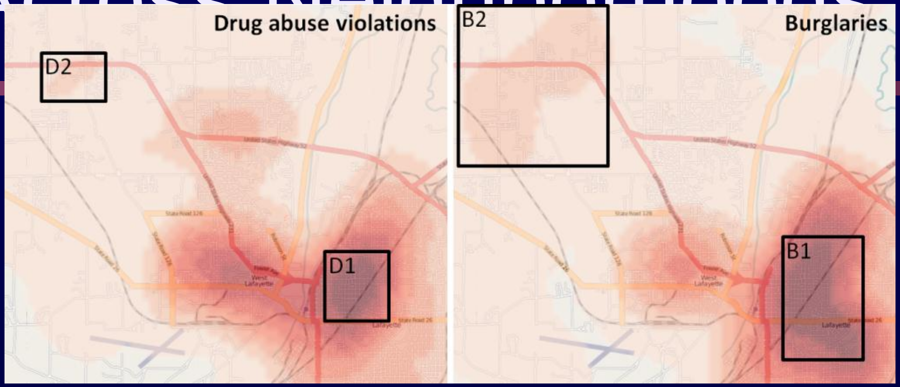
VALET Feature Highlight: Correlating Datasets at Multiple Granularities

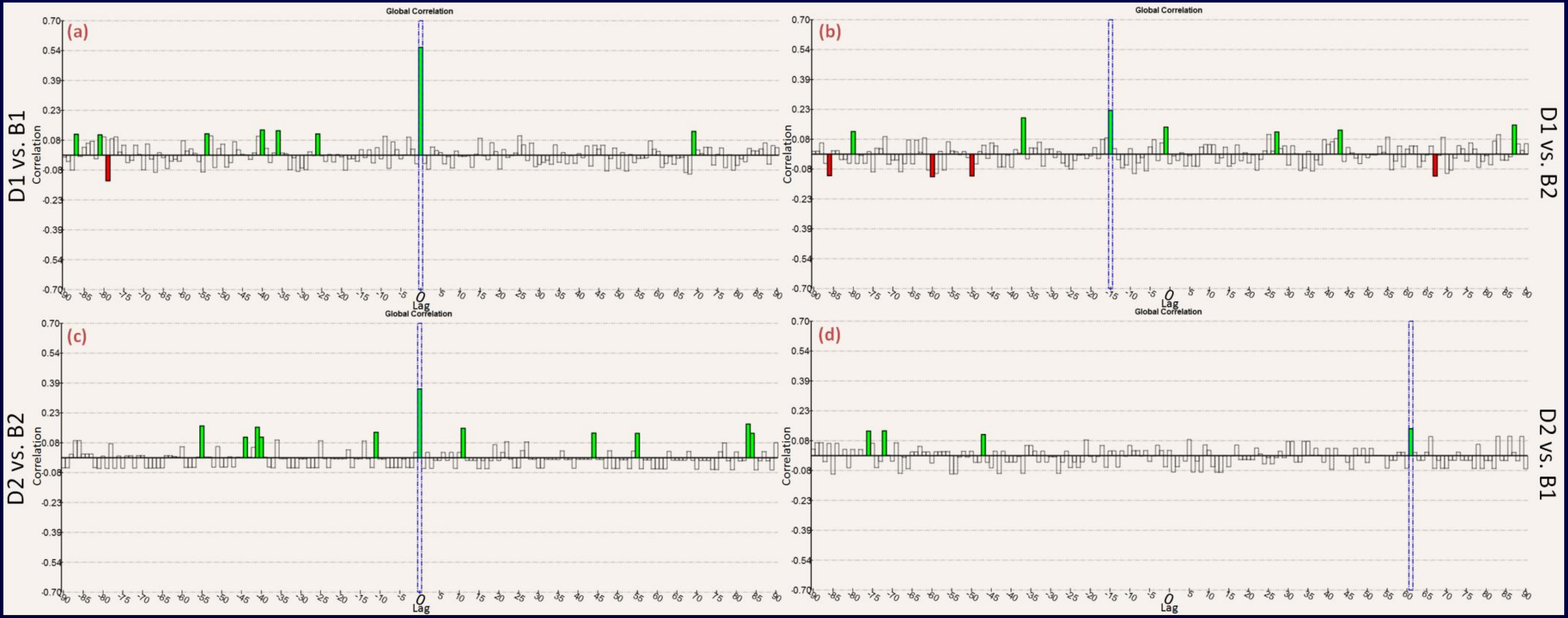
- Correlative analysis framework for exploring correlations among multi-variate/-source data
- Determining forces that determine nature of relationships among datasets
- Detection of periodic properties among different data variables
- Allows correlations at multiple spatial and temporal granularities

VACCINE

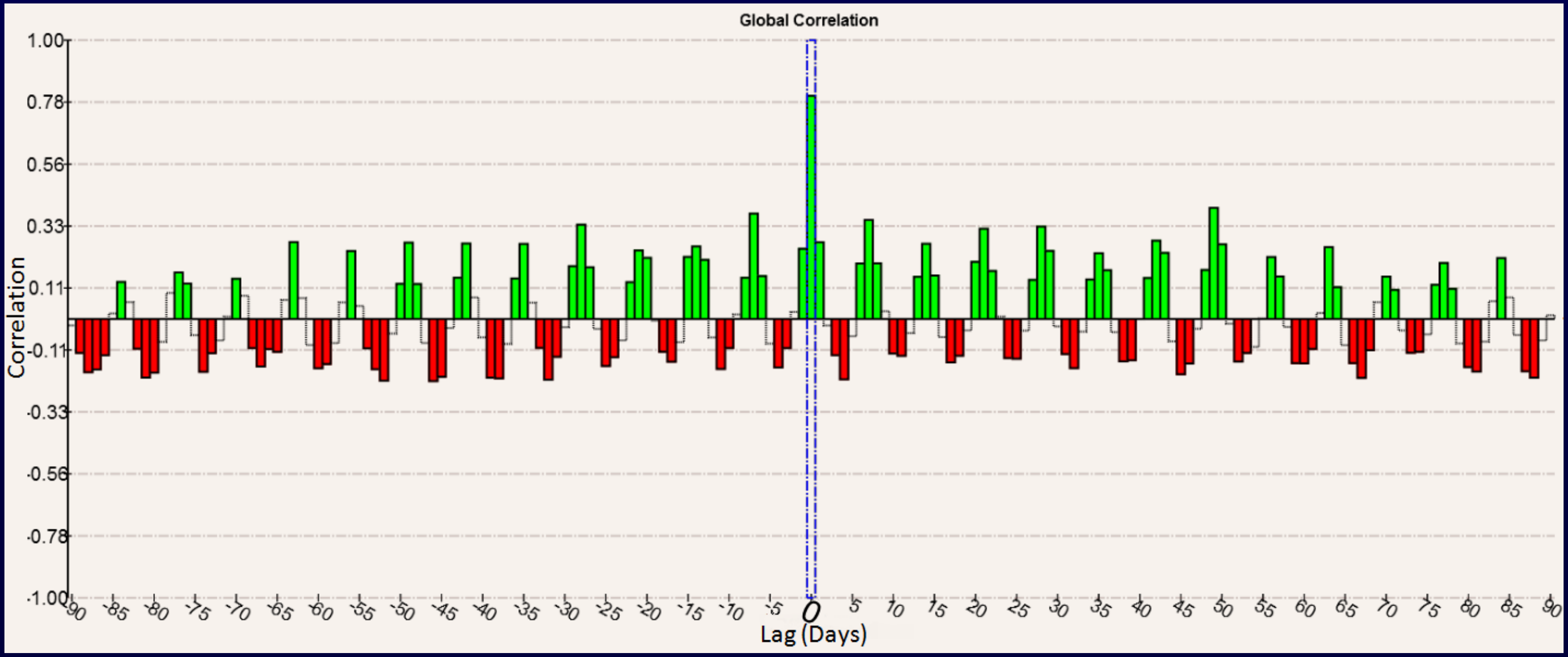
Comparing Drug Abuse Violations & Burglaries: Within and Across Neighborhoods

- Tippecanoe County, IN





Drunkenness/public intoxication and noise complaints vs. traffic accidents



Drunkenness/public intoxication/noise complaints and traffic accidents repeat periodically every 7 days

VACCINE

VALET Feature Highlight:
Real-time Twitter Monitoring and Integration

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud

Grand Prix Weekend, Purdue University

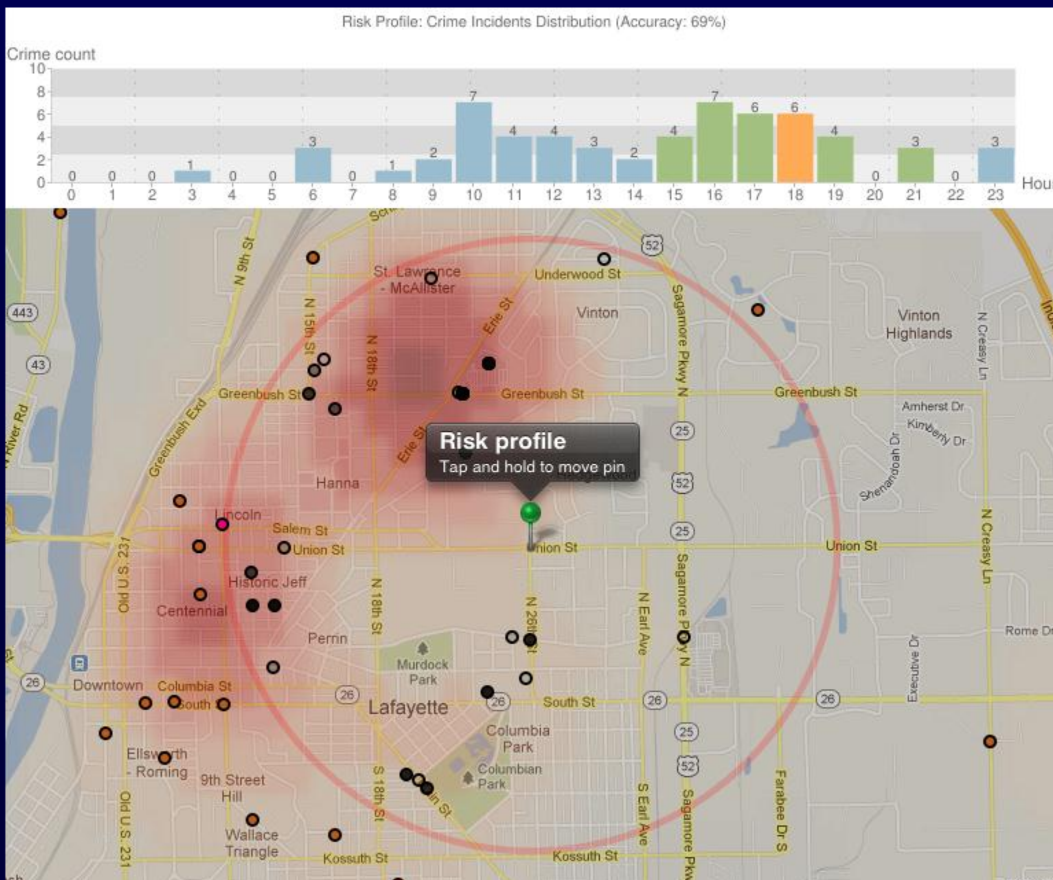
VACCINE

iVALET

- Explore criminal, traffic and civil data on-the-go
- Risk assessment
- Linked views to explore the multivariate spatiotemporal dataset

VACCINE

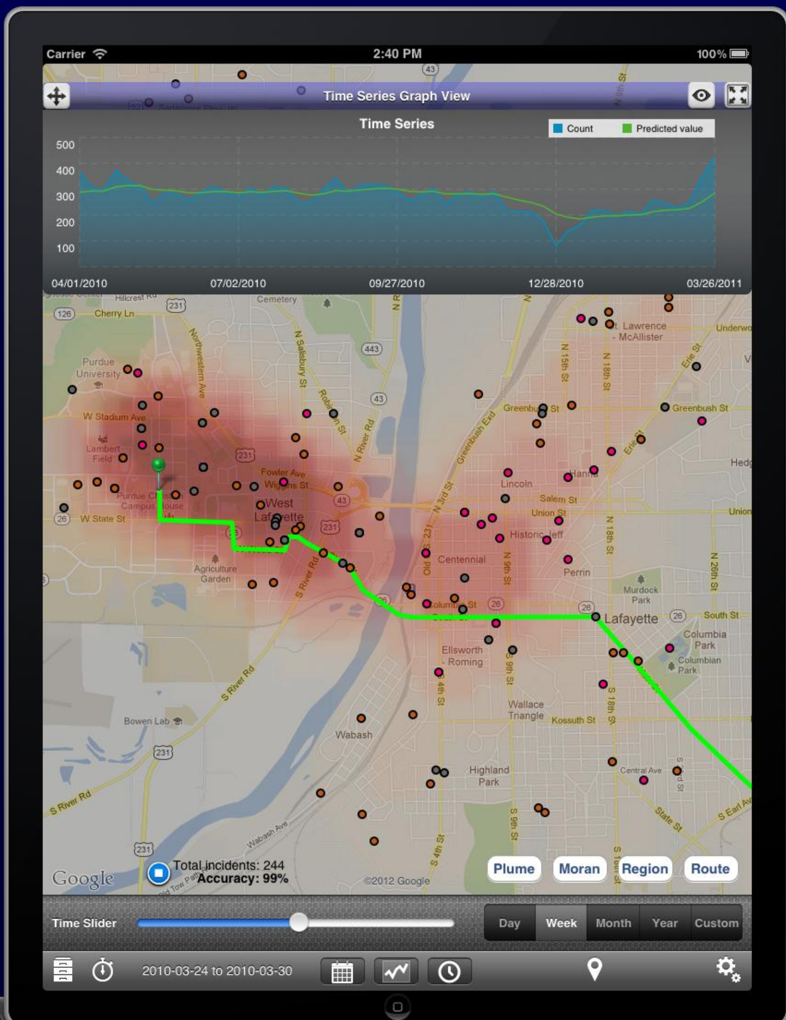
iVALET – Risk profile – Where & When



- Use current spatial & temporal context of user into situational assessment and risk analysis

VACCINE

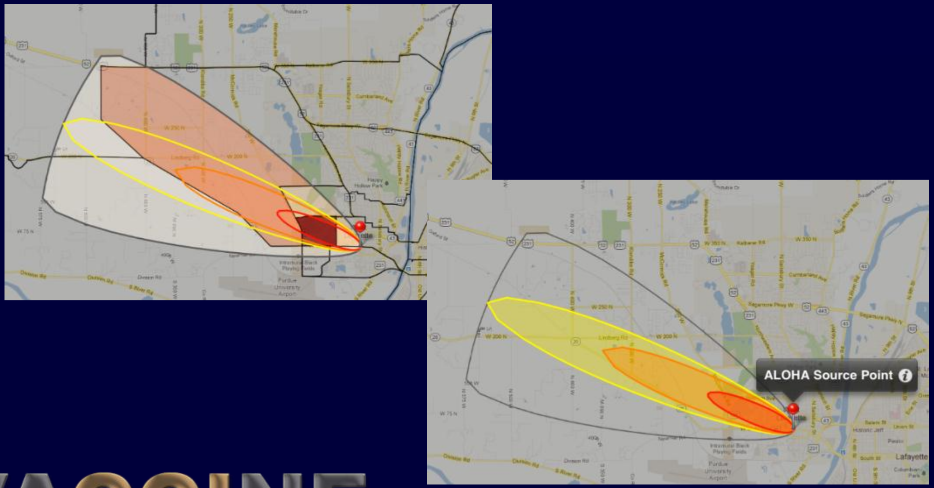
iVALET

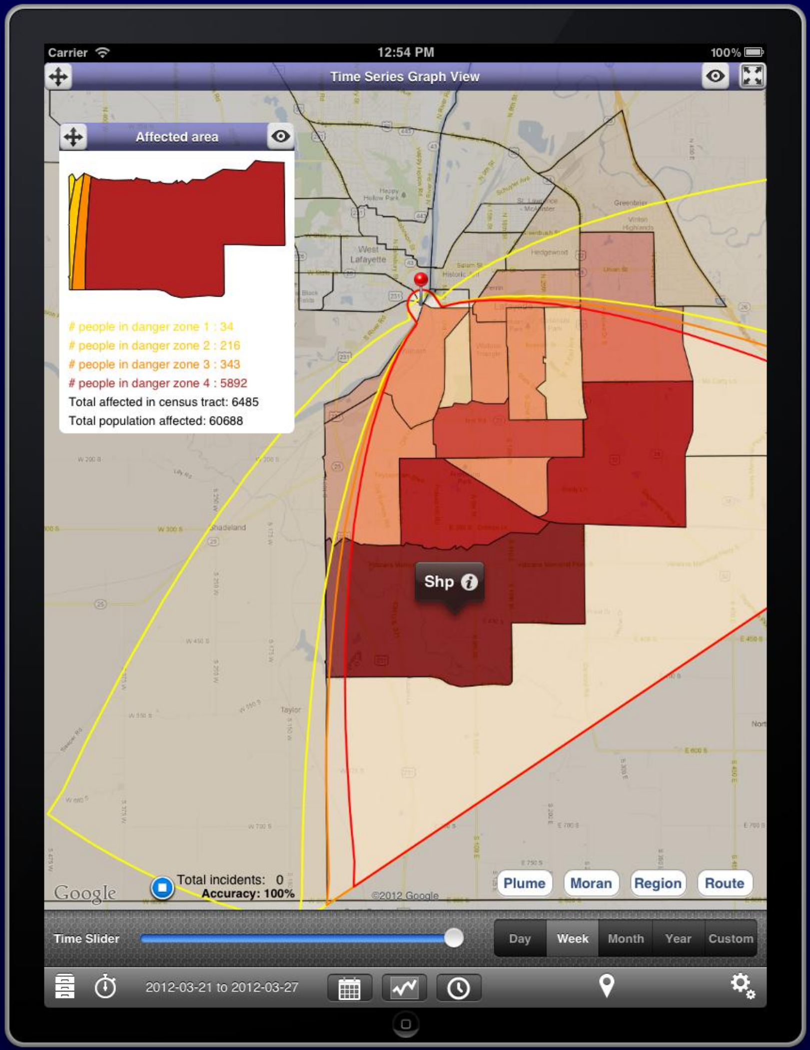


- Display a path between 2 (user-specified) points to show risks along intended route

iVALET – ALOHA plume model integration

- The chemical release plume model identifies census tracts with the highest number of expected people affected.

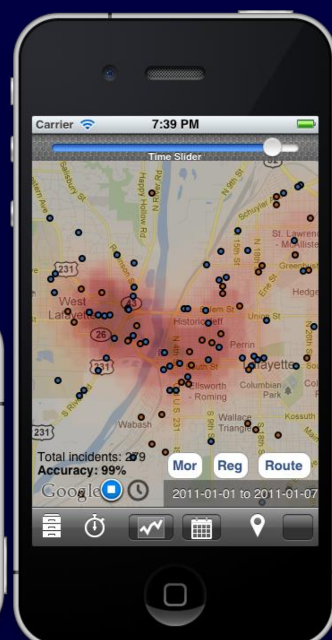


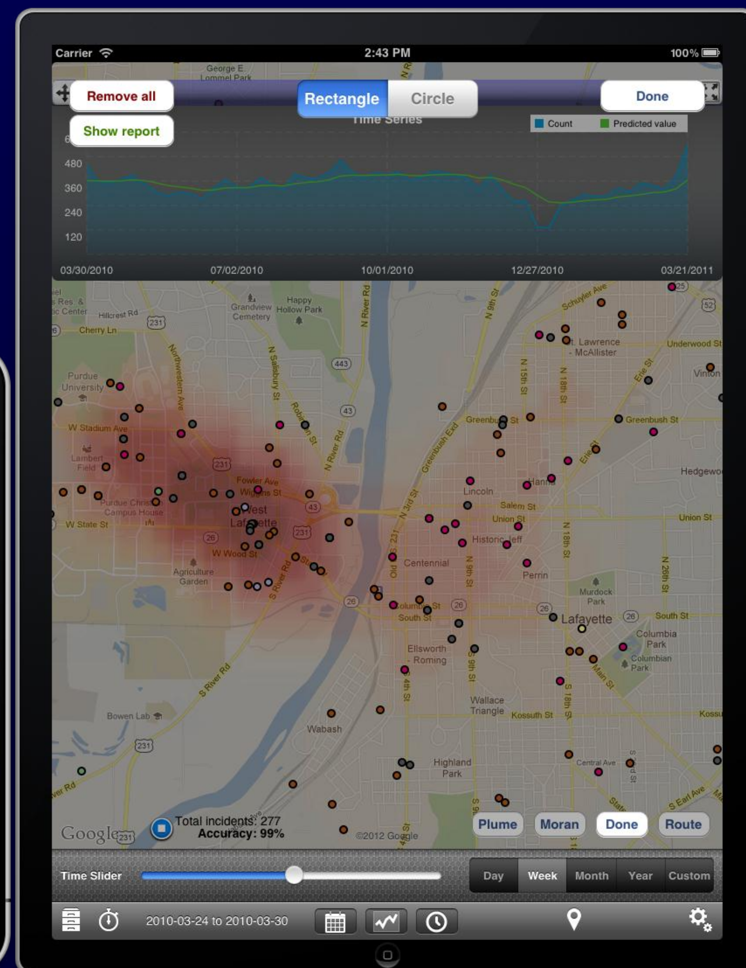


VACCINE

iVALET

- Works on iPhone/iPad
- Readily downloadable at iVALET website





Jigsaw

Visualization for Investigative Analysis across Document Collections

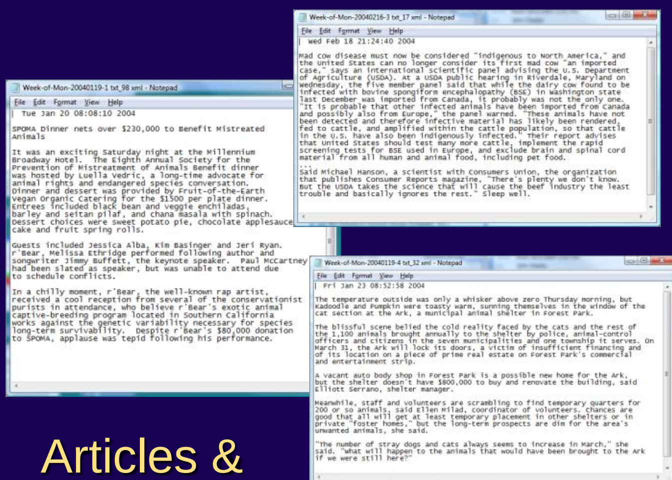
- Law enforcement & intelligence community
- Fraud (finance, accounting, banking)
- Academic research
- Journalism & reporting
- Consumer research

"Putting the pieces together"

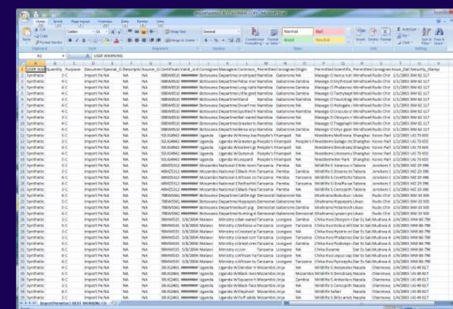
information interfaces

Problem Addressed

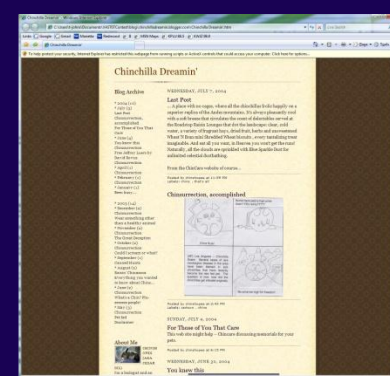
Help "investigators" explore, analyze and understand large document collections



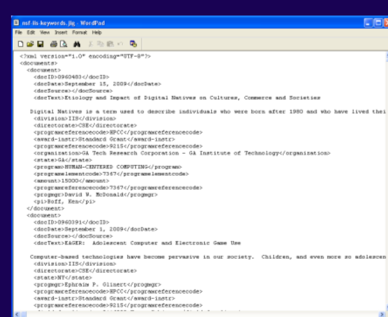
Articles & reports



Spreadsheets



Blogs



XML documents

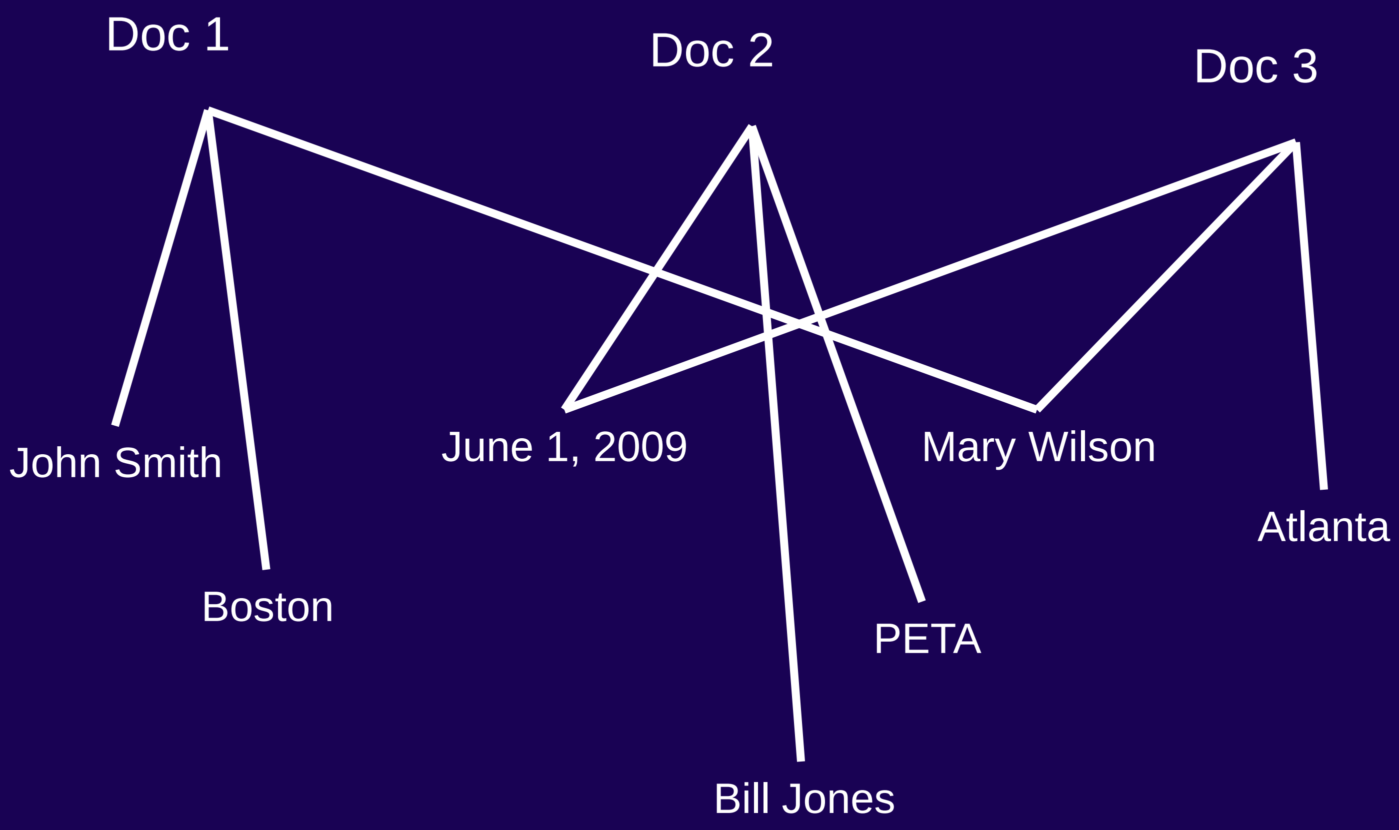
Example Documents

1) Report Date: 12 July, 2004. A routine customs inspection was performed on a package that was sent by a person named Pieter Dopple, 22 Hoveniersstraat, Antwerp, Belgium. This package was addressed to A. Hijazi, 1212 Lyons Ave., Newark, NJ. The customs form stated that this package contained two decorative clocks having a commercial value of $250.00. The package instead contained 111 polished diamonds, whose value is estimated to be $47,000. Discussions with the FBI resulted in a decision not to pursue a customs violation charge against HIjazi. The reason is that HIjazi is currently under FBI surveillance. This package was resealed and delivered to HIjazi by USPS.

2) Report Date: 22 February, 2003. Surveillance report on Cesar Arze, whose residence is 77 Avenue Francis, Santo Domingo, Dominican Republic. Arze, who moved from Havana, Cuba to Santo Domingo in 1992, works as a medical technician in Santo Domingo. Arze is under surveillance because of information that he is associated with Cuban intelligence services. Arze was photographed in company with a man identified as Hector Lopez in Bogota, Columbia on 23 January, 2003. Lopez, a known representative of FARC, has conducted narcotics distribution activities throughout South and Central America and the Caribbean.

Analytic Focus

- Entities within the documents
  - Person, place, organization, phone number, date, license plate, etc.
- Thesis: A story/narrative/plot/threat within the documents will involve a set of entities in coordination


Doc 1
Doc 2
Doc 3
John Smith
Boston
June 1, 2009
Bill Jones
PETA
Mary Wilson
Atlanta

Sample Document 1

Report: 20040510-4_16
May 14 2004

VANCOUVER, British Columbia - A Canadian immigration panel is considering whether accused environmental saboteur Tre Arrow can apply for refugee status in Canada.

Arrow, 30, who is wanted for fire bombing logging and cement trucks in Oregon, asked the Canadian authorities to remain in Canada as a political refugee at a hearing in Vancouver on Tuesday.

A key issue will be whether Arrow is affiliated with a terrorist group, which would immediately disqualify him from receiving refugee status in Canada, authorities said.

The Immigration and Refugee Board is scheduled to decide by May 31 whether Arrow is affiliated with the Earth Liberation Front, a group the FBI considers a terrorist organization responsible for scores of attacks on property over the past dozen years.

Entities Identified

Source:
Date: May 14, 2004

VANCOUVER, British Columbia - A Canadian immigration panel is considering whether accused environmental saboteur Tre **Arrow** can apply for refugee status in Canada.

Arrow, 30, who is wanted for fire bombing logging and cement trucks in Oregon, asked the Canadian authorities to remain in Canada as a political refugee at a hearing in Vancouver on Tuesday.

A key issue will be whether **Arrow** is affiliated with a terrorist group, which would immediately disqualify him from receiving refugee status in Canada, authorities said.

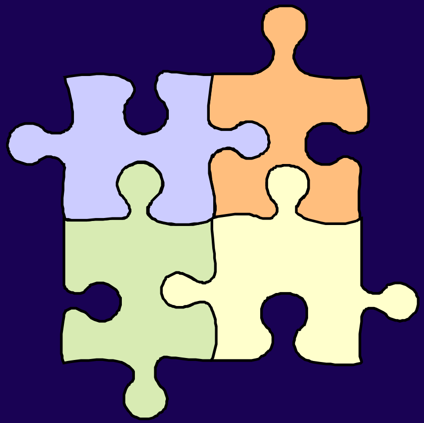
The Immigration and Refugee Board is scheduled to decide by May 31 whether **Arrow** is affiliated with the Earth Liberation Front, a group the FBI considers a terrorist organization responsible for scores of attacks on property over the past dozen years.

Connections

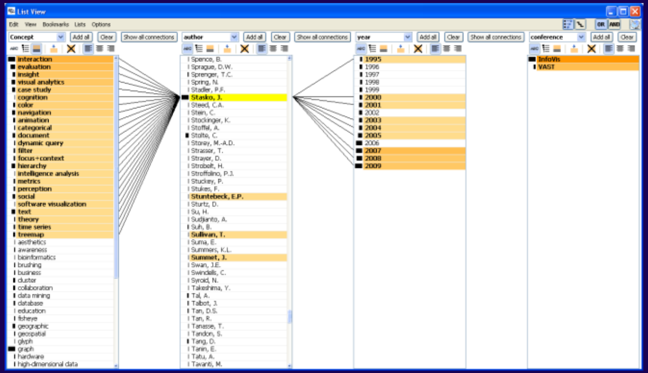
- Entities relate/connect to each other to make a larger "story"
- Connection definition:
  - Two entities are connected if they appear in a document together
  - The more documents they appear in together, the stronger the connection

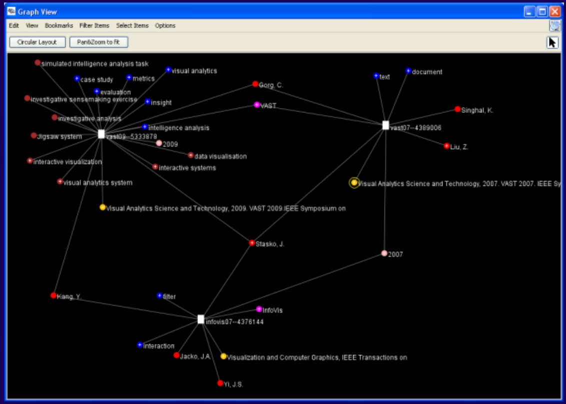
Jigsaw

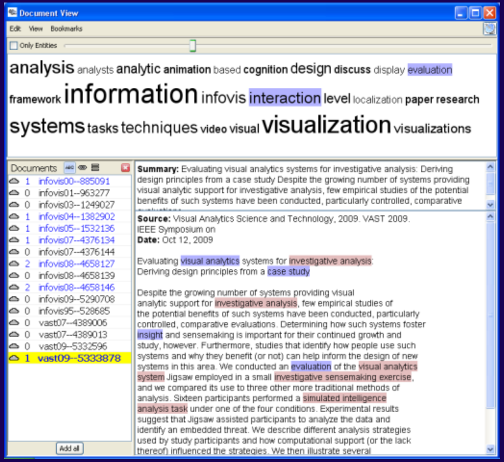
- Computational analysis of document text
  - Entity identification, document similarity, clustering, summarization, sentiment
- Multiple visualizations (views) of documents, analysis results, entities and their connections
- Views are highly interactive and coordinated

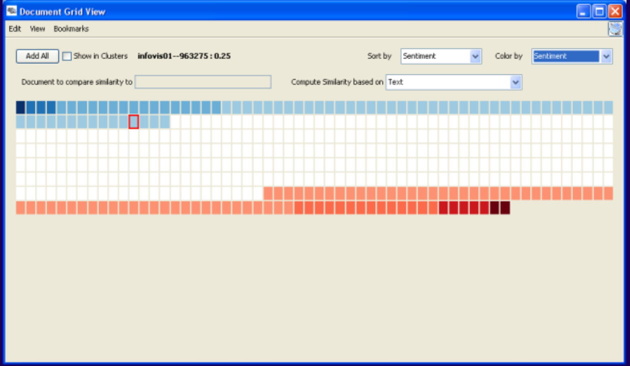


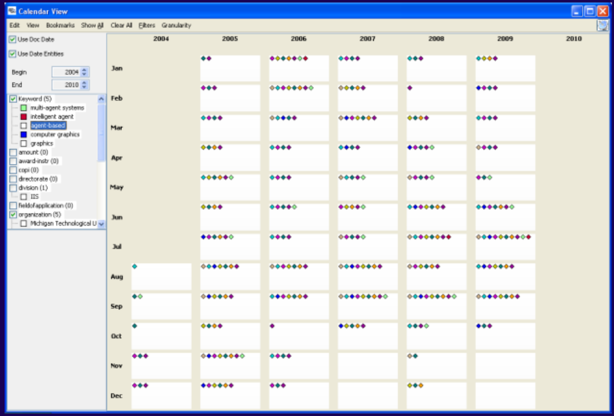
System Views

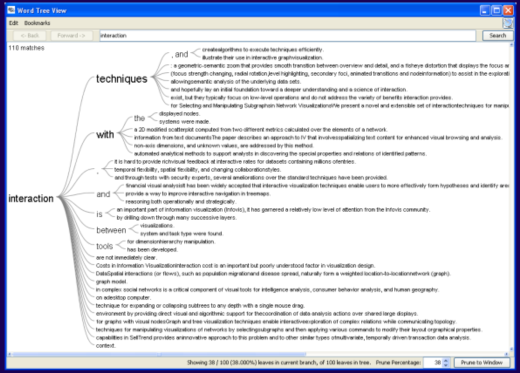


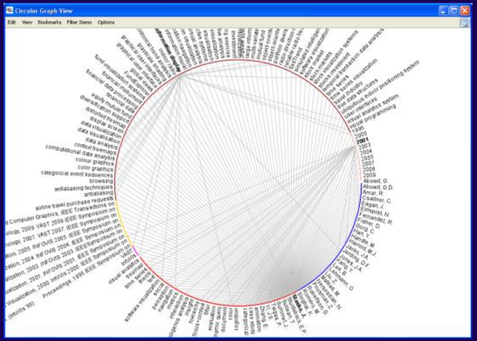






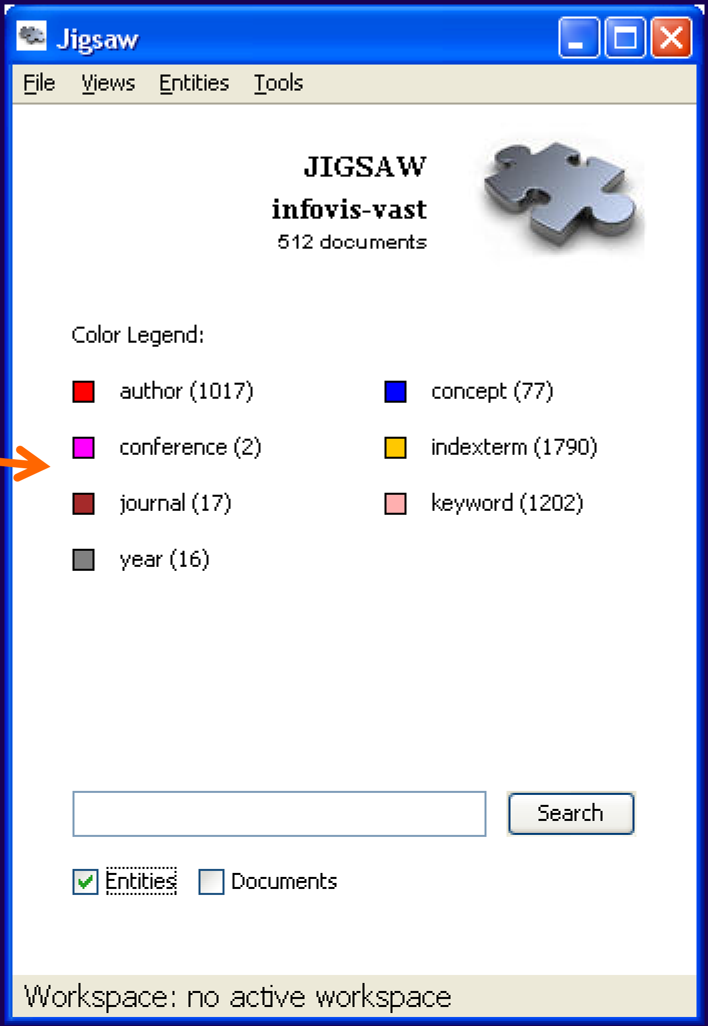




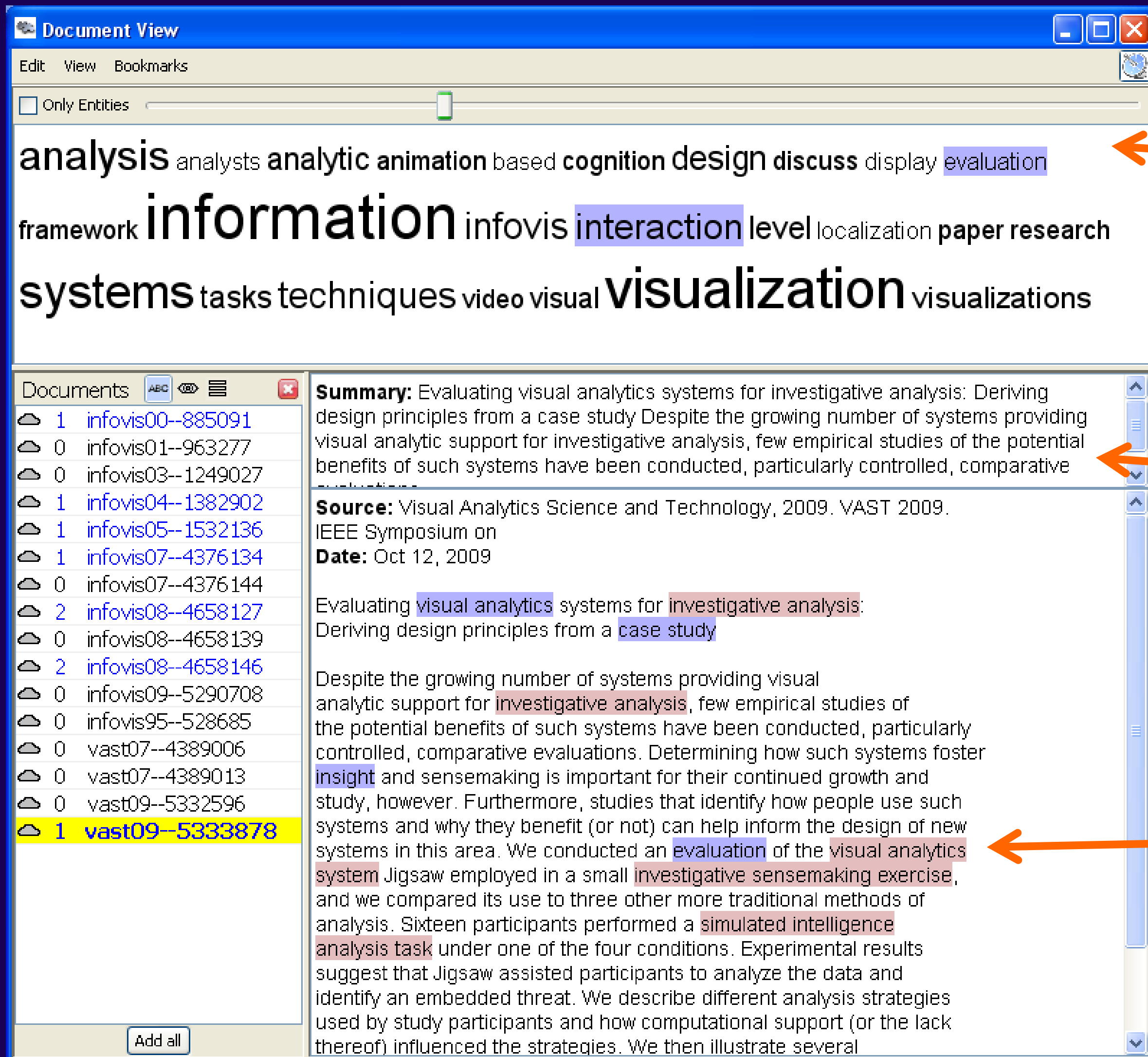


Console

Entity types



Document View



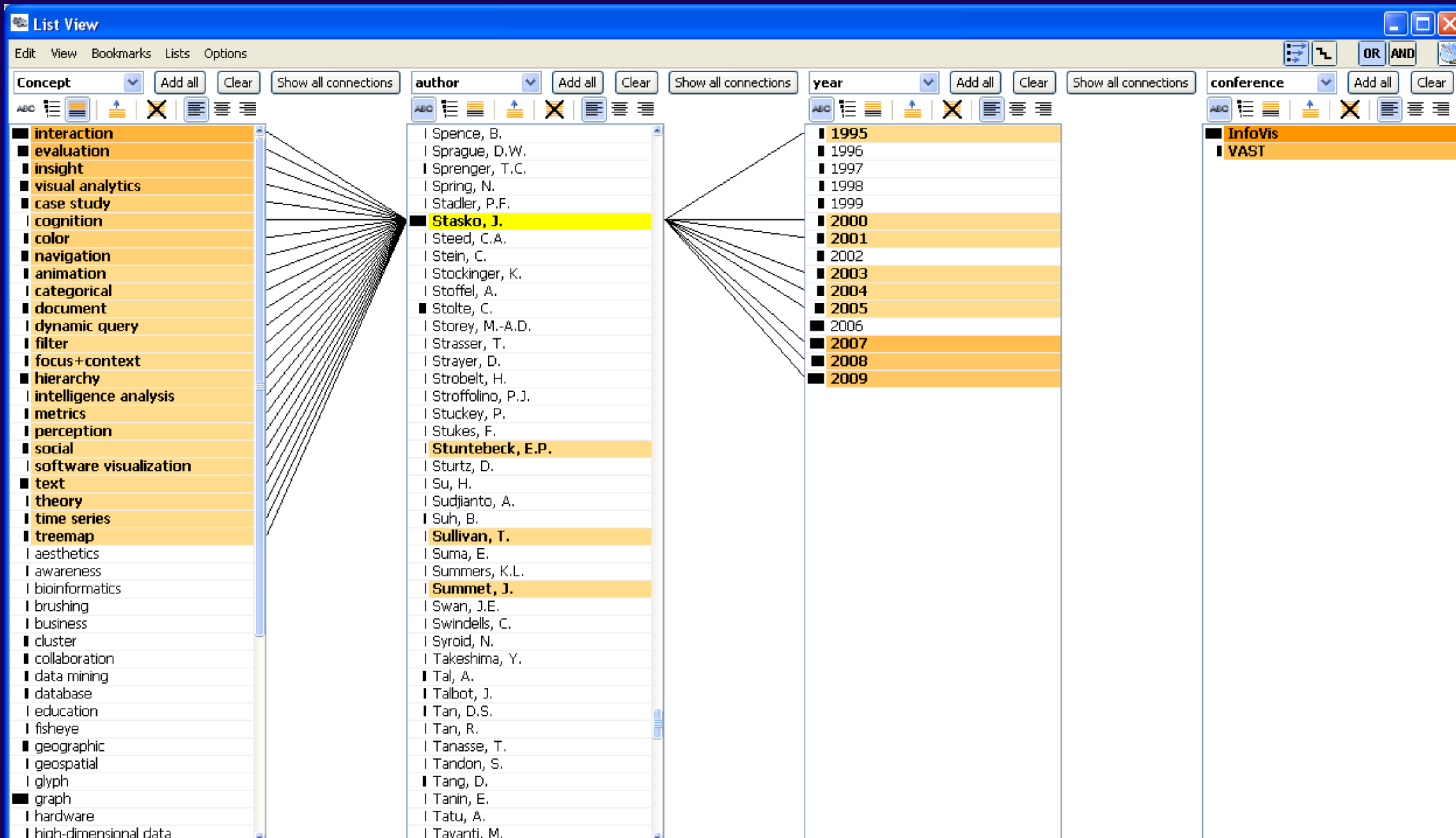
Important words in loaded docs

Automatic summary

Entities identified

List View

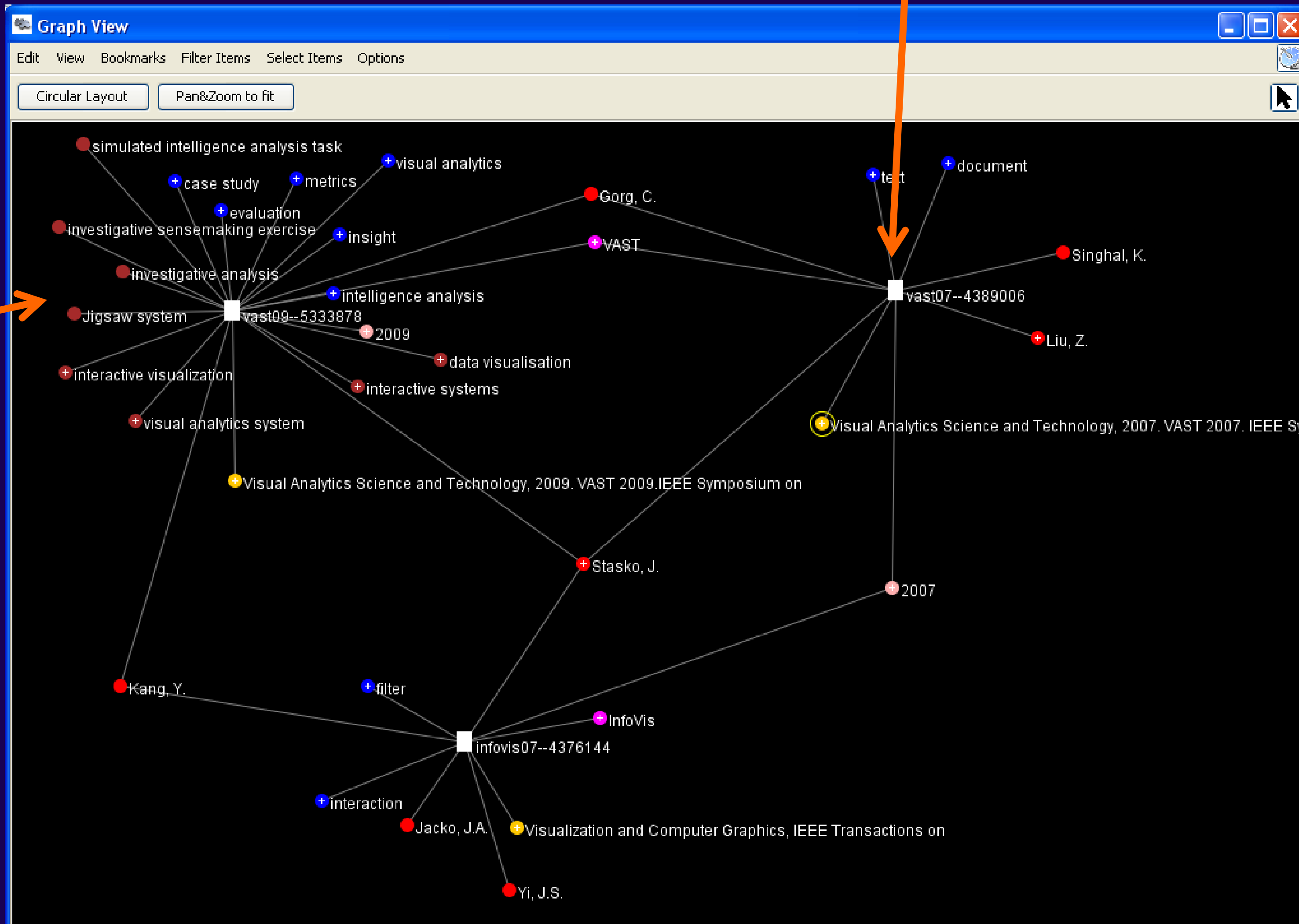
Lists of entities by type
Connections highlighted



Graph View

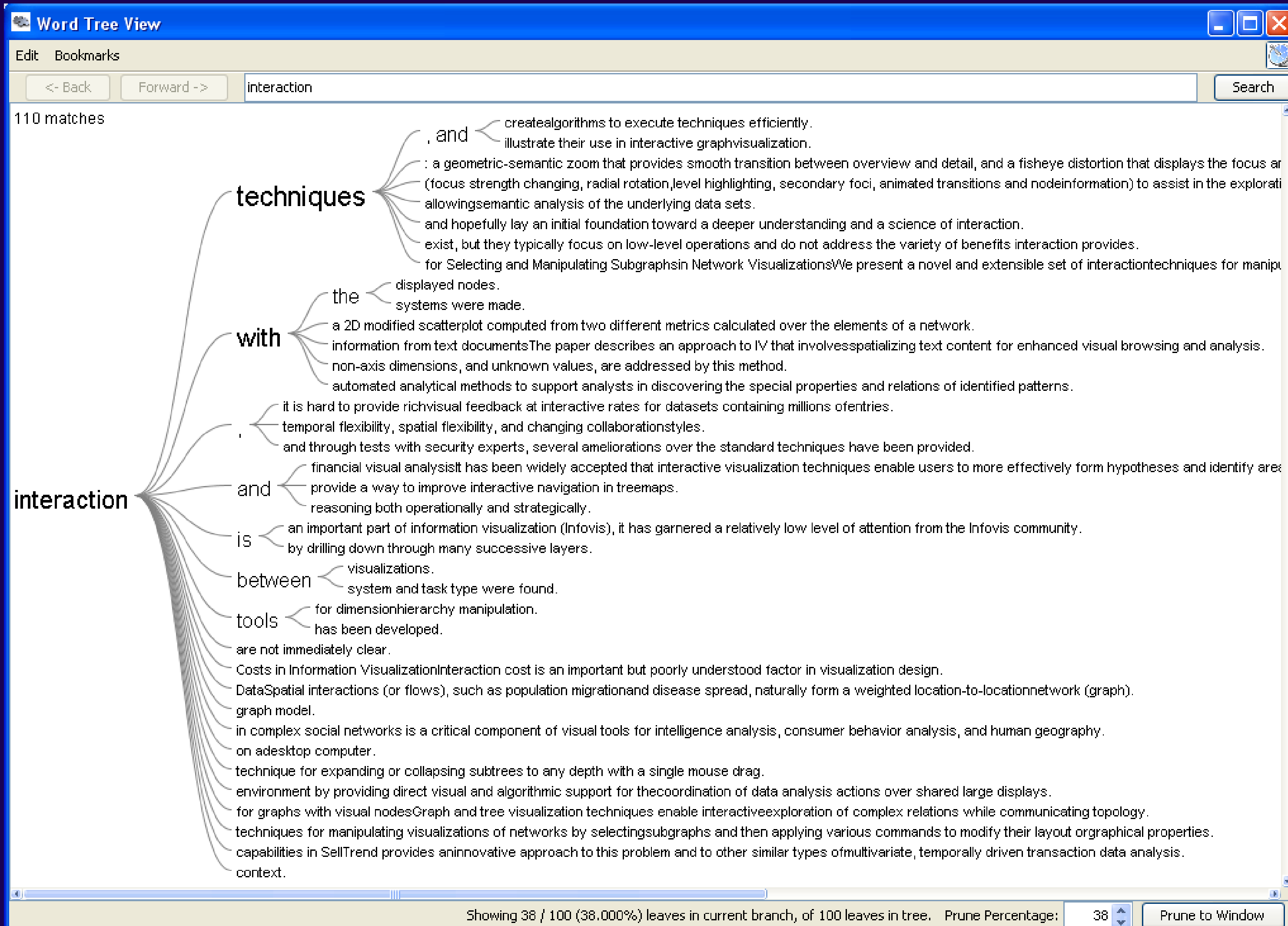
Document

Entities



WordTree View

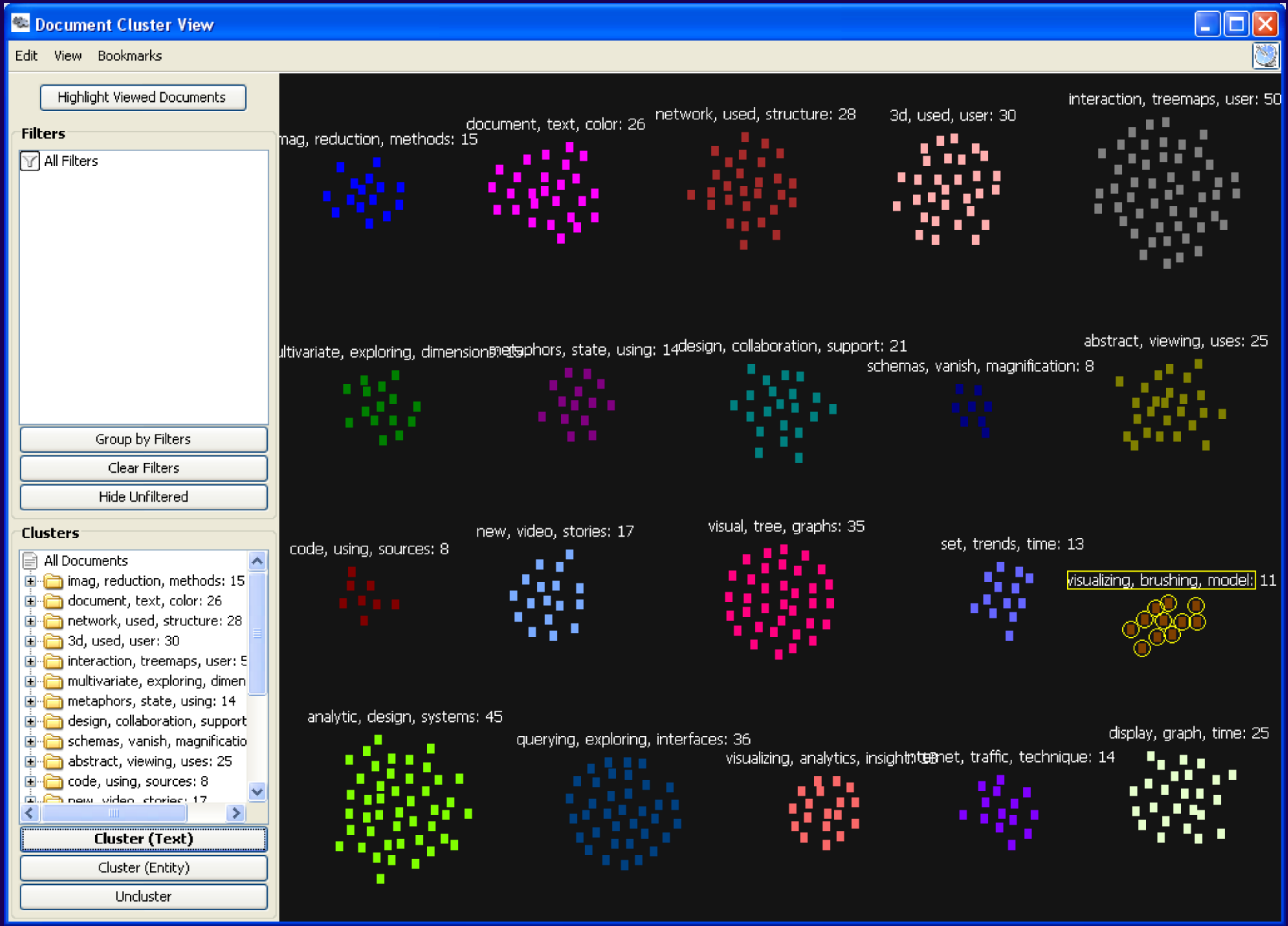
Context of a word
in the collection



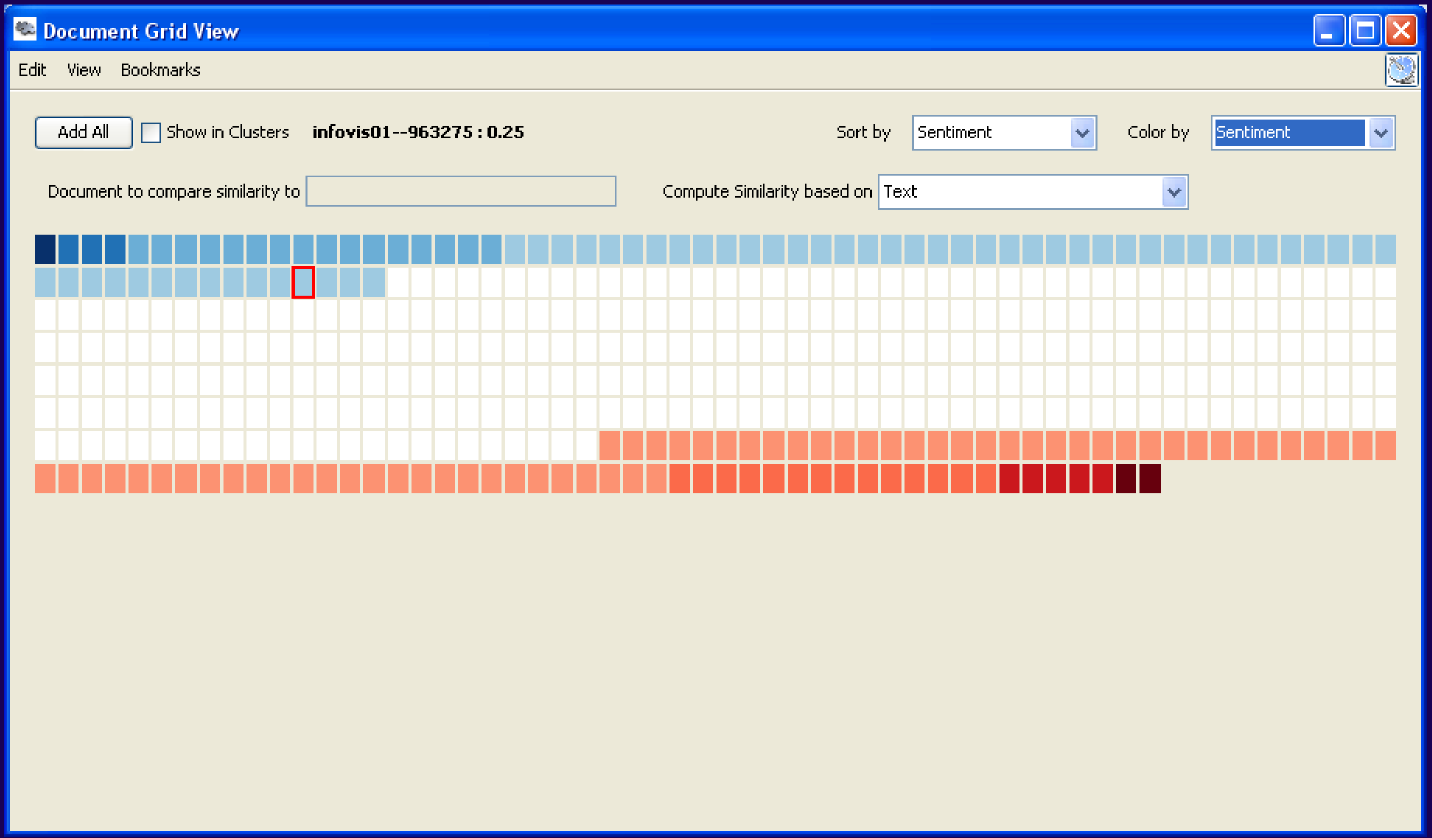
Document Cluster View

Clustered by document text or by entities

Summarized by three words



Document Grid View

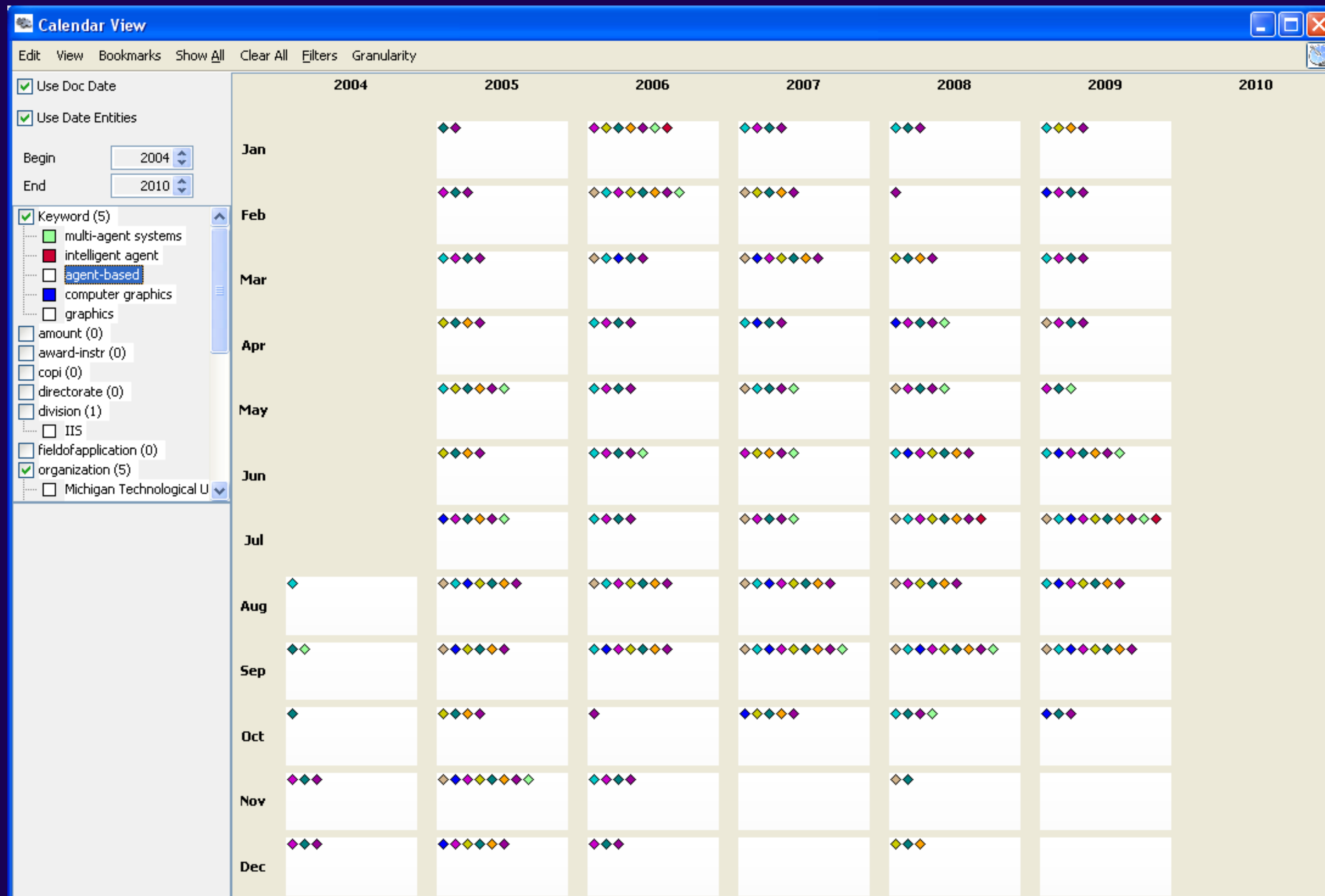


User controls order and color

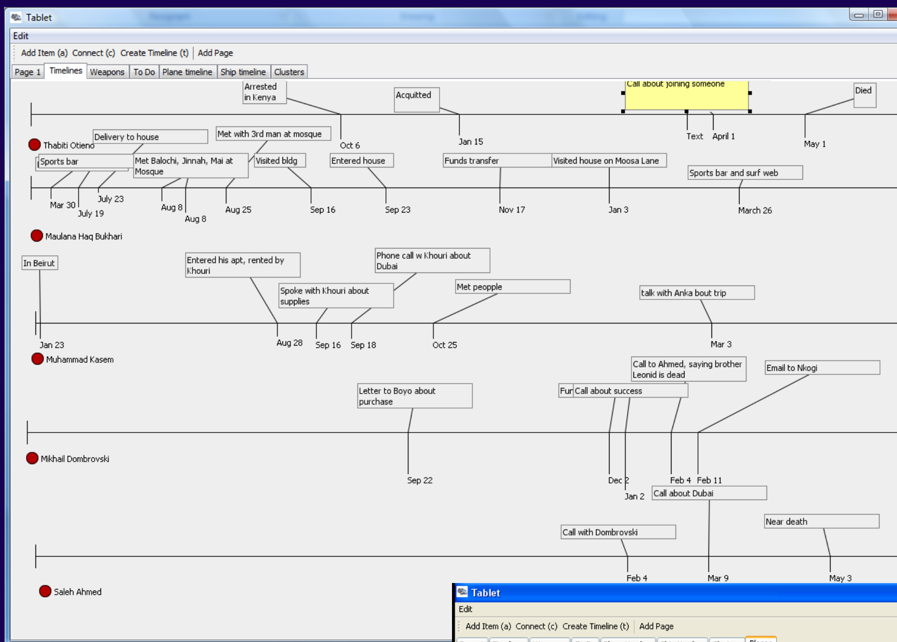
Sentiment analysis shown here

Calendar View

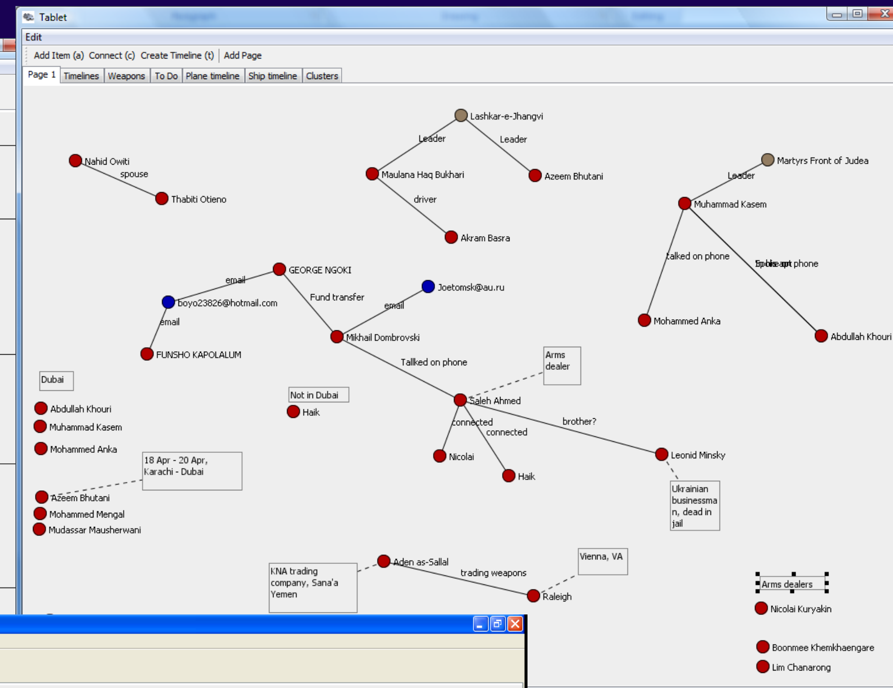
Showing connections between entities and dates



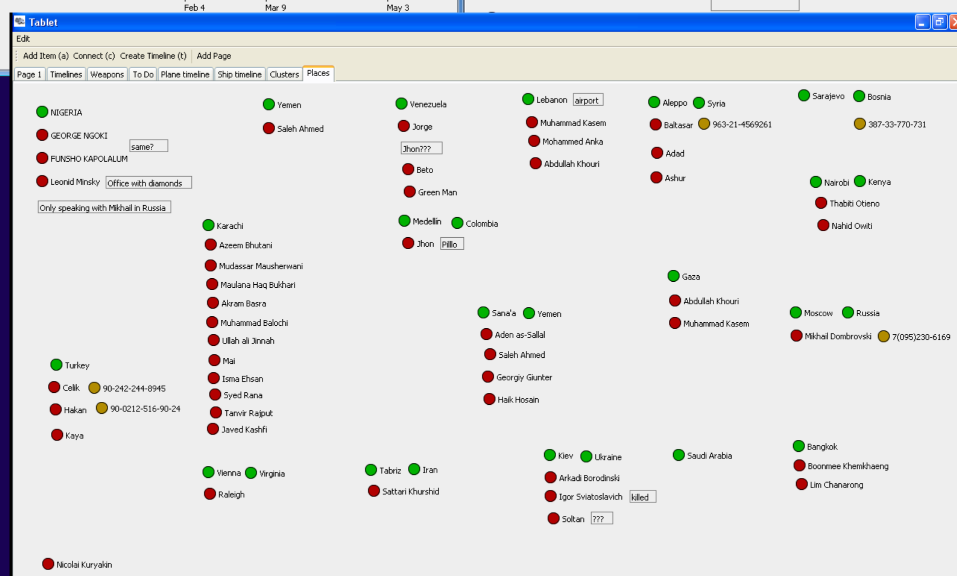
Tablet



Tablet
Edit
Add Item (a) Connect (c) Create Timeline (t) Add Page
Page 1
Timelines
Weapons
To Do
Plane timeline
Ship timeline
Clusters
Arrested in Kenya
Acquitted
Call about joining someone
Died
Thabiti Otieno
Delivery to house
Met with 3rd man at mosque
Oct 6
Jan 15
Text
April 1
May 1
Sports bar
Met Balochi, Jinnah, Mai at Mosque
Visited bldg
Entered house
Funds transfer
Visited house on Moosa Lane
Sports bar and surf web
Mar 30
July 19
July 23
Aug 8
Aug 8
Aug 25
Sep 16
Sep 23
Nov 17
Jan 3
March 26
Maulana Haq Bukhari
In Beirut
Entered his apt, rented by Khouri
Phone call w Khouri about Dubai
Spoke with Khouri about supplies
Met peopple
talk with Anka bout trip
Jan 23
Aug 28
Sep 16
Sep 18
Oct 25
Mar 3
Muhammad Kasem
Call to Ahmed, saying brother Leonid is dead
Email to Nkogi
Letter to Boyo about purchase
Call about success
Mikhail Dombrovski
Sep 22
Dec 2
Jan 2
Feb 4
Feb 11
Call about Dubai
Near death
Call with Dombrovski
Saleh Ahmed
Feb 4
Mar 9
May 3



Tablet
Edit
Add Item (a) Connect (c) Create Timeline (t) Add Page
Page 1
Timelines
Weapons
To Do
Plane timeline
Ship timeline
Clusters
Lashkar-e-Jhangvi
Leader
Leader
Nahid Owiti
spouse
Thabiti Otieno
Maulana Haq Bukhari
Azeem Bhutani
driver
Akram Basra
Martyrs Front of Judea
Leader
Muhammad Kasem
talked on phone
Spoke on phone
Mohammed Anka
Abdullah Khouri
GEORGE NGOKI
email
boyo23826@hotmail.com
email
FUNSHO KAPOLALUM
Fund transfer
Joetomsk@au.ru
email
Mikhail Dombrovski
Tallked on phone
Arms dealer
Dubai
Abdullah Khouri
Muhammad Kasem
Mohammed Anka
Not in Dubai
Haik
Saleh Ahmed
connected
connected
brother?
Nicolai
Haik
Leonid Minsky
18 Apr - 20 Apr, Karachi - Dubai
Azeem Bhutani
Mohammed Mengal
Mudassar Mausherwani
Ukrainian businessman, dead in jail
Aden as-Sallal
trading weapons
Vienna, VA
KNA trading company, Sana'a Yemen
Raleigh
Arms dealers
Nicolai Kuryakin
Boonmee Khemkhaengare
Lim Chanarong



Tablet
Edit
Add Item (a) Connect (c) Create Timeline (t) Add Page
Page 1
Timelines
Weapons
To Do
Plane timeline
Ship timeline
Clusters
Places
NIGERIA
GEORGE NGOKI
same?
FUNSHO KAPOLALUM
Leonid Minsky
Office with diamonds
Only speaking with Mikhail in Russia
Yemen
Saleh Ahmed
Venezuela
Jorge
Jhon???
Beto
Green Man
Medellin
Colombia
Jhon
Pillo
Lebanon
airport
Muhammad Kasem
Mohammed Anka
Abdullah Khouri
Aleppo
Syria
Baltasar
963-21-4569261
Adad
Ashur
Sarajevo
Bosnia
387-33-770-731
Nairobi
Kenya
Thabiti Otieno
Nahid Owiti
Karachi
Azeem Bhutani
Mudassar Mausherwani
Maulana Haq Bukhari
Akram Basra
Muhammad Balochi
Ullah ali Jinnah
Mai
Isma Ehsan
Syed Rana
Tanvir Rajput
Javed Kashfi
Turkey
Celik
90-242-244-8945
Hakan
90-0212-516-90-24
Kaya
Gaza
Abdullah Khouri
Muhammad Kasem
Sana'a
Yemen
Aden as-Sallal
Saleh Ahmed
Georgiy Giunter
Haik Hosain
Moscow
Russia
Mikhail Dombrovski
7(095)230-6169
Vienna
Virginia
Raleigh
Tabriz
Iran
Sattari Khurshid
Kiev
Ukraine
Arkadi Borodinski
Igor Sviatoslavich
killed
Soltan
???
Saudi Arabia
Bangkok
Boonmee Khemkhaeng
Lim Chanarong
Nicolai Kuryakin

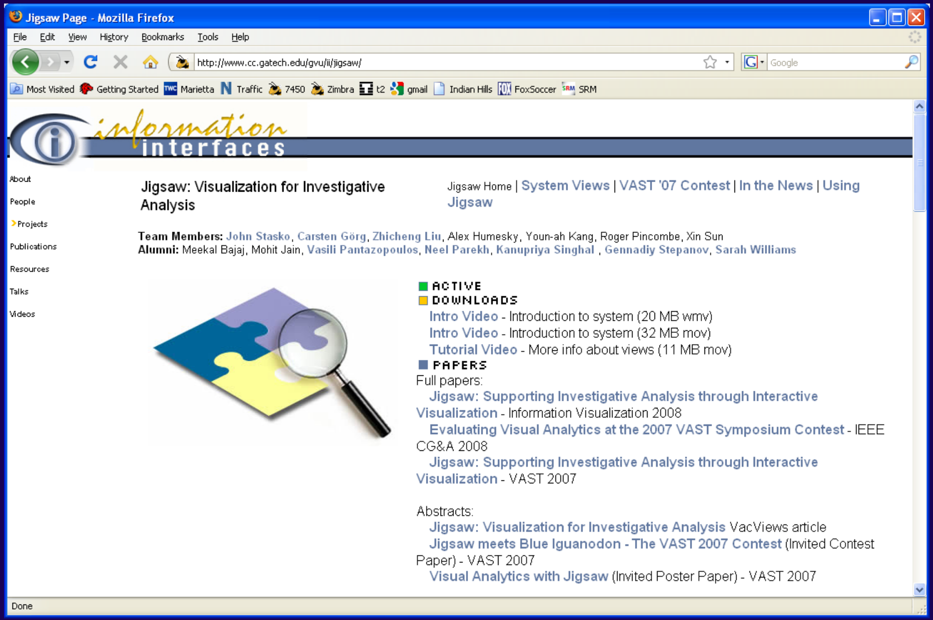
To Learn More about Jigsaw & Availability

Table of Contents

http://www.gvu.gatech.edu/ii/jigsaw

Available for (free) trial use

Send email to:
stasko@cc.gatech.edu



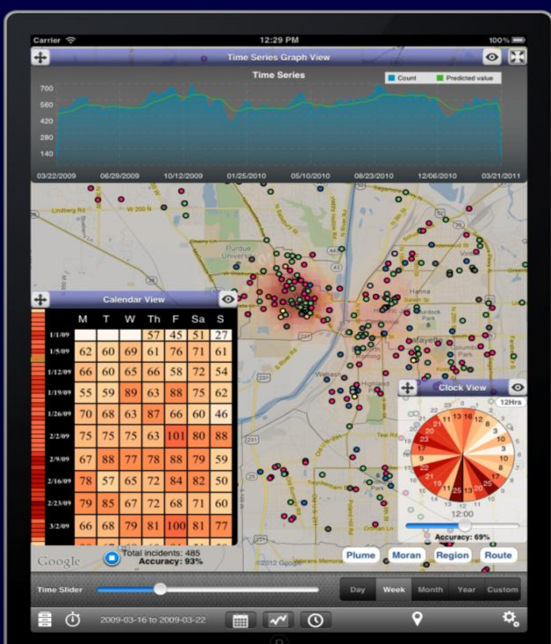
VACCINE

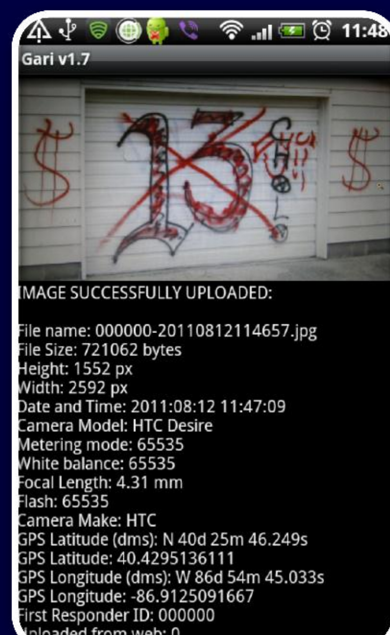
Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

Table of Contents

Spatiotemporal Visualization and Analysis of Gang-related Criminal Activity Using Mobile Imaging





David S. Ebert

Purdue University

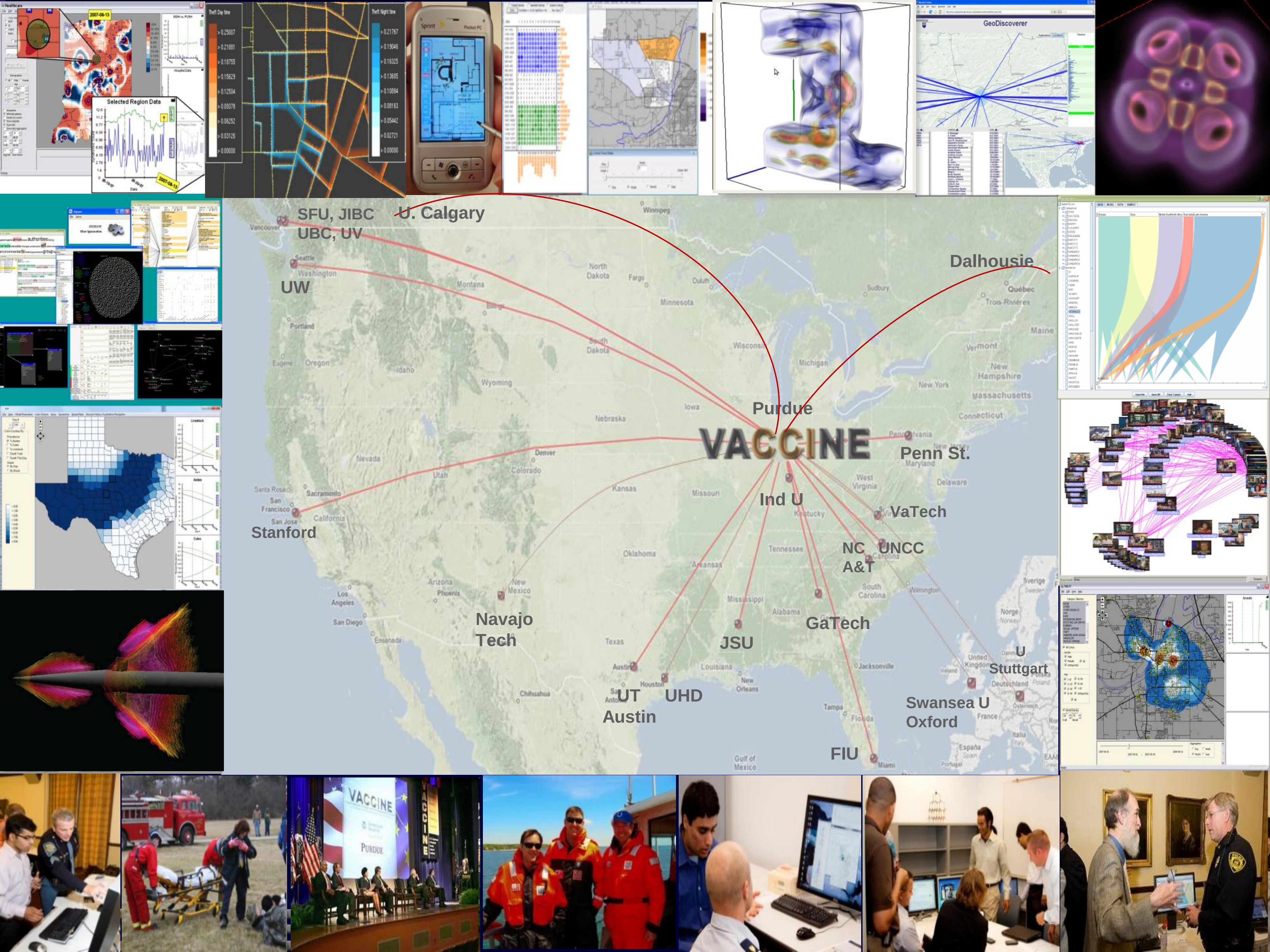
CVADA

Center for Visual Analytics and Data Analytics

A U.S. Department of Homeland Security Center of Excellence

Outline

- Background or “Why am I here?”
  - What is VACCINE?
- Public health visual analytics examples
- Extending syndromic surveillance to law enforcement:
  - Crime analytics
  - Mobile gang graffiti app



SFU, JIBC
UBC, UV
U. Calgary
UW
Dalhousie
Purdue
VACCINE
Penn St.
Ind U
VaTech
NC A&T
UNCC
Stanford
Navajo Tech
JSU
GaTech
U Stuttgart
UT Austin
UHD
Swansea U
Oxford
FIU
GeoDiscoverer
Selected Region Data

Who We Are:
International Team of Experts
- 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

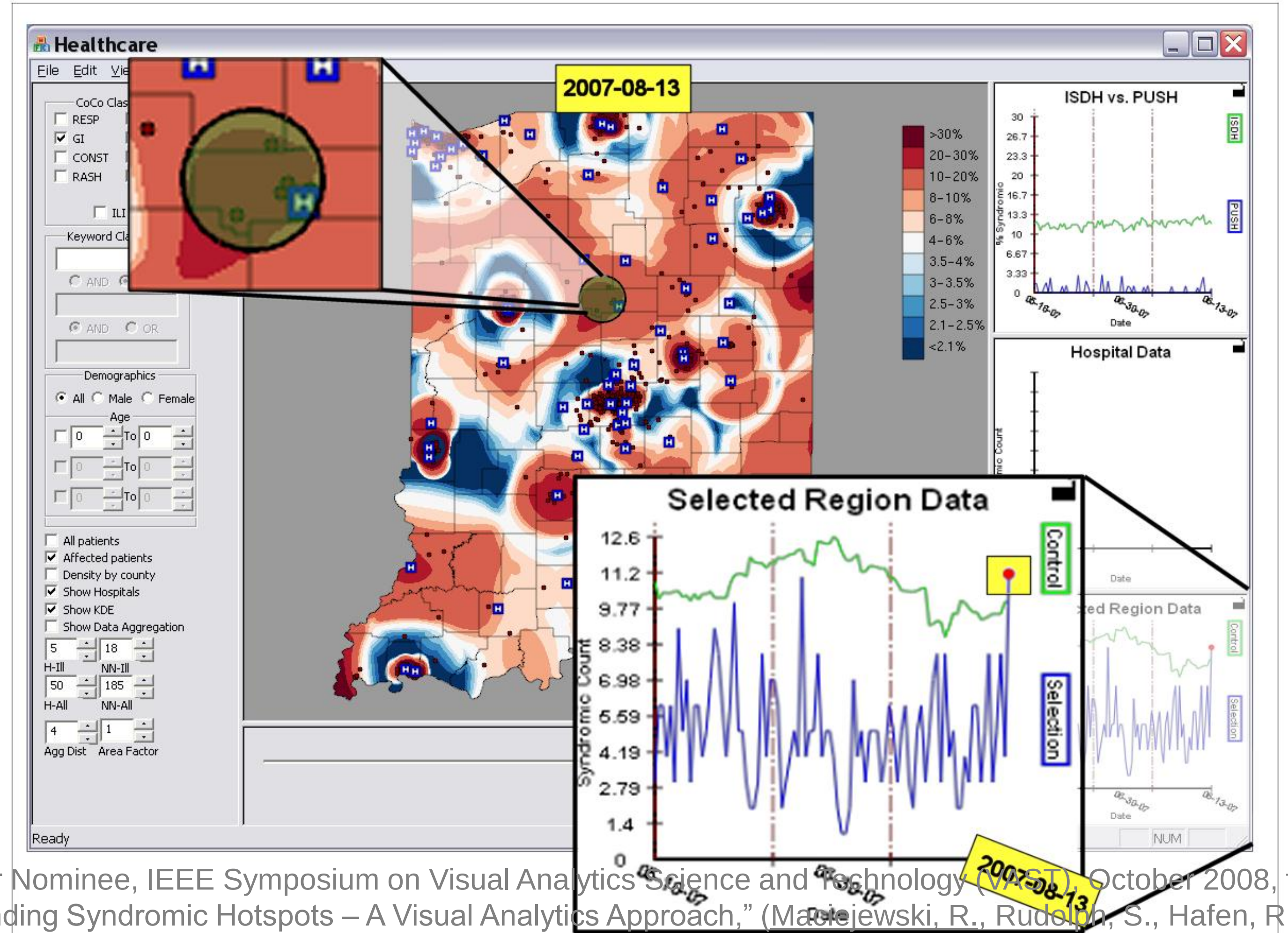
VACCINE Benefit

Enable users to be more effective through innovative interactive visualization, analysis, and decision making tools

- Provide the right information, in the right format within the right time to solve the problem
- Turn data deluge into a pool of relevant, actionable knowledge
- Enable user to be more effective from planning to detection to response to recovery
- Enable effective communication of information

Approach: Partner-driven solutions and research

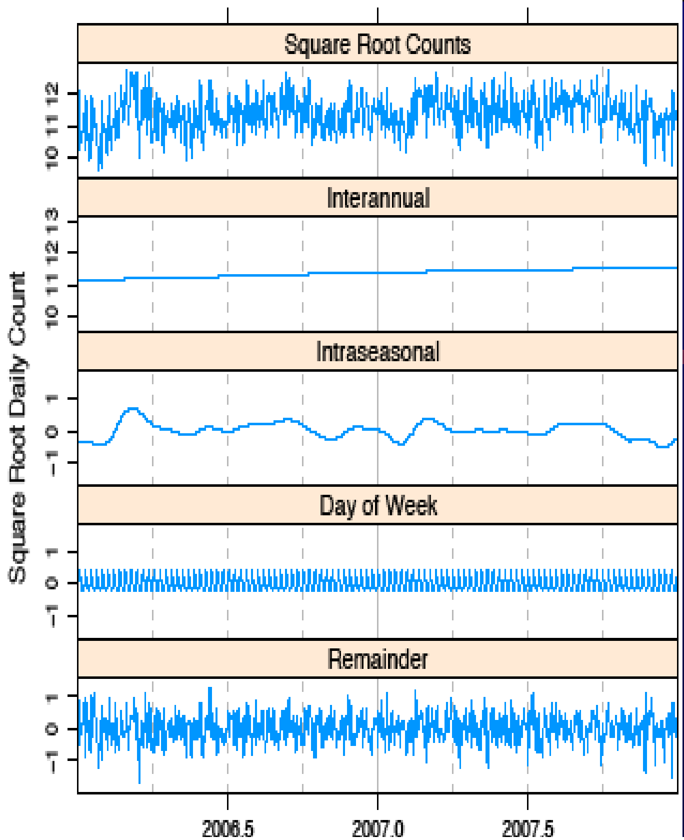
Visual Analytics for Syndromic Surveillance: Hypothesis Generation and Exploration

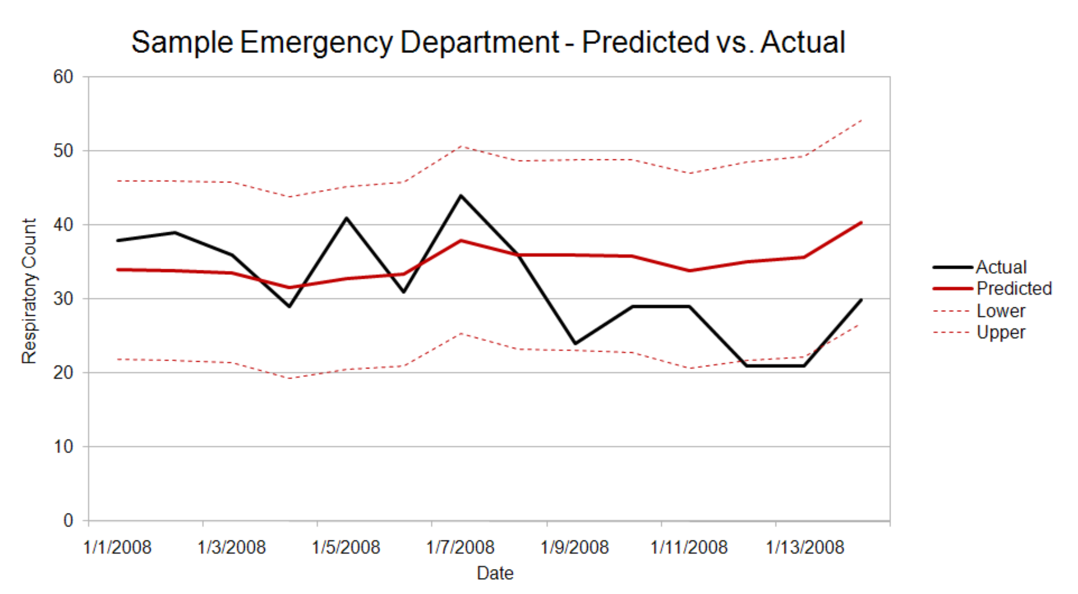


Best Paper Nominee, IEEE Symposium on Visual Analytics Science and Technology (VAST), October 2008, for "Understanding Syndromic Hotspots – A Visual Analytics Approach," (Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W., Grannis, S., Wade, M., Ebert, D.).

Predictive Visual Analytics

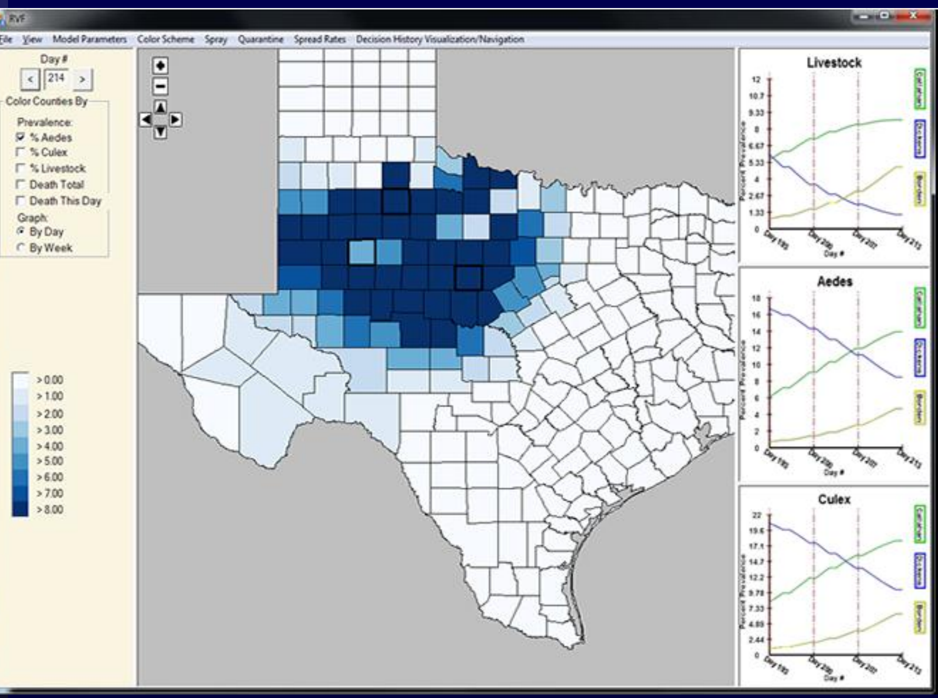
- Time Series Modeling:
  - Seasonal-Trend Decomposition Based on Loess
    - Time series ~ sum of multiple trend components
    - For each data signal, components are extracted
    - Can then analyze correlation between components
- **Predictive Visual Analytics using STL**

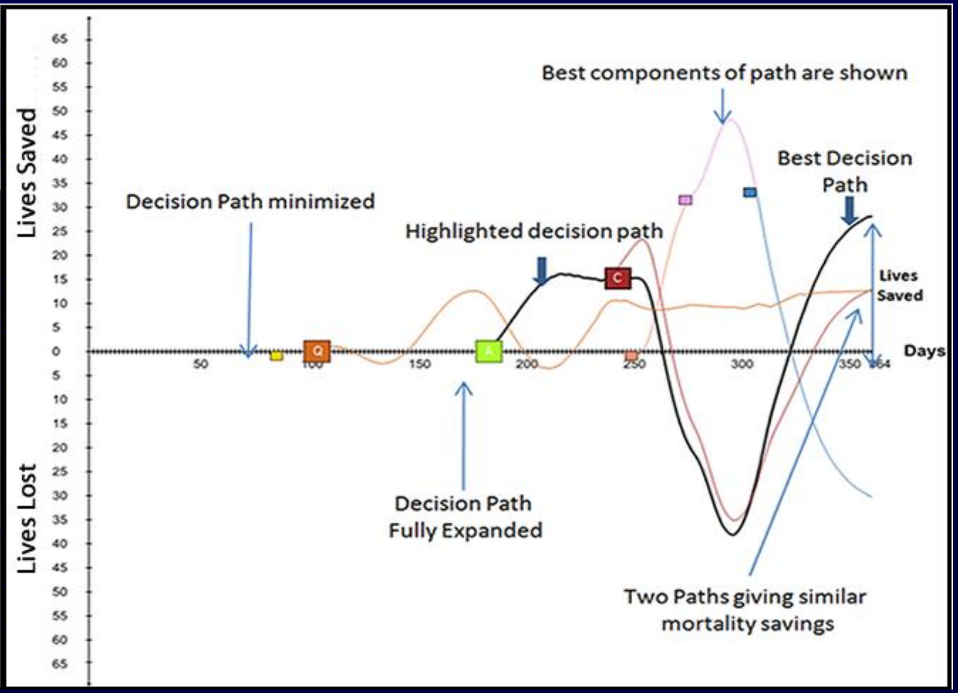




Hafen, R., Anderson, D., Cleveland, W., Maciejewski, R., Ebe
Visualizing, and Monitoring Disease Counts," *BMC Medical In*

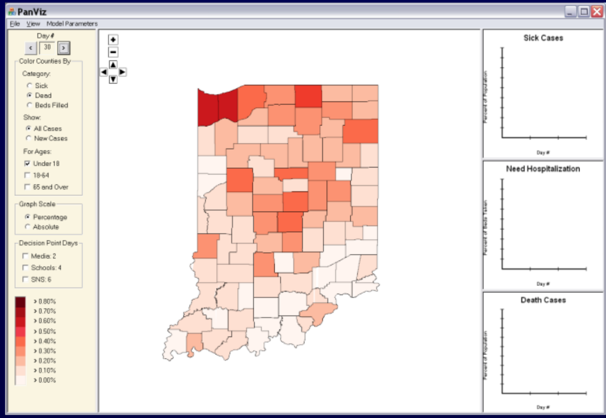
Advanced Decision Support Tools: Rift Valley Fever



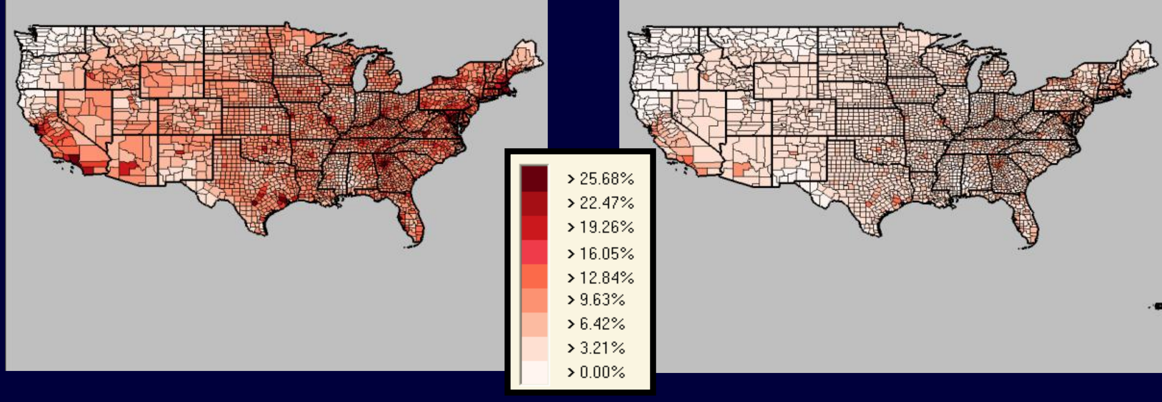


Modeling a Pandemic

- Pandemic Influenza Planning Tool
- Models user specified:
  - Pandemic influenza characteristics
  - County population, demographics, hospital beds
- Decision measures
  - Strategic National Stockpile deployment
  - School Closures
  - Media Alerts



Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.

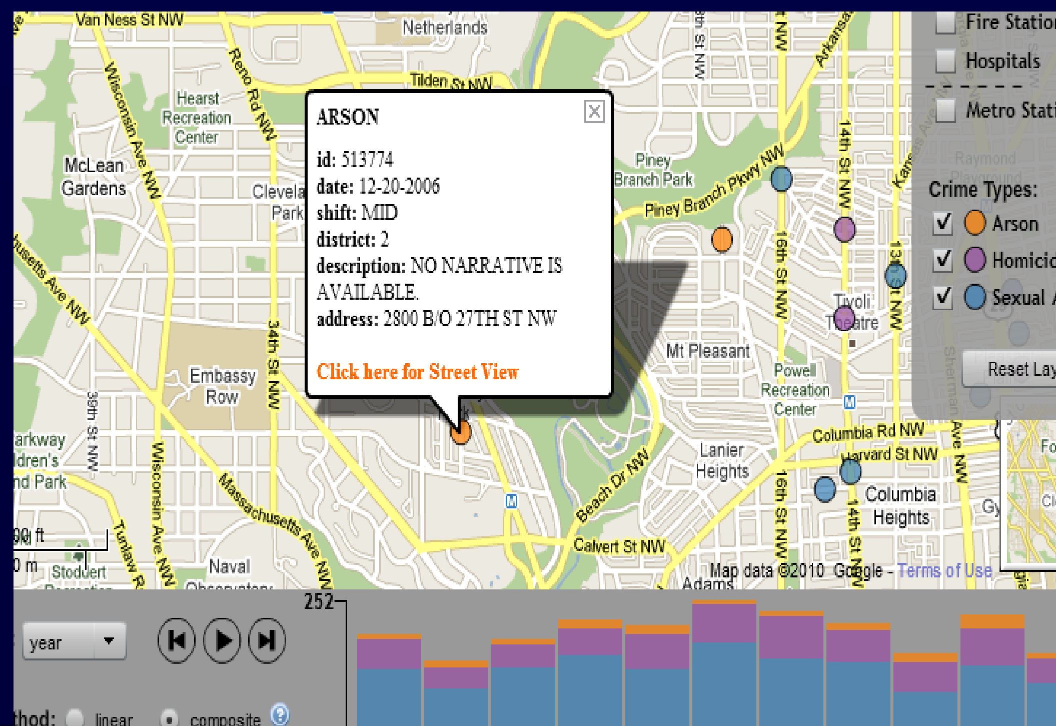


Pandemic spread on day 37 with no decision measures implemented

Pandemic spread on day 37 with all decision measures implemented

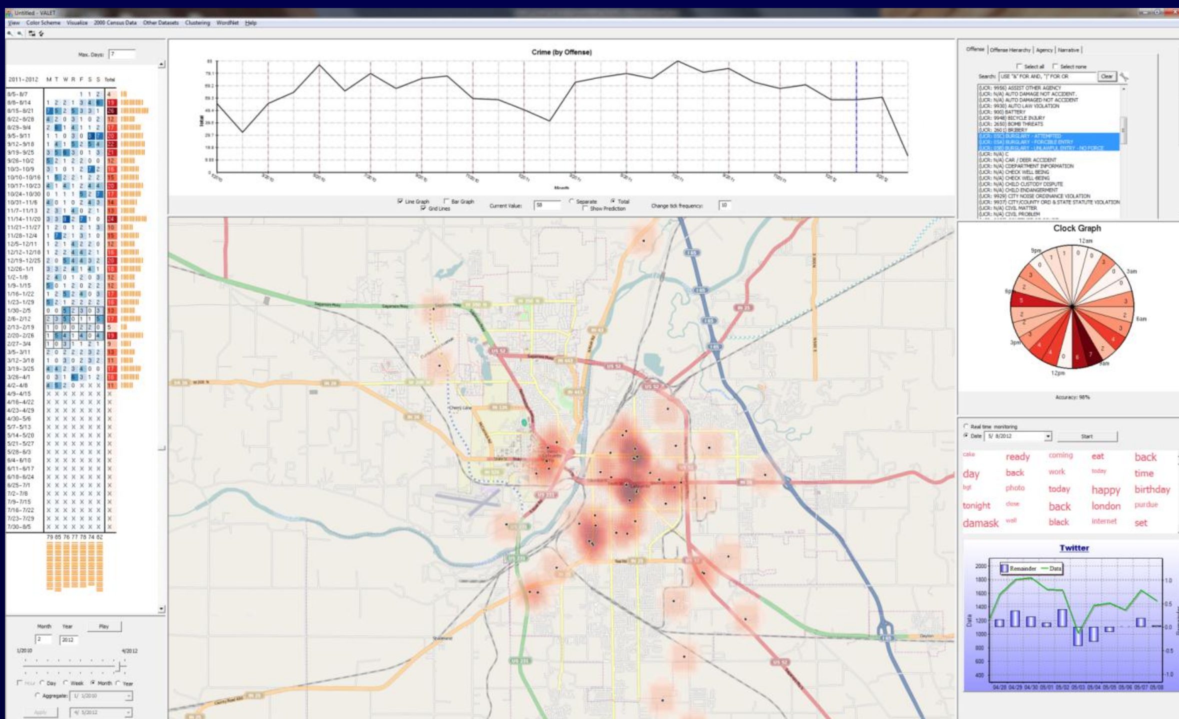
Visual Analytics Uses for Law Enforcement

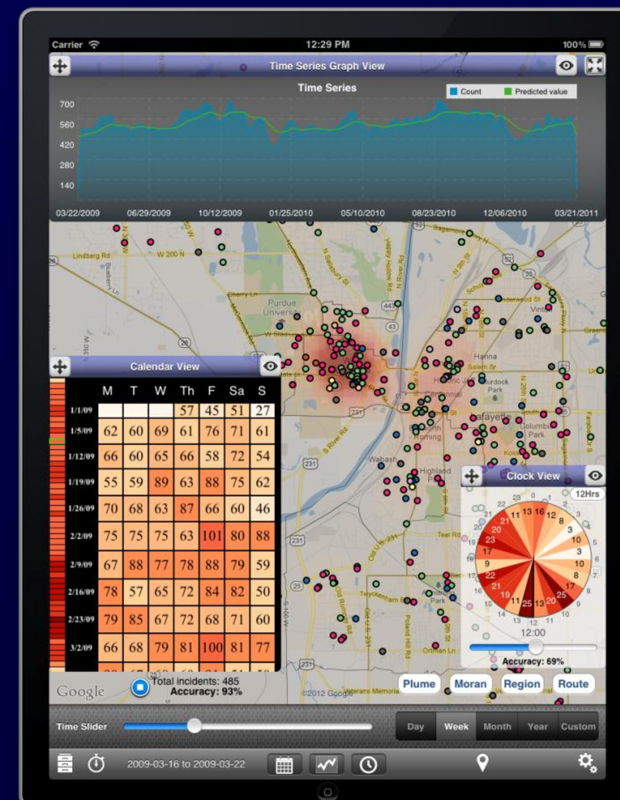
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis
- Purpose:
  - Planning for resiliency
  - Long-term analysis
  - Predictive analytics
  - Training
  - Detection
  - Investigation
  - Response
  - Recovery, remediation



VACCINE

Visual Analytics Law Enforcement Toolkit (i)VALET



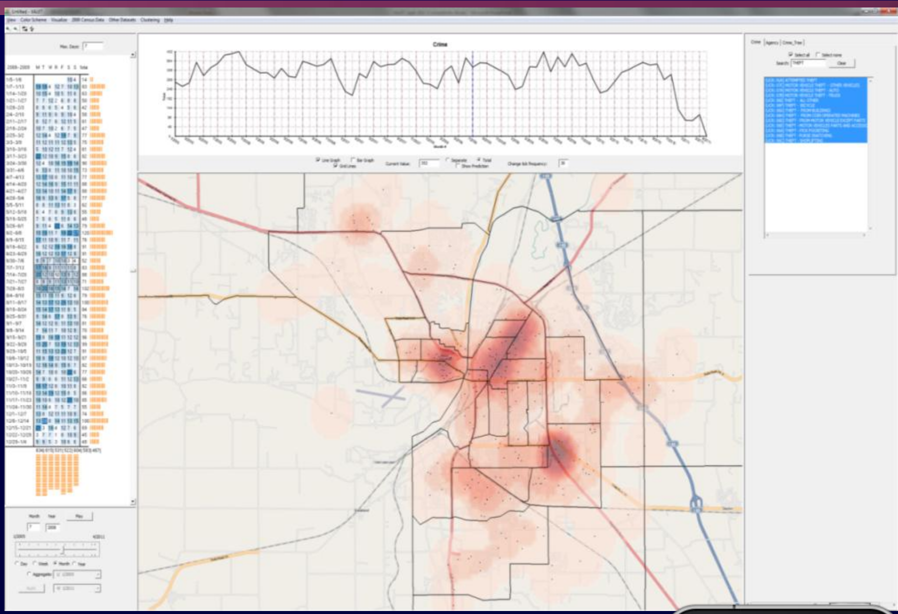


VACCINE

Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

Impacts:

- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (September 2012) for community-based policing
- Investigating correlation of bus routes and crime, street lights and crime
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning

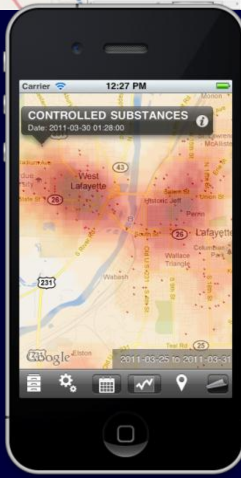


VALET delivered:

- Spring 2011: WL, Lafayette Police

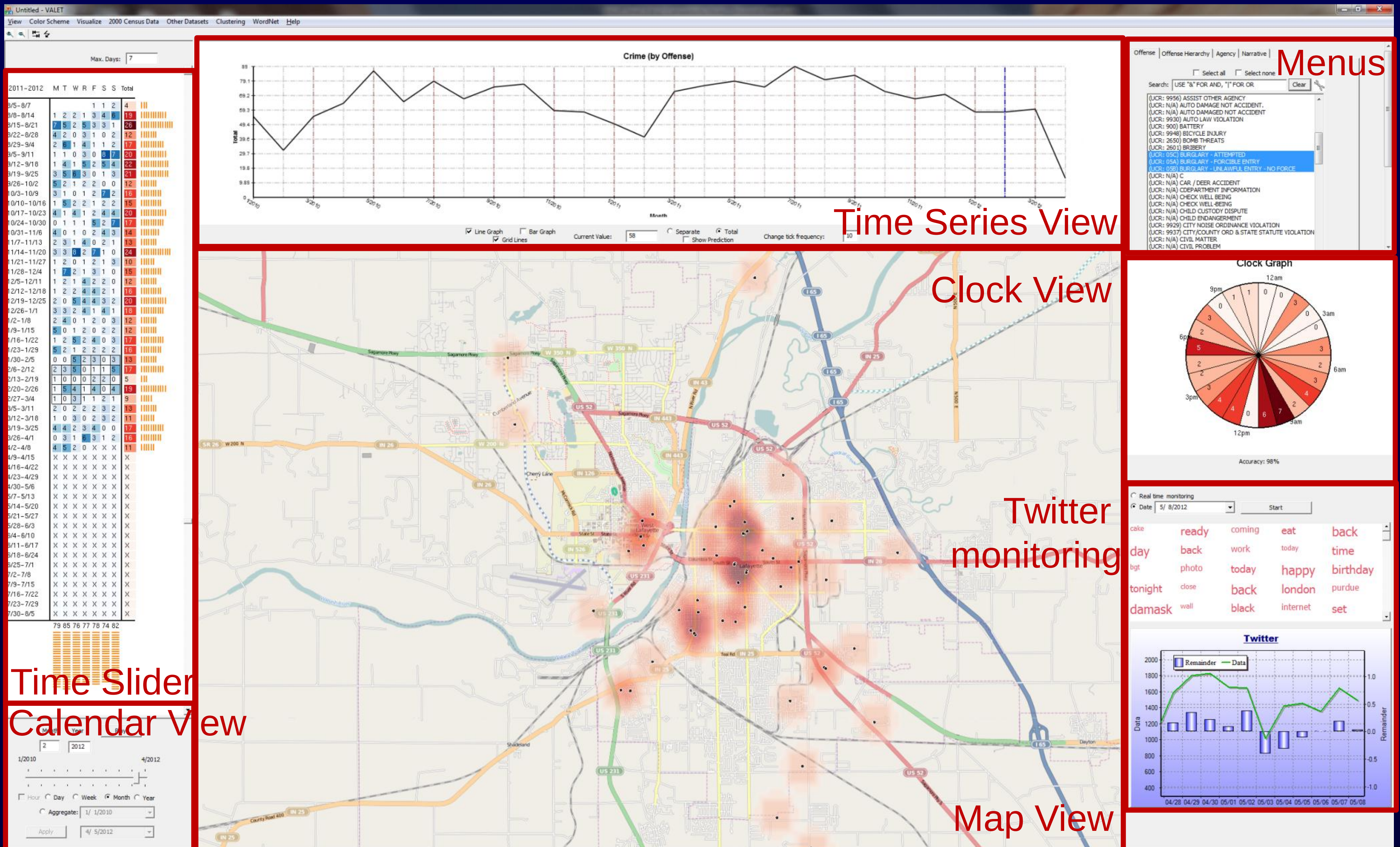
iVALET delivered:

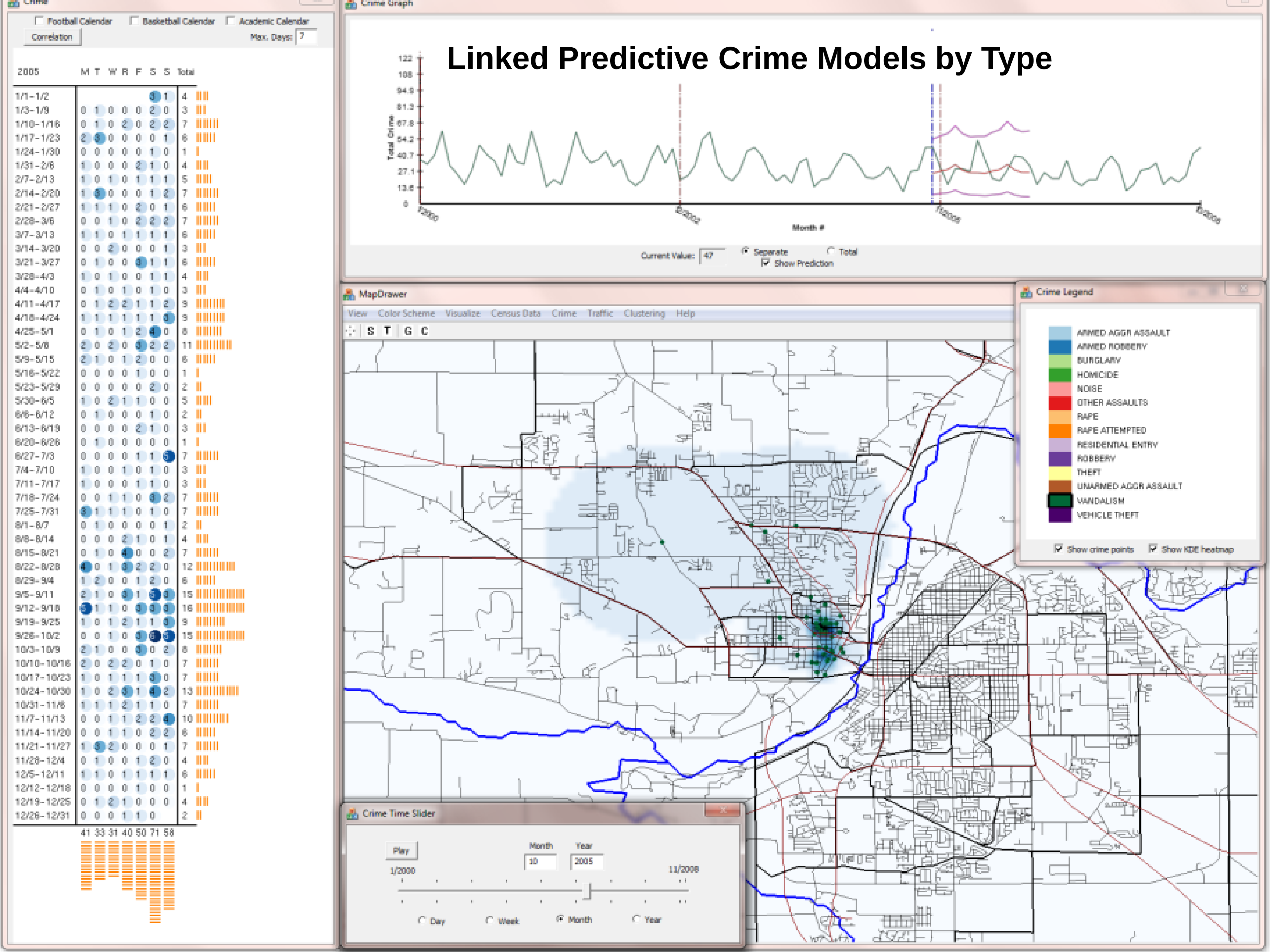
- October 2011: Purdue, WL Police



VACCINE

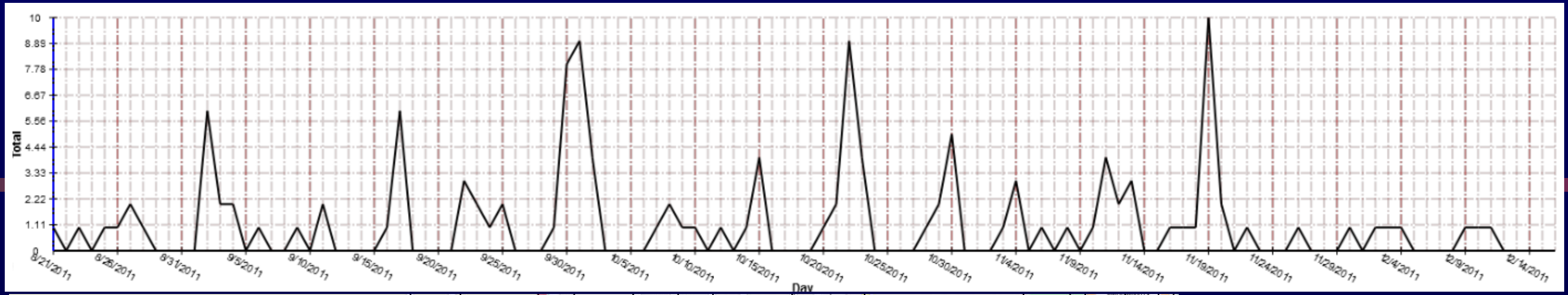
VALET Overview

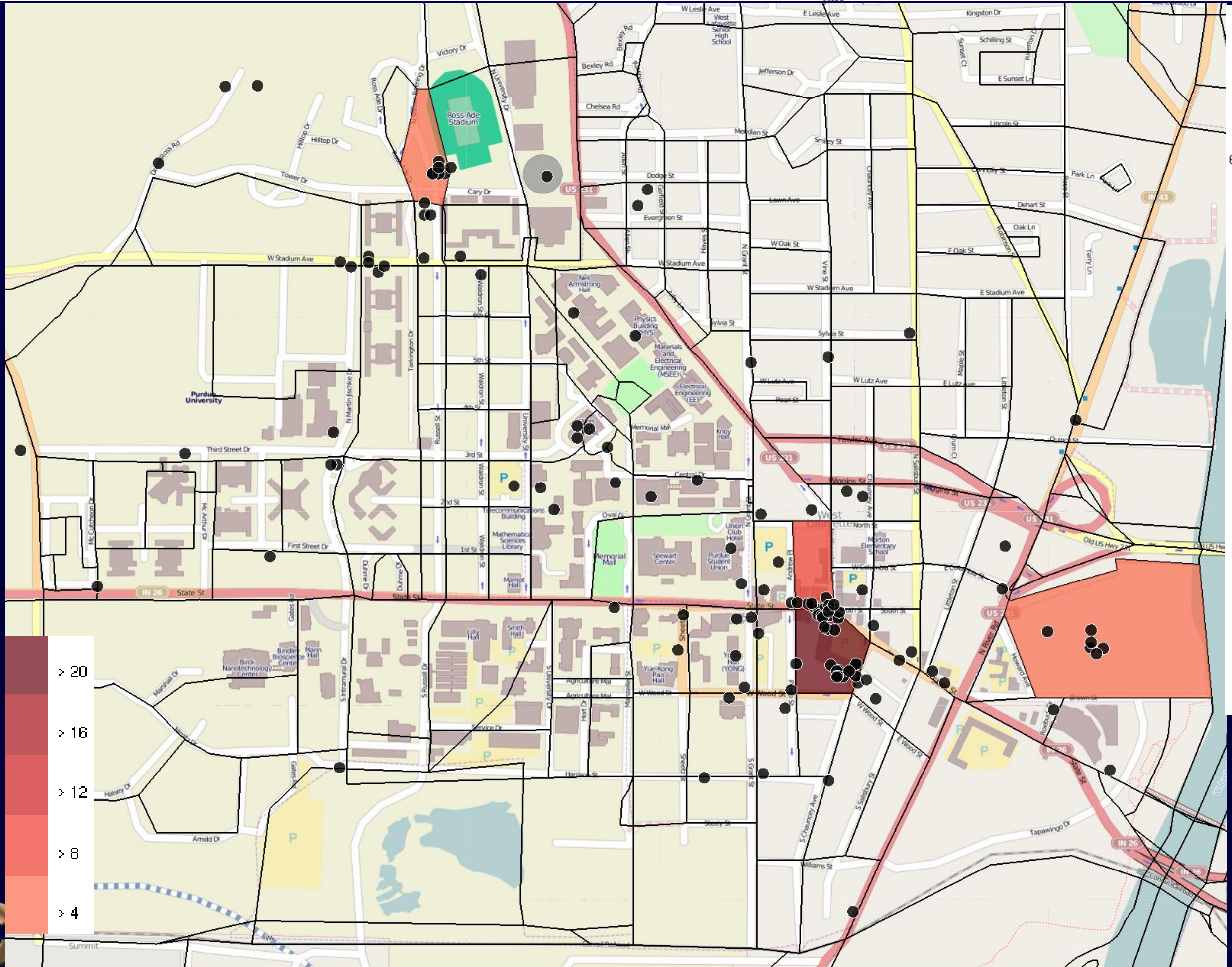


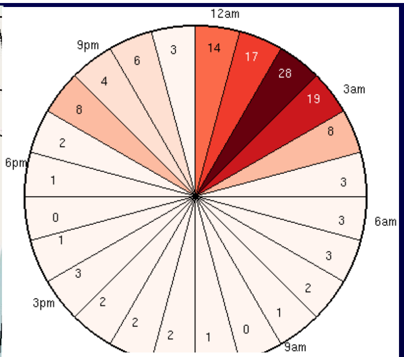


Linked Predictive Crime Models by Type
Crime
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
Max. Days: 7
2005
M T W R F S S Total
Crime Graph
Total Crime
Month #
Current Value: 47
Separate
Total
Show Prediction
MapDrawer
View Color Scheme Visualize Census Data Crime Traffic Clustering Help
S T G C
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Show crime points
Show KDE heatmap
Crime Time Slider
Play
Month 10
Year 2005
1/2000
11/2008
Day
Week
Month
Year

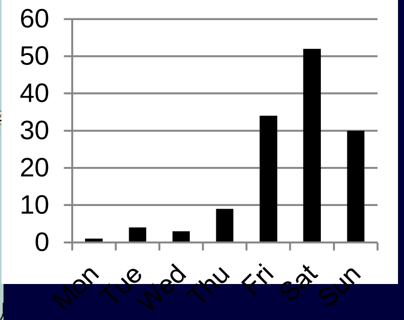
Example: Drunkenness / Public Intoxication





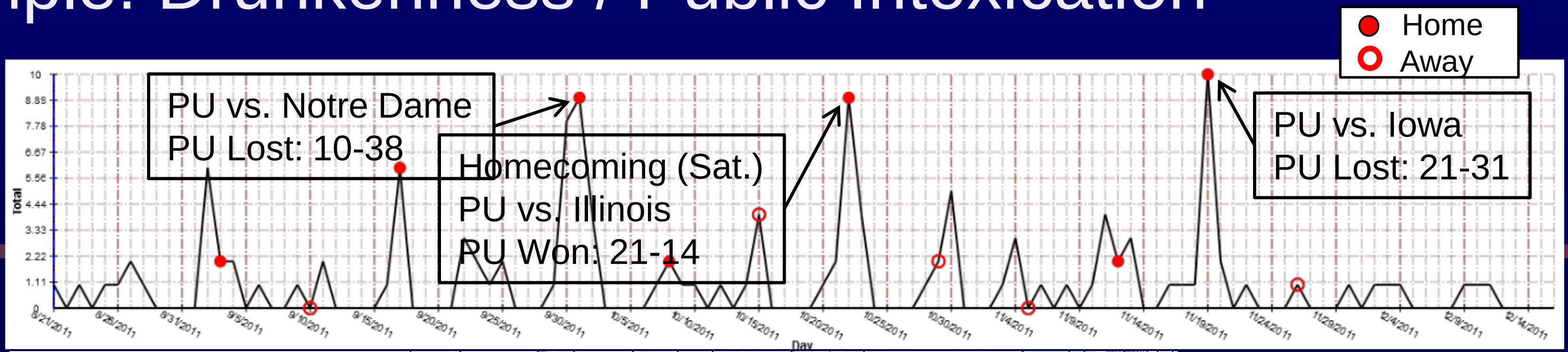


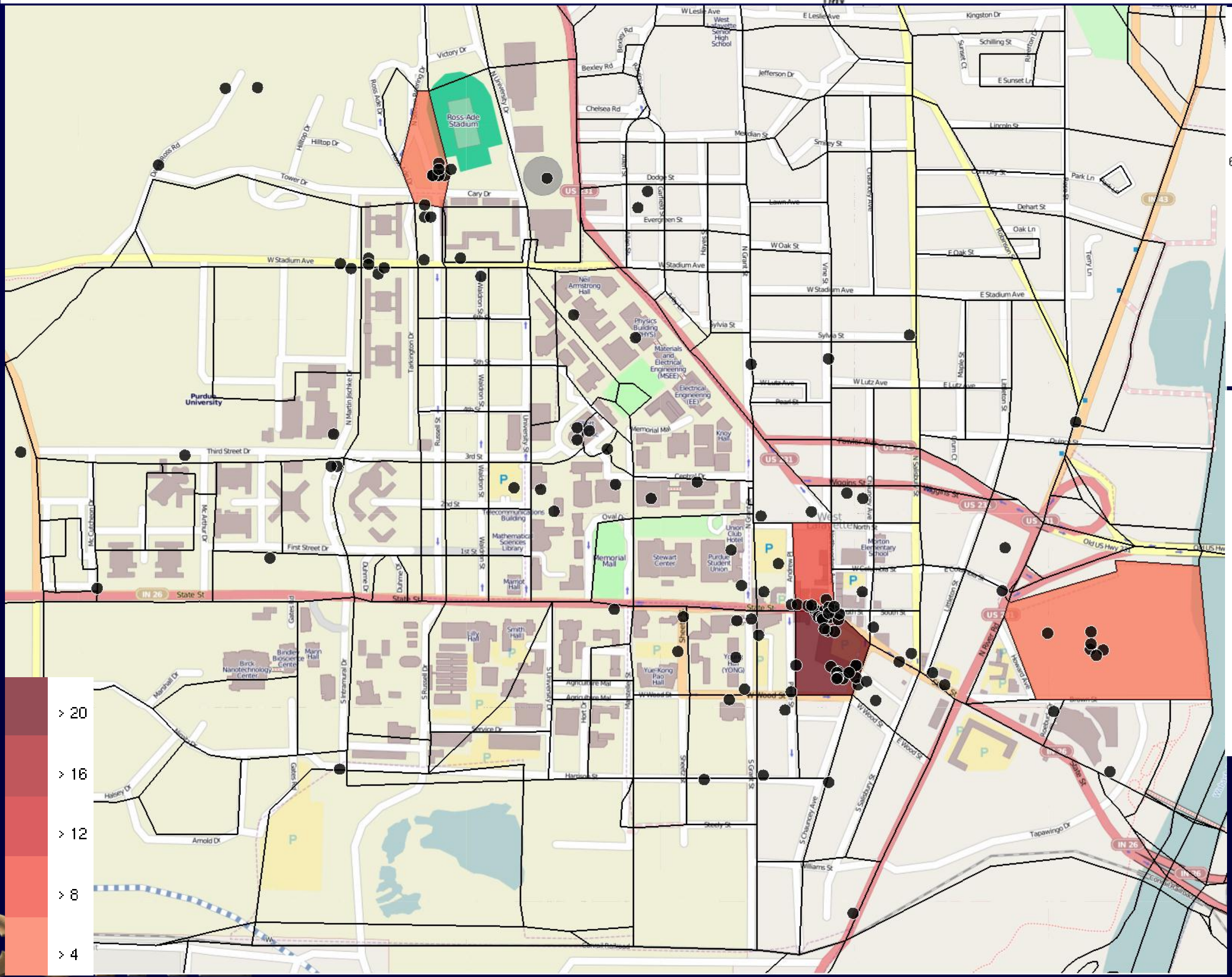
Day-of-the-Week

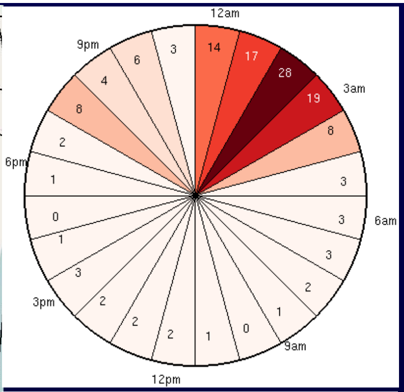


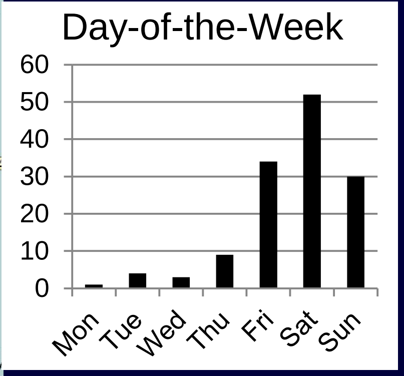
Example: Drunkenness / Public Intoxication

Football season



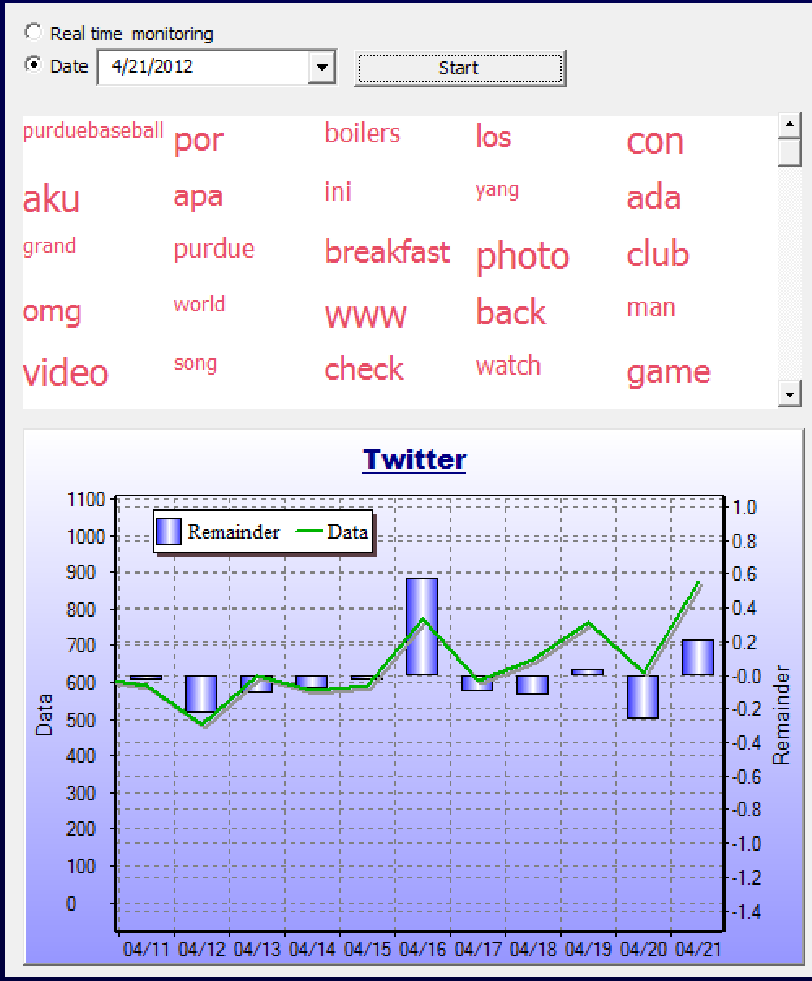




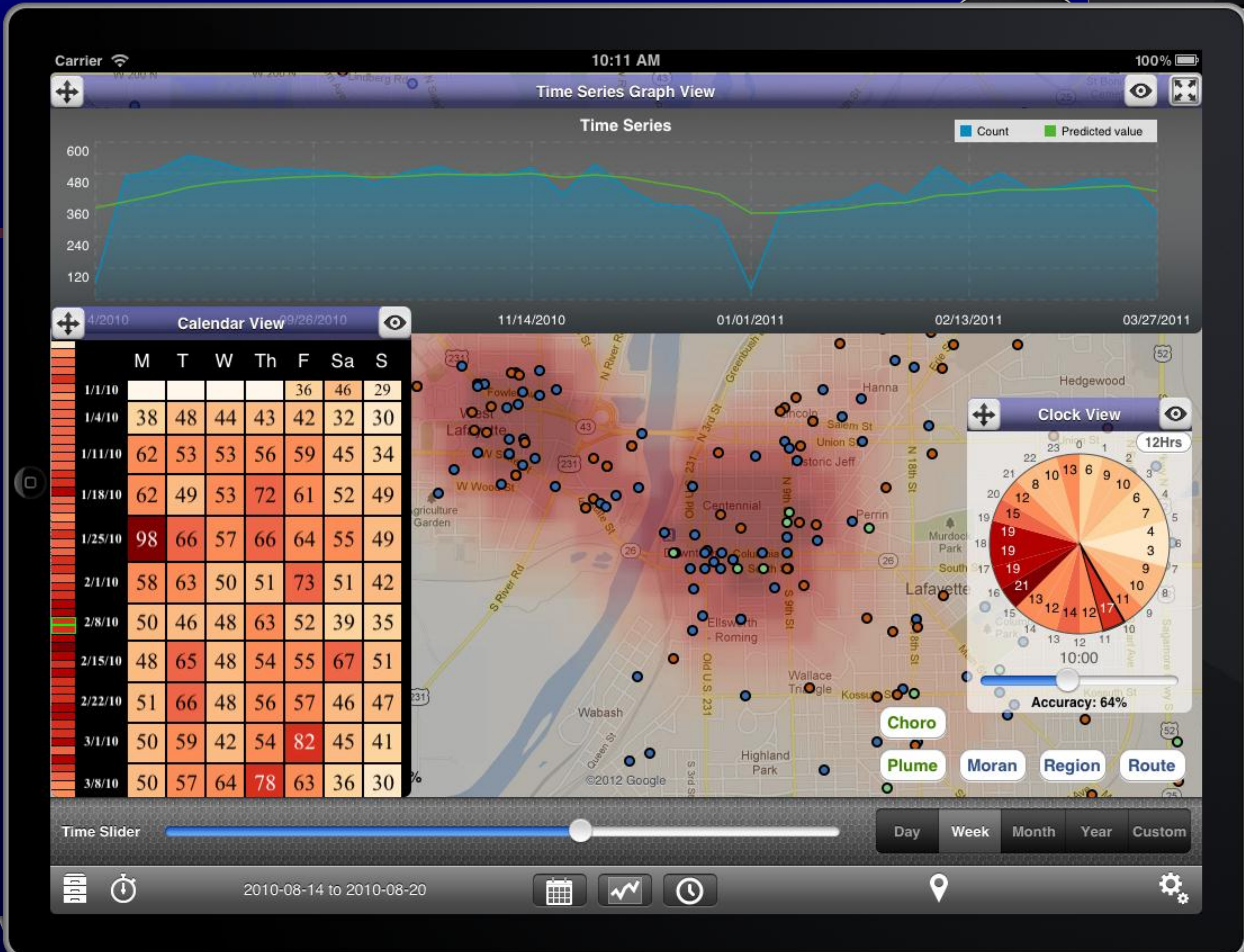


VALET Feature:
Real-time Twitter Monitoring and Integration

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud

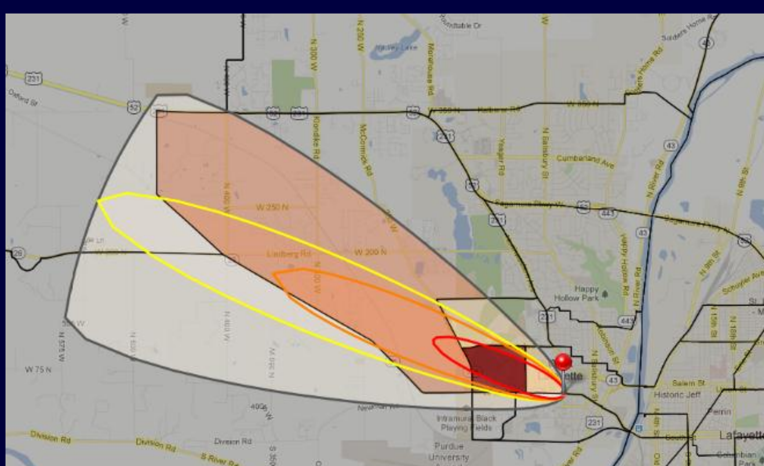


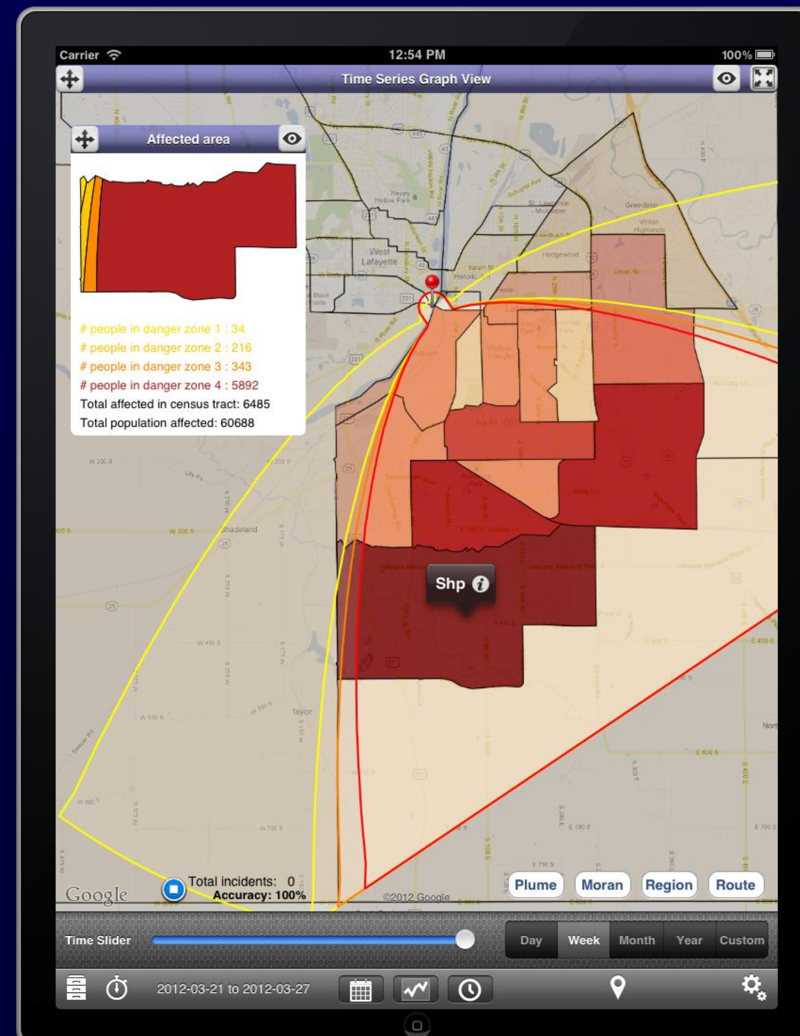
Grand Prix Weekend, Purdue University


Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
11/14/2010
01/01/2011
02/13/2011
03/27/2011
Calendar View
M T W Th F Sa S
1/1/10 36 46 29
1/4/10 38 48 44 43 42 32 30
1/11/10 62 53 53 56 59 45 34
1/18/10 62 49 53 72 61 52 49
1/25/10 98 66 57 66 64 55 49
2/1/10 58 63 50 51 73 51 42
2/8/10 50 46 48 63 52 39 35
2/15/10 48 65 48 54 55 67 51
2/22/10 51 66 48 56 57 46 47
3/1/10 50 59 42 54 82 45 41
3/8/10 50 57 64 78 63 36 30
Clock View
12Hrs
10:00
Accuracy: 64%
Choro
Plume
Moran
Region
Route
©2012 Google
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20

MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

- Chemical release plume modeling identifies census tracts with the highest number of expected people affected





VACCINE

Gang Graffiti Recognition and Analysis (GARI)
PI: Ed Delp

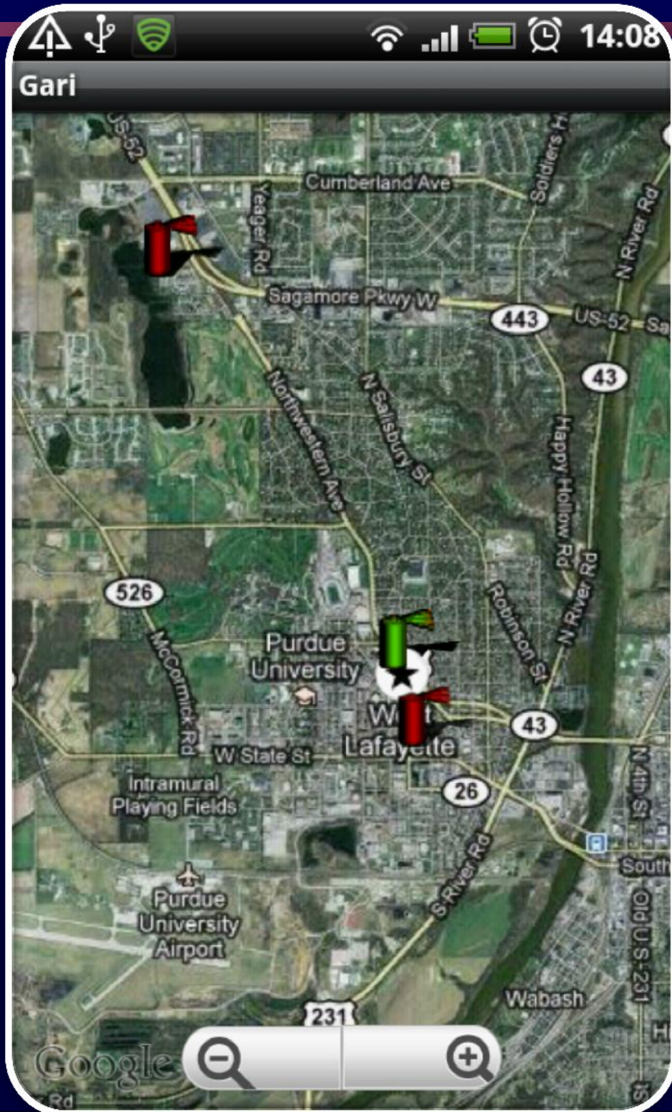
VACCINE

There's An App For That

- What is GARI?
  - A mobile device application that analyzes gang graffiti
- How does GARI work?
  - User takes an image of the graffiti
  - User receives an analysis
    - Geographic locations
    - Graffiti colors, shapes, meaning
  - Database search of similar tags

Why It Works

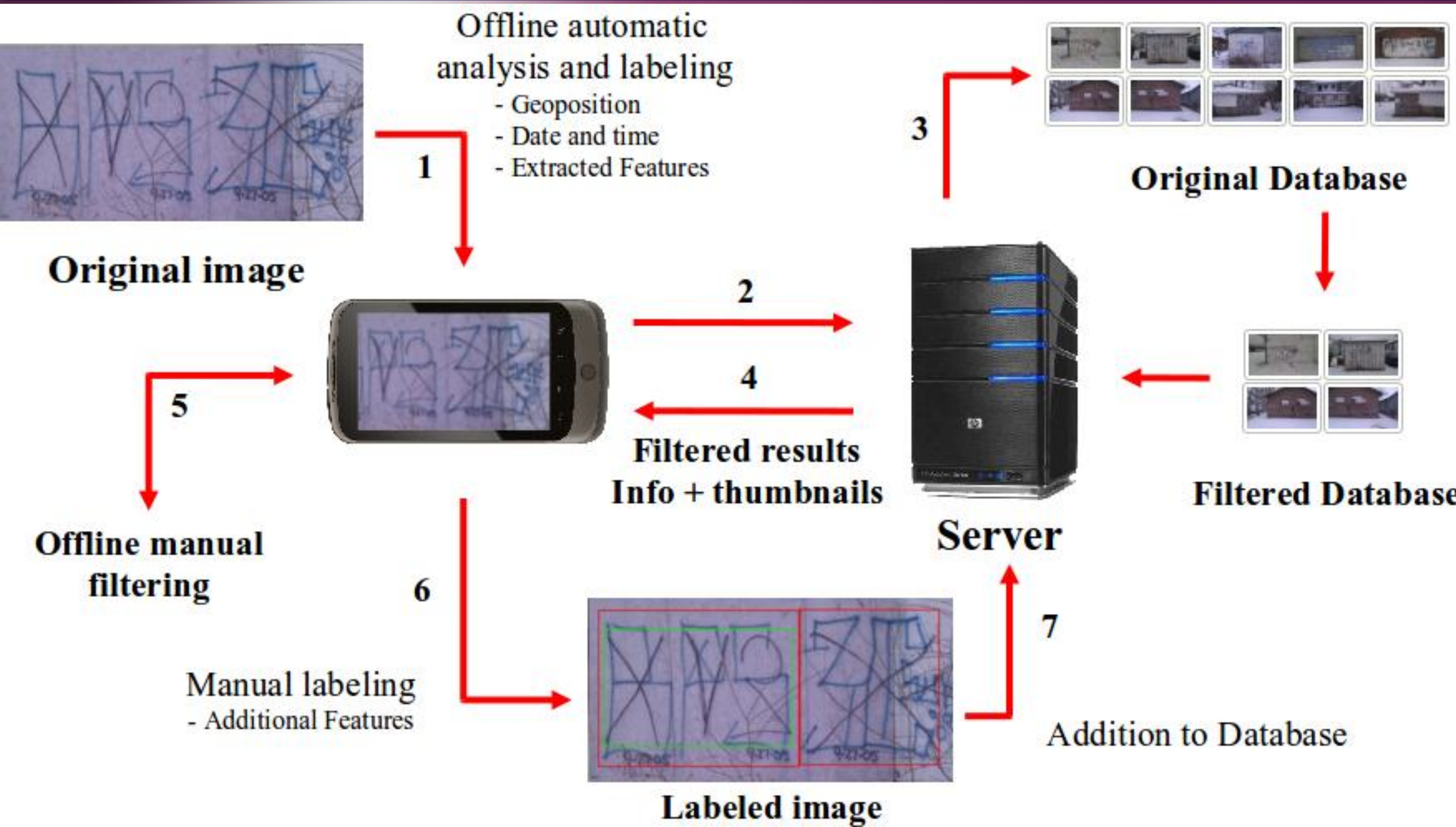
- Catalogs and analyzes graffiti images
  - More than 500 graffiti images acquired (plus 506 for research testing)
- Search of related image
- Search of local area
- Partially interprets graffiti images
  - Colors, symbols
  - Recorded gang analyst notes
- Future analysis
  - Image meaning
  - Which gang
- Helps law enforcement track and identify gang activity throughout a region



GARI On The Street

- Over 50 law enforcement officials are using or field testing GARI
- Delivered to:
  - IMPD gang detectives (Summer 2011)
  - IMPD at large, Indiana Fusion Gang Task Force (August 2011)
  - Gang detectives across Indiana (October 2011)
- Homeland Security applications

GARI System Overview

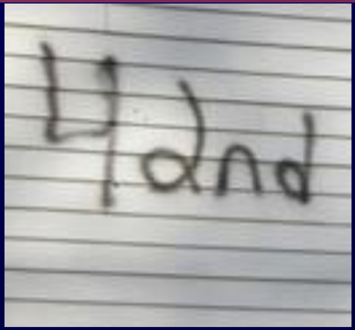


Project Status

- Deployed on Android Phone
- Main phone application is image database system and browsing (acquisition tool) with searching and some limited analysis capabilities
- Desktop backend
  - Browse by radius
  - Upload images
  - View and edit image details
  - Interact with map
  - Searching for similar images

Gang Graffiti Interpretation

SHAPE
Simple, Straightforward



NUMBERS
42nd street gang

SYMBOLS
6-point star, pitchforks

COLOR
Goon Squad: Red/Black

Gang Graffiti Interpretation

LETTERS

East Side Gang

TIME

Black: 18 ST (18th Street Gang)

Red: 13 SUR (Sureños 13)

POSITION/ALIGNMENT

Letters at star points

Numbers in the middle

Letters at the bottom

Pitchforks upright

Image Analysis – Color Recognition

MEXICANOS MALDITOS
SUREÑOS 13

18 STREET GANG

SUREÑOS 13

Image Analysis – Scene Analysis

GANG RIVALRY TRACKING

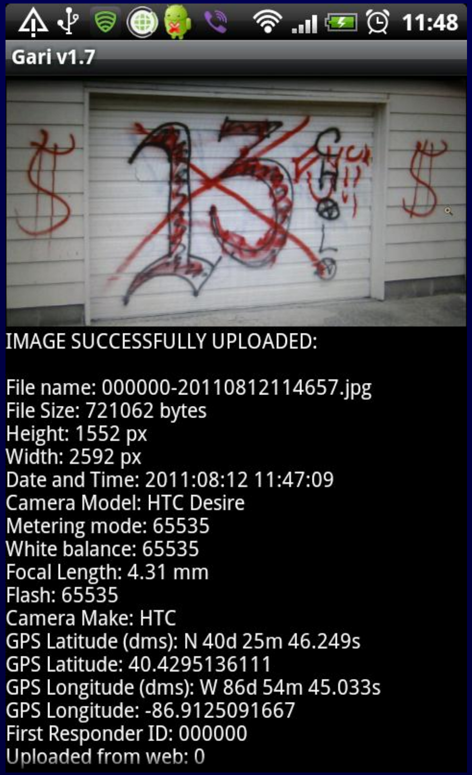
- Date: 08/19/2010
- Time: 3.25 PM
- Geo: 41.387917, 2.169919

- Date: 01/03/2011
- Time: 5.11 PM
- Geo: 41.387917, 2.169919

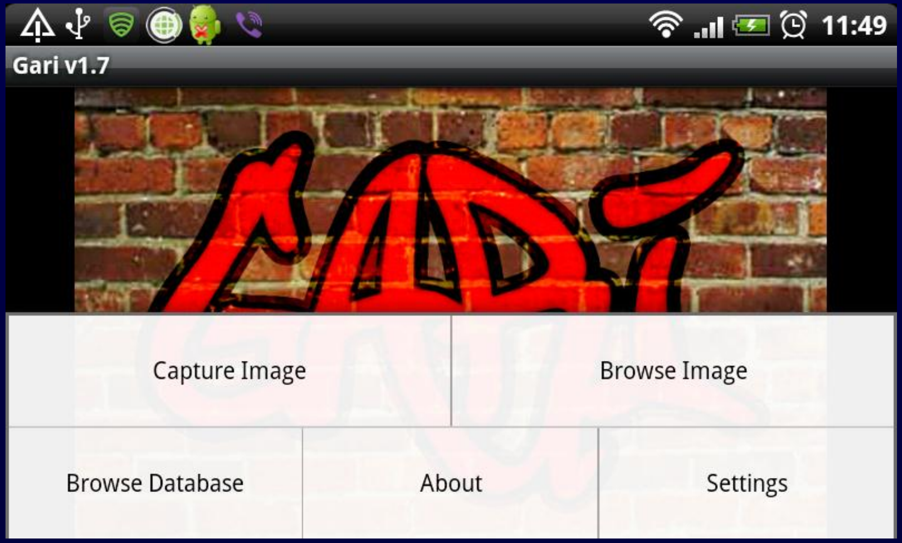
VACCINE

GARI App Features

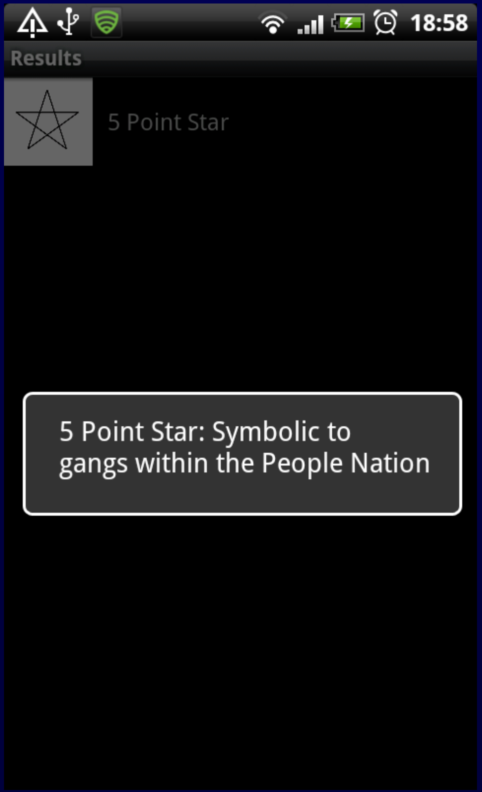
Gang Graffiti –User Interface



Send to server



Main menu



Analysis

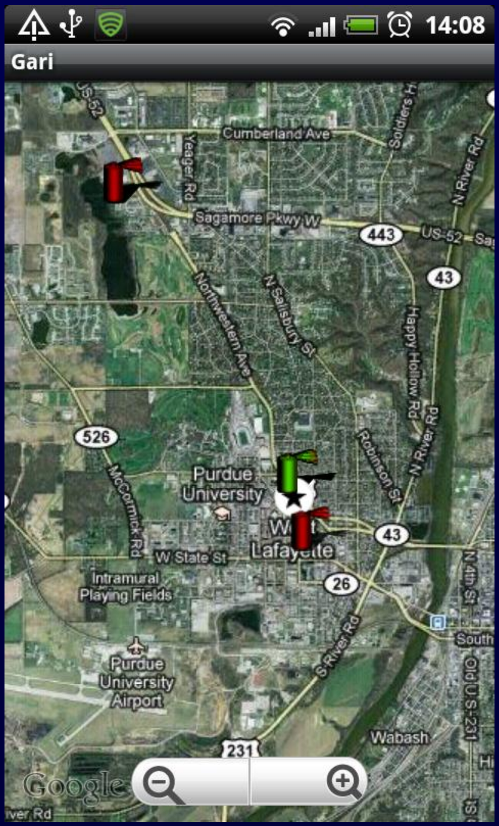
Gang Graffiti – Browse Database

- Browse by radius from current position
- Download images and information from server (image EXIF tags, gang related information)
  - Compare results
  - Track graffiti
- Browse map (Google Maps API)

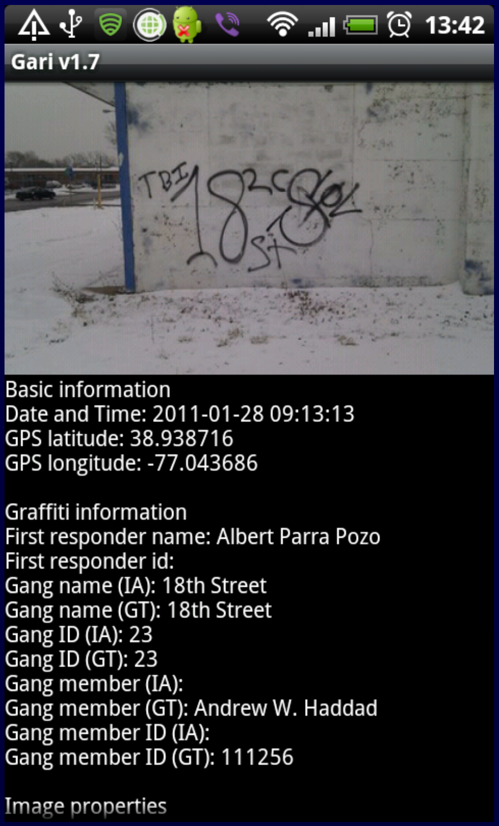
- Network connection required

Gang Graffiti – Browse Database

Browse general results



Show graffiti on map



Inspect specific graffiti

GARI Progress and Directions

- Detective Steve Shafer, IPD adding annotations to database
- Shape and tag analysis always improving
- Transitioning to Indiana State Police / IIFC Gang Task Force
- Creating linked, satellite database and tool for Navajo Nation (using Navajo Technical College)
- Investigating use by ICE, CBP, FBI
- Two-way linkage with VALET / iVALET

How to Get The App?

- Contact gari@ecn.purdue.edu
  - We assign user IDs and initial passwords
  - Database server

VACCINE

Table of Contents

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu

ebertd@purdue.edu

VACCINE

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

Table of Contents

Teaching Visual Analytics:

Leveraging Multidisciplinarity

David S. Ebert
Niklas Elmqvist
ebertd@purdue.edu

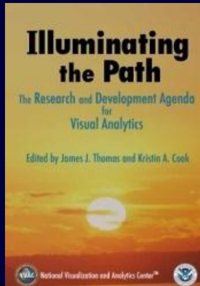
Definition

Visual Analytics[1] is the science of analytical reasoning facilitated by interactive visual interfaces

People use visual analytics tools and techniques to

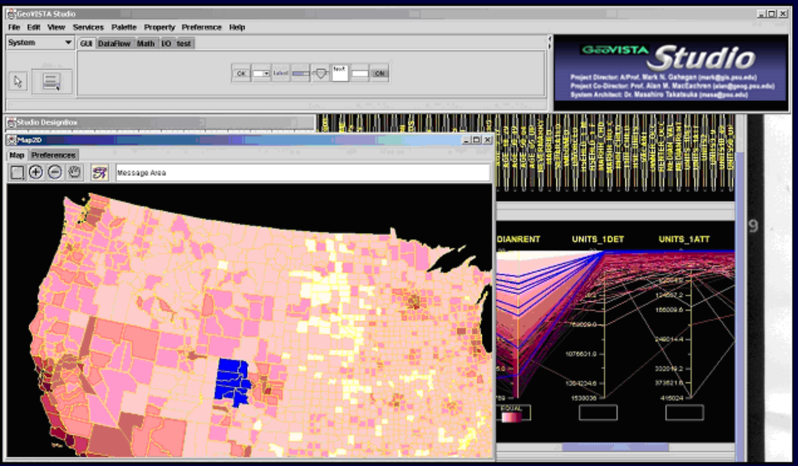
- Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
- Detect the expected and discover the unexpected
- Provide timely, defensible, understandable assessments
- Communicate assessment effectively for action

1. *Illuminating the Path: The R&D Agenda for Visual Analytics, Editors: Thomas and Cook*



Motivation

To solve today's and tomorrow's problems requires exploring, analyzing, and reasoning with massive, multisource, multiscale, heterogeneous, streaming data



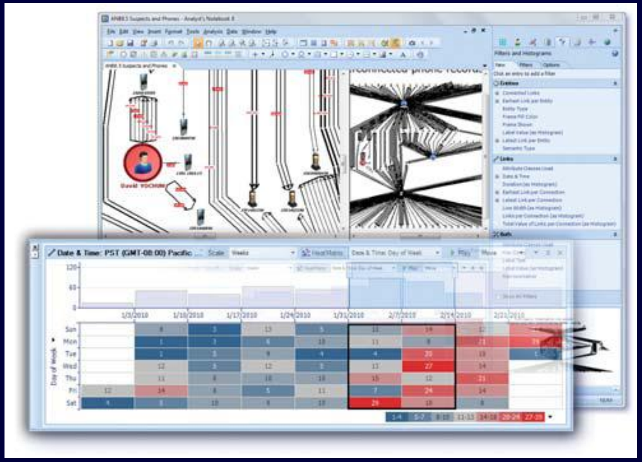
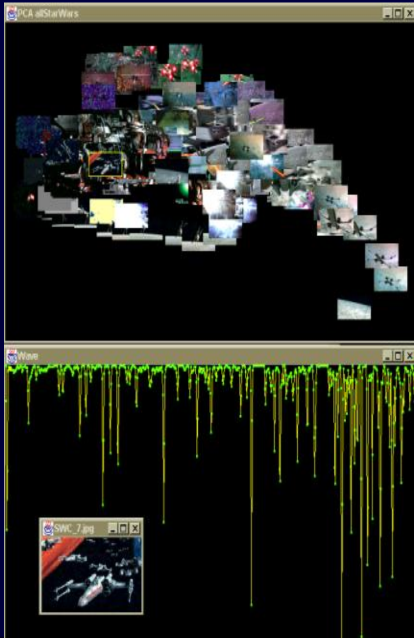


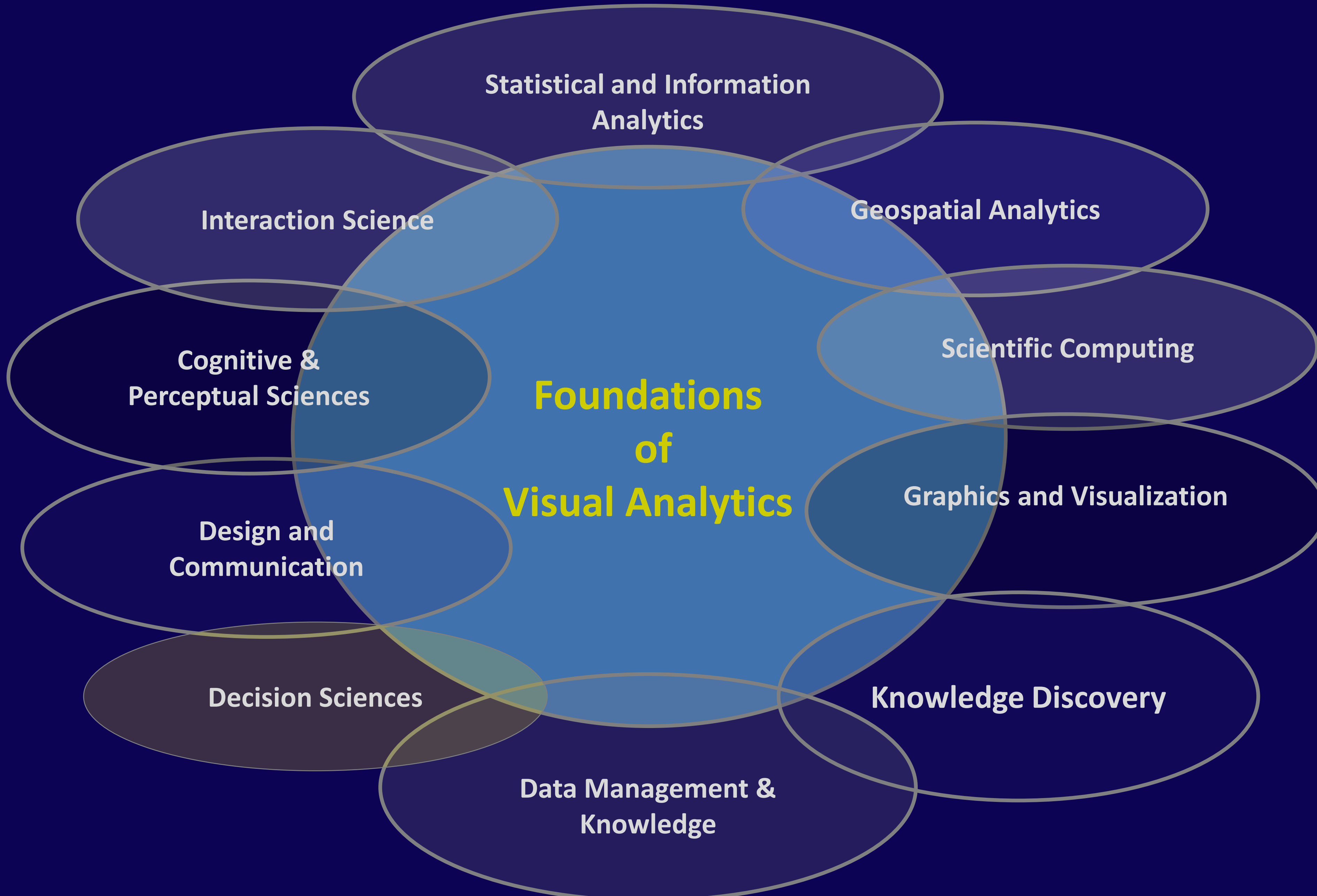
Image of Analyst's Notebook



VACCINE

VA Goals – Solution Driven

- Enable effective decision making through interactive visual analytic environments
- Enable effective communication of information
- Provide quantitative, reliable, reproducible evidence
- Enable user to be more effective from planning to detection to response to recovery
- Enable proactive and predictive visual analytics
- Enable effective situational awareness (perception, comprehension, projection)


Statistical and Information Analytics
Interaction Science
Geospatial Analytics
Cognitive & Perceptual Sciences
Scientific Computing
Foundations of Visual Analytics
Graphics and Visualization
Design and Communication
Decision Sciences
Knowledge Discovery
Data Management & Knowledge

Course Outline

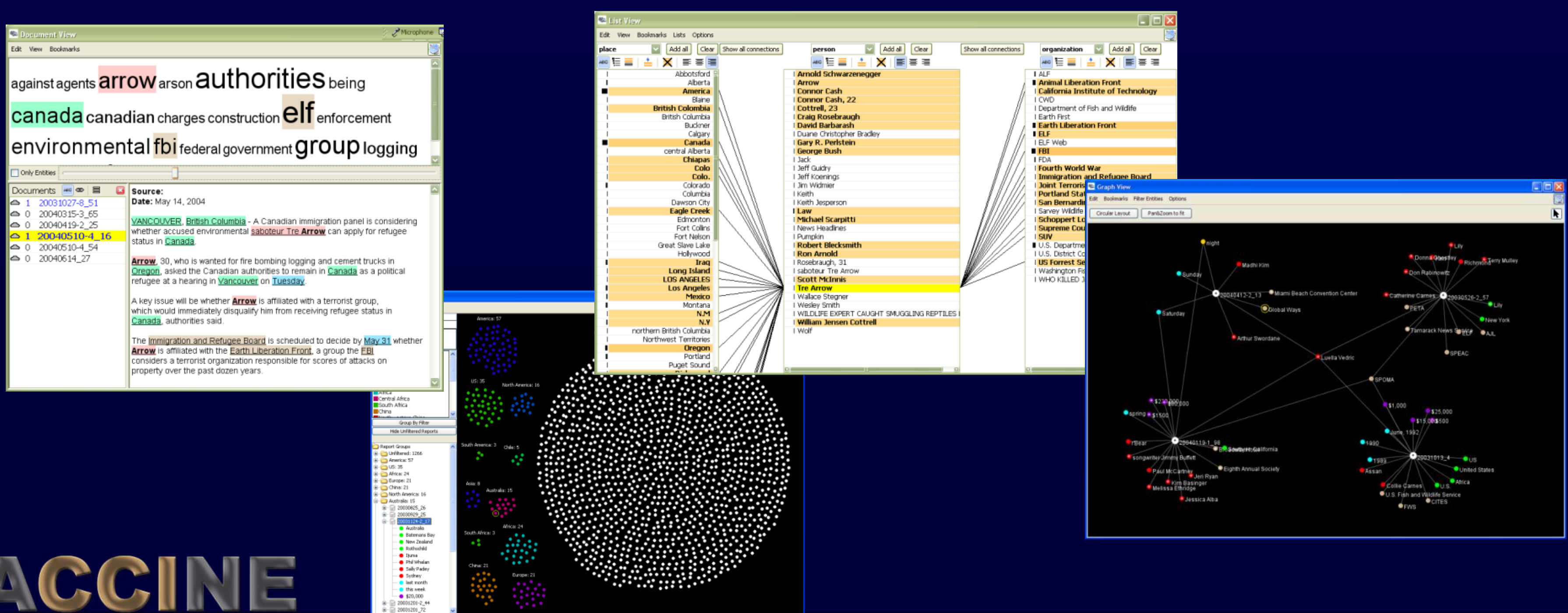
Week	Topic	Contents
1	Introduction	Analytical exercise
2–3	Analytical reasoning	The analysis process, critical thinking, sensemaking, and situation awareness
4	Perception	Human perception, preattentiveness, color, shape, and texture
5	Cognition	Cognitive theory
6–7	Data management	Representations, transformations, and statistics (temporal and spatial)
9	Visual representations	Visualization techniques
11	Interaction	Interaction techniques
12	Communication	Production, presentation, and dissemination
13	Collaboration	Collaborative VA
14	Evaluation	Evaluating VA
15	Advanced topics	Conducting VA research, novel computing platforms, and mobile VA

Jigsaw:
Visual Analytics for Investigative Analysis and Exploration of Document Collections

"Putting the pieces together"

Goal: Assist investigators with understanding, sense-making, and analysis of large, unstructured and structured document collections

Approach: Provide multiple visual perspectives on the documents and entities within them, highlighting connections between entities



VACCINE

Course Outline

Week	Topic	Contents
1	Introduction	Analytical exercise
2–3	Analytical reasoning	The analysis process, critical thinking, sensemaking, and situation awareness
4	Perception	Human perception, preattentiveness, color, shape, and texture
5	Cognition	Cognitive theory
6–7	Data management	Representations, transformations, and statistics (temporal and spatial)
9	Visual representations	Visualization techniques
11	Interaction	Interaction techniques
12	Communication	Production, presentation, and dissemination
13	Collaboration	Collaborative VA
14	Evaluation	Evaluating VA
15	Advanced topics	Conducting VA research, novel computing platforms, and mobile VA

VA Course Challenges

1. Broad field to cover in 15 weeks
2. Students with diverse backgrounds
3. Teacher expertise in normally only portion of course topics

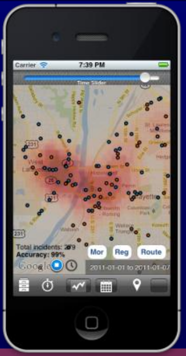
To Program or Not to Program….

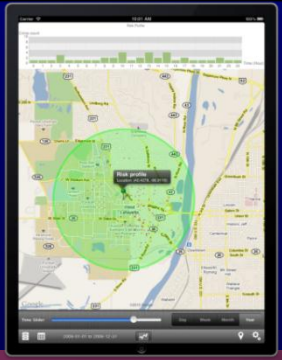
- We design course without programming expertise required
- Challenges:
  - Evaluating students results fairly across projects of different types(design studies, application use, new techniques)
  - Overcoming faculty biases

Current Results

- Graduate class offered twice with 10-15 students per semester
- Variety of student backgrounds
- At least 5 projects have lead to conference submissions

iVALET

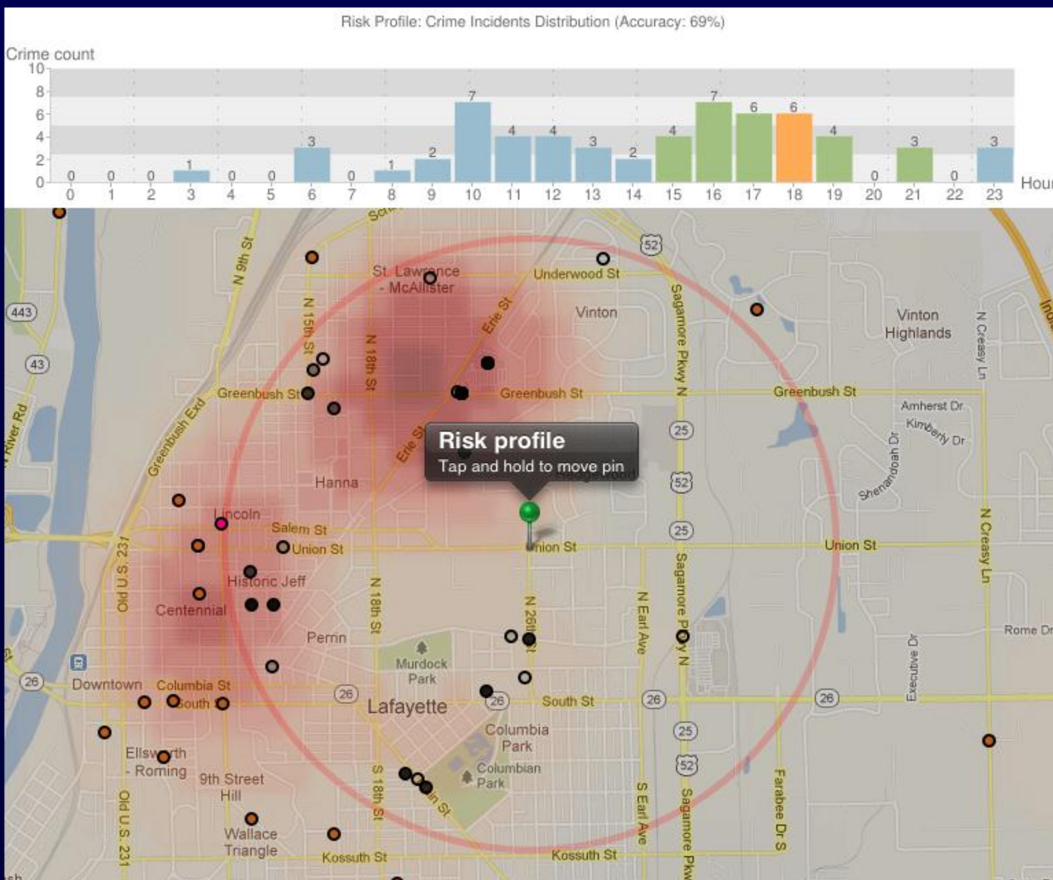




- Explore criminal, traffic and civil data on-the-go
- Risk assessment
- Use current spatial + temporal context into analysis

VACCINE

iVALET



- Linked views to explore multivariate spatiotemporal dataset
- Analytical tools to help explore data

VACCINE

VA Course Challenges

1. Broad field to cover in 15 weeks
2. Students with diverse backgrounds
3. Teacher expertise in normally only portion of course topics

Table of Contents

For Further Information

ebertd@purdue.edu

elm@purdue.edu

Table of Contents

Transitioning GARI to Law Enforcement and ICE

Dr. David S. Ebert
Director, Visualization Sciences Co-Lead (VACCINE)
CVADA
Purdue University

Homeland Security

Science and Technology

Homeland Security
Science and Technology

GARI - Gang Graffiti Automatic Recognition and Interpretation

Enabling tracking, analysis, and interpretation of gang graffiti / activity

- **Problem:** How to use gang graffiti for improved policing and safety
  - Complex language difficult to interpret
  - Encodes threats, information, gang activity
  - Most officers aren’t trained to interpret it
- **Solution:** A mobile device application that analyzes gang graffiti
  - User takes a picture and receives some analysis
  - Finds similar images, locations, and officer notes on the images
- **Key Users:** Indianapolis PD, Indiana Gang Intelligence Network (INGANG), Indiana State Police, Indiana Intelligence Fusion Center, Illinois State Police (in process), Cook County, IL Sherriff’s department (in process)

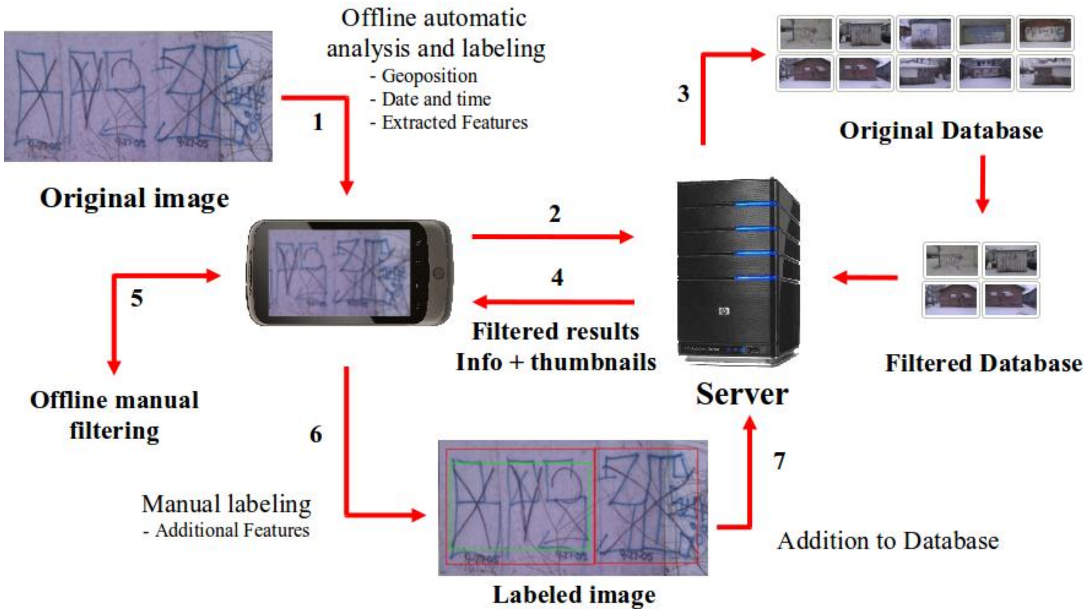
GARI

- **When and Where:**
  - Summer 2012 – Indianapolis - Pilot – deployed to gang detectives IPD, lead detective Steve Shafer
  - "GARI can really help the street officers because it's available 24/7--it's quick and instantaneous" - Detective Steve Schafer, Criminal Gang Unit, Indianapolis Police
  - Share gang expertise / force multiplier
  - Fall 2012 – Indiana – deployed to multiple agencies across Indiana – hundreds of images uploaded and analyzed
    - Over 800 images, 72 users

Indianapolis gang threats

GARI

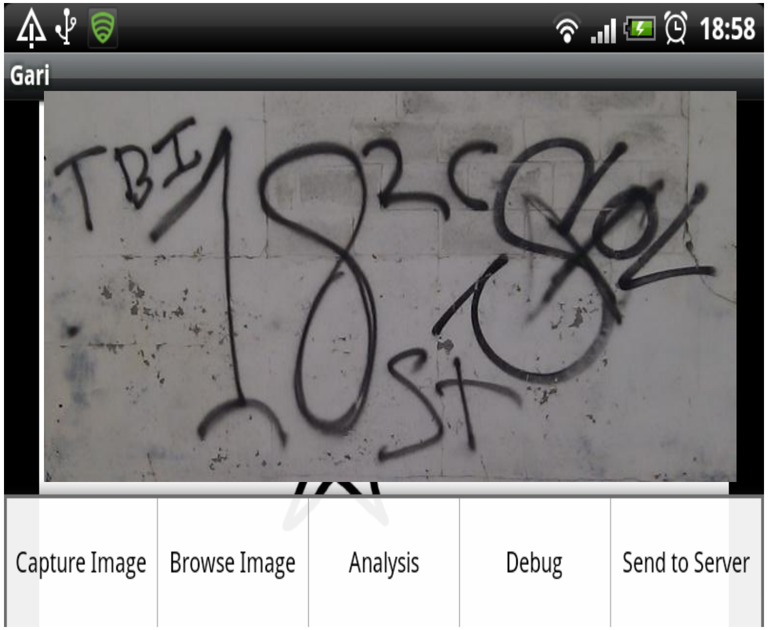
- **Winter 2013 – INGANG deployment** – all Indiana law enforcement agencies – server will be transferred to Indiana State Police (April 2013)
  - Adding prison tattoos to linked database for Indiana Department of Corrections
- **Spring 2013 - Illinois State Police**
  - “This does more than the RFP we are about to issue for software to capture images,” “Sign us up now,” Aaron Kustermann, ISP Chief of Intelligence
- **Spring 2013 – Cook County, IL**
  - Will upload thousands of images
  - Adding capability to record graffiti removal, background type (e.g., garage door, picket fence, brick)



Homeland Security
Science and Technology

GARI

- **How:** Transitioning to Sustainable Deployment
  - **Key:** Close interaction with officers from the start – Purdue police, Tippecanoe police departments, Indiana police
  - Tight Integration with Indiana State Police INGANG network (Capt. Scott Beamon)
  - Server transitioning to ISP
  - ISP will vet law enforcement officers nationwide who want GARI and agree to their INGANG terms of use
  - Midwest regional deployment is underway with Ohio State Highway Patrol, Ohio DHS, and Illinois partners

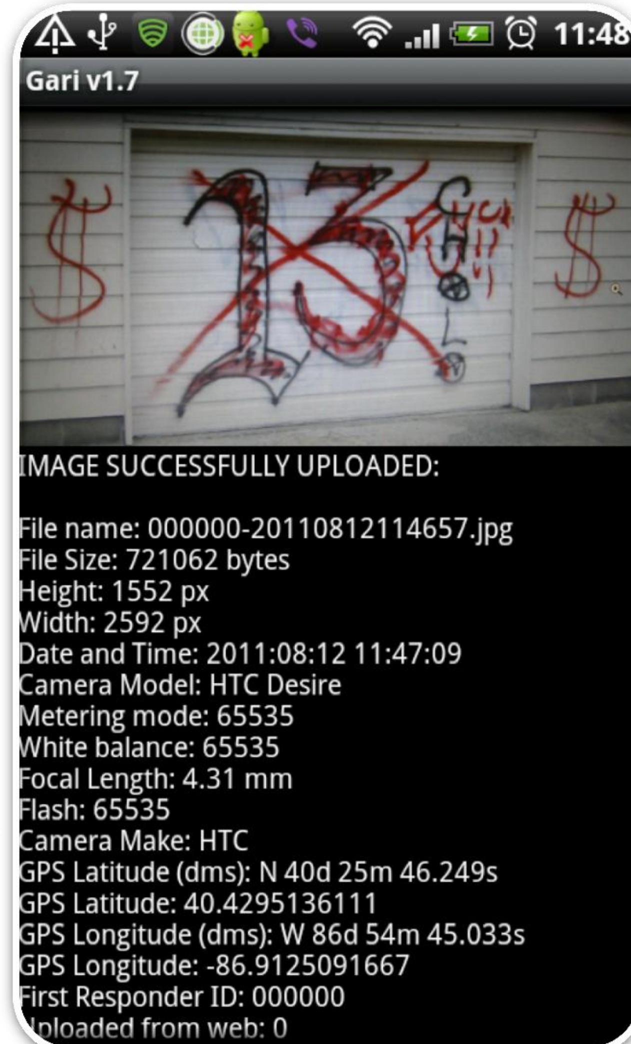


GARI app interface, showing multiple tags

GARI

- **Transition Cycle to Date**
  - Question from Chief John Cox at monthly public safety consortium meeting – “Can the Rosetta Phone be changed to help with this picture I just took of graffiti in the parking garage?”
  - Prototype with a handful of officers and 1 gang graffiti expert
  - Statewide deployment
  - Midwest regional deployment

Purdue Rosetta Phone



GARI

Table of Contents

Homeland Security

Science and Technology

Table of Contents

VACCINE

A DHS S&T Center of Excellence

USCG Innovation Showcase

June 2012

Dr. David Ebert
VACCINE, Director

Homeland Security

VACCINE Overview

- **VACCINE Mission Statement**

  Create methods and tools to analyze and manage vast amounts of information for all mission areas of homeland security.

- **Mission Relevance to USCG**

  VACCINE develops tools to enable more effective and efficient decision making both tactically, operationally, and strategically.

  We have developed a visual analytic tool allowing USCG to analyze and understand their SAR, LE, etc. data for D9, D5 and D1.

  Enabling formulation of various strategies to balance operational staffing levels with the risk/reward impact on response time, potential lives saved and property lost.

VACCINE Partners

Principal Partners	Areas of Expertise/Core Capabilities
Purdue University	Visualization of Structured, Unstructured and Streaming Data; Scalable Filtering and Dissemination; Video and Image Processing; Mobile, Lightweight Information Analytics; Public Safety Operations; Undergraduate/Graduate Education; Professional Training
Georgia Institute of Technology	Information Visualization; Visual Analytics; Peripheral Awareness; Undergraduate/Graduate Education;
Pennsylvania State University	Spatial cognition; Geo-information Representation, Spatial Analysis, Cartography; Undergraduate/Graduate Education; Professional Training
University of NC-Charlotte	3D Multimodal Interaction; Bioinformatics Visualization; Virtual Environments; Visual Reasoning; Interactive Visualization of Large-scale Information Spaces; Undergraduate/Graduate Education; Professional Training

Extended Partner Network
Florida International University*; University of Houston-Downtown*; University of Texas, Austin; California State University Dominguez Hills*; Morgan State University*; Virginia Polytechnic Institute and State University; Navajo Technical College*; Simon Fraser University, Canada; University of British Columbia, Canada; Oxford University, UK; Swansea University, UK; Stuttgart University, Germany; Stanford University; University of Washington; Tennessee State University; Middlesex University, UK; Dalhousie University, Canada; University of Manitoba; University of Victoria; Indiana University; Jackson State University*; University of Calgary, Canada; Carleton University; Justice Institute of British Columbia, Canada; Ontario Institute of Technology, Canada. * Indicates MSI

Homeland Security

Science and Technology

VACCINE Research Areas

Research Area	Approaches	Expected Uses	Customers	COE Partners
1. Indiana and Ohio State Law Enforcement and Public Safety	▪ Integrate and improve public safety's analysis of data ▪ Promote improved operations in the law enforcement, fire, emergency management and EMS fields	▪ Large-scale field tests on technologies including VALET, TRIP, iVALET, MERGE and GARI. ▪ Uses include intelligence analysis, social media fusion, asset and resource management, hazmat operations and investigations.	▪ Law Enforcement, Emergency Management ▪ TSA Rail Inspectors ▪ Community Watch	Seed Proposals to be considered from current and extended partners that meet objectives of this E2E project.
2. Operations and Analysis Visual Analytics Tools: Tactical and Strategic Decision Making	▪ Phase I will be the extension and adaption of cgSARVA to meet the needs of the USCG at the tactical and strategic level, including their proposed Coastal Operations and Analysis Suite of Tools (COAST).	▪ To inform senior CG leaders' decisions regarding asset capabilities, acquisitions, and allocations; unit locations; policies and concepts of operations; and mission tradeoffs. ▪ COAST is intended to measure the operational effectiveness of all 11 Coast Guard missions.	▪ USCG Headquarters and District Operations and Planning Personnel ▪ D1, D5, D9	Seed Proposals to be considered from current and extended partners that meet objectives of this E2E project.
3. TSA Airport Database Evaluation Visualization	▪ Analysis, assessment and visualization of TSA databases across airport environment	▪ Phase I of this project will provide TSA with an understanding of the database linkages, redundancies, efficiencies and inefficiencies across their airport environment.	▪ TSA	Seed Proposals to be considered from current and extended partners that meet objectives of this E2E project.

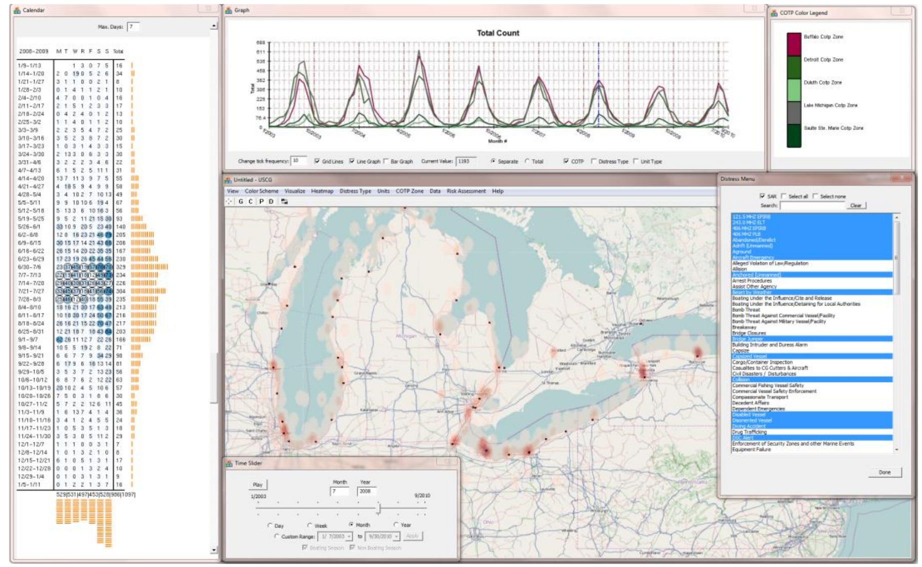
Homeland Security

Science and Technology

VACCINE –USCG Related Projects/Technologies

cgSARVA/Visual Analytics

- Coast Guard Search and Rescue Visual Analytics - Presents a risk analysis system that enables the interactive visualization, analysis, and assessment of SAR and other missions completed by each Coast Guard station. Through use of advanced GIS visualization and statistical research, cgSARVA provides a spatial and temporal overview of the USCG operations as well as a comprehensive maritime risk assessment and analysis on a strategic and operational level.
- Leveraged for COAST project



MERGE/Visual Analytics

- Mobile Emergency Response GuidE - The MERGE system is an electronic version of the Emergency Response Guidebook with many new features and capabilities. These new capabilities include the use of image analysis methods to automatically determine the type of Hazardous Materials present based on an image taken of the sign/placard.

Table of Contents

Homeland Security

Science and Technology

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

Table of Contents

Visual Analytic Applications for Law Enforcement

David S. Ebert

Edward Delp

www.VisualAnalytics-CCI.org

Who We Are:
International Team of Experts
- 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

VACCINE

What We Offer

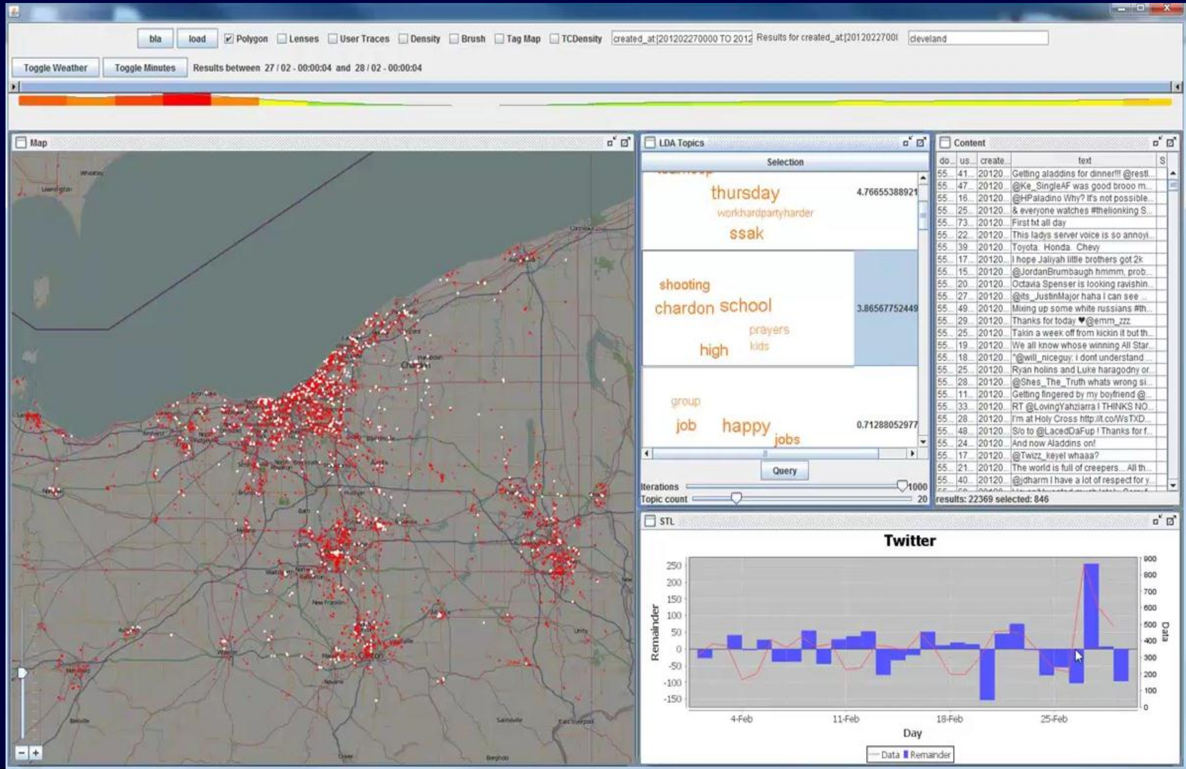
- **Innovative Fielded Solutions:** We provide innovative visual analytic and scalable solutions to the extended homeland security community
- **Improved Effectiveness**: We enable users to be more effective through innovative, interactive visualization, analysis, and decision making tools
  - Provide the right information, in the right format, within the right time to solve the problem
  - Enable user to be more effective from planning to detection to response to recovery
  - Enable effective communication of information
- **People and Partnerships**
  - Interdisciplinary world-leading team of researchers and students
  - Actively Engaged Partners –
    We define and extend the new science of visual analytics driven by real-world, real-scale problems of engaged partners (local, state, federal)

VADM Robert Parker with VACCINE student researchers (cgSARVA, COAST, iOPAR)

"cgSARVA has proven its worth time and again, providing key analytic information for decision makers for large scale projects…"
VADM Robert Parker, 2012 MRS Keynote Address

Visual Analytics Uses for Crisis Management and Response

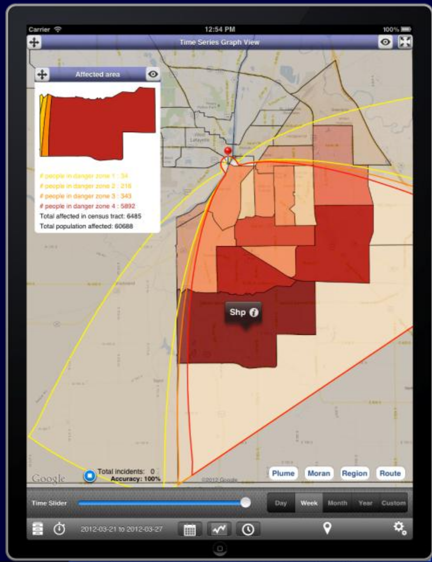
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis
- Purpose:
  - Planning for resiliency
  - Long-term analysis
  - Predictive analytics
  - Training
  - Detection
  - Investigation
  - Response
  - Recovery, remediation



VACCINE

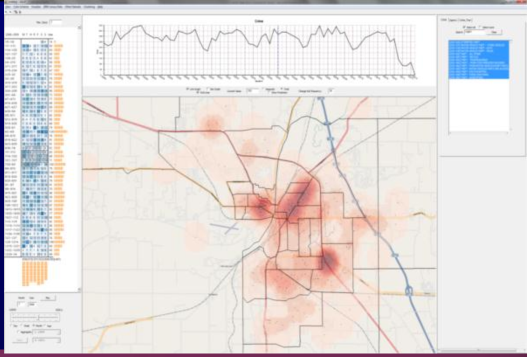
Visual Analytics Uses for Public Safety

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive and multi-day investigations
- Multisource, multimedia massive data integration & analysis
- Visual standardization and interoperability



VACCINE

Visual Analytics Uses for Law Enforcement

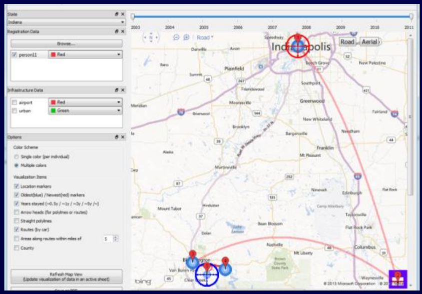


- Situational Awareness and Assessment from massive data
  - In-field and desktop historical and current incident/crime information
  - Innovative spatiotemporal visualization, analysis, prediction
  - Image analysis tools
  - Integrated statistical forecasting and simulation models
  - Advanced social media analytics
- Tools for intelligence analysis
- Planning and response support
  - Integrated simulation, resource allocation, census, weather data

Current Public Safety Projects: Law Enforcement

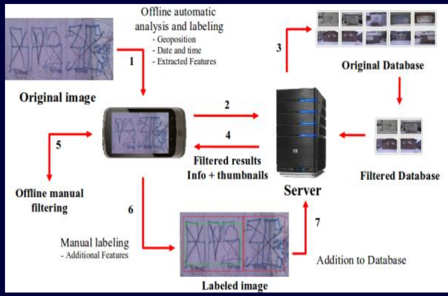
Desktop crime analytics and intelligence analytics

- Jigsaw
  - Delivered to West Lafayette PD, IMPD 09/2010, TSA 06/2012
- TRIP
  - Delivered to Indiana Intelligence Fusion Center May 2012
- VALET & CrimeViz
  - Delivered to Lafayette PD 2011, WL PD Spring 2012
  - Delivered to Harrisburg Bureau of Police Spring 2011



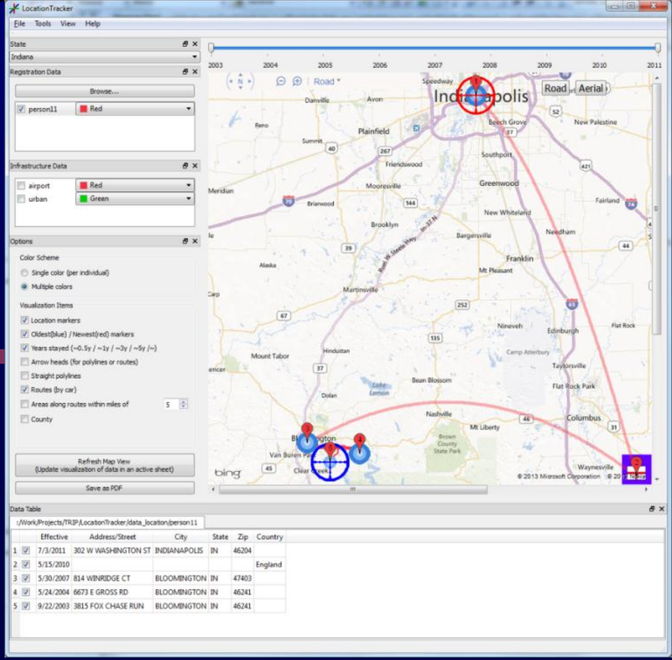
In-field crime analytics

- Gang Graffiti Mobile Application (Gari)
  - Delivered to IPD 09/2011, in use by 15 different agencies in Indiana
- iValet
  - Delivered to WLPD, LPD, Purdue PD Fall 2011

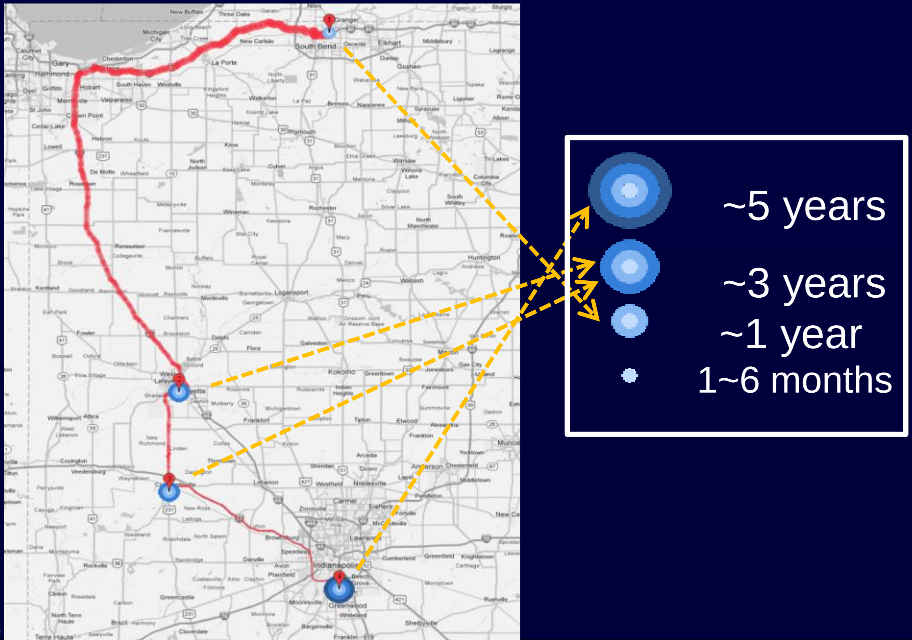


VACCINE

TRIP: Travel Response Investigative Profiler

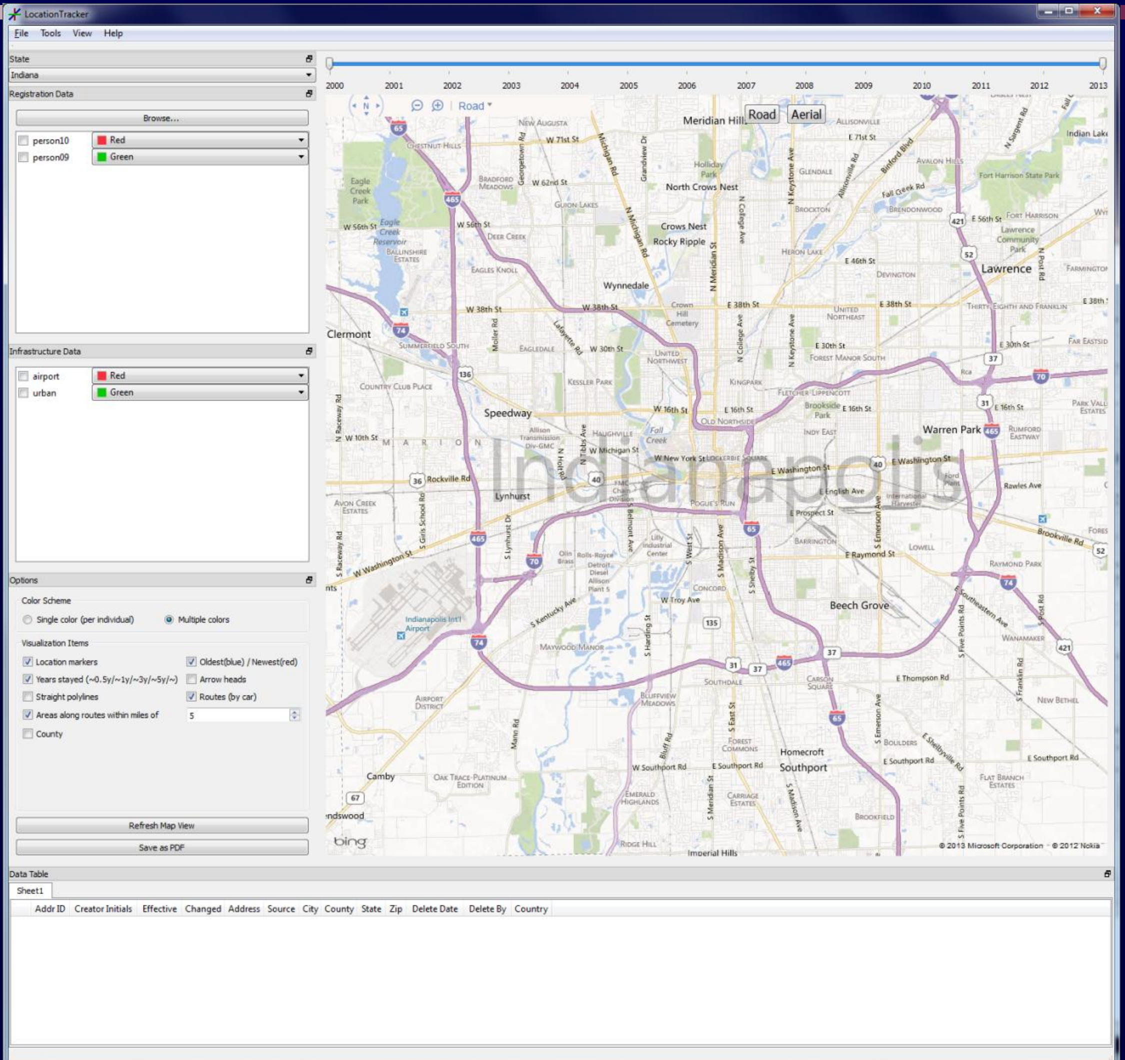


- Understanding movement behavior and spatiotemporal patterns of individuals
- Integrated spatiotemporal visualization, exploration and analysis of multiple individuals' movement history
  - Geocoding
  - Routes
  - Reachable areas
  - Temporal cues
  - Various visualization items
  - Support various address formats
  - International address visualization



System Overview

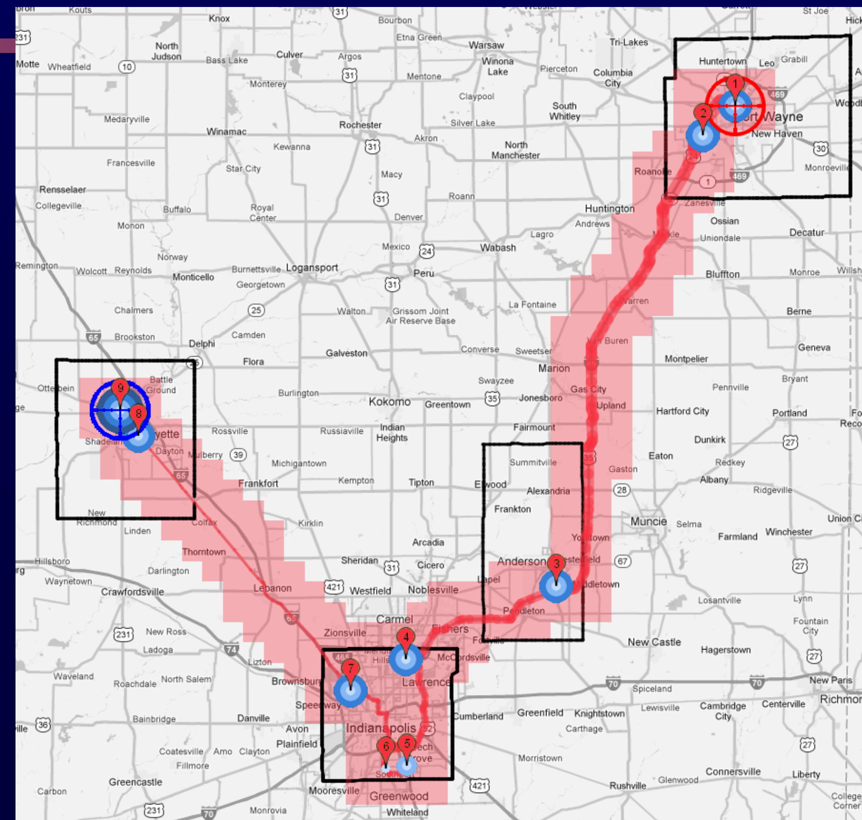
1. State selector
2. Data browse button
3. Data list view
5. Infrastructure data list view
6. Color scheme selector
7. Visualization options
8. Refresh & Save PDF
9. Table view



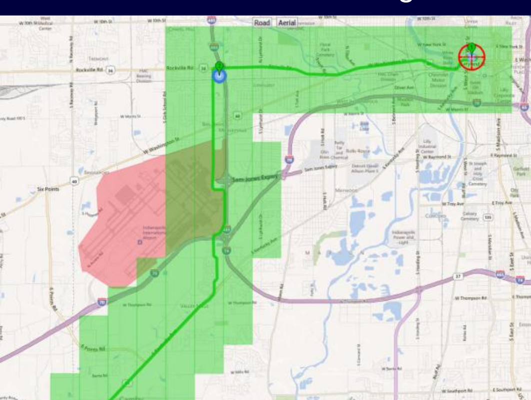
VACCINE

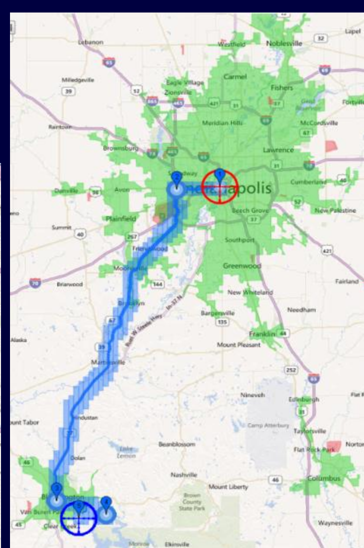
Various Visualization Items

- Locations
- Oldest / newest locations
- County
- Routes
- Reachable areas
- Infrastructures



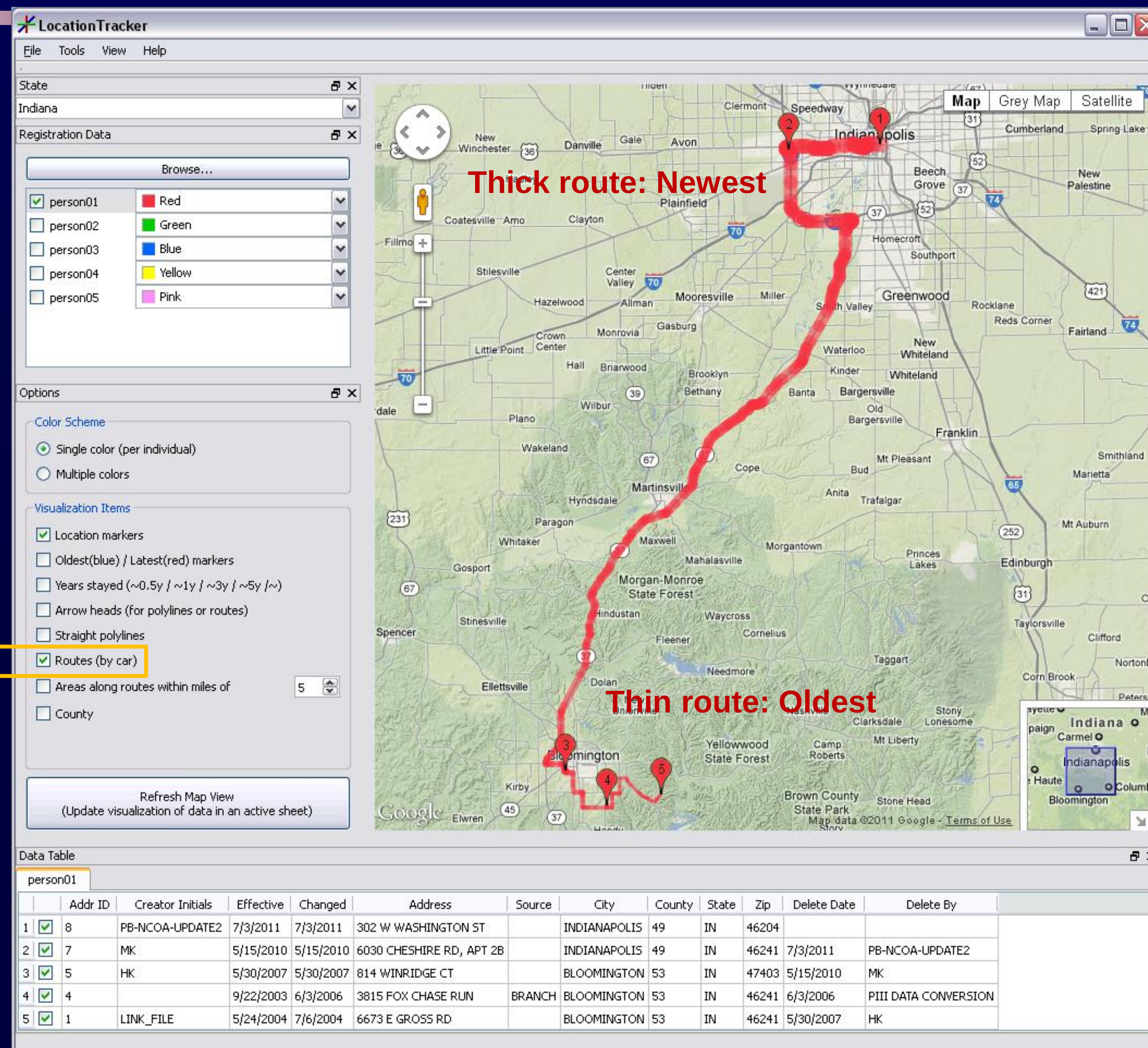
Infrastructure and Metro regions





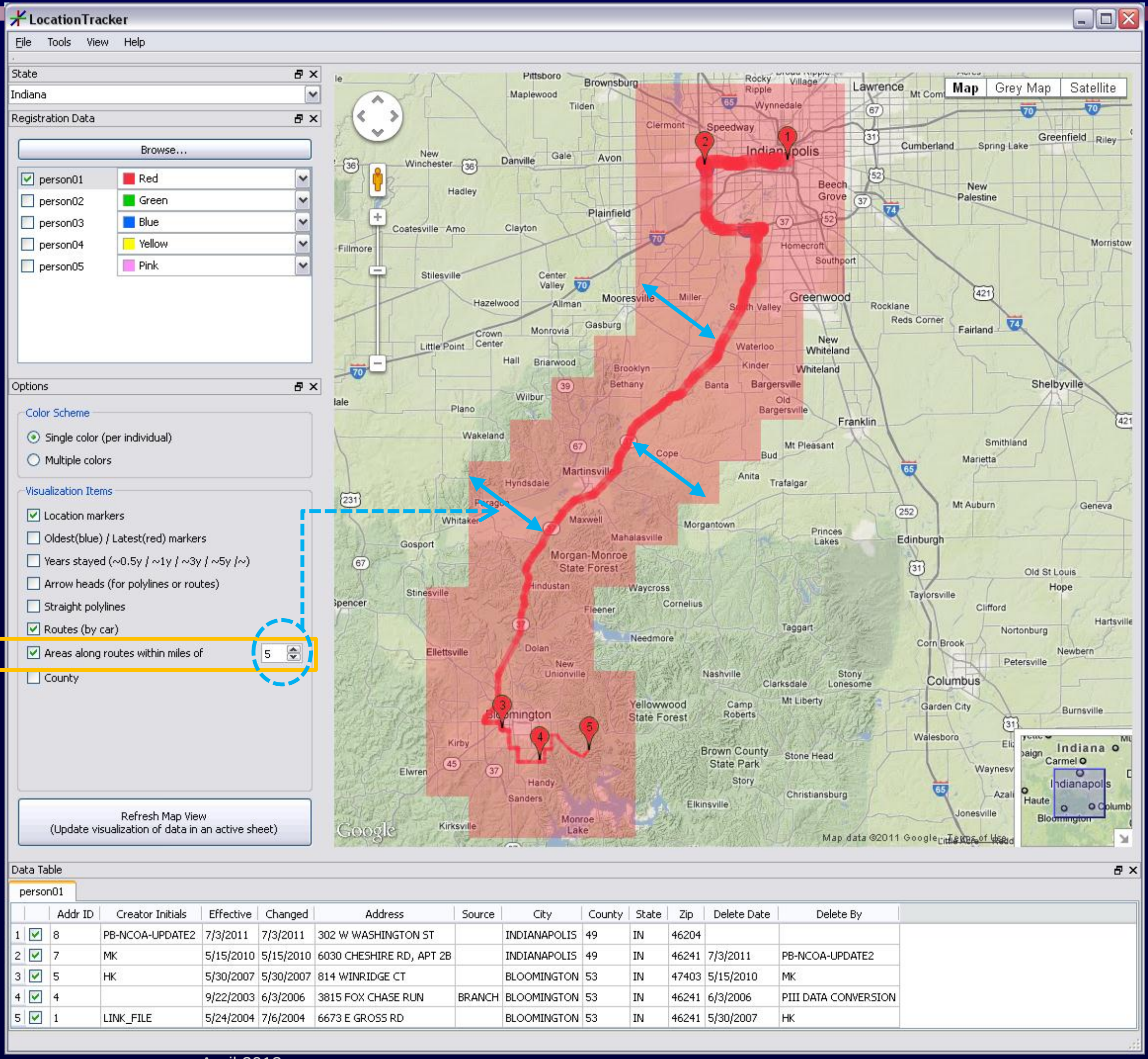
Routes

- Driving routes
- Route thickness
  - Temporal cue from the oldest to the newest

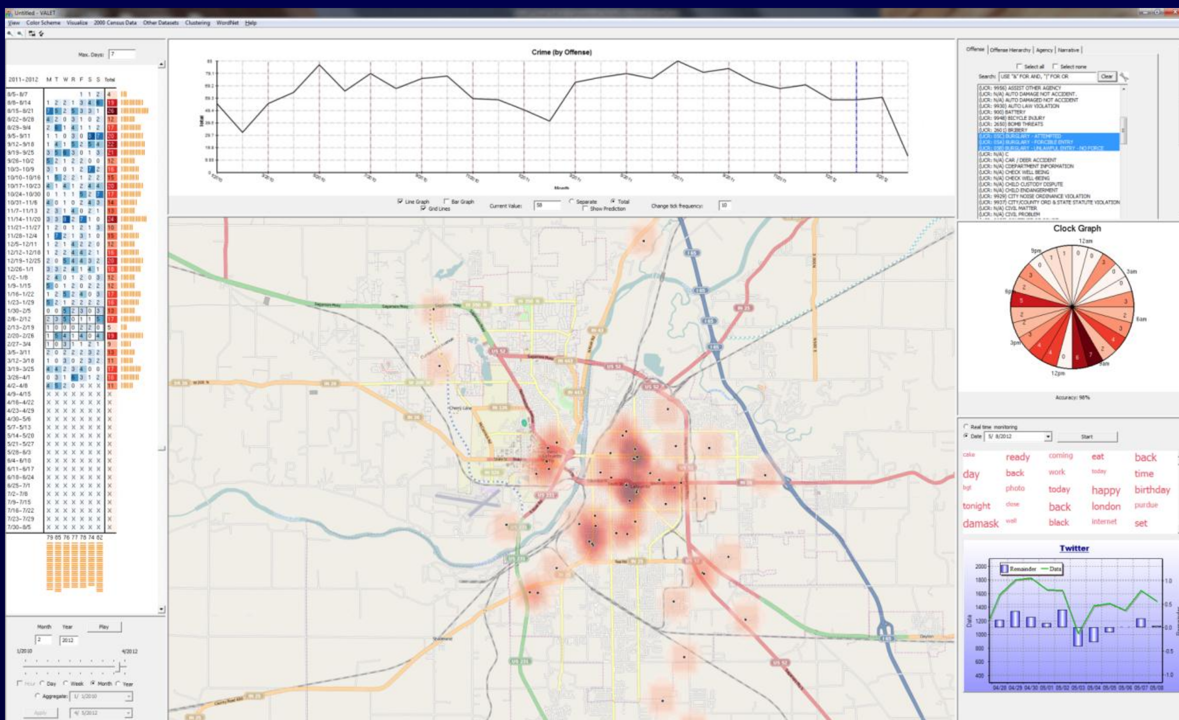


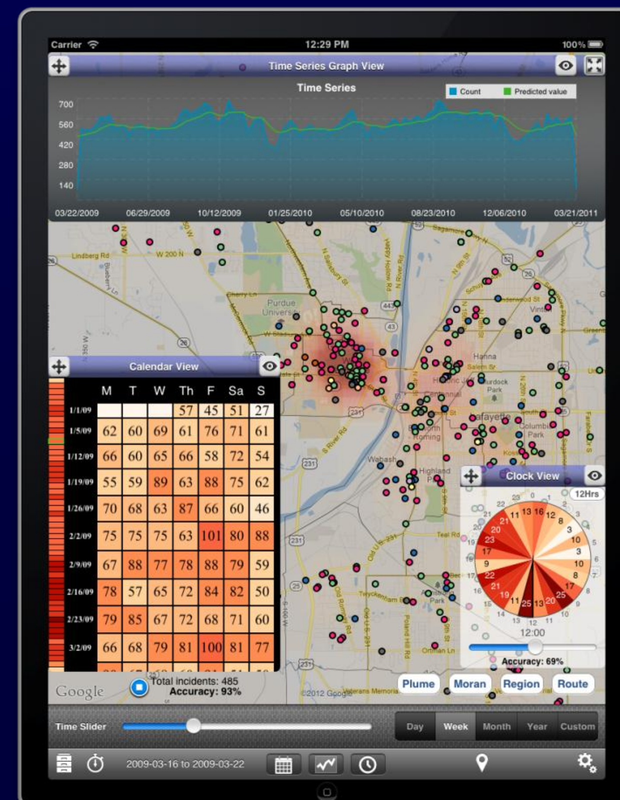
Reachable Areas

- Polygonal geofences
  - Along the routes
  - Given a distance that can be traveled.



Visual Analytics Law Enforcement Toolkit
(i)VALET

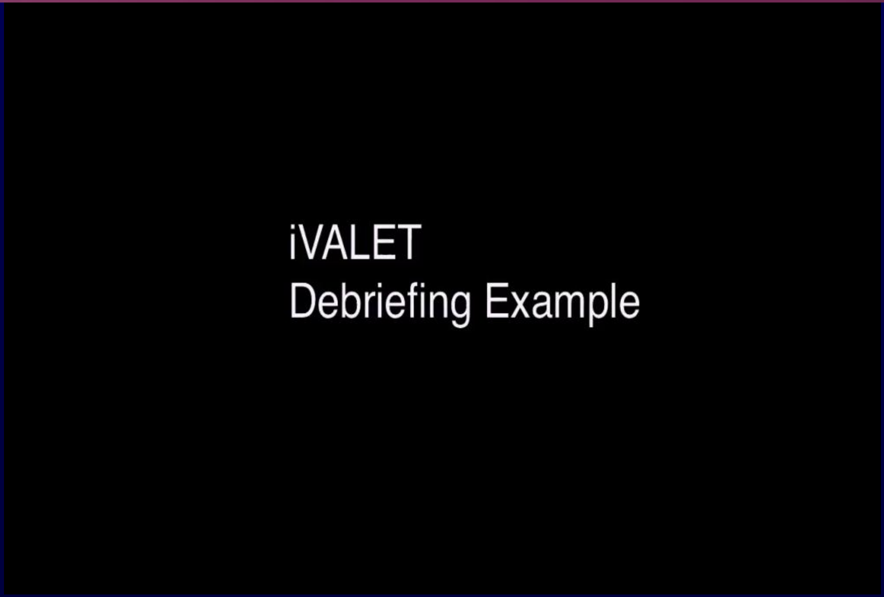




Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

Impacts:

- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (February 2013) for community-based policing
- Investigating correlation factors
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning
- Incorporating predictive alerts

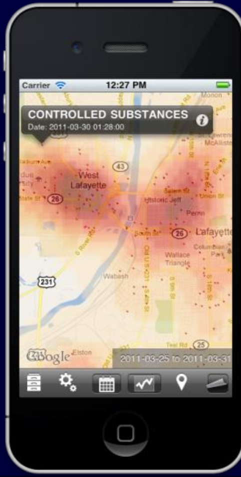


VALET delivered:

- Spring 2011: WL, Lafayette Police

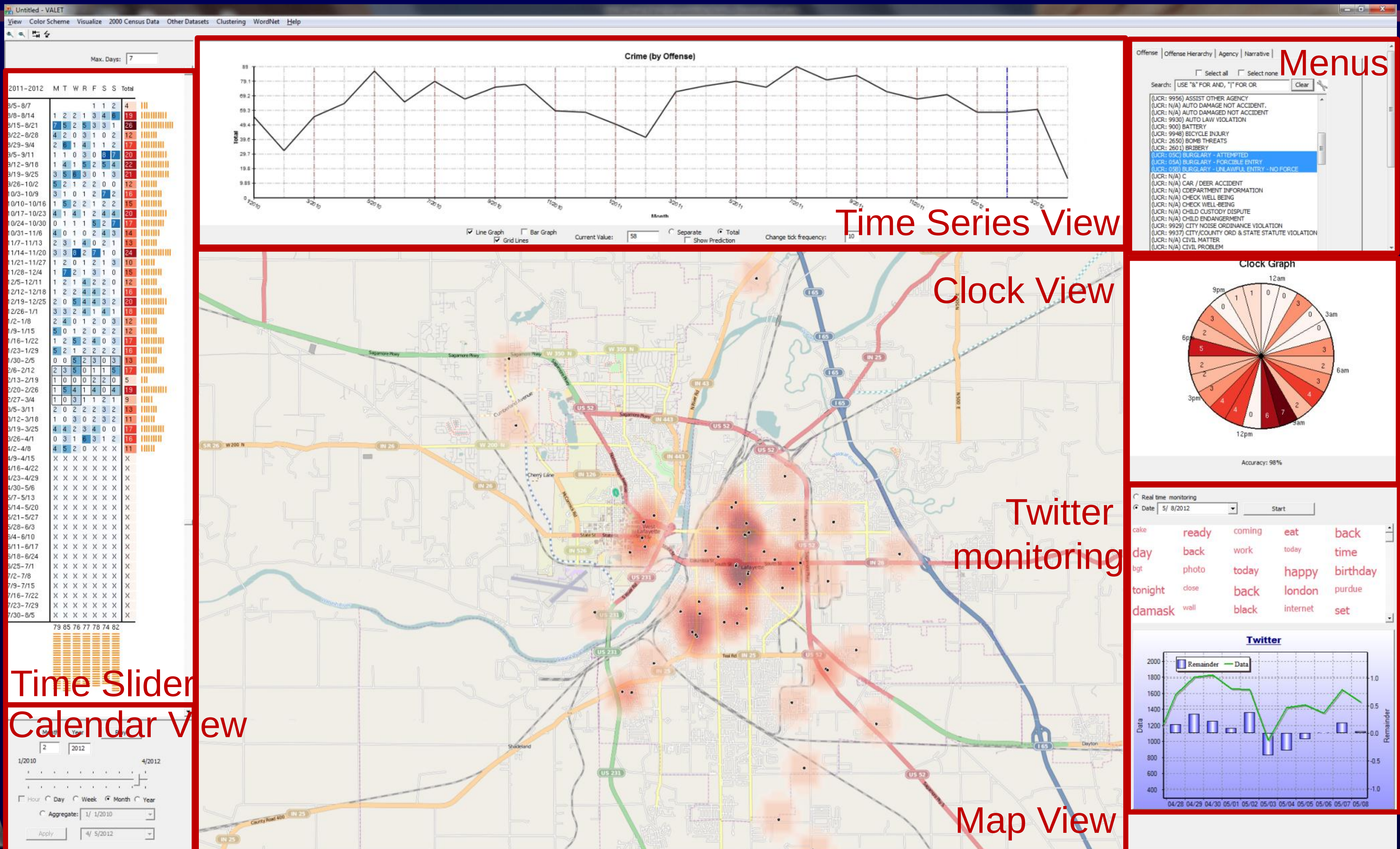
iVALET delivered:

- October 2011: Purdue, WL Police

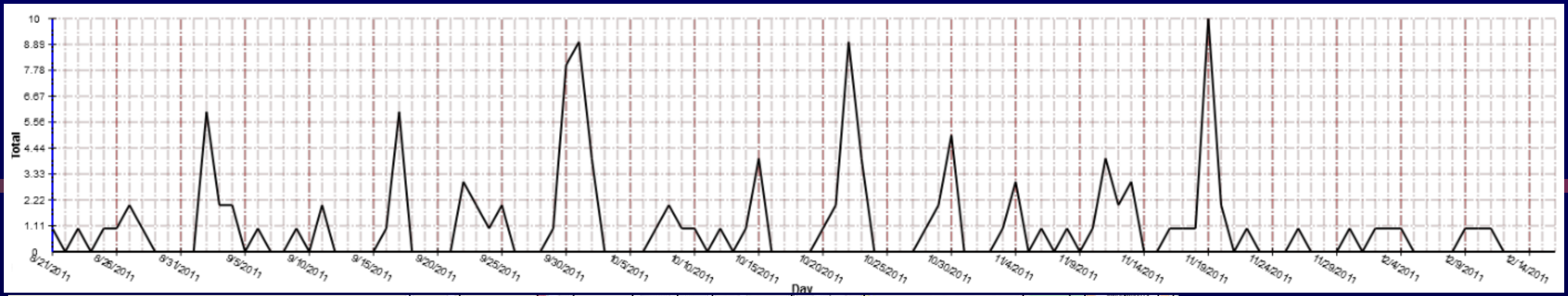


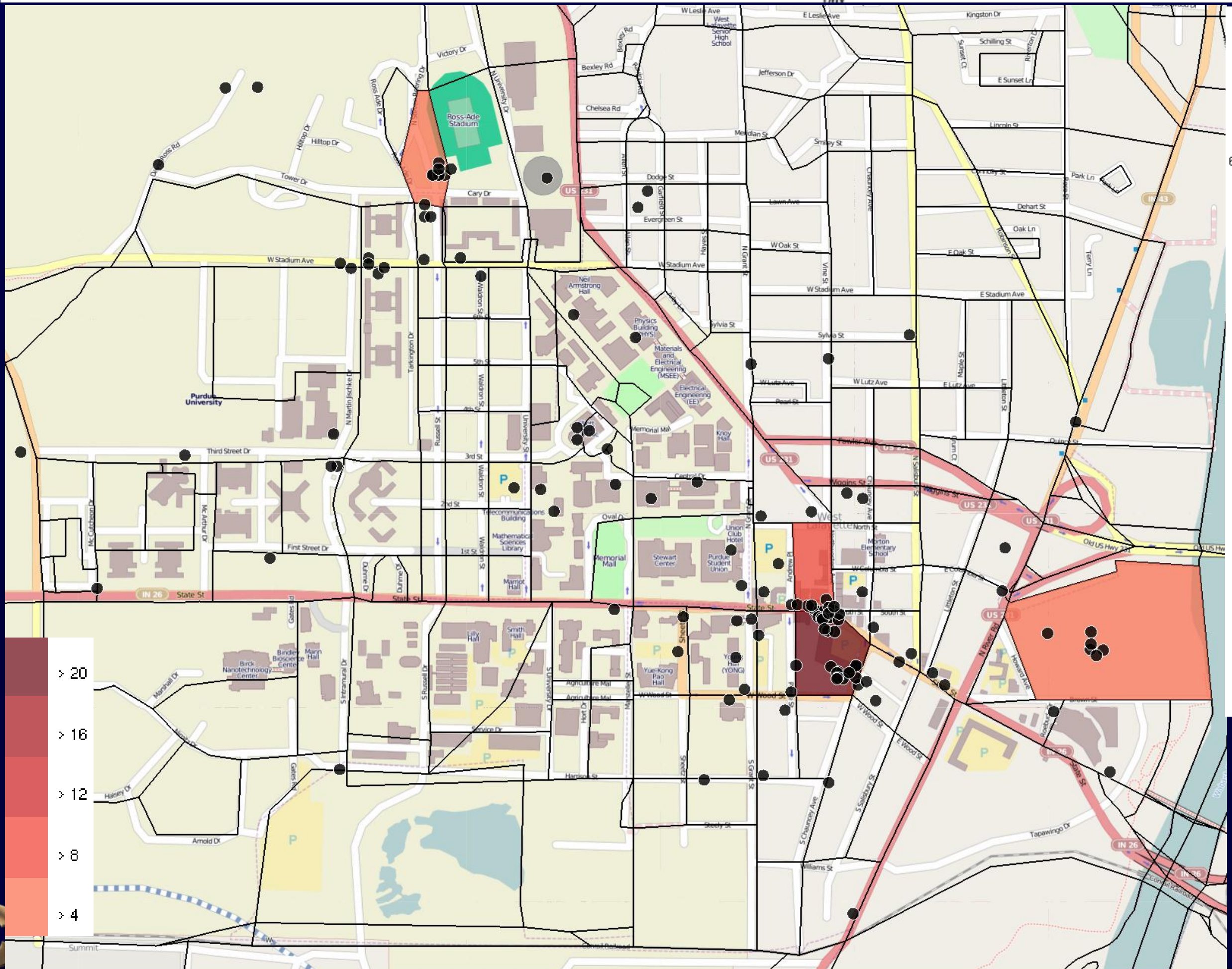
VACCINE

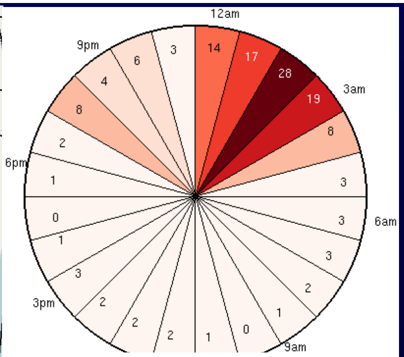
VALET Overview



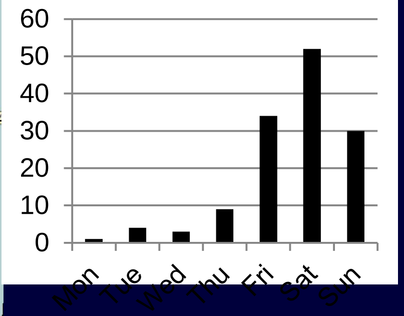
Example: Drunkenness / Public Intoxication



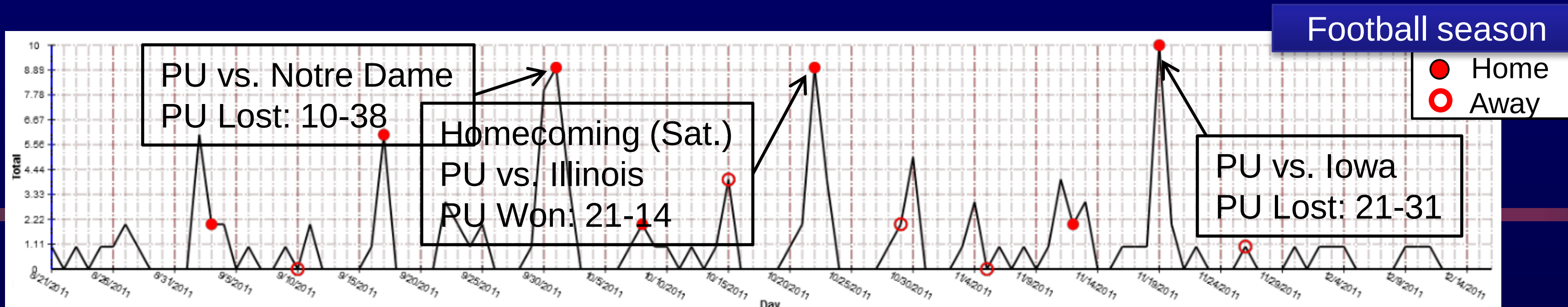


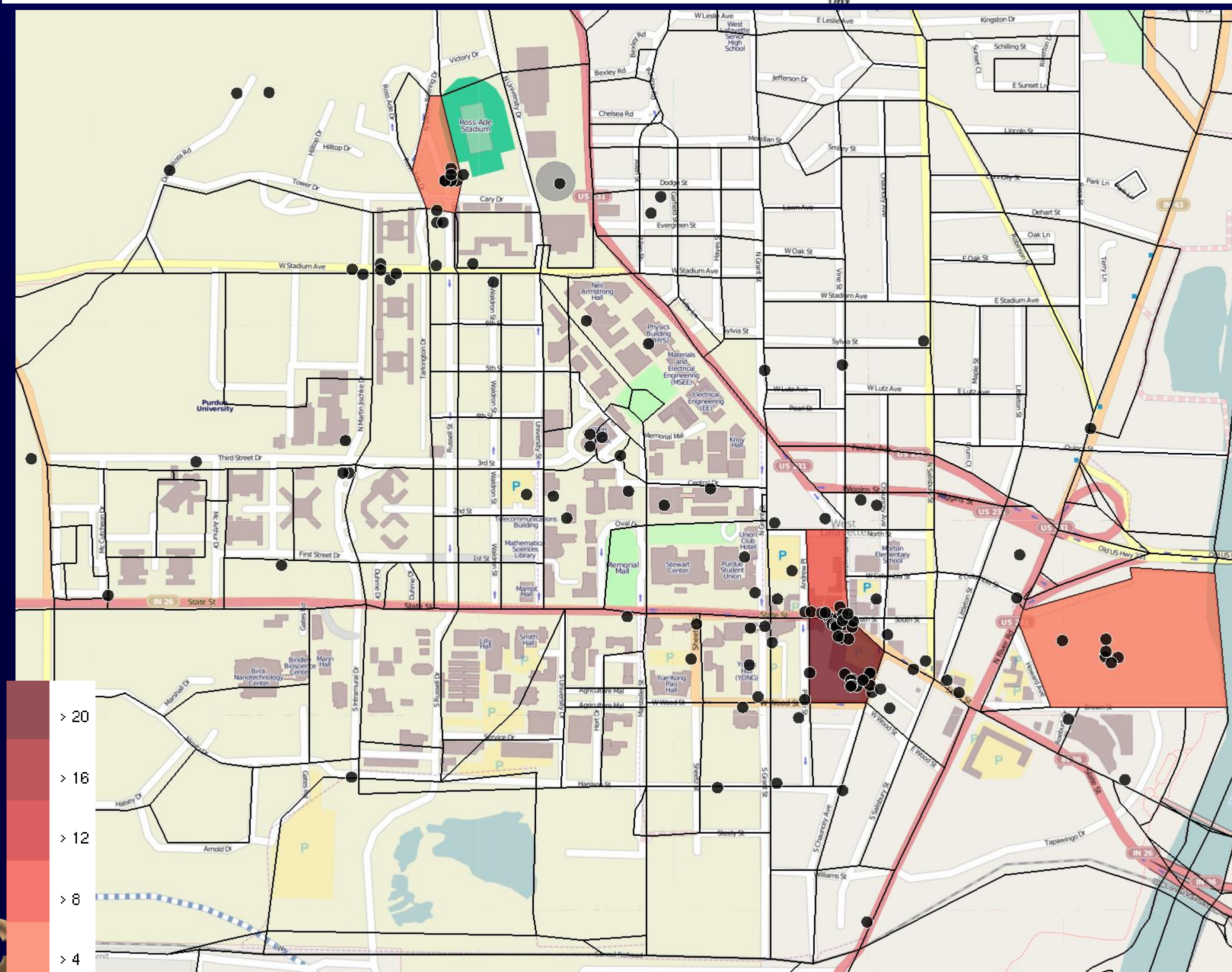


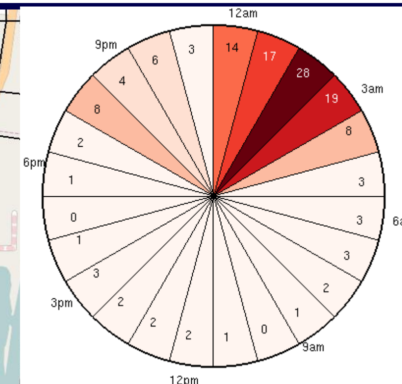
Day-of-the-Week

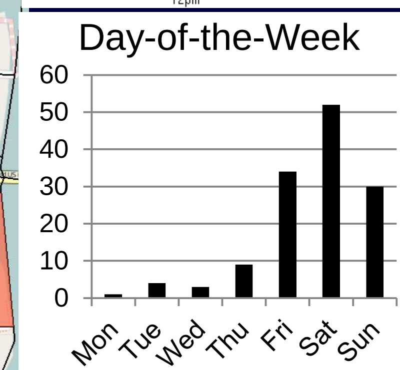


Example: Drunkenness / Public Intoxication





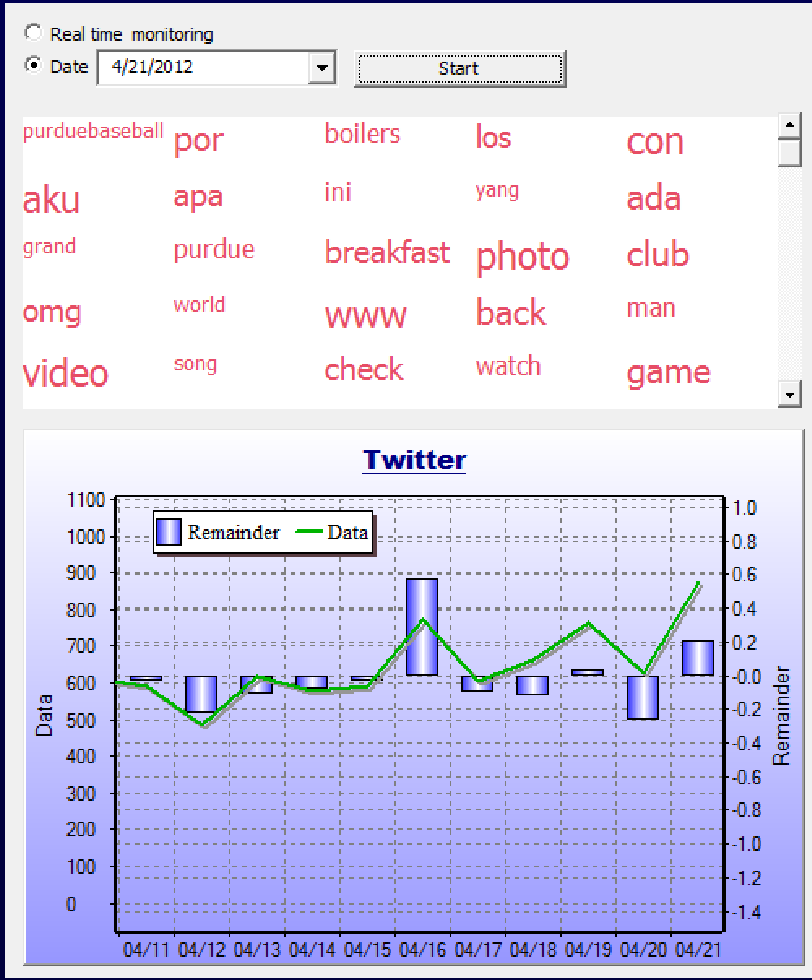




Social Media:Real-time Twitter Monitoring and Integration into Tools

(Purdue, Stuttgart, Penn St.)

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud

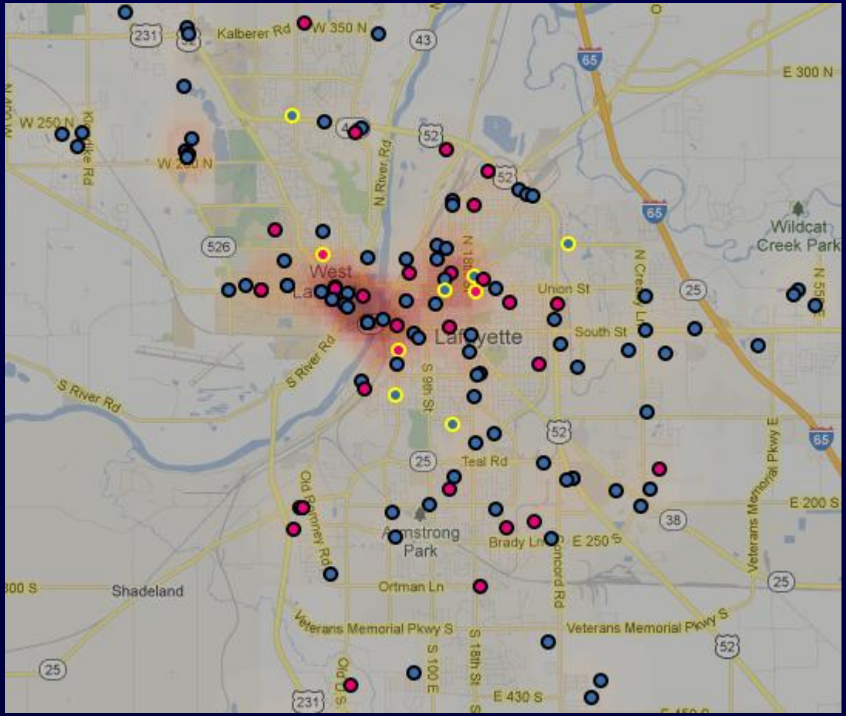


Grand Prix Weekend, Purdue University

VACCINE

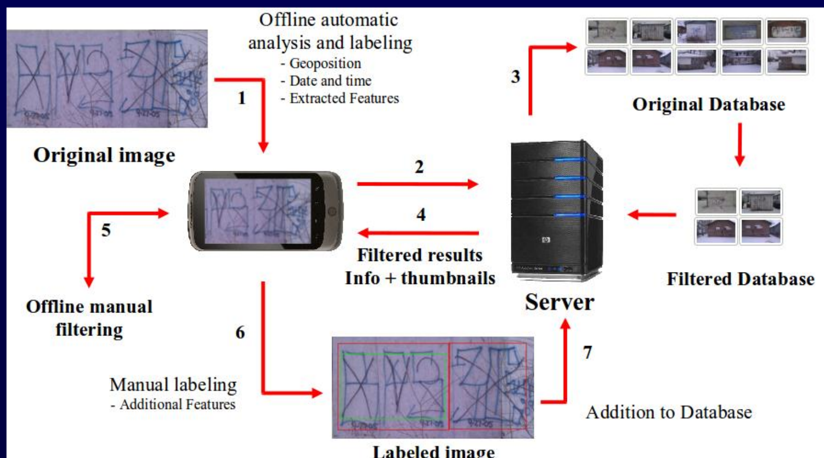
Top 10 Hot Incidents

- Identify unusual localized high-frequency patterns of crimes (near repeats)
- Each data entry is checked for other crimes with similar properties within a 1 block radius of the incident location and a 14-day time period;
- Top 10 incidents with the most number of related incidents in this space-time window are highlighted




Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
11/14/2010
01/01/2011
02/13/2011
03/27/2011
Calendar View
M T W Th F Sa S
1/1/10 36 46 29
1/4/10 38 48 44 43 42 32 30
1/11/10 62 53 53 56 59 45 34
1/18/10 62 49 53 72 61 52 49
1/25/10 98 66 57 66 64 55 49
2/1/10 58 63 50 51 73 51 42
2/8/10 50 46 48 63 52 39 35
2/15/10 48 65 48 54 55 67 51
2/22/10 51 66 48 56 57 46 47
3/1/10 50 59 42 54 82 45 41
3/8/10 50 57 64 78 63 36 30
Clock View
12Hrs
10:00
Accuracy: 64%
Hedgewood
Hanna
Centennial
Perrin
Lafayette
Wabash
Highland Park
©2012 Google
Choro
Plume
Moran
Region
Route
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20

Gang Graffiti Analysis and Recognition Using a Mobile Telephone (GARI)



IMPACT:

- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be “interpreted”
- More than 75 users and 800 graffiti images acquired

GARI delivered:

- Summer 2011:
  - IPD gang detectives
- August 2011:
  - IPD at large
  - Ind Fusion Gang Task Force
- September 2011:
  - Gang detectives across Indiana

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu

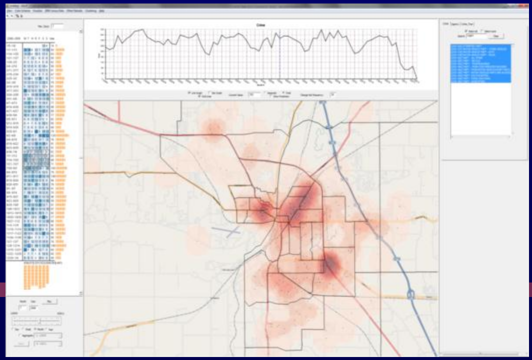
ebertd@purdue.edu

VACCINE

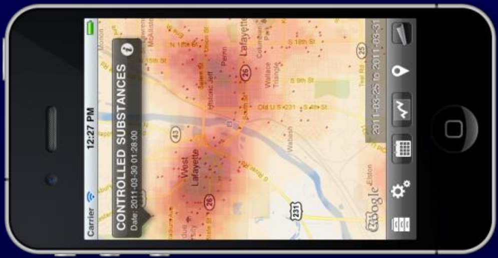
VACCINE

Table of Contents

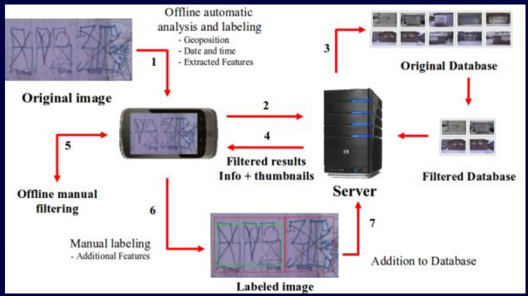
Public Safety & Law Enforcement



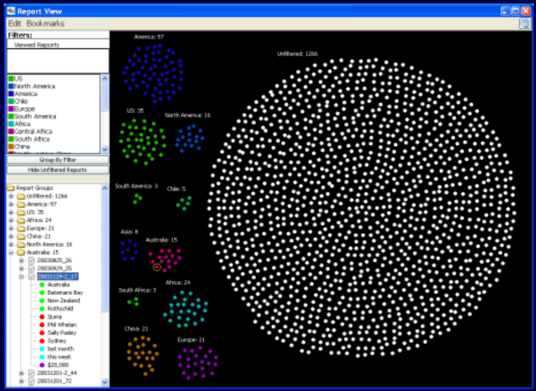
- Law enforcement visual analytics
  - VALET, iVALET (iPhone/iPad), CrimeViz
    - Visual Analytics Law Enforcement Toolkit
    - Analyzing crime patterns and time of day problems



- Gang activity analytics
  - GARI
    - Gang Graffiti Recognition and Interpretation using a mobile telephone
    - Allowing police to catalog and analyze gang graffiti images, better track and determine gang activities



- Document visual analytics
  - JIGSAW
    - Visualization for investigative analysis
    - Discovery of hidden relationship and threats across documents



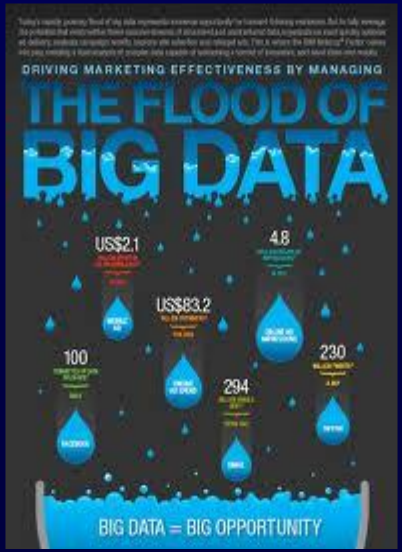
Visual Analytics: A Lifeboat In The Big Data Ocean

Table of Contents

David S. Ebert
ebertd@purdue.edu

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

Big Data Is The Rage



Big Data Challenges

- Heterogeneity
- Unprecedented volume and/or complexity
- Multivariate, Multidimensional
- Streaming
- Qualitative, Quantitative, Categorical, Temporal,…
- Distributed Sources: sensors, text, images, maps, …
- Inconsistent and Uncertain
- Different Scales and Cross Scale

The Curse of Dimensionality (Bellman 1961)

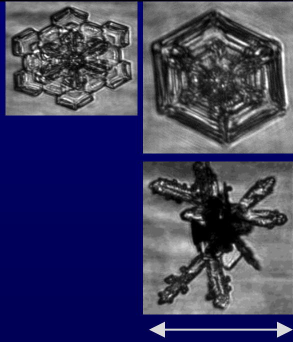
- Problems associated with multivariate data analysis as the dimensionality increases the available data becomes sparse
- Sparsity is a problem for any method that requires statistical significance
- In visualization, we are also limited with screen space and the number of available visual variables, so choosing the most appropriate dimensions is key

Richard Ernest Bellman; Rand Corporation (1957). *Dynamic programming*. Princeton University Press.

Big Data Isn't Just Large Data!

- All of the above characteristics lead to big, complex data
- Example 1: A few million crime records with uncertain, missing data, dozens of fields, associated risks, related social media data, images, …
- Example 2: Basic question of understanding why and when a cumulus cloud will produce precipitation…

Example Challenging Problem - Atmospheric Science: Multi-scale Interactions (in the words of a cloud physicist)



1 mm
1 kHz

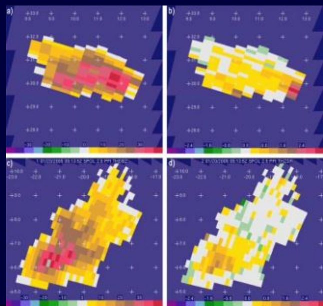
No observing platform can measure the quantities of interest over all needed spatial and temporal scales needed

No numerical model can simulate the quantities of interest over all needed spatial and temporal scales

→We observe/simulate over a subset of the pertinent scales, using different instruments/models, and must assimilate these results to understand the "big picture"

1km
5min

Visual analytics is crucial for this task

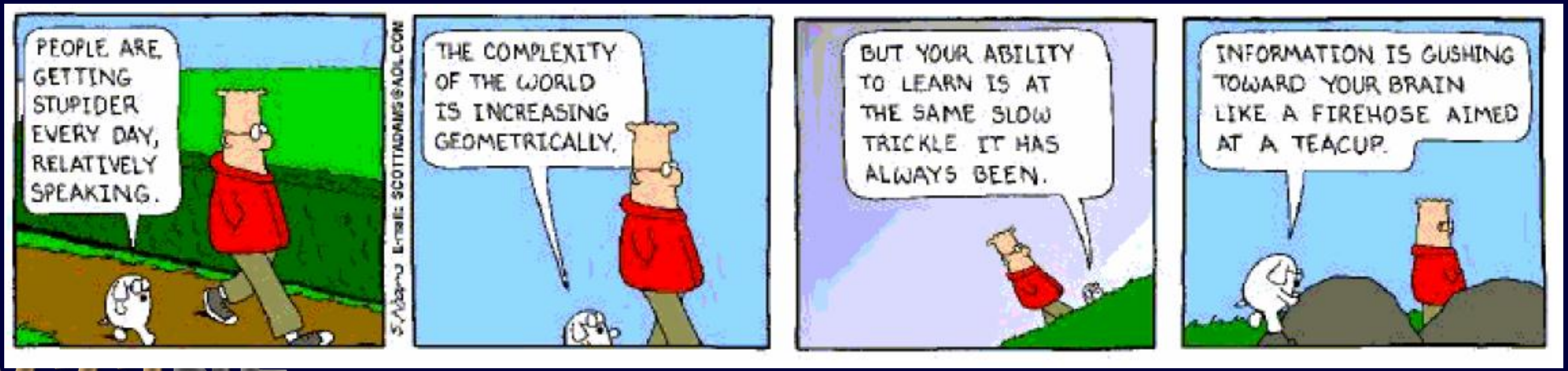


Issues: Multi-scale, multi-system, multisource, massive, data & simulations

VACCINE

Data Overload

- How can we make sense of this data?
- How can we harness this data in the decision making process?
- How do we avoid being overwhelmed by all of this?
- **How to turn data into actionable knowledge?**



VACCINE

The Power of Visualization

- The goal is to take all of this *data* and transform it into *information and knowledge*
- How many terabytes of data we have collected doesn't matter, it's how many petaflops of *insights* we can generate from this data
- "*The purpose of computing is insight, not numbers.*" RW Hamming (1971)
- We need to make the data understandable to people and a key way of doing this is through *visualization*

VA Approaches to Big Data

- Don’t display all the data!!
- Extract relevant information at the appropriate natural scales using analytics and statistical models to reduce data space so that it is mappable to the visual space
- User interaction to guide above process
- Should we use big iron or cell phones?

Six Challenges for Proactive & Predictive Big Data Visual Analytics

1. Creating computer-human visual cognition environments
2. Integrating interactive simulation and analytics
3. Solving specific scale issues and cross-scale issues
4. Managing uncertainty and time
5. Enabling risk-based decision making environments
6. Developing the Science of Interaction for Visual Analytics

Path: Where Do We Want to Go?

- Solving these challenges will enable dramatic changes in information-driven risk-based decision making, discoveries, science, and engineering!

VACCINE

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu
ebertd@purdue.edu

VACCINE

Path: Where Do We Want to Go?

- Answers and challenges
- An example
- Illustrative examples
  - Science visual analytics
  - Health visual analytics
  - Public safety and law enforcement
- Visual analytic challenges and future

What We Do With Visual Analytics

- Enable effective decision making through interactive visual analytic environments
- Enable effective communication of information
- Provide quantitative, reliable, reproducible evidence
- Enable user to be more effective from planning to detection to response to recovery
- Enable proactive and predictive visual analytics
- Enable effective situational awareness (perception, comprehension, projection)

What's Needed for Proactive and Predictive Visual Analytics?

Table of Contents

- Reliable and reproducible models and simulation
- Understanding of the data
  - Distribution and skewness, errors, appropriate analysis techniques
- Understanding of the sources and types of data
- Comparable or Correlative sources of data
  - Appropriate transformations applied to enable meaningful comparison and correlation
- Understanding of the use and problem to be solved!

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

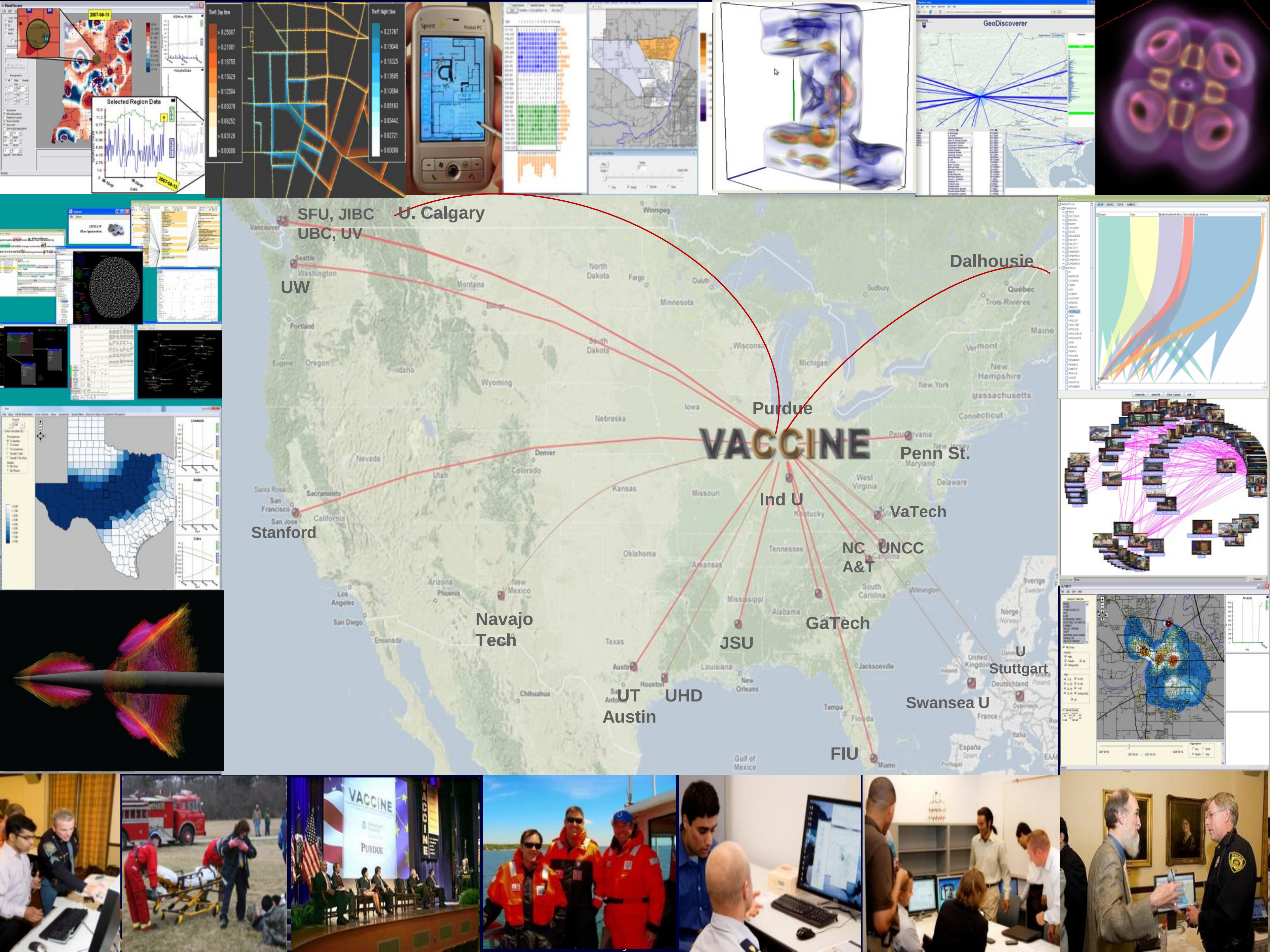
Table of Contents

Visual Analytics:

From Situational Awareness to Risk-Based Decision Making

David S. Ebert

ebertd@purdue.edu



SFU, JIBC
UBC, UV
U. Calgary
UW
Dalhousie
Purdue
VACCINE
Penn St.
Ind U
VaTech
NC
A&T
UNCC
Stanford
Navajo
Tech
JSU
GaTech
UT
Austin
UHD
FIU
U
Stuttgart
Swansea U
GeoDiscoverer
Selected Region Data

Who We Are:
International Team of Experts
- 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

Some VACCINE Regional and Corporate Partners

- Indiana Department of Homeland Security
- Indiana State Department of Health
- Indianapolis Public Safety
- Charlotte Mecklenburg Police Department
- Lafayette, West Lafayette, Purdue Police Departments
- Tippecanoe County EMA & Sherriff's Department
- Harrisburg, PA Police
- Coast Guard Sector Boston, Seattle, LA
- Joint Harbor Operations Command Center
- Port of Seattle
- Indianapolis Public Safety
- USCG District 9, D5, D1
- Motorola
- Harris Corporation
- Boeing
- Next Wave Systems, LLC
- Banfield, The Pet Hospital
- Raytheon
- MacDonald, Dettwiler and Associates
- Oculus Info Inc.
- Kx Systems
- Bank of America
- Duke Energy
- World Vision International
- Gates Foundation
- Kimberly Clark
- Hallmark

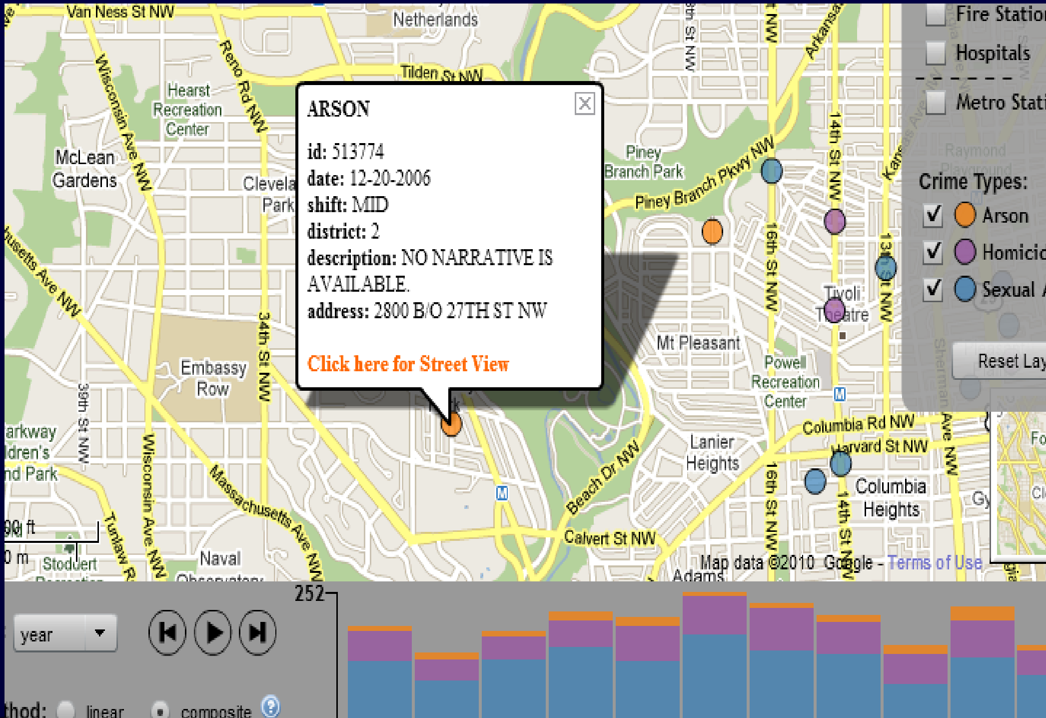
VACCINE

Some VACCINE Government Partners

- U.S. Coast Guard
- Federal Emergency Management Agency
- Customs and Border Patrol
- National Geospatial Intelligence Agency
- National Science Foundation
- Army Research Office
- Department of Defense
- Department of Health and Human Services
- U.S. Dept. of State
- National Institute of Justice
- National Maritime Intelligence Center
- Defence Research & Development Canada
- Foreign Broadcast Information Service
- US-CERT
- Oak Ridge National Laboratories
- US Army Corps of Engineers-ERDC
- Pacific Northwest National Laboratory
- ARL CERDEC & ARDEC
- US Attorney's Office (David Capp)

Visual Analytics Uses for Public Safety and Situational Awareness

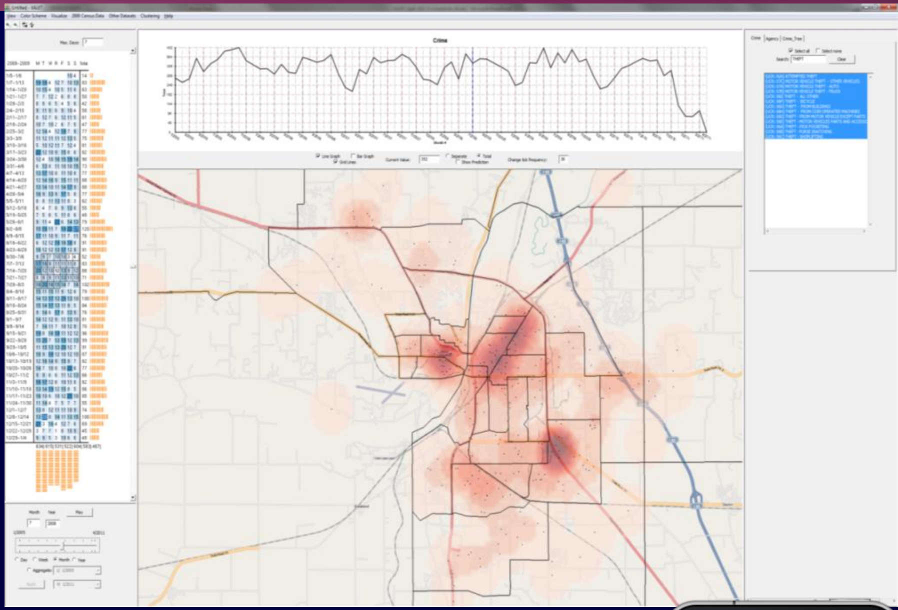
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis
- Purpose:
  - Planning for resiliency
  - Long-term analysis
  - Predictive analytics
  - Training
  - Detection
  - Investigation
  - Response
  - Recovery, remediation



VACCINE

Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

End-Users: Lafayette, WL Police, Indiana Fusion Center, Indiana State Police, Ohio Fusion Center (initiated)



Impacts:

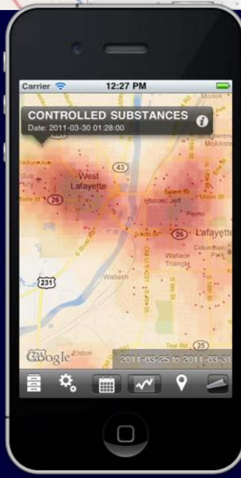
- In use to analyze crime patterns in Lafayette, Indiana and connect strings of activities
- Mobile version being released to public (November 2011) for community-based policing
- Investigating correlation of bus routes and crime, street lights and crime
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning

VALET delivered:

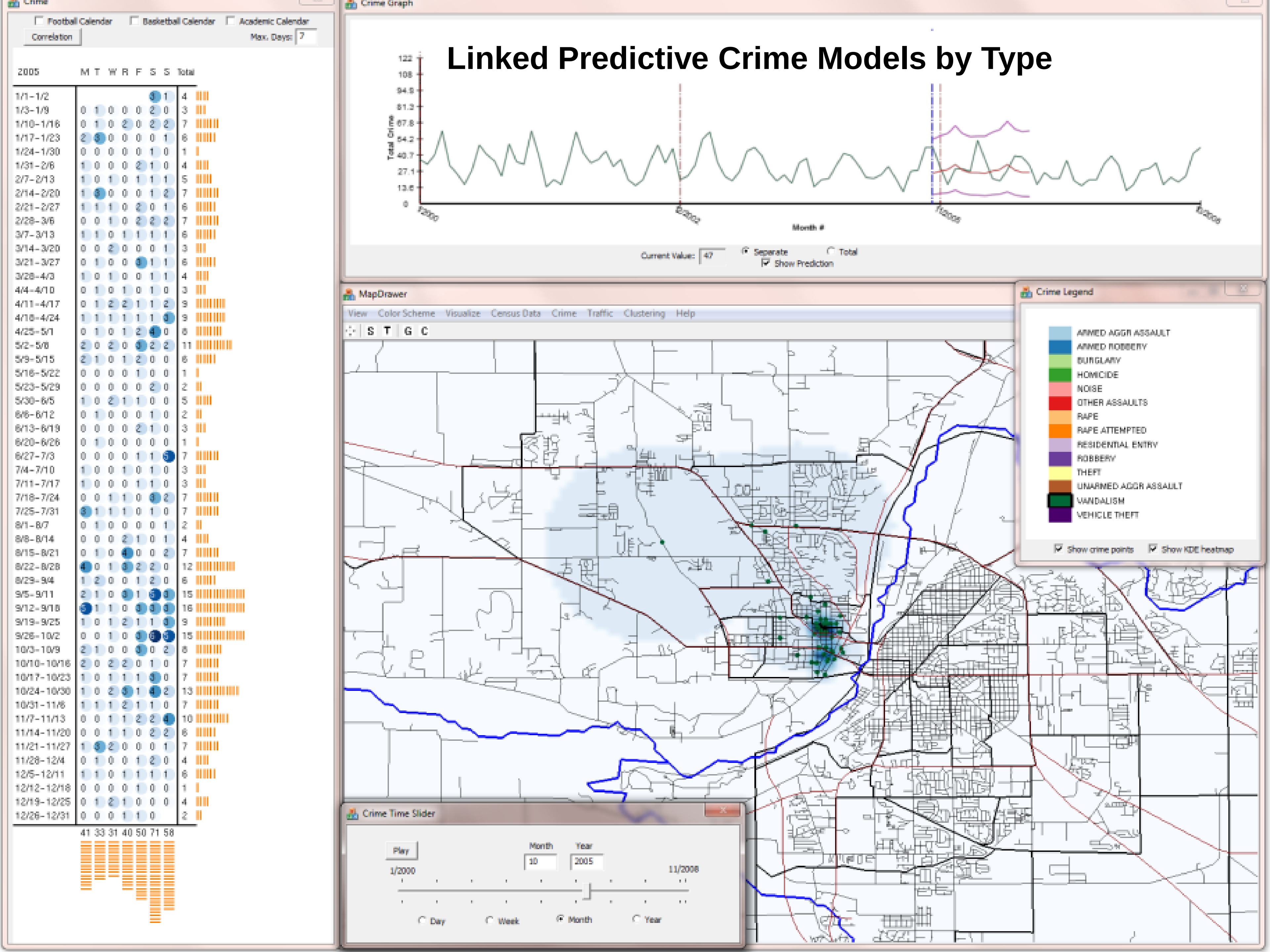
- Spring 2011: WL, Lafayette Police

iVALET delivered:

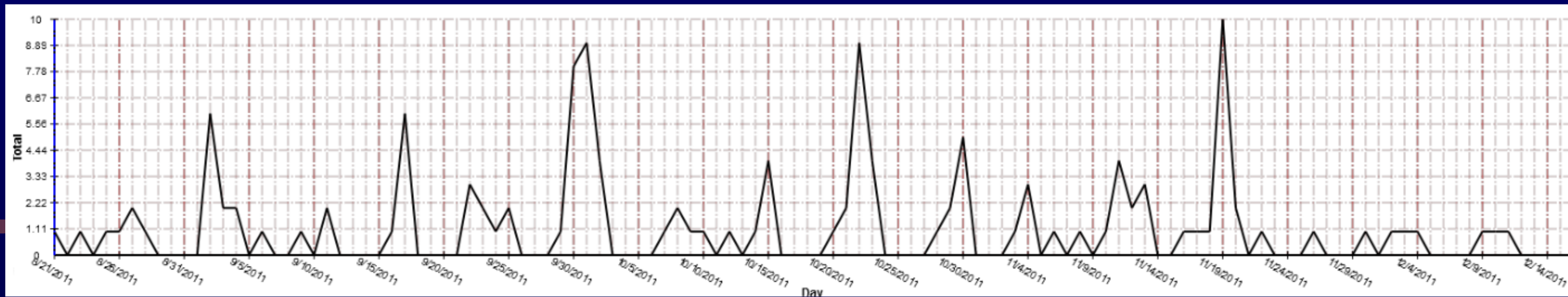
- October 2011: Purdue, WL Police

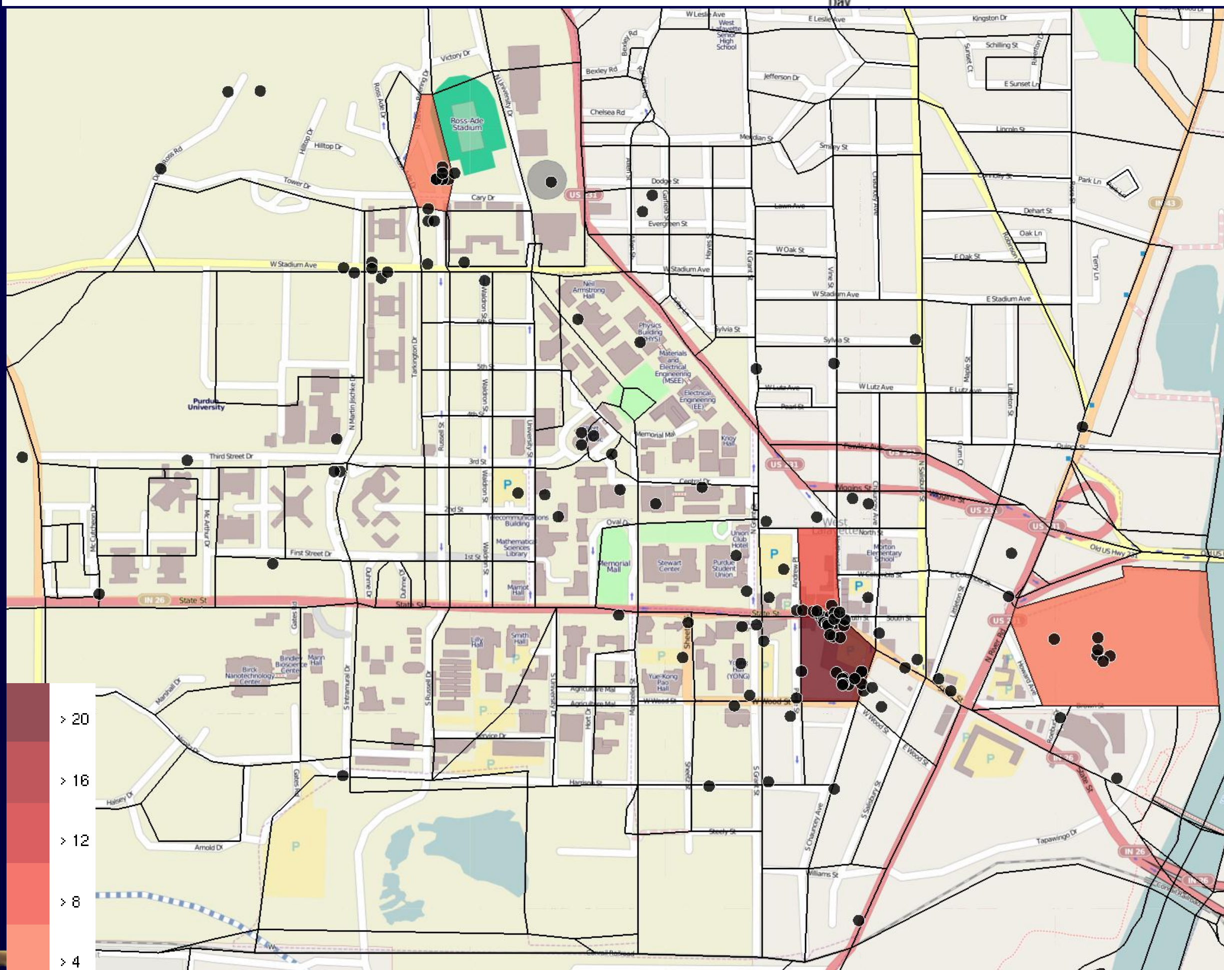


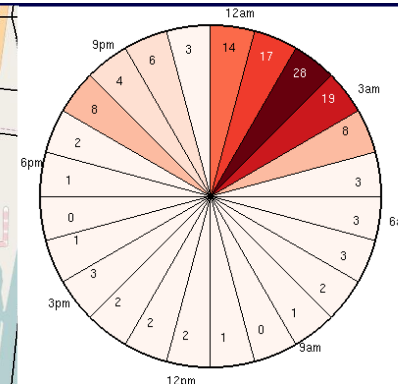
VACCINE

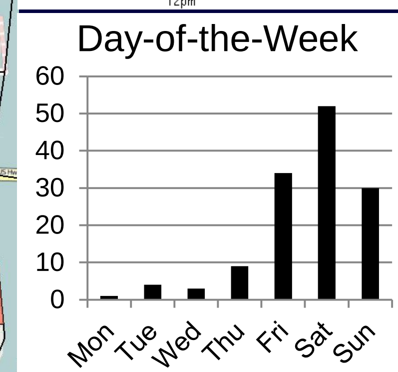

Linked Predictive Crime Models by Type
Crime
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
Max. Days: 7
2005
M T W R F S S Total
Crime Graph
Total Crime
Month #
Current Value: 47
Separate
Total
Show Prediction
MapDrawer
View
Color Scheme
Visualize
Census Data
Crime
Traffic
Clustering
Help
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Show crime points
Show KDE heatmap
Crime Time Slider
Play
Month
Year
10
2005
1/2000
11/2008
Day
Week
Month
Year

Drunkenness / Public Intoxication

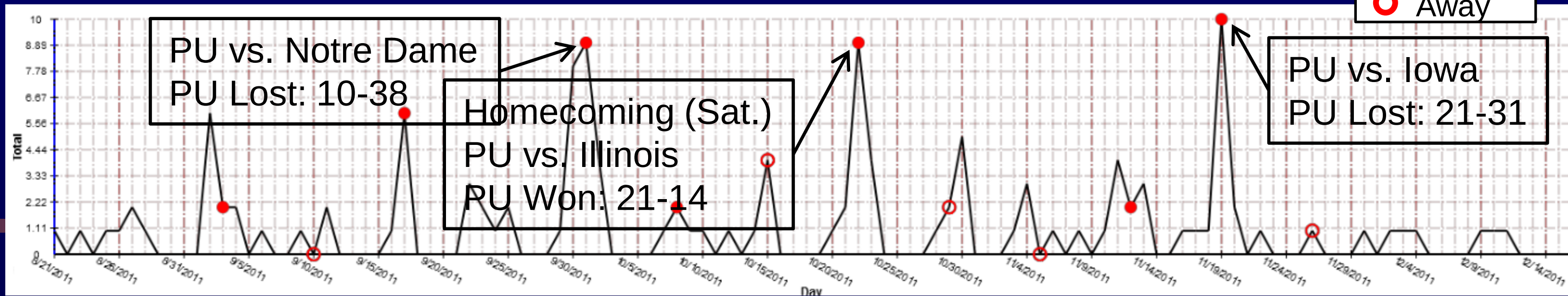


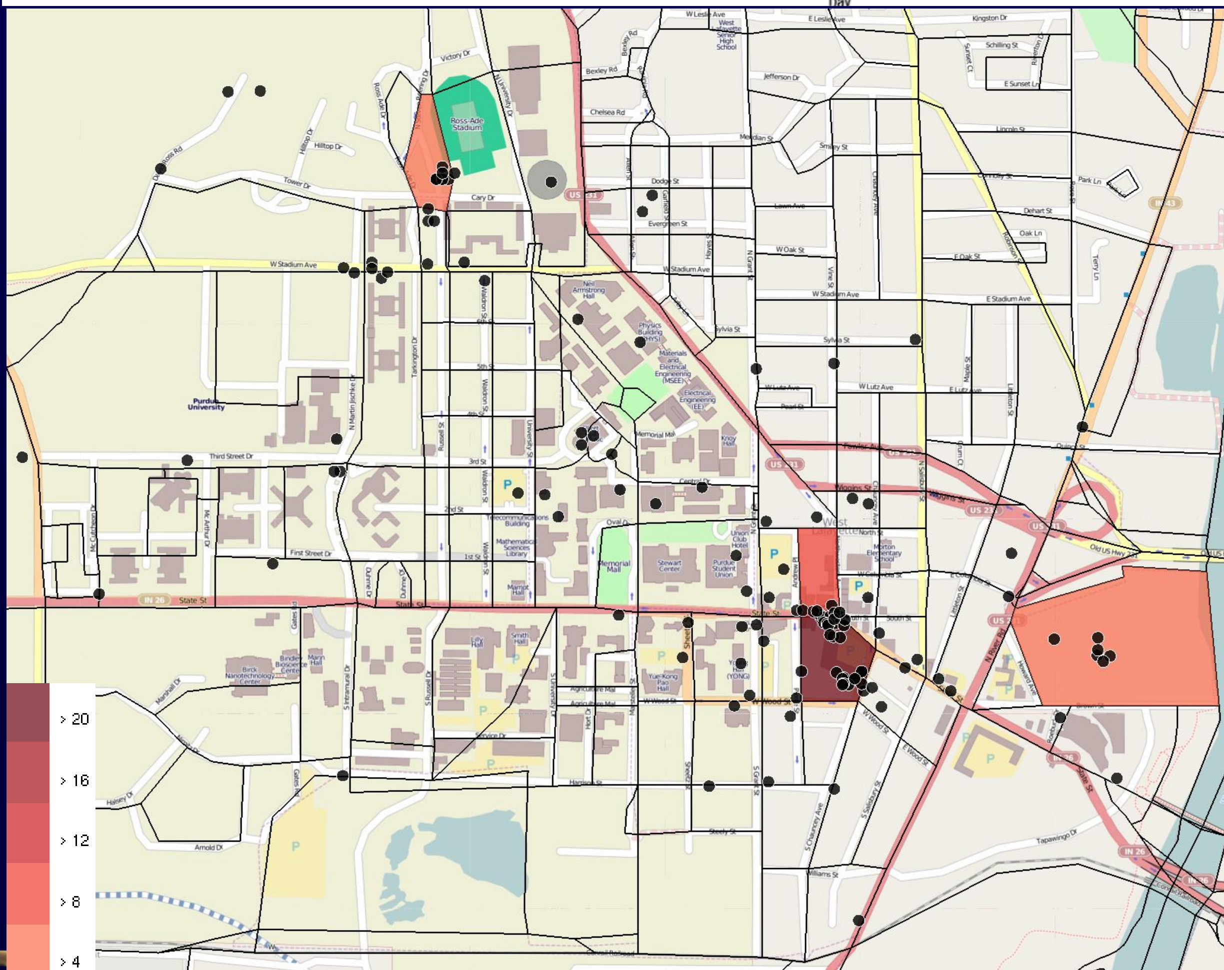


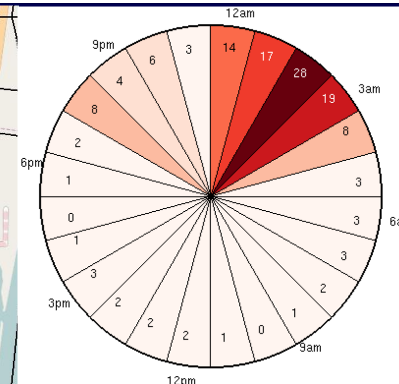


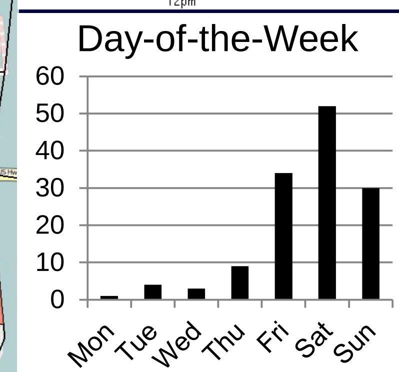


Drunkenness / Public Intoxication









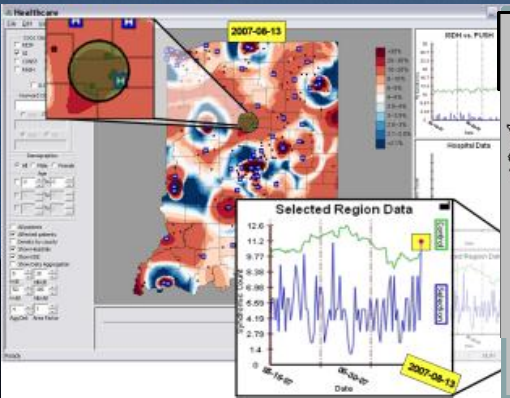
VACCINE

Example: Public Health Visual Analytics

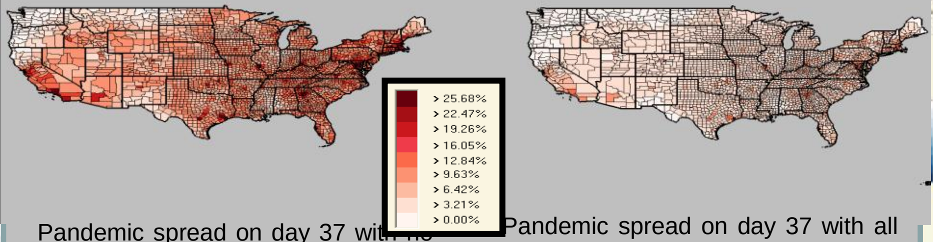
Health Surveillance

Pandemic Simulation and Planning

Zoonoses Planning and Response

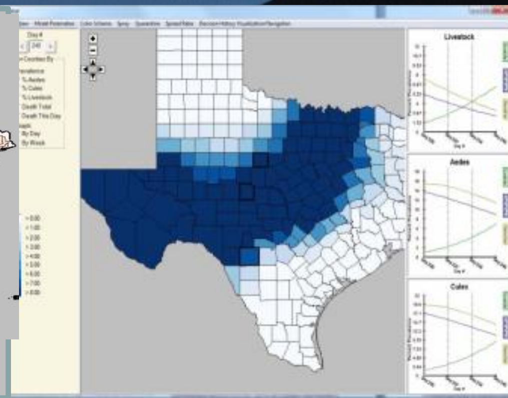


Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.

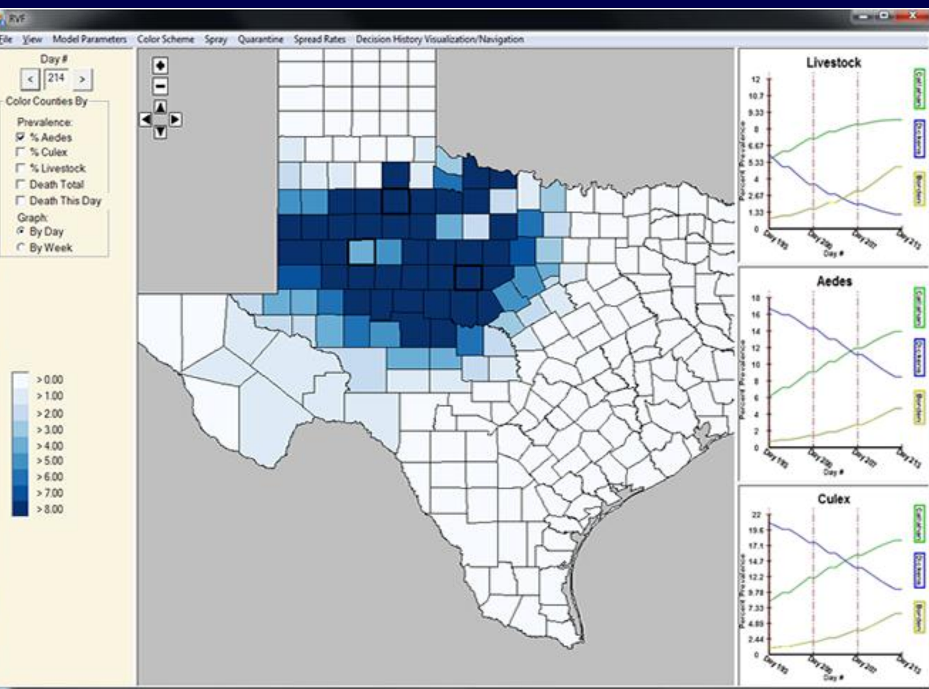


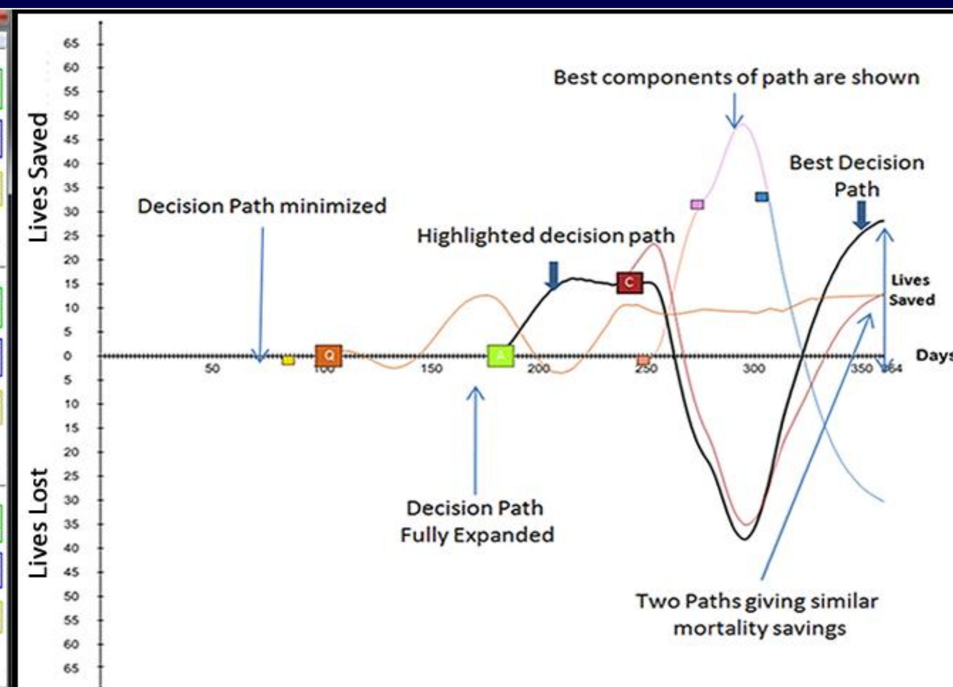
Pandemic spread on day 37 with no decision measures implemented

Pandemic spread on day 37 with all decision measures implemented



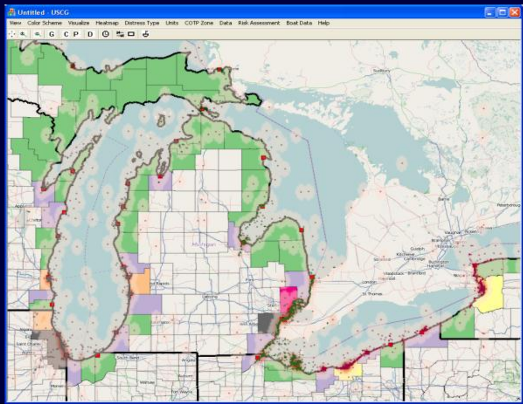
Advanced Decision Support Tools: Rift Valley Fever



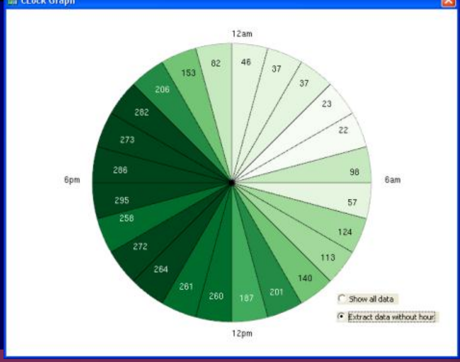


Visual Analytics Uses for Risk-Based Decision Making

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis

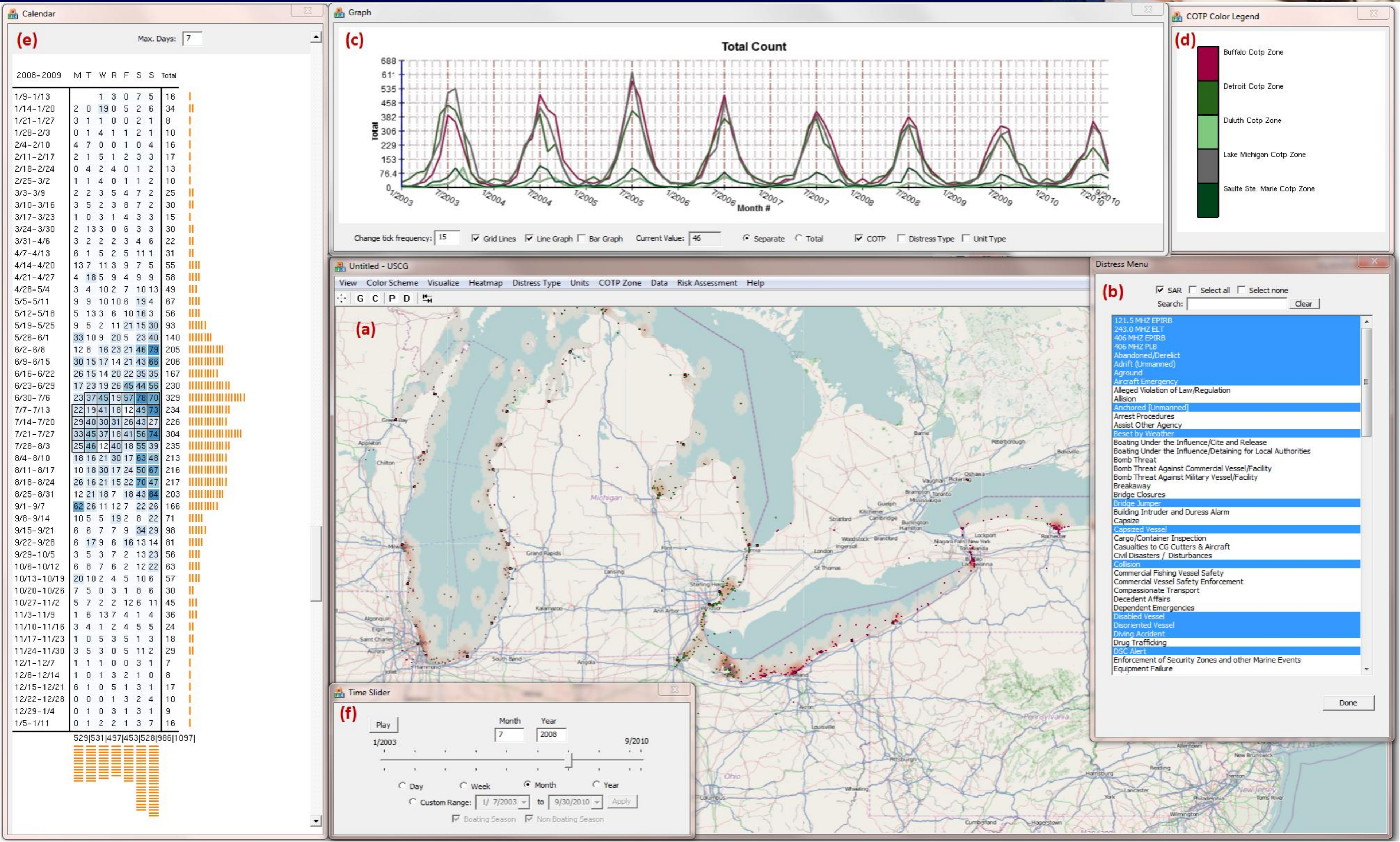


Risk-Based Allocations



- Comparative visual analysis of mission cases/hours vs. staffing hours
- Comparative visualization of resources vs. risk
- Trend visual analytics
  - Increase/decrease in resource allocation
  - Increase/decrease in risk (total, mitigated, residual)
  - Increase/decrease in incidents
- Exploration of alternatives and effect on risk
- Predictive analytics based on historical data (STL and EWMA)

Example: USCG D9 Search And Rescue Operational Analysis

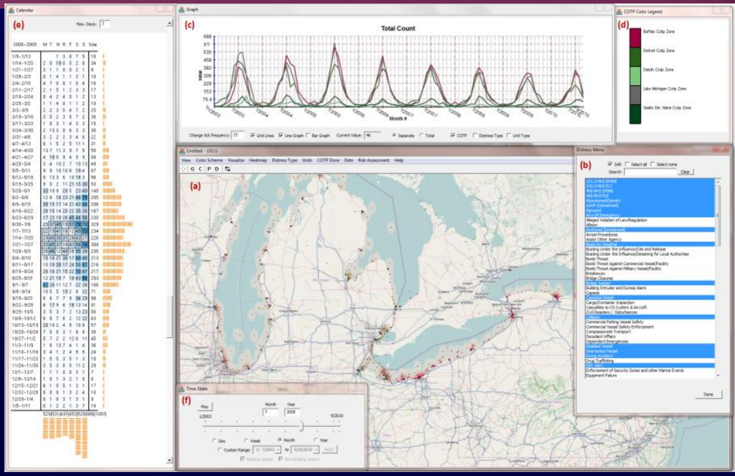


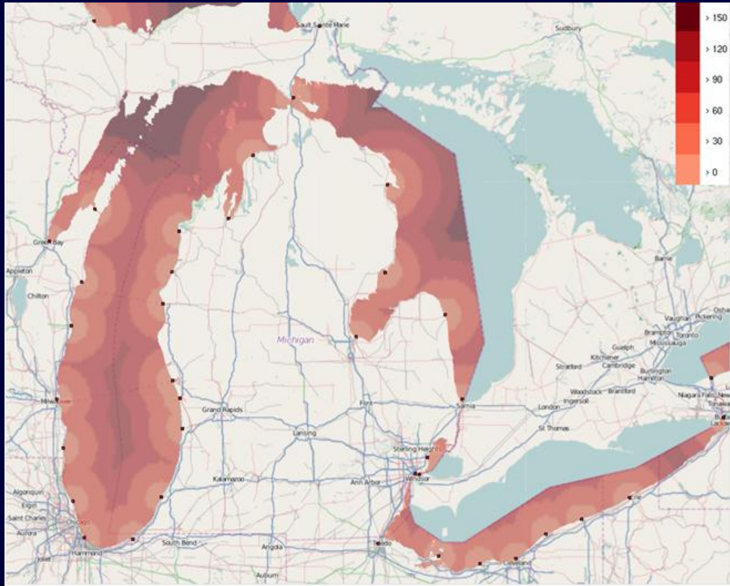
U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7 (Operational Analysis) , USCG D9, USCG D5

IMPACTS:

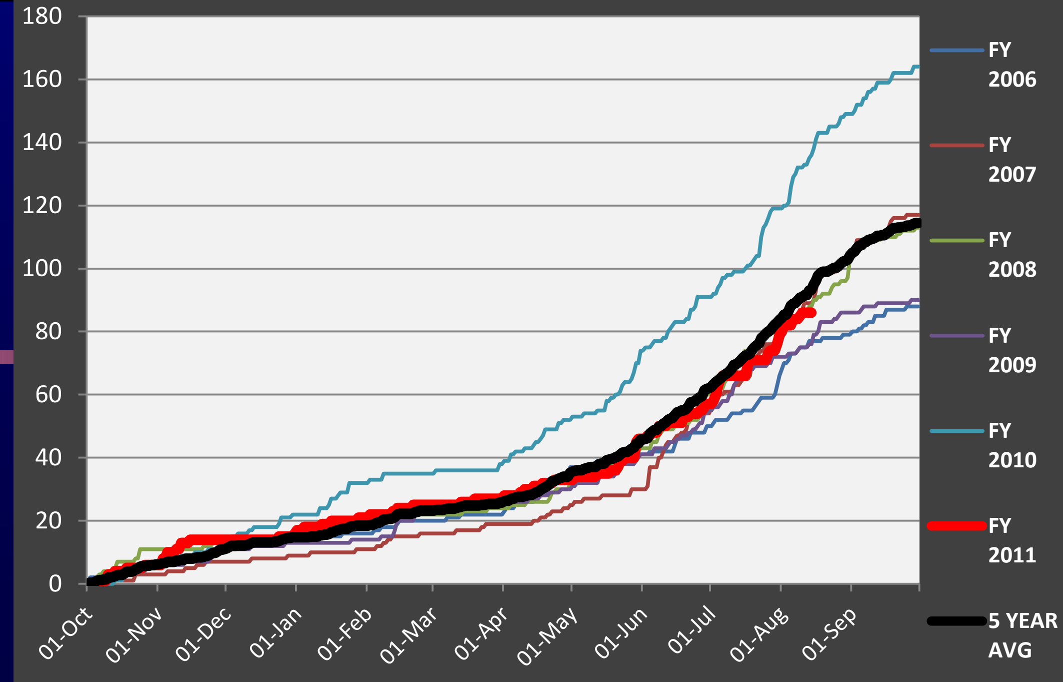
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided evidence of temporal and spatial patterns used in planning – new insights to SAR mission
- Hurricane Irene resource allocation decision based on cgSARva analysis and visualization
    - Highest SAR workload that weekend for D9





VACCINE

U.S. Coast Guard Swimmer Death Analysis

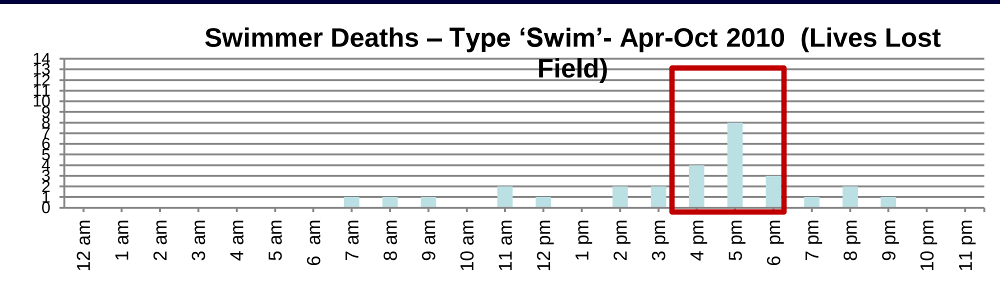


Impact:

- Analyzed spatial and temporal patterns of shore-based and boat-based swimmer deaths to understand death dramatic increase in D9 in Summer 2010
- Provided information and visualizations used for public information campaign 2011 and for patrols 2011
- Significant decrease in deaths in 2011

Findings:

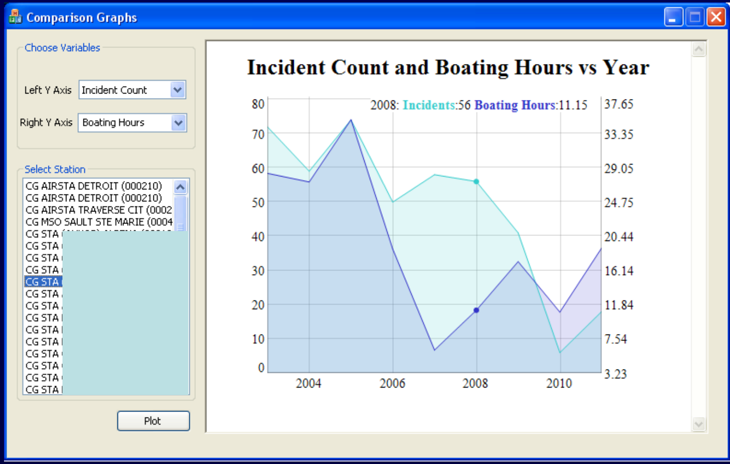
- Swimmer deaths
  - August highest frequency
  - Late afternoon highest frequency
  - Lake Michigan (south and west shore) have high concentration
- Boating deaths
  - Fri, Sat, Sun account for almost all deaths
  - Mid July to Mid August have highest frequency (only 1 week significantly high)
- 2009-2010 from MISLE Data
  - Large increase on Mon, Thu, Fri, Sun
  - Early and late season increase

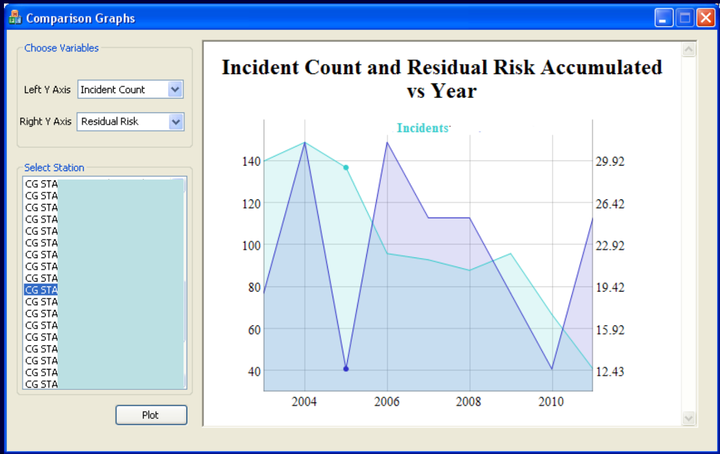


VACCINE

Resource Allocation and Risk-Based Decision Making

- Explore risk-based decision making and utilize historical data for analysis and prediction
  - Total Risk, Mitigated Risk, Residual Risk
  - Explore 11 different USCG missions
  - Explore allocation of assets with different capabilities
  - Explore staffing, utilization, assets vs. risk measures
  - Perform What-If scenarios





VACCINE

USCG Port Closure Economic Impact VA

Partners: USC CREATE, USCG RDC, USCG D7, USCG LANT

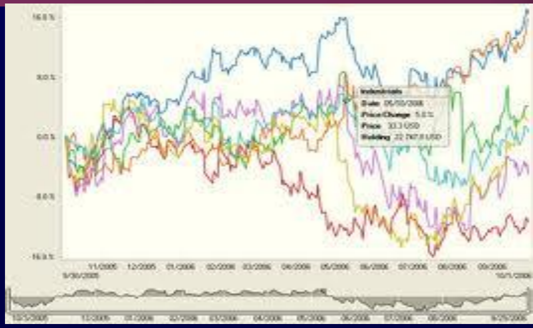
IMPACT:

- Provided tool for use analysis and planning for impact of port closure in Port Arthur, Tx
- Economic sector impact, local and national impact
- Impact and effectiveness of alternative mitigation strategies

VACCINE

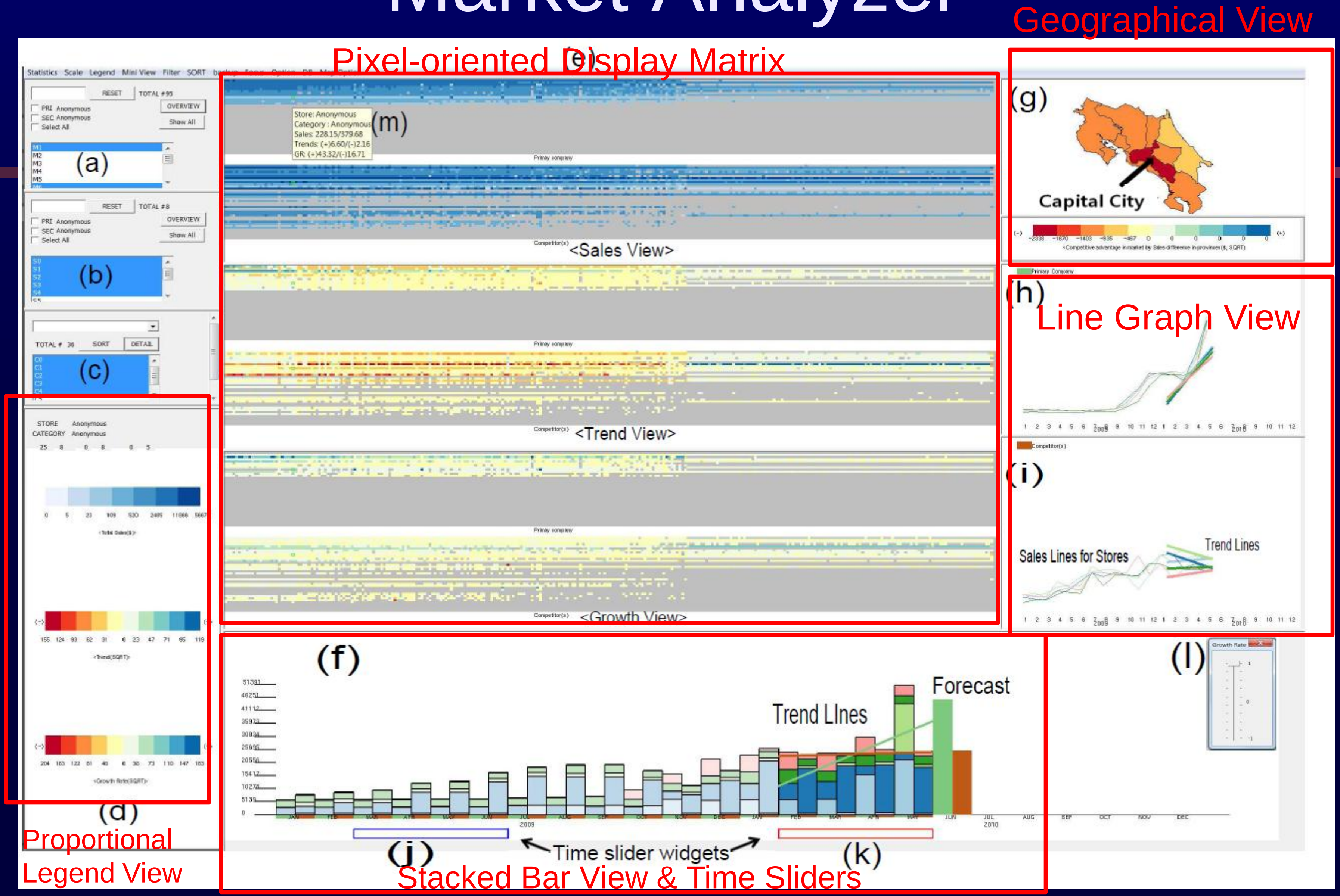

Financial Risk VA

Competitive Business Intelligence Based On Point of Sales Data



- Characteristics of point of sale data
  - Multivariate
    - Large # of dimensions
      - E.g., 38 categories(products) in 288 stores
  - Temporal : 18 months
  - Spatial
    - Stores located all over Costa Rica
  - Requirements
    - Supporting easy comparison among companies
      - E.g., visualize all data at once, sorting by importance
      - Enable geographical comparison
    - Easy recognition of any change in sales
      - E.g., proportional legends
    - Forecast for decision-making

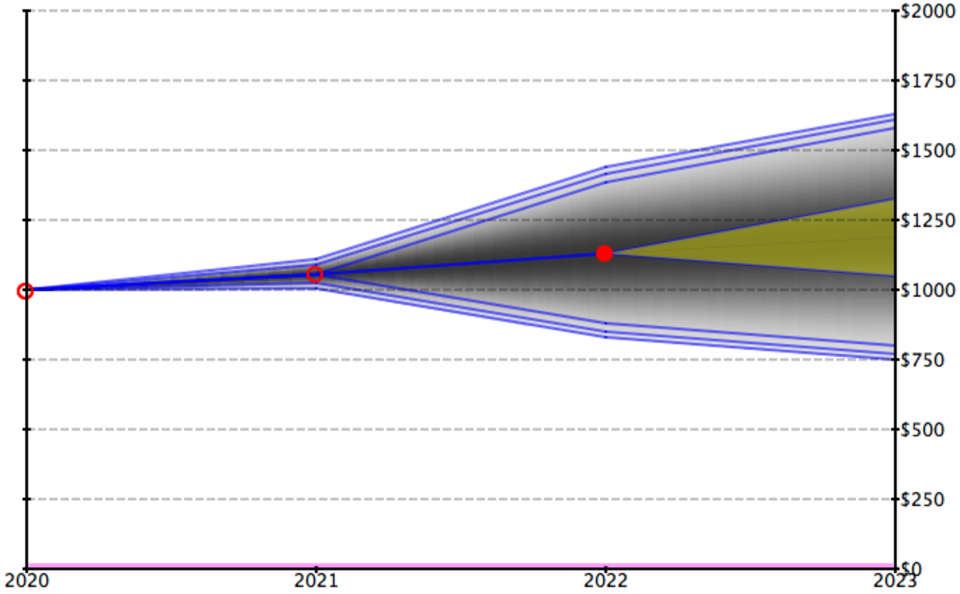
Market Analyzer



VACCINE

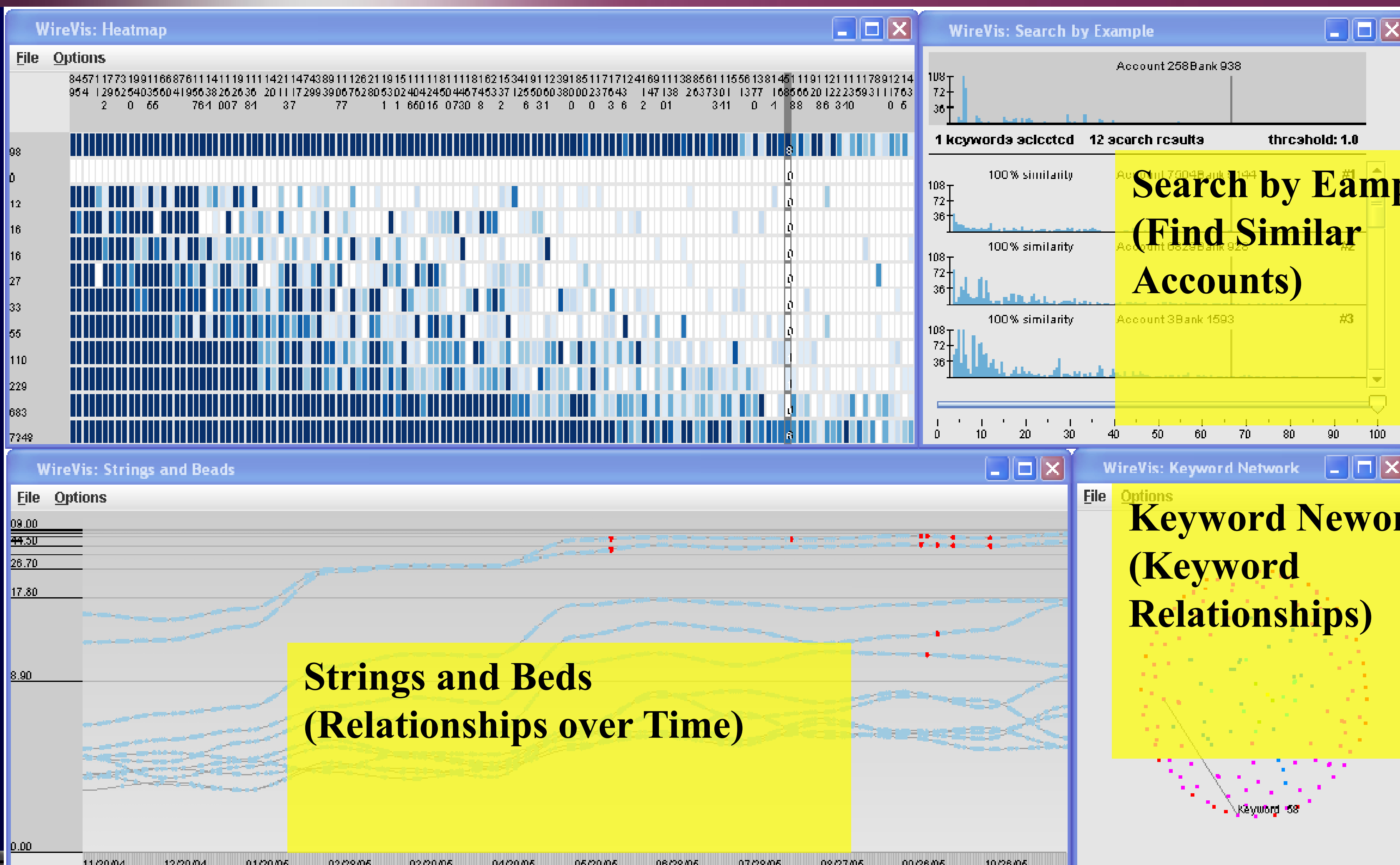
FinVis: Applied Visual Analytics for Personal Financial Planning

- First visual analytics work to address portfolio planning for casual users
  - *General public could benefit from this tool*
- Decisions are better when using FinVis
  - *Quantified using experimental economics*
  - *Improved*
    - *Decisions*
    - *Learning*
    - *exploration,*
    - *confidence*



WireVis – Streaming Data - Multiple Linked Views

- Temporal, geospatial, theme, cluster, list views with association linkages between views



Path

Introduction to our Center - VACCINE

Motivation and introduction to visual analytics

Example solutions we’ve developed

- Health visual analytics
- Public safety
- Resource allocation

Visual analytic challenges and future

~~Five~~ Six Challenges for Proactive & Predictive Visual Analytics

1. Creating computer-human visual cognition environments
2. Integrating interactive simulation and analytics
3. Solving specific scale issues and cross-scale issues
4. Managing uncertainty and time
5. Enabling risk-based decision making environments
6. Developing the Science of Interaction for Visual Analytics

Table of Contents

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu

ebertd@purdue.edu

VACCINE

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

Table of Contents

Visual Analytics: Powering Discovery, Innovation, and Decision Making

(Much more than Big Data Analytics plus Visualization)

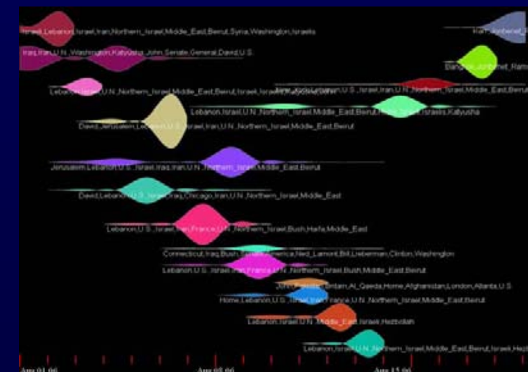
David S. Ebert
ebertd@purdue.edu

Path

- Introduction to our Center - VACCINE
- Motivation and introduction to visual analytics
- Example solutions we’ve developed
  - Health visual analytics
  - Maritime public safety and law enforcement
  - Public safety
  - Resource allocation
- Visual analytic challenges and future


SFU, JIBC
UBC
UW
Stanford
Purdue
VACCINE
Penn St.
Ind U
VaTech
NC A&T
UNCC
Navajo Tech
JSU
GaTech
U Stuttgart
Swansea U
UT Austin
UHD
FIU
GeoDiscoverer
Selected Region Data

VACCINE Mission



- Provide visual analytic and scalable solutions for all 2.3 million extended homeland security personnel
  - 185,000 DHS personnel, 350,000 law enforcement personnel, 750,000 homeland security practitioners, 2.3 million extended personnel

- Achieve excellence in visual analytics and visualization sciences

- Educate homeland security stakeholders and the next generation of talent

VACCINE

Who We Are:
International Team of Experts
- 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

Some VACCINE Regional and Corporate Partners

- Indiana Department of Homeland Security
- Indiana State Department of Health
- Indianapolis Public Safety
- Charlotte Mecklenburg Police Department
- Lafayette, West Lafayette, Purdue Police Departments
- Tippecanoe County EMA & Sherriff's Department
- Harrisburg, PA Police
- Coast Guard Sector Boston, Seattle, LA
- Joint Harbor Operations Command Center
- Port of Seattle
- Indianapolis Public Safety
- USCG District 9, D5, D1
- Motorola
- Harris Corporation
- Boeing
- Next Wave Systems, LLC
- Banfield, The Pet Hospital
- Raytheon
- MacDonald, Dettwiler and Associates
- Oculus Info Inc.
- Kx Systems
- Bank of America
- Duke Energy
- World Vision International
- Gates Foundation
- Kimberly Clark
- Hallmark

Some VACCINE Government Partners

- U.S. Coast Guard
- Federal Emergency Management Agency
- Customs and Border Patrol
- National Geospatial Intelligence Agency
- National Science Foundation
- Army Research Office
- Department of Defense
- Department of Health and Human Services
- U.S. Dept. of State
- National Institute of Justice
- National Maritime Intelligence Center
- Defence Research & Development Canada
- Foreign Broadcast Information Service
- US-CERT
- Oak Ridge National Laboratories
- US Army Corps of Engineers-ERDC
- Pacific Northwest National Laboratory
- ARL CERDEC & ARDEC
- US Attorney's Office (David Capp)

What We Do

- Enable effective decision making through interactive visual analytic environments
- Enable effective communication of information
- Provide quantitative, reliable, reproducible evidence
- Enable user to be more effective from planning to detection to response to recovery
- Enable proactive and predictive visual analytics
- Enable effective situational awareness (perception, comprehension, projection)

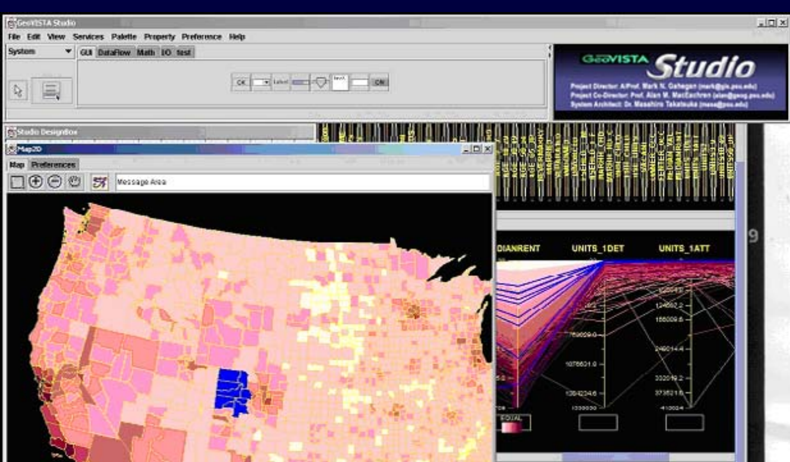
VACCINE

Path

- Introduction to our Center - VACCINE
- Motivation and introduction to visual analytics
- Example solutions we’ve developed
  - Health visual analytics
  - Maritime public safety and law enforcement
  - Public safety
  - Resource allocation
- Visual analytic challenges and future

Motivation

To solve today’s and tomorrow’s problems requires exploring, analyzing, and reasoning with massive, multisource, multiscale, heterogeneous, streaming data



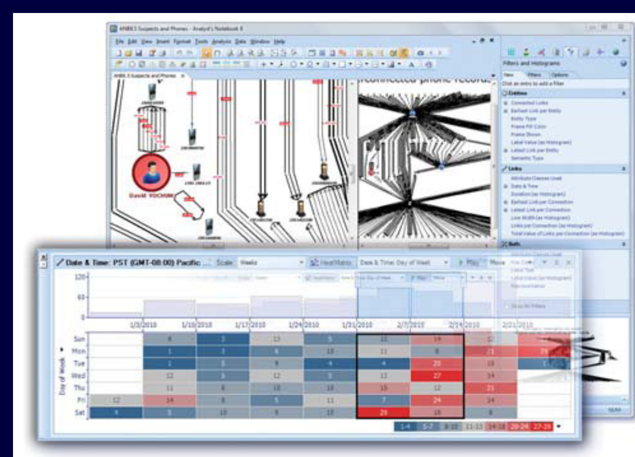
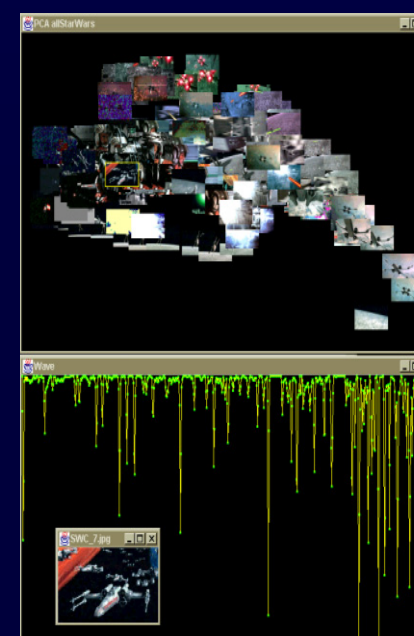


Image of Analyst’s Notebook



VACCINE

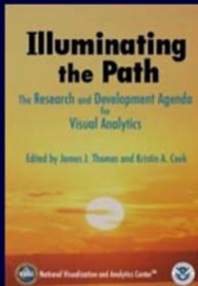
Definition

Visual Analytics[1] is the science of analytical reasoning facilitated by interactive visual interfaces

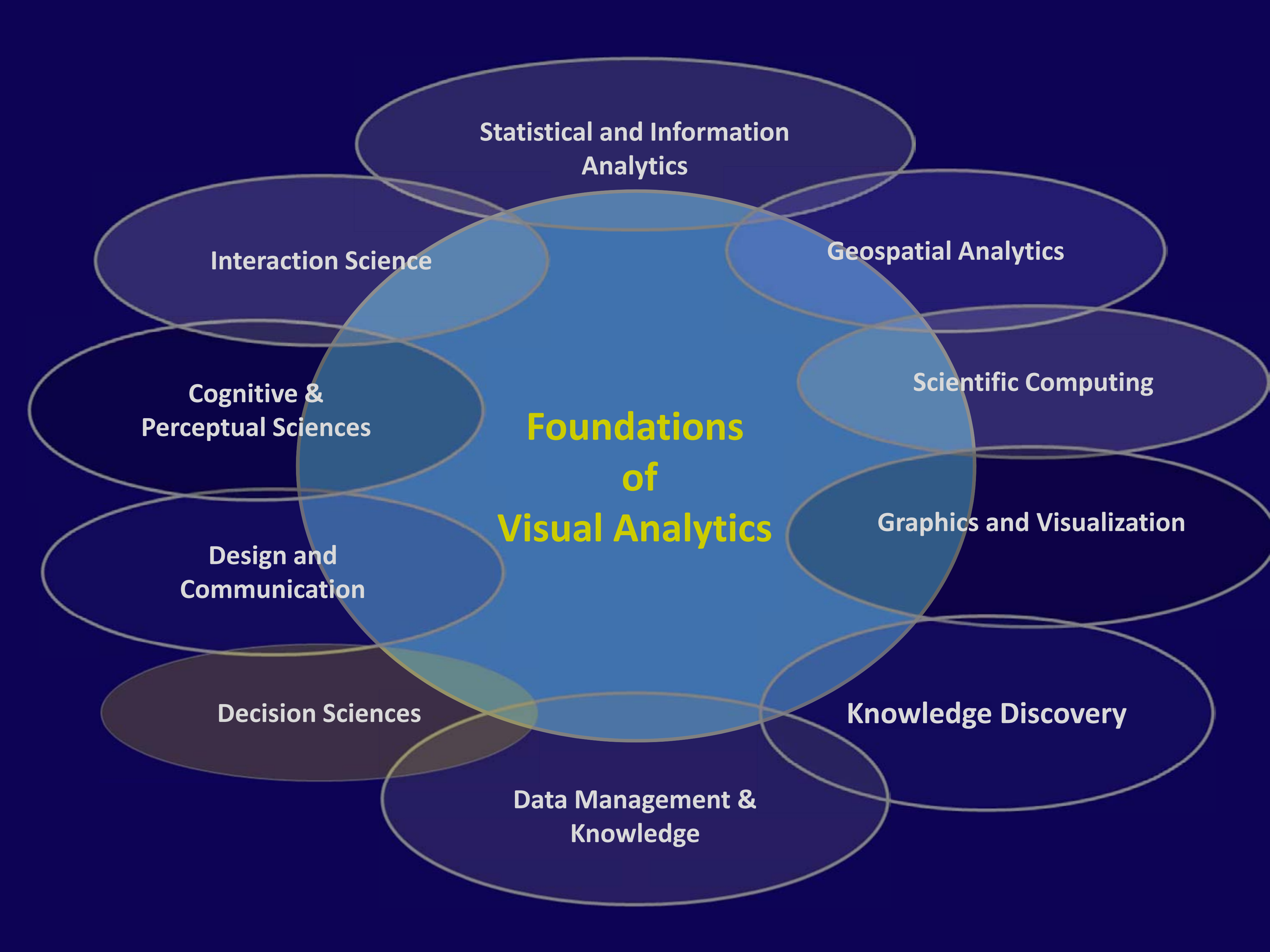
People use visual analytics tools and techniques to

- Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
- Detect the expected and discover the unexpected
- Provide timely, defensible, understandable assessments
- Communicate assessment effectively for action

1. *Illuminating the Path: The R&D Agenda for Visual Analytics, Editors: Thomas and Cook*



VACCINE


Statistical and Information Analytics
Interaction Science
Geospatial Analytics
Scientific Computing
Cognitive & Perceptual Sciences
Foundations of Visual Analytics
Graphics and Visualization
Design and Communication
Decision Sciences
Knowledge Discovery
Data Management & Knowledge

What's Needed for Proactive and Predictive Visual Analytics?

- Reliable and reproducible models and simulation
- Understanding of the data
  - Distribution and skewness, errors, appropriate analysis techniques
- Understanding of the sources and types of data
- Comparable or Correlative sources of data
  - Appropriate transformations applied to enable meaningful comparison and correlation
- Understanding of the use and problem to be solved!

How to Harness Visual Representations

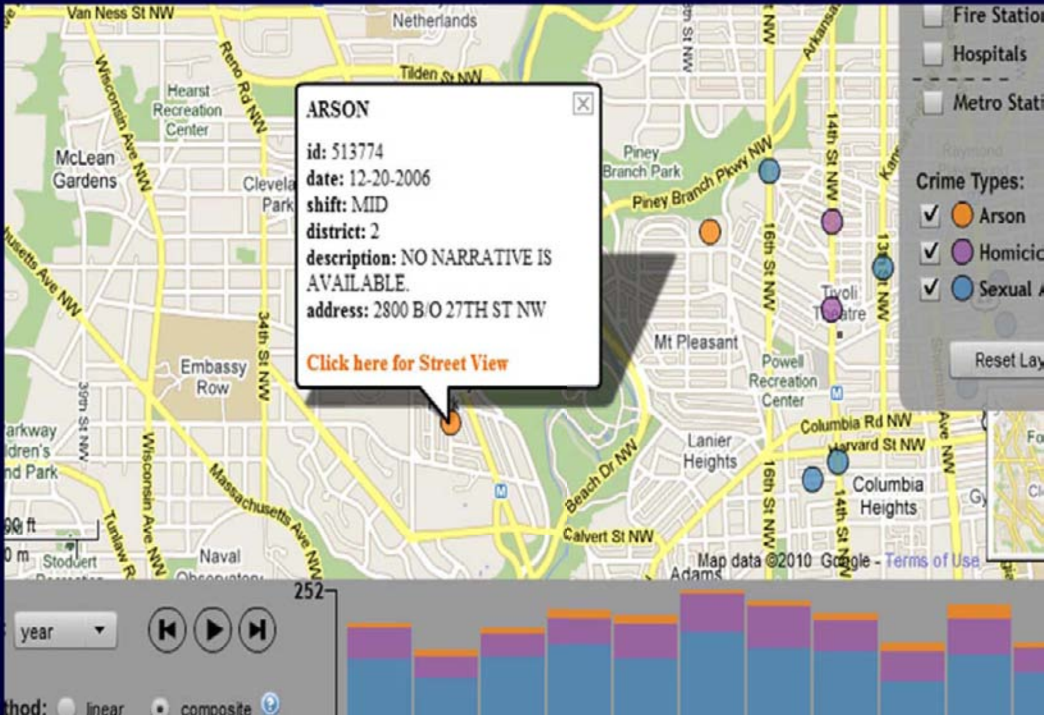
- Visual representations translate data into a visible form, highlighting important features, such as commonalities and anomalies
- Visual representations make it easy for users to perceive salient aspects of their data quickly
- These visual representations augment the cognitive reasoning process with perceptual reasoning which enhances the analytical reasoning process

1 – J. J. Thomas and K. A. Cook (Ed.). *Illuminating the Path: The R&D Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.

VACCINE

Why Was I Invited?
Visual Analytics Uses
for Crisis Management and Response?

- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis
- Purpose:
  - Planning for resiliency
  - Long-term analysis
  - Predictive analytics
  - Training
  - Detection
  - Investigation
  - Response
  - Recovery, remediation



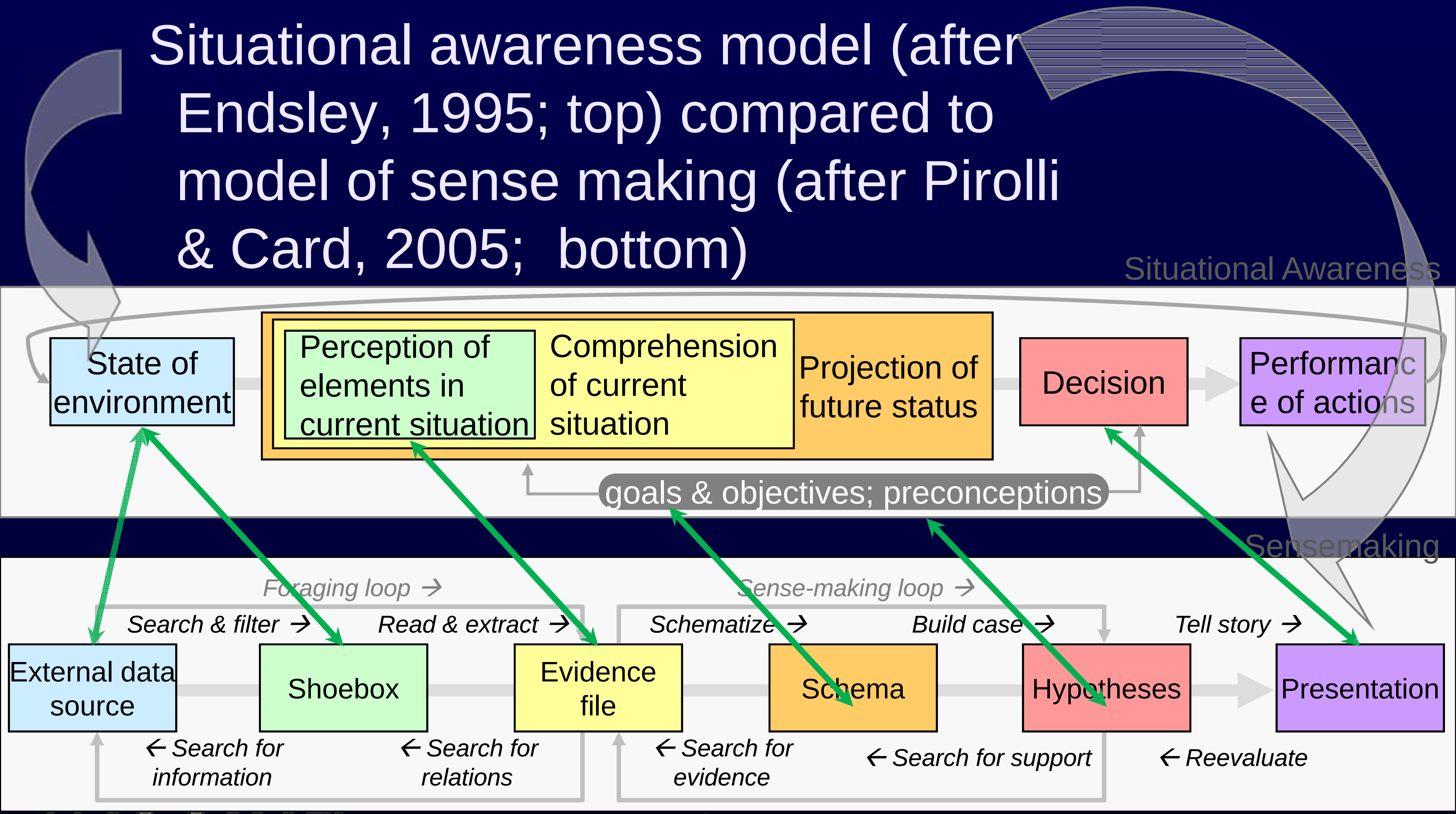
VACCINE

Visual Analytics vs. Situational Awareness

Situational Awareness ←→ Sensemaking

situational awareness as a sensemaking process (courtesy Alan MacEachren)

- Situational awareness model (after Endsley, 1995; top) compared to model of sense making (after Pirolli & Card, 2005; bottom)



Path

Introduction to our Center - VACCINE

Motivation and introduction to visual analytics

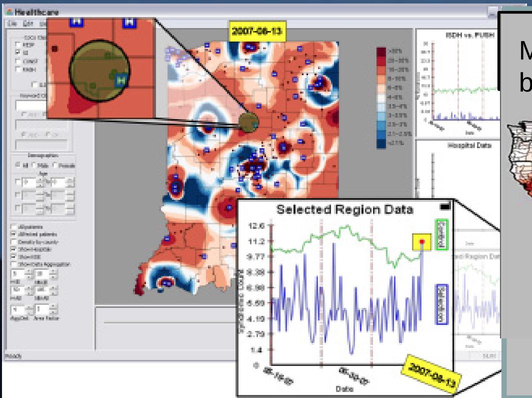
Example solutions we've developed

- Health visual analytics
- Maritime public safety and law enforcement
- Public safety
- Resource allocation and business intelligence

Visual analytic challenges and future

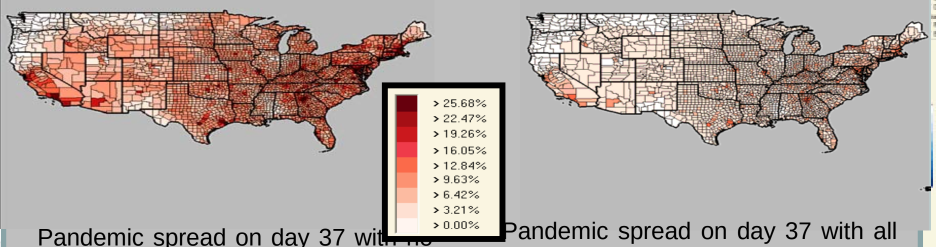
Example: Public Health Visual Analytics

Health Surveillance



Pandemic Simulation and Planning

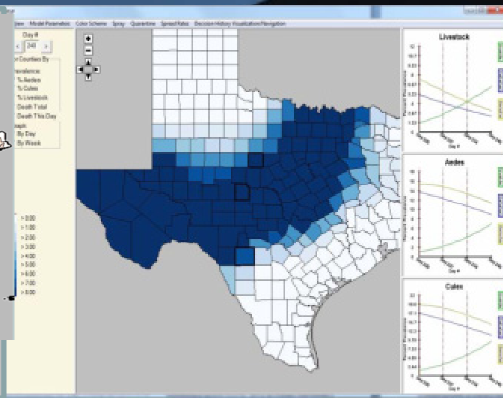
Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.



Pandemic spread on day 37 with no decision measures implemented

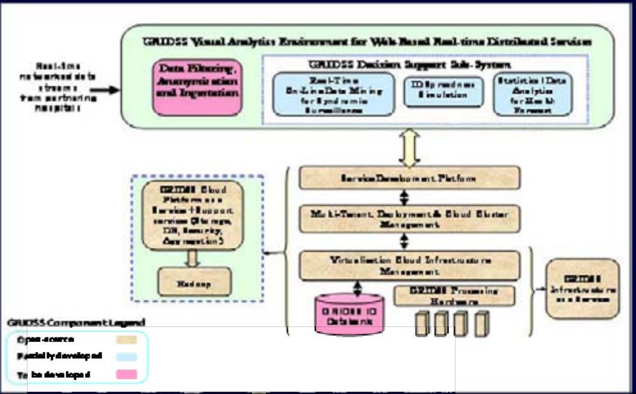
Pandemic spread on day 37 with all decision measures implemented

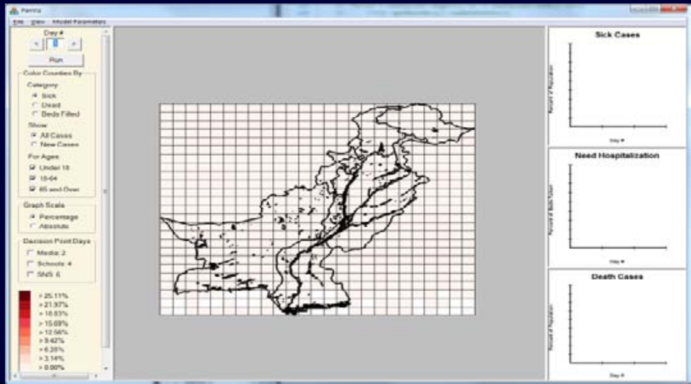
Zoonoses Planning and Response



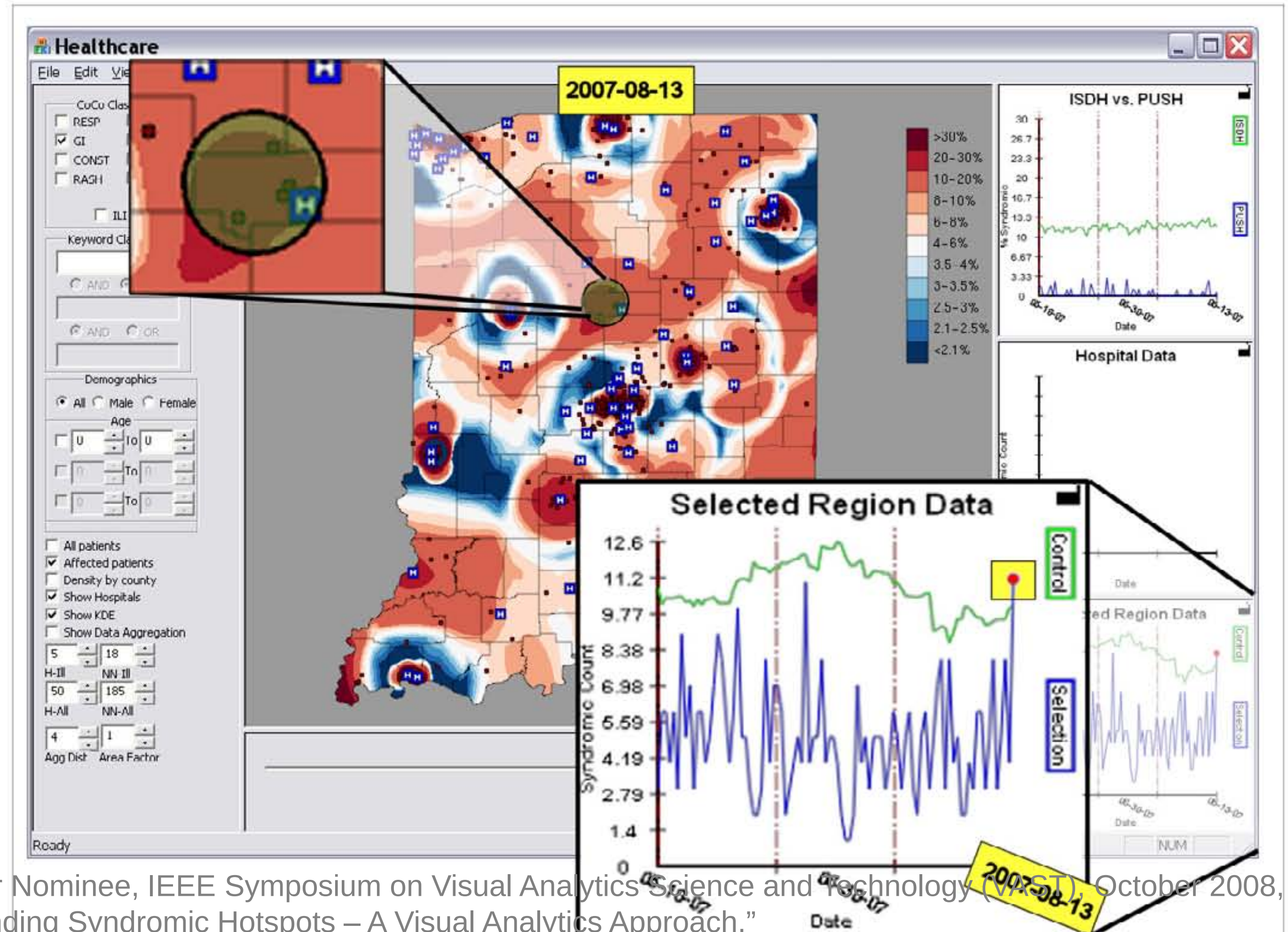
Visual Analytics For Syndromic Surveillance

- Syndromic surveillance: detection of adverse health events focusing on pre-diagnosis information to improve response time
- Pre-diagnosis information can consist of multiple data sources:
  - Over the counter medicine sales
  - News reports on emerging diseases
  - Pro-med news feeds
  - Emergency department chief complaints





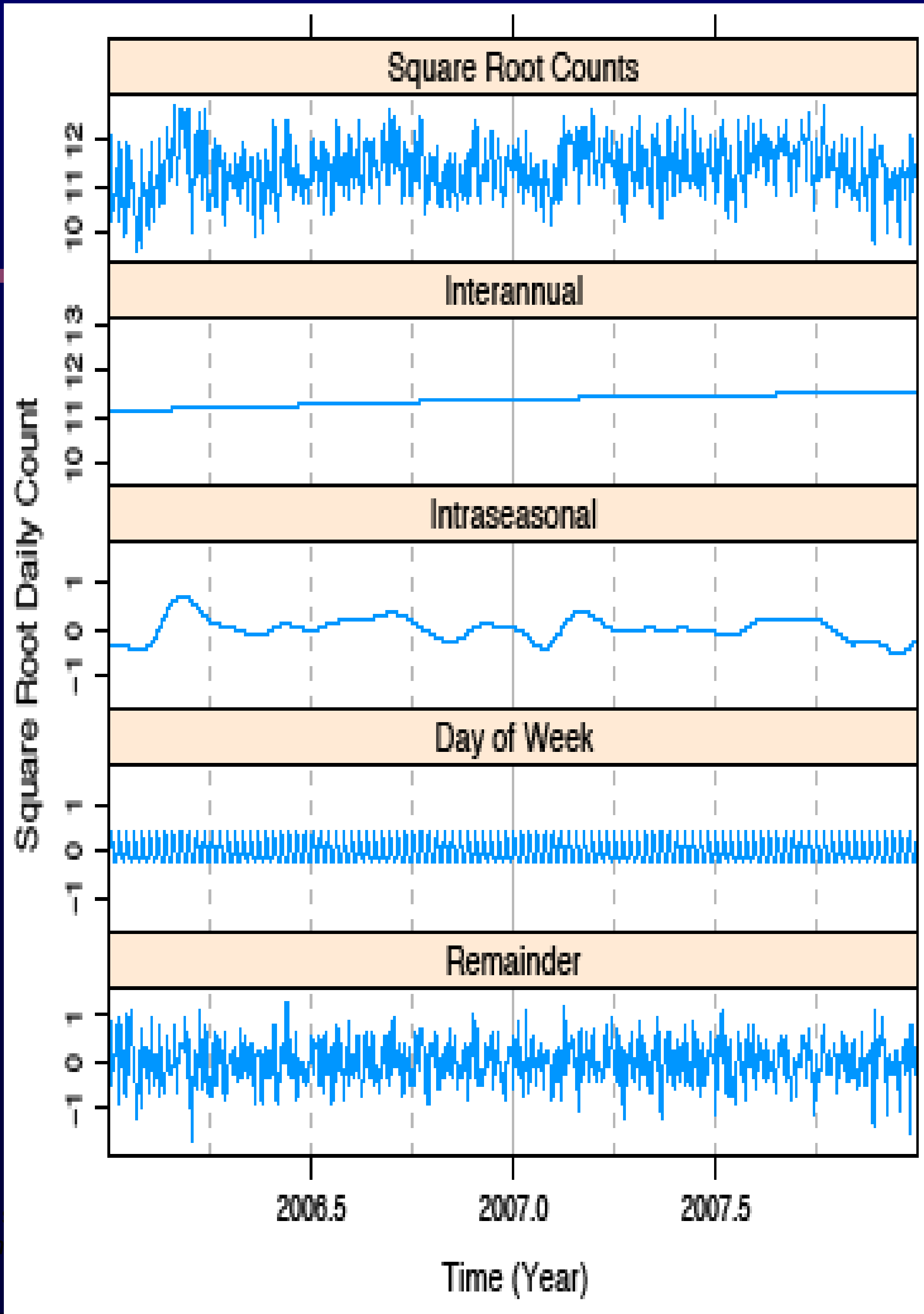
Hypothesis Generation and Exploration



Best Paper Nominee, IEEE Symposium on Visual Analytics Science and Technology (VAST), October 2008, for "Understanding Syndromic Hotspots – A Visual Analytics Approach," (Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W., Grannis, S., Wade, M., Ebert, D.)

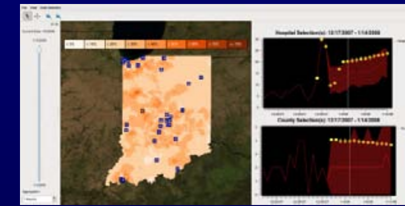
Time Series Modeling

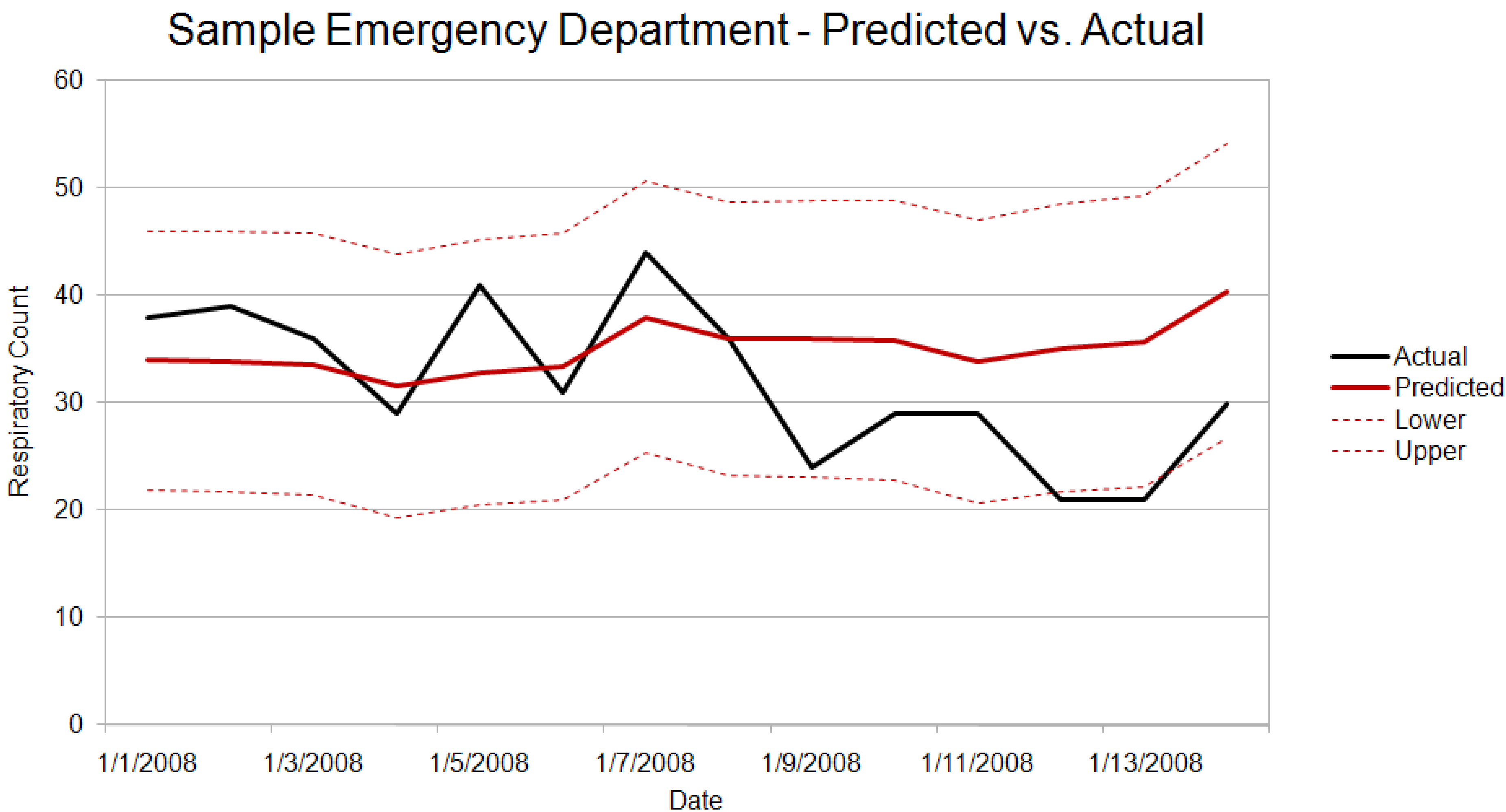
- Seasonal-Trend Decomposition Based on Loess
  - Time series can be viewed as the sum of multiple trend components
  - For each data signal, components are extracted
  - Can then analyze correlation between components



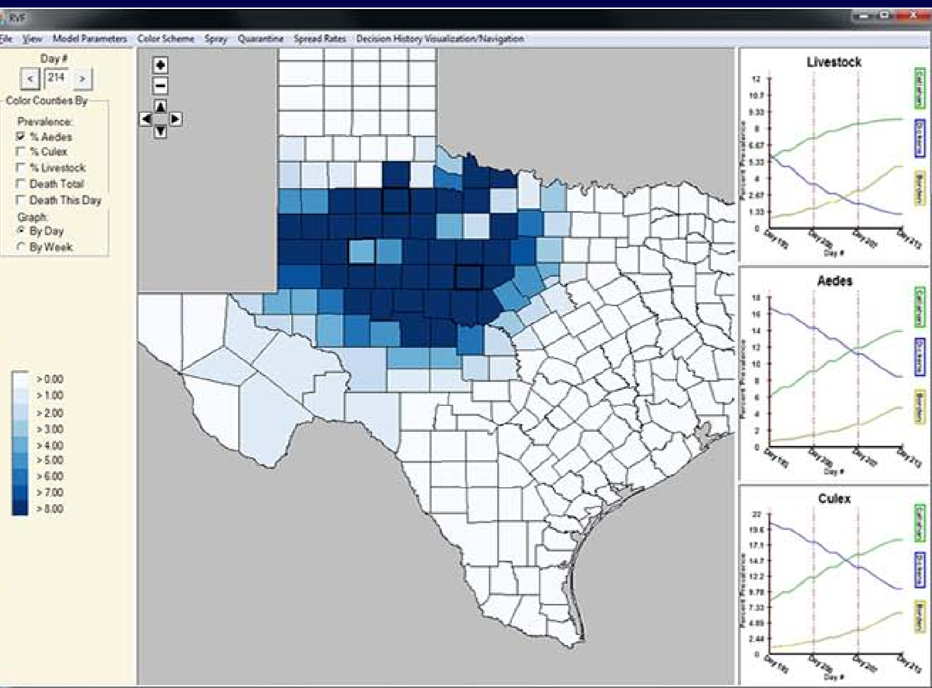
Hafen, R., Anderson, D., Cleveland, W., Maciejewski, R., Ebert, D., Abu
Modeling, Visualizing, and Monitoring Disease Counts," *BMC Medical In*

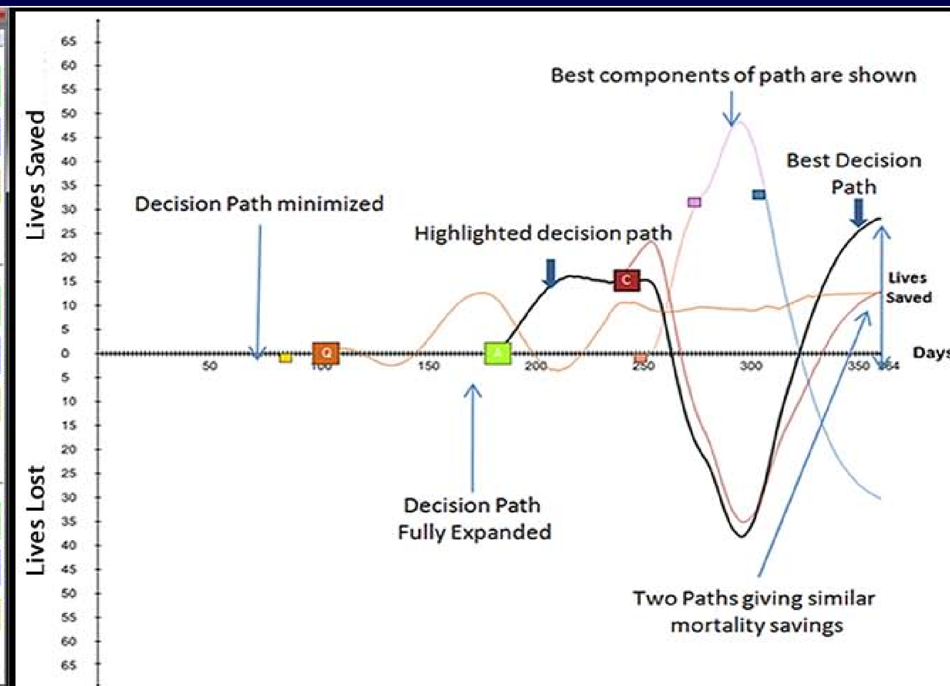
Predictive Visual Analytics





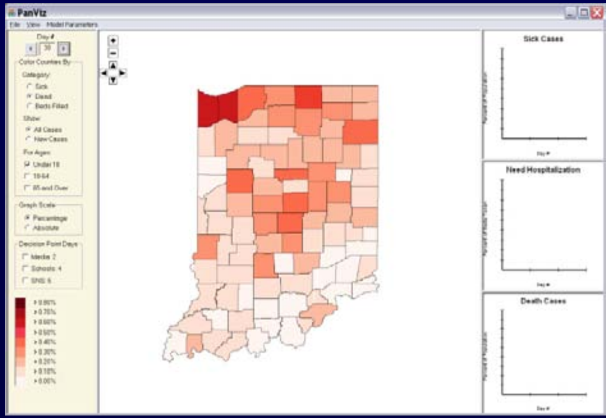
Advanced Decision Support Tools: Rift Valley Fever



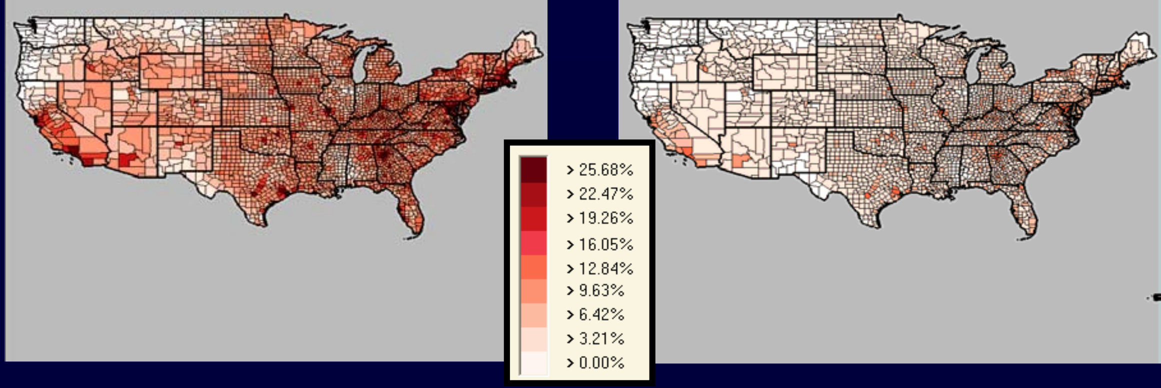


Modeling a Pandemic

- Pandemic Influenza Planning Tool
- Models user specified:
  - Pandemic influenza characteristics
  - County population, demographics, hospital beds
- Decision measures
  - Strategic National Stockpile deployment
  - School Closures
  - Media Alerts



Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.

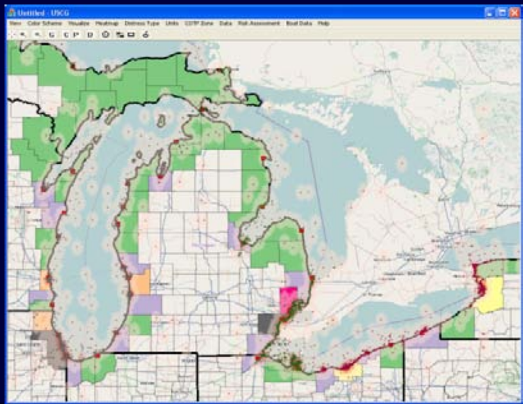


Pandemic spread on day 37 with no decision measures implemented

Pandemic spread on day 37 with all decision measures implemented

Visual Analytics Uses for Public Safety

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



VACCINE

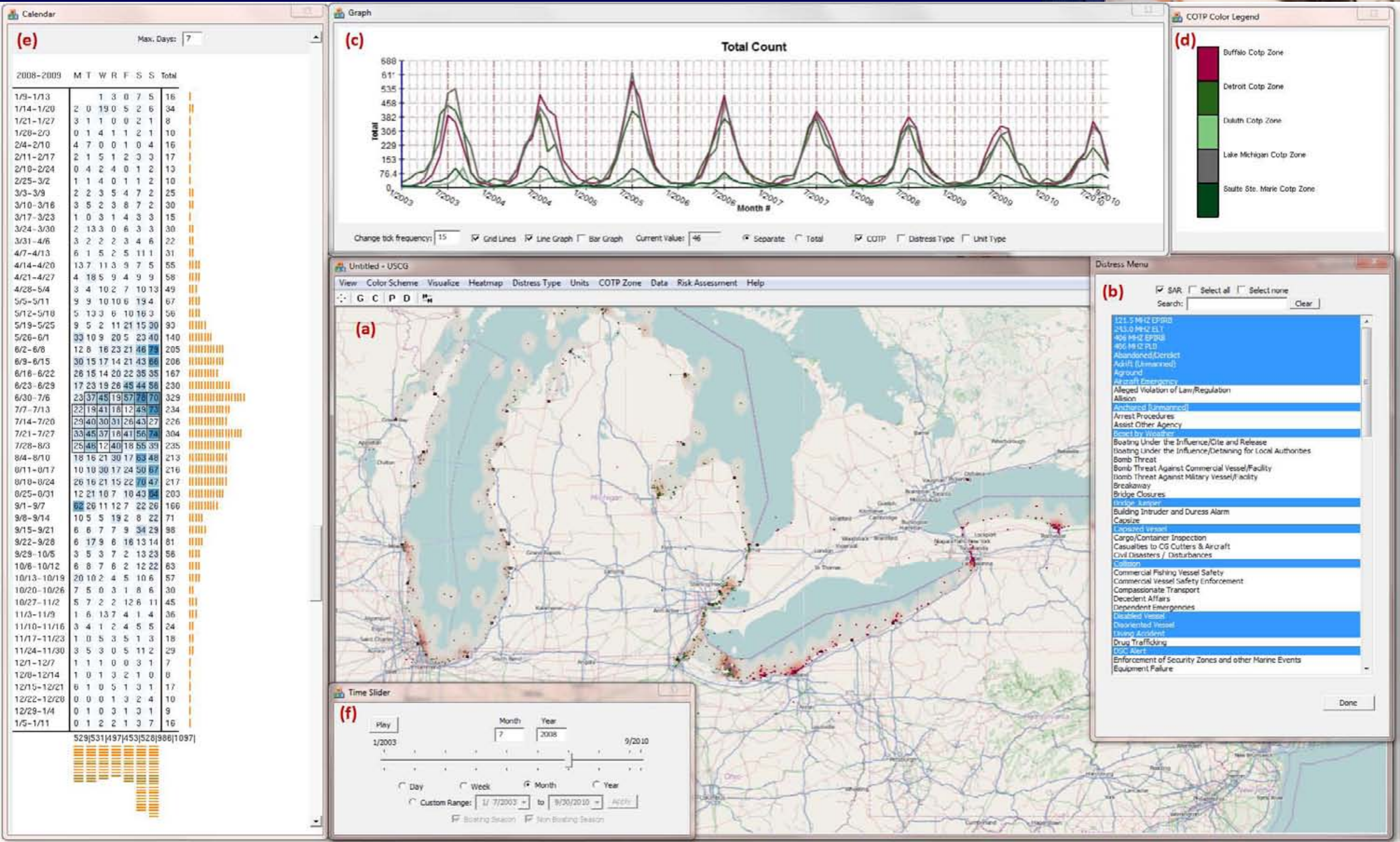
Example Maritime Public Safety Projects

- Search and rescue resource allocation
- Swimmer death analysis
- PWCS analysis
- Economic impact analysis
- Resource allocation and risk-based decision making

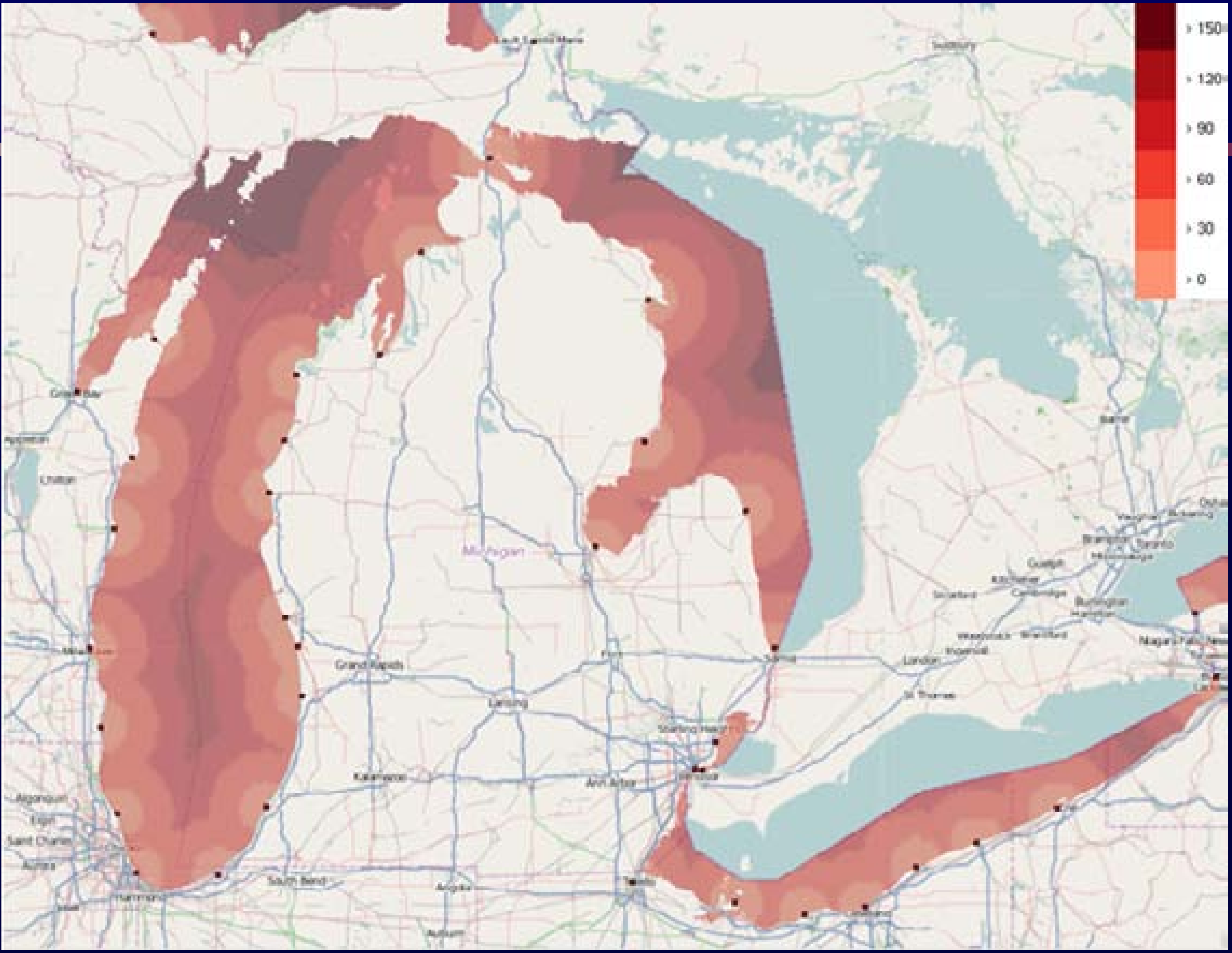
Example: Visual Analytics Environment

- Supports decision making and risk assessment
- Interactive exploration and analysis of trends, patterns and anomalies
- Allows analysis of risks associated with
  - Closing one or more Coast Guard stations
    - Find optimal stations that absorb work load of the closing station
  - Allocating new resources
    - Impact on safety and efficiency of operation
- Currently being used by analysts at the U.S. Ninth District, HQ, and Atlantic Commands

Example: USCG D9 Search And Rescue Operational Analysis



System Features: SAR Risk Profile



Time taken by CG stations to deploy an asset to the Great Lakes to respond to a SAR incident.

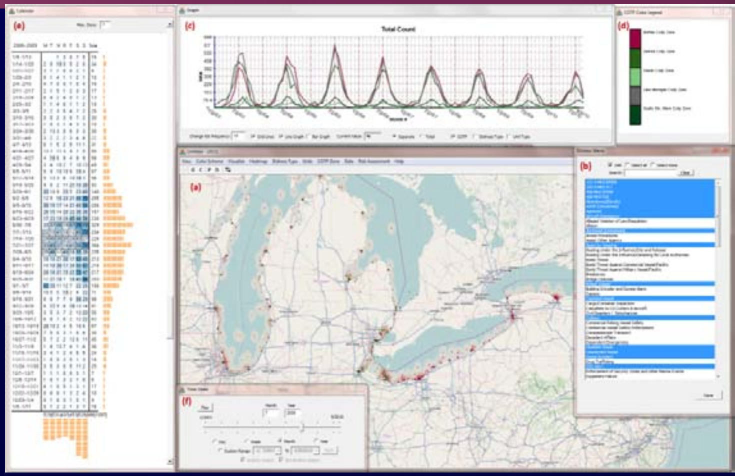
VACCINE

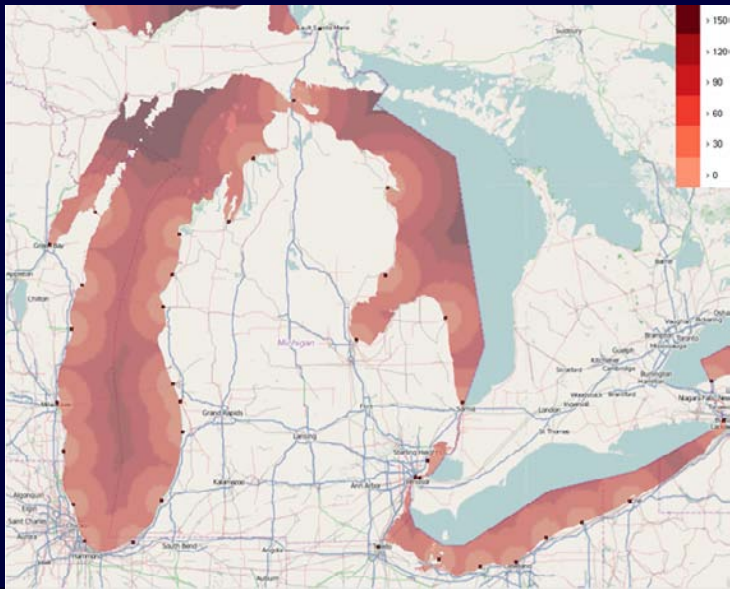
U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7 (Operational Analysis) , USCG D9, USCG D5

IMPACTS:

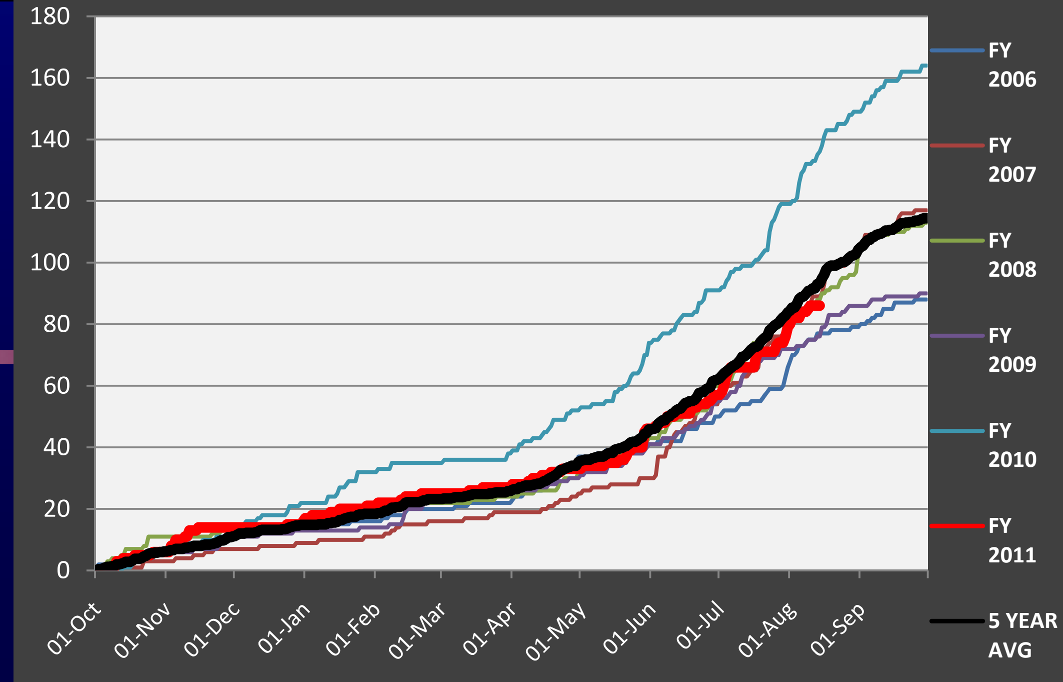
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided evidence of temporal and spatial patterns used in planning – new insights to SAR mission
- Hurricane Irene resource allocation decision based on cgSARva analysis and visualization
  - Highest SAR workload that weekend for D9





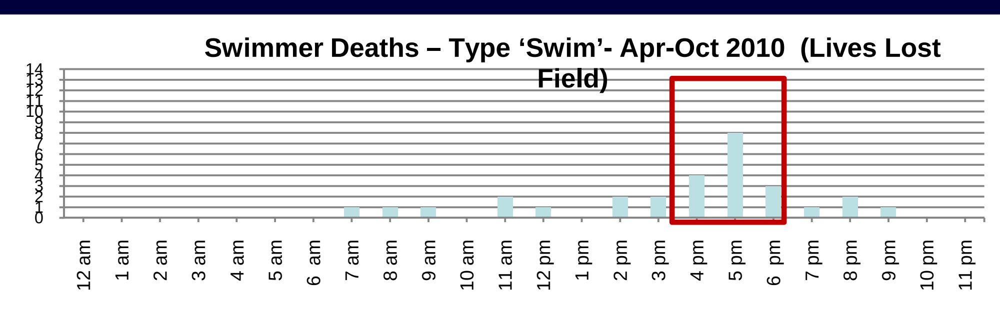
VACCINE

U.S. Coast Guard Swimmer Death Analysis



Impact:

- Analyzed spatial and temporal patterns of shore-based and boat-based swimmer deaths to understand death dramatic increase in D9 in Summer 2010
- Provided information and visualizations used for public information campaign 2011 and for patrols 2011
- Significant decrease in deaths in 2011



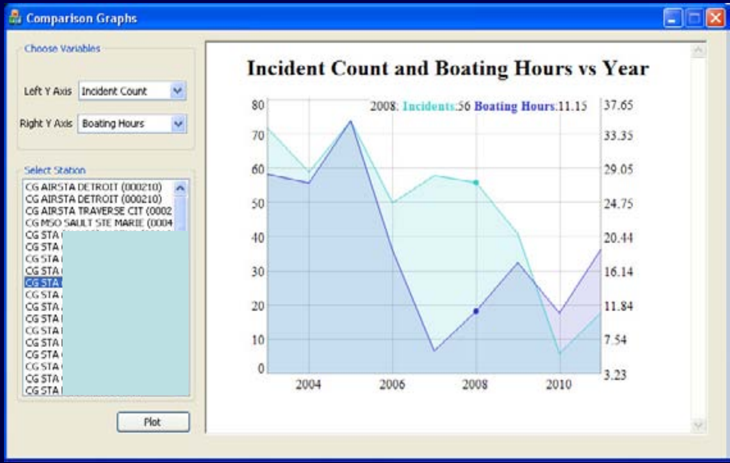
Findings:

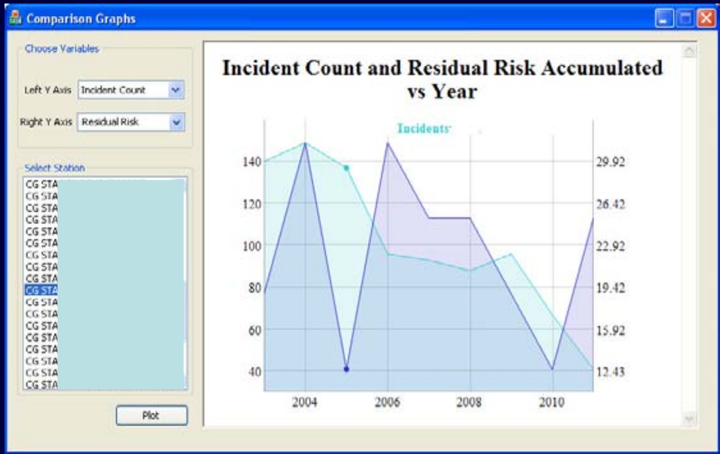
- Swimmer deaths
  - August highest frequency
  - Late afternoon highest frequency
  - Lake Michigan (south and west shore) have high concentration
- Boating deaths
  - Fri, Sat, Sun account for almost all deaths
  - Mid July to Mid August have highest frequency (only 1 week significantly high)
- 2009-2010 from MISLE Data
  - Large increase on Mon, Thu, Fri, Sun
  - Early and late season increase

VACCINE

Resource Allocation and Risk-Based Decision Making

- Explore risk-based decision making and utilize historical data for analysis and prediction
  - Total Risk, Mitigated Risk, Residual Risk
  - Explore 11 different USCG missions
  - Explore allocation of assets with different capabilities
  - Explore staffing, utilization, assets vs. risk measures
  - Perform What-If scenarios



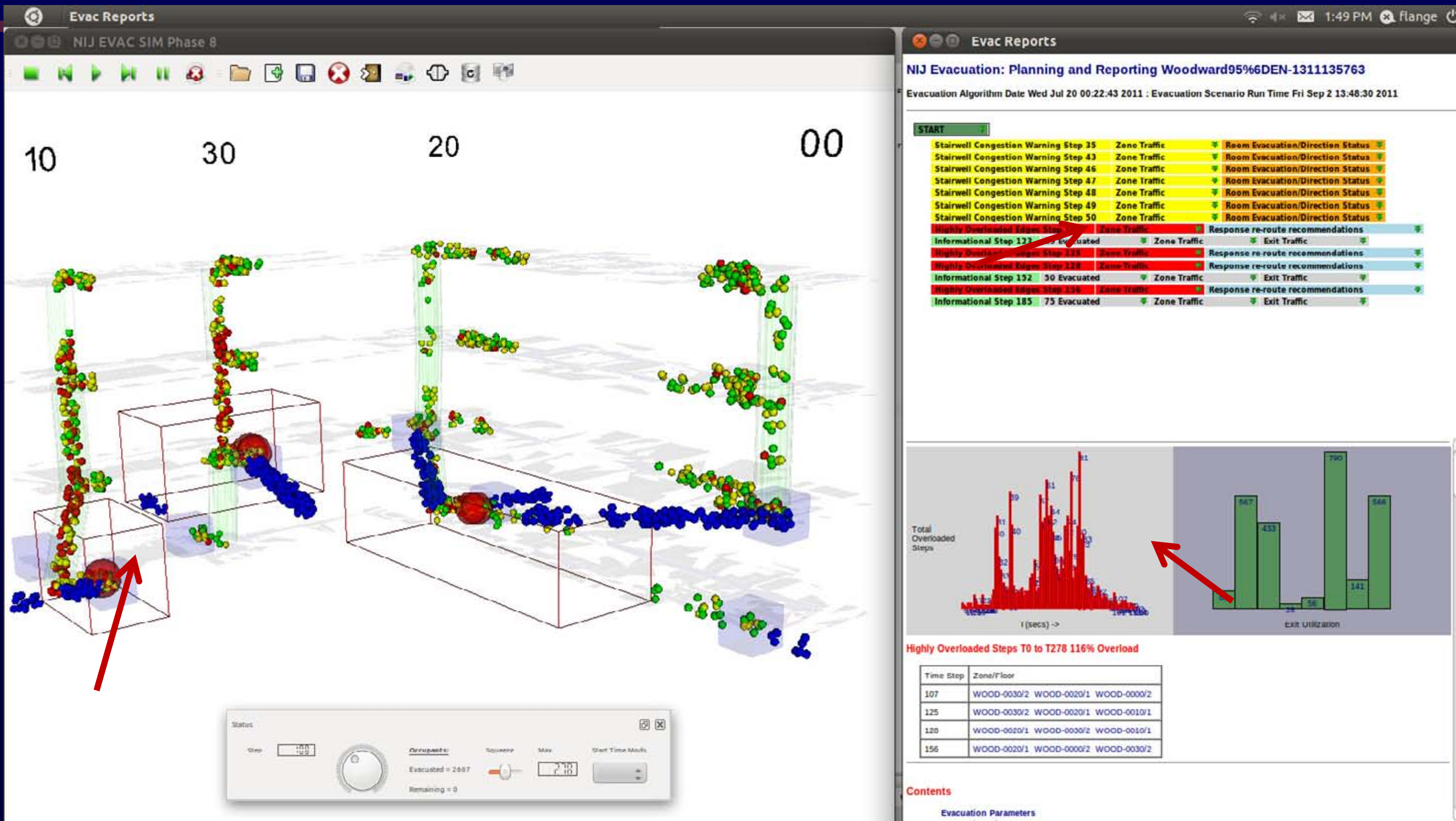


VACCINE

Example Public Safety Projects:

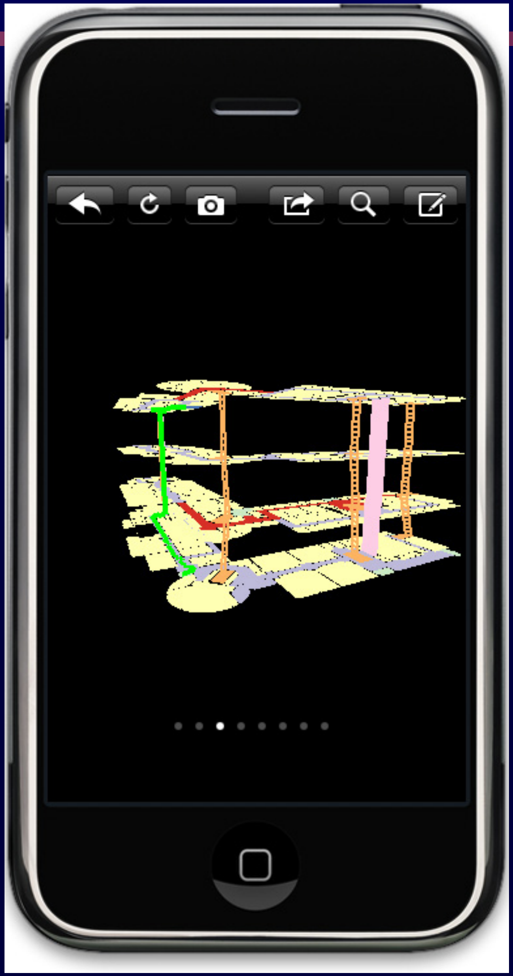
- Evacuation planning and management
- Hazardous Materials - mobile imaging
- Law enforcement visual analytics
- Additional projects
  - GARI – gang graffiti
  - Jigsaw

Large Scale Evacuation Model: Decision Support (Ribarsky et al. UNCC)

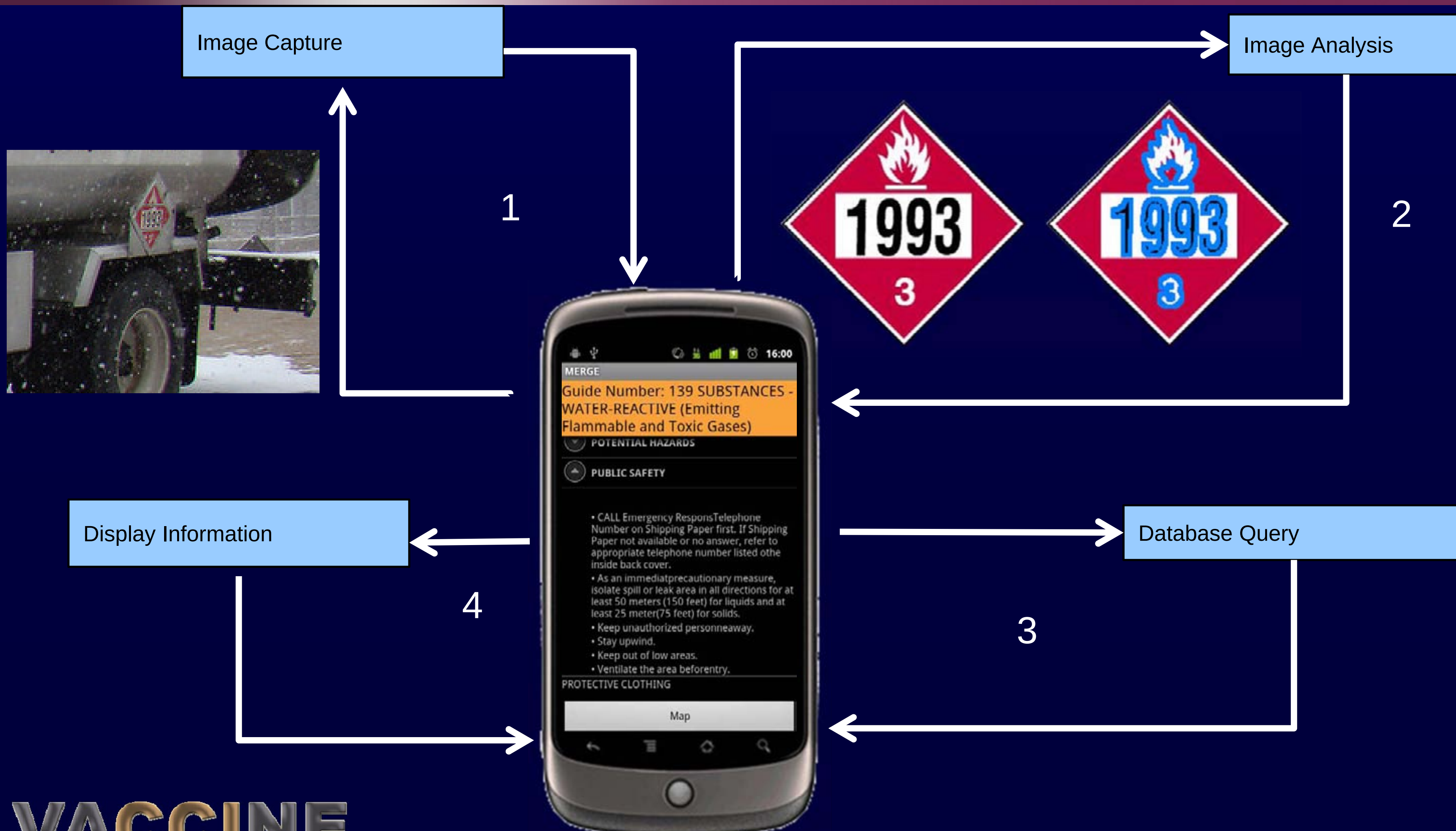


VACCINE

Mobile Interface

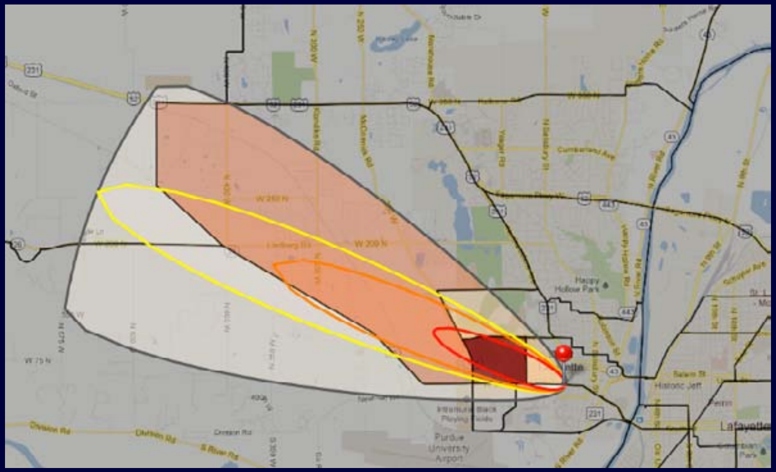


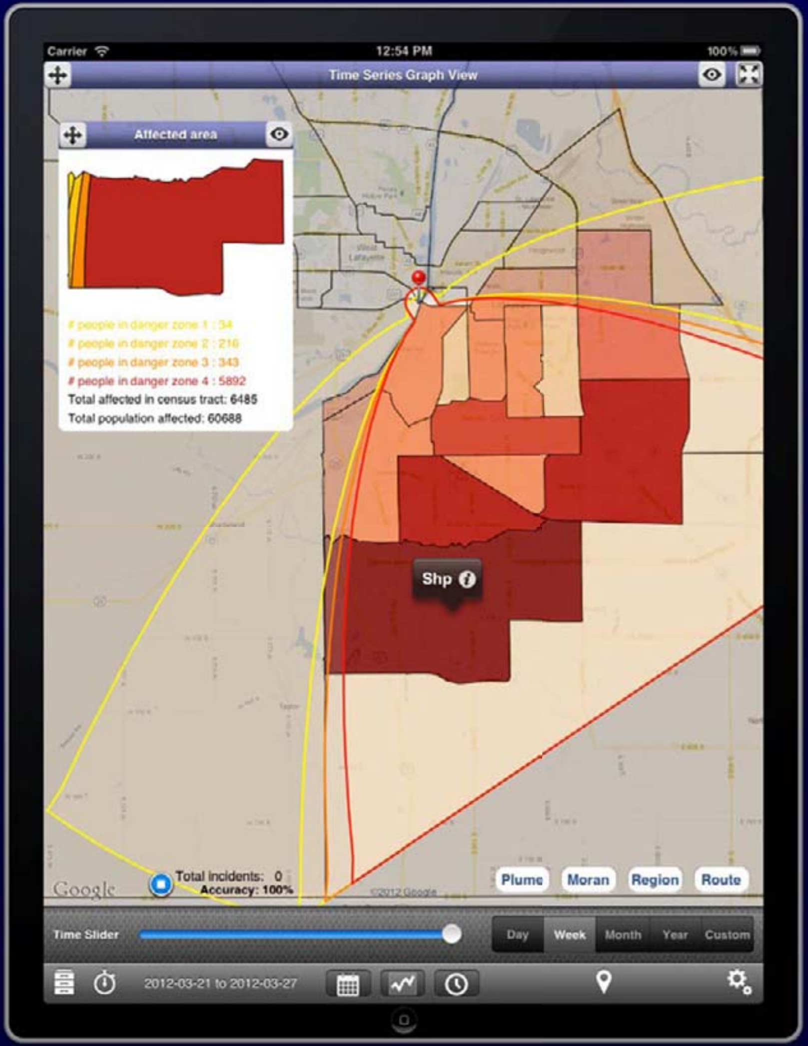
MERGE: Mobile Emergency Response Guide – System Overview



MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

- Chemical release plume modeling identifies census tracts with the highest number of expected people affected

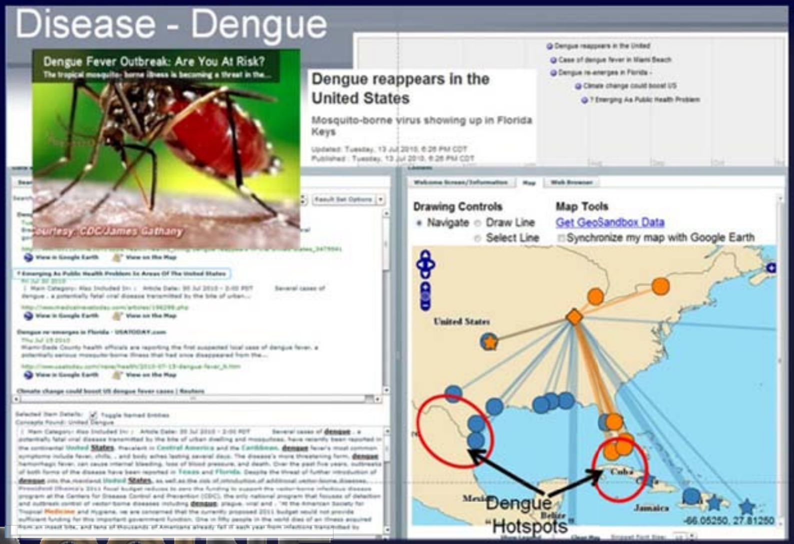


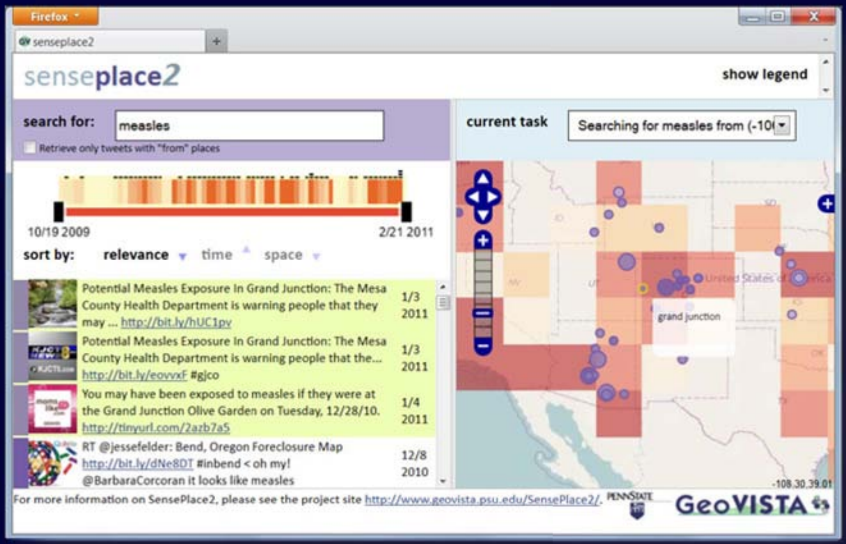


VACCINE

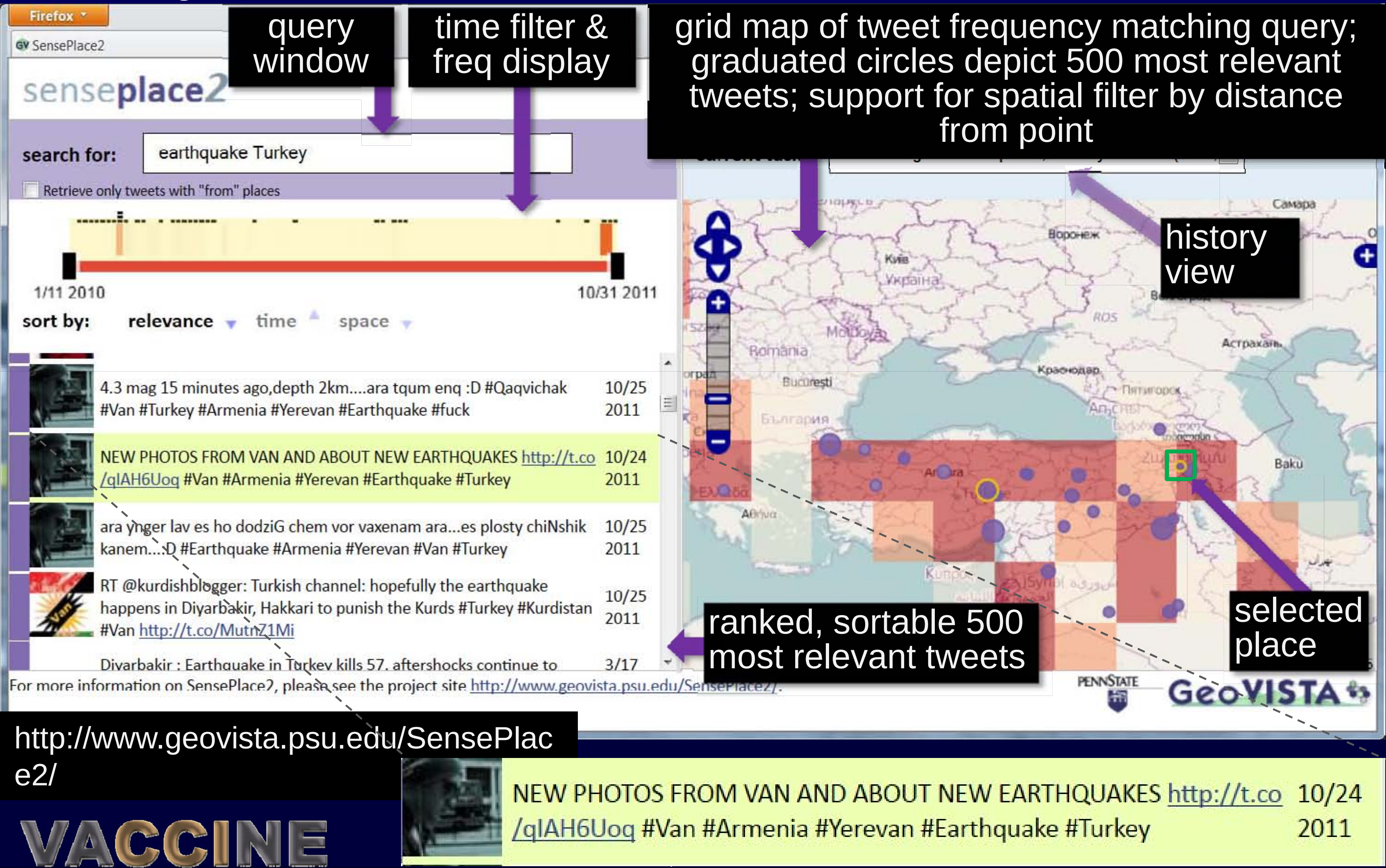
Geovisual Analytics Support For *Sensemaking* In Public Health And Crisis Management
(GeoVista Center, PSU)

- SensePlace – leveraging news to support infectious disease modeling and delivery of services in a developing country context
- SensePlace2 – leveraging social media to support situational awareness for health and crises





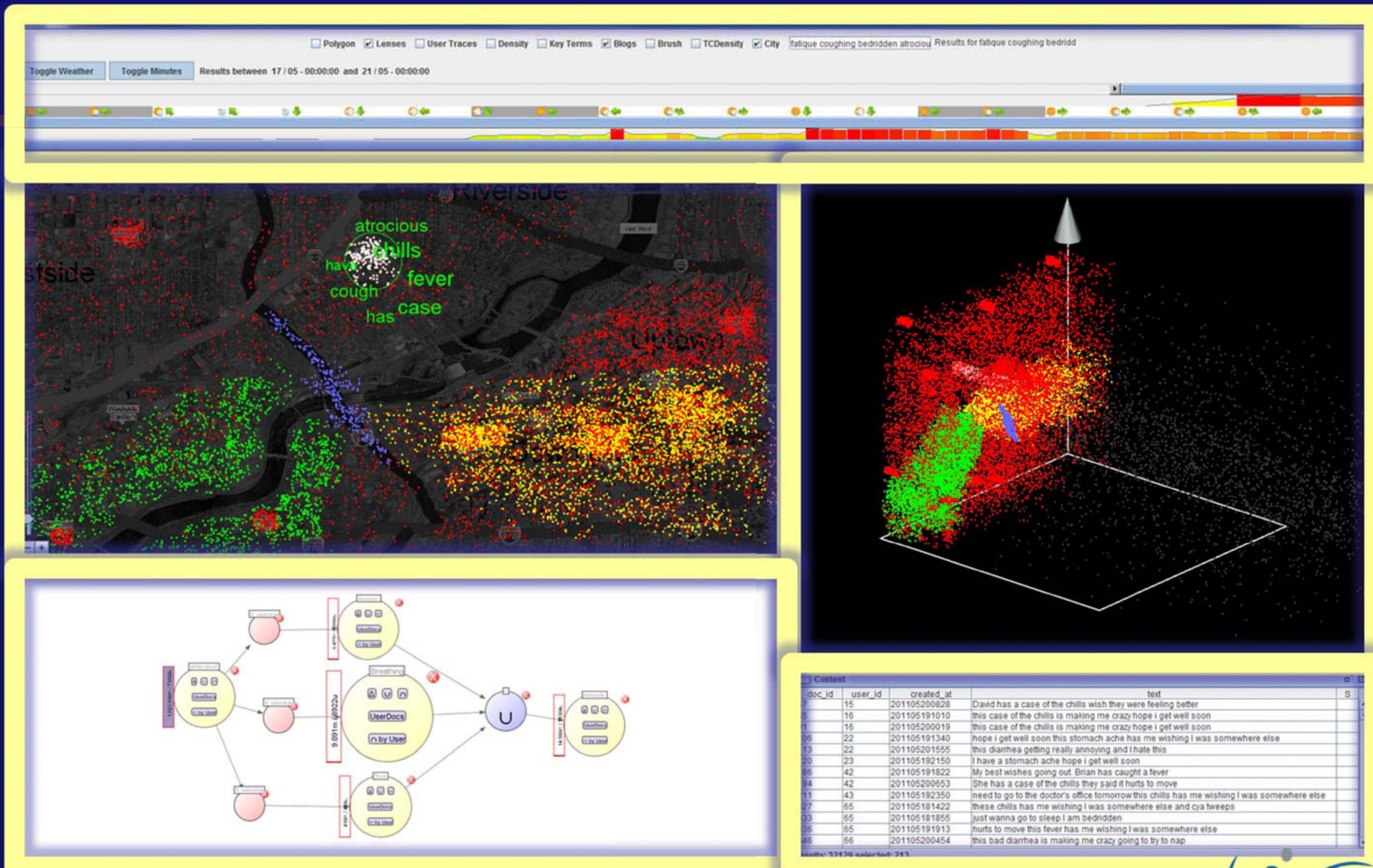
SensePlace2: Place-Time-Concept Analytics



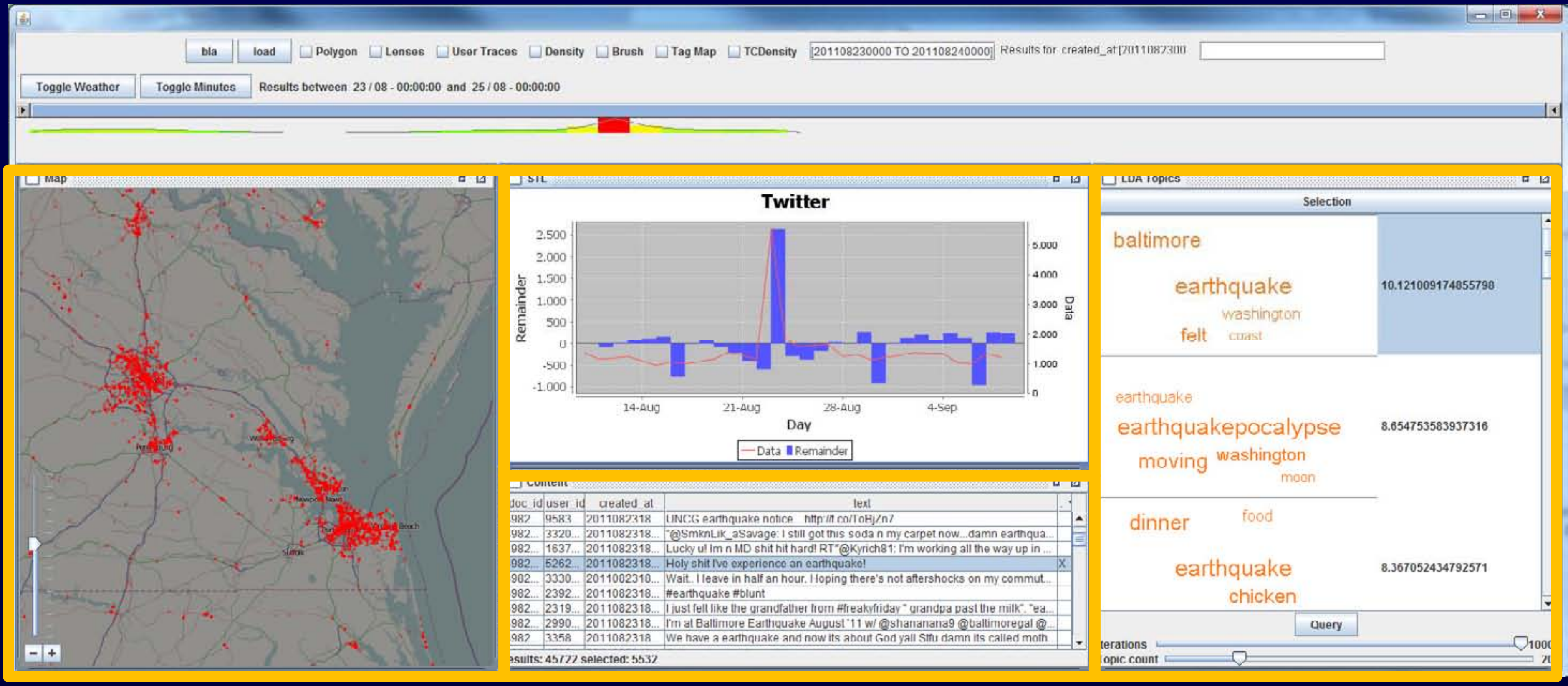
http://www.geovista.psu.edu/SensePlace2/

VACCINE

Scatterblogs: Geo-Spatial Document Analysis

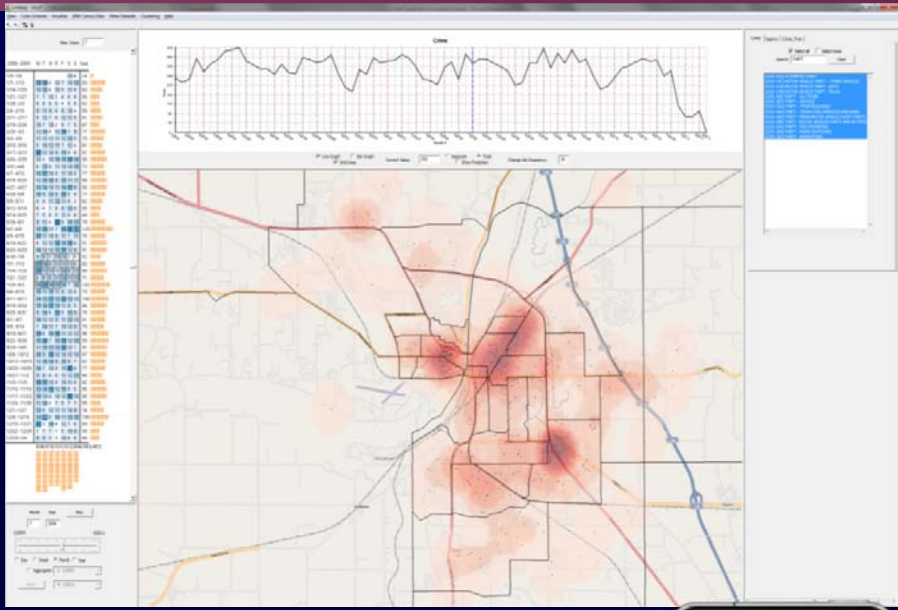


Spatiotemporal Social Media Analytics for Abnormal Event Detection



Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

End-Users: Lafayette, WL Police, Indiana Fusion Center, Indiana State Police, Ohio Fusion Center (initiated)



Impacts:

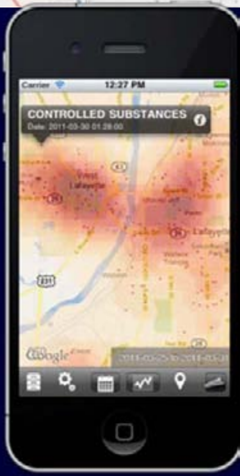
- In use to analyze crime patterns in Lafayette, Indiana and connect strings of activities
- Mobile version being released to public (November 2011) for community-based policing
- Investigating correlation of bus routes and crime, street lights and crime
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning

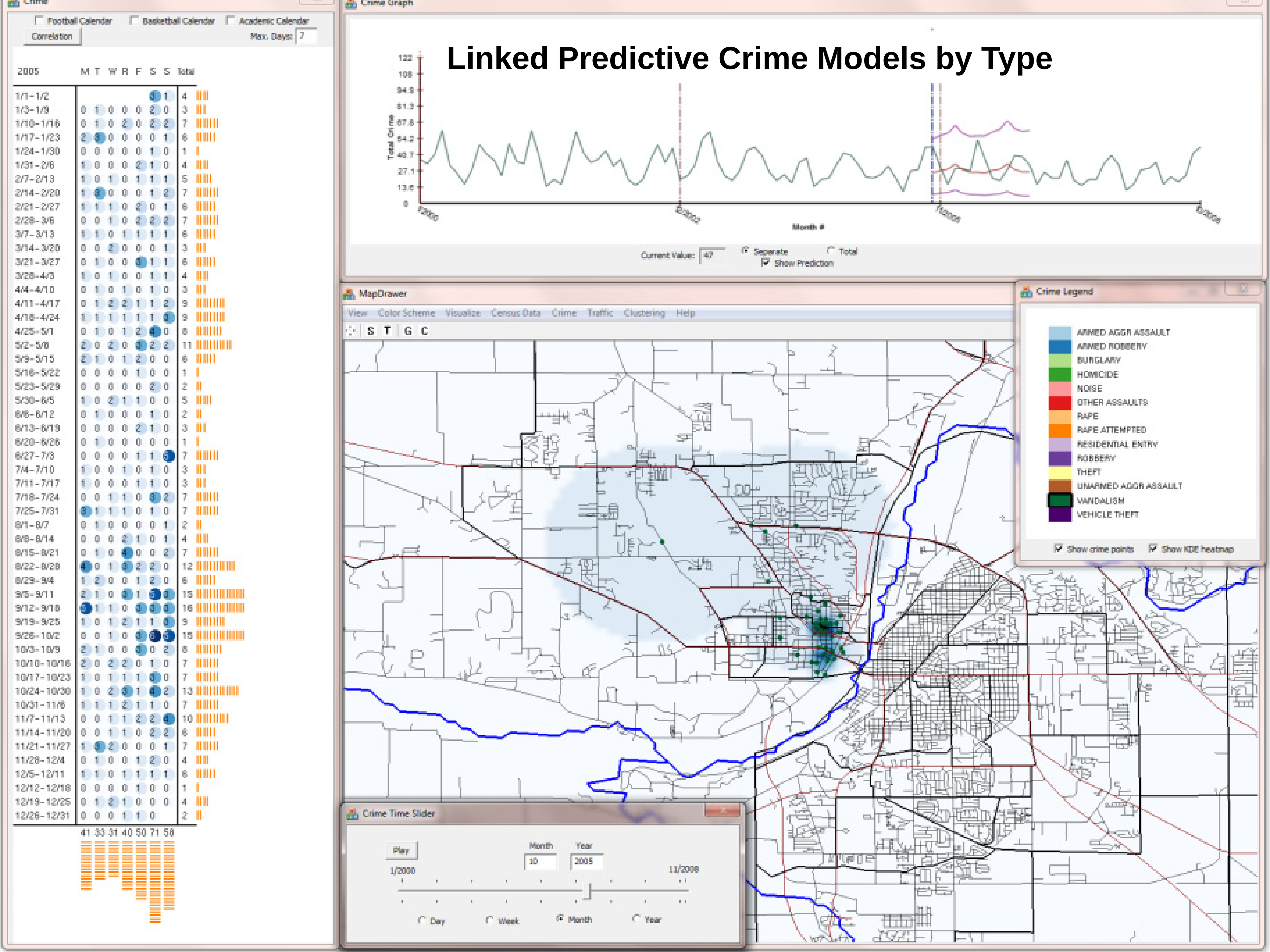
VALET delivered:

- Spring 2011: WL, Lafayette Police

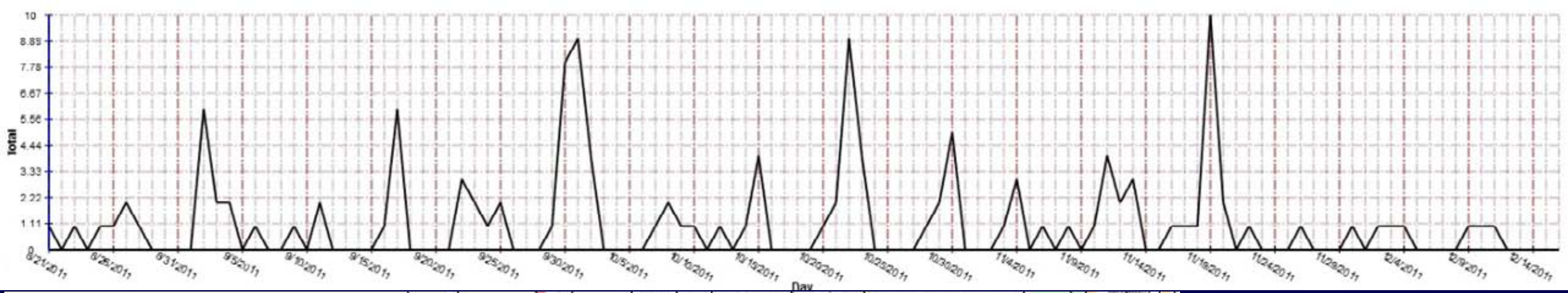
iVALET delivered:

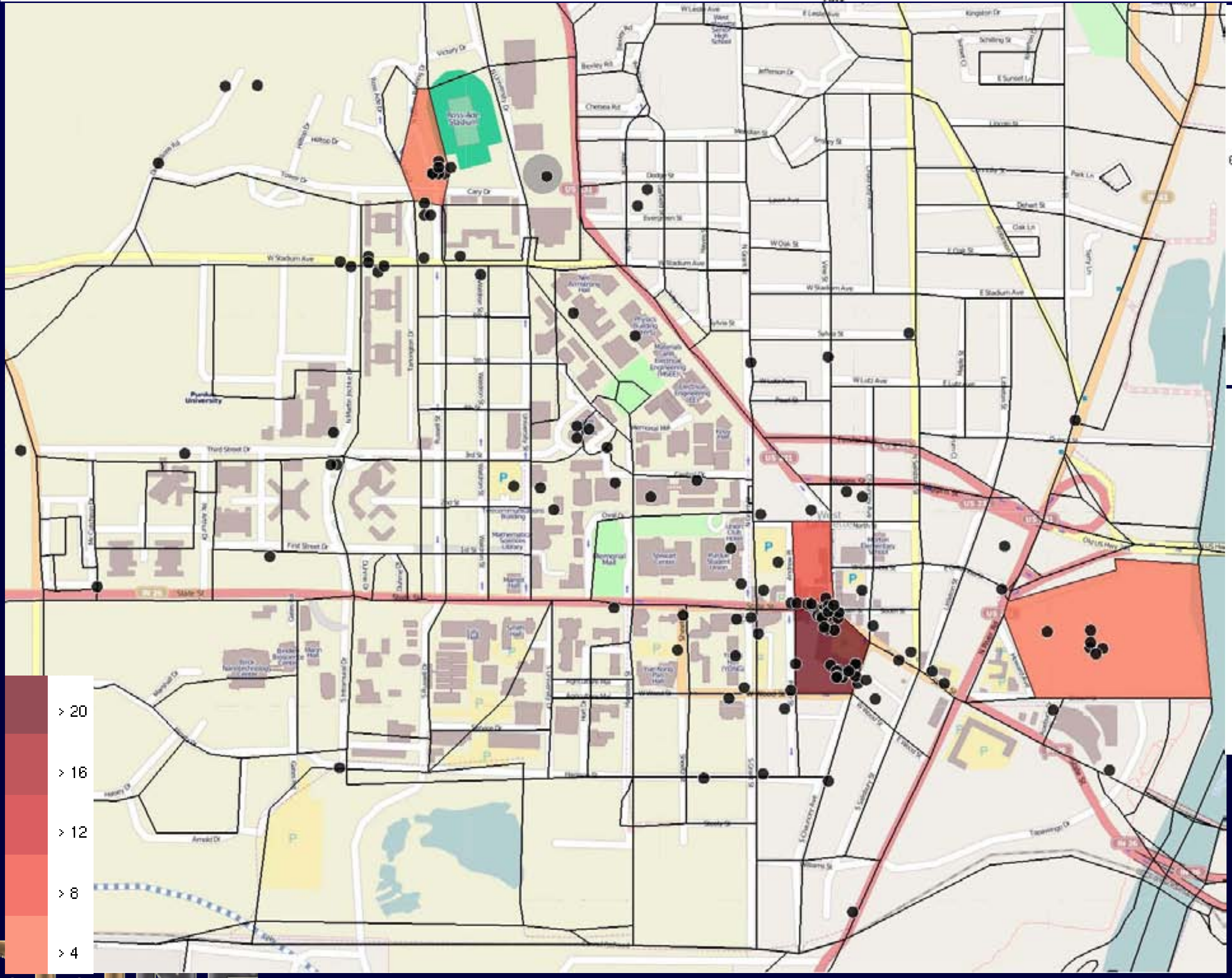
- October 2011: Purdue, WL Police

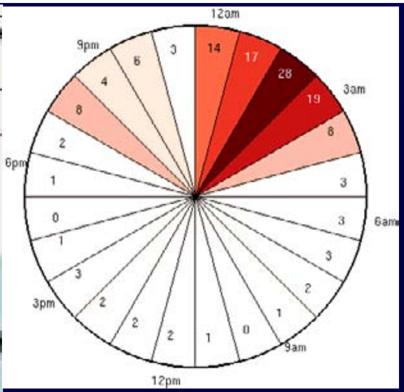


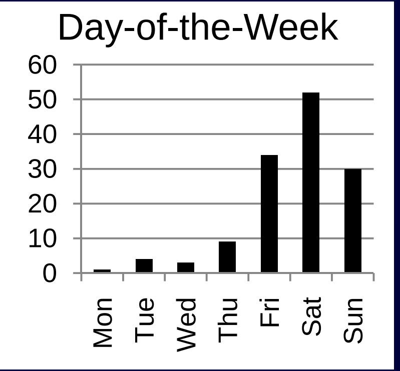

Linked Predictive Crime Models by Type
Crime
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
Max. Days: 7
2005
M T W R F S S Total
Crime Graph
Total Crime
Month #
Current Value: 47
Separate
Total
Show Prediction
MapDrawer
View Color Scheme Visualize Census Data Crime Traffic Clustering Help
S T G C
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Show crime points
Show KDE heatmap
Crime Time Slider
Play
Month
Year
10
2005
1/2000
11/2008
Day
Week
Month
Year

Drunkenness / Public Intoxication

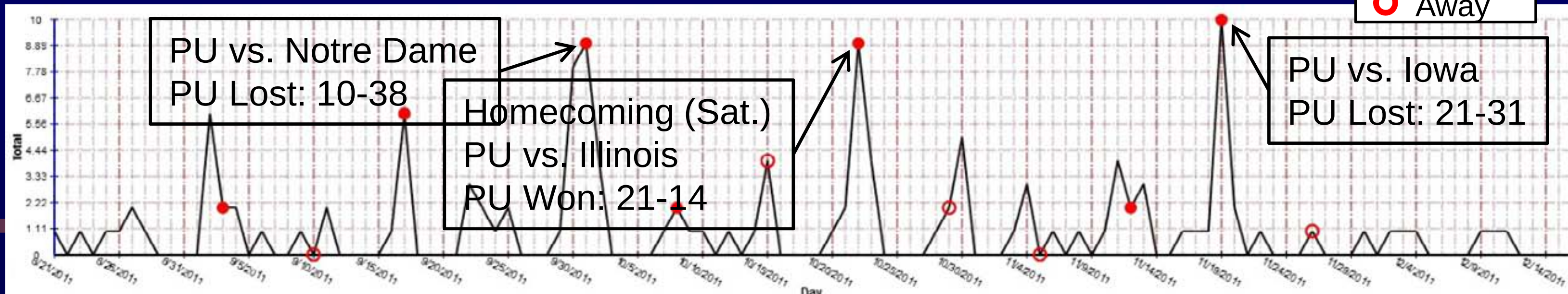


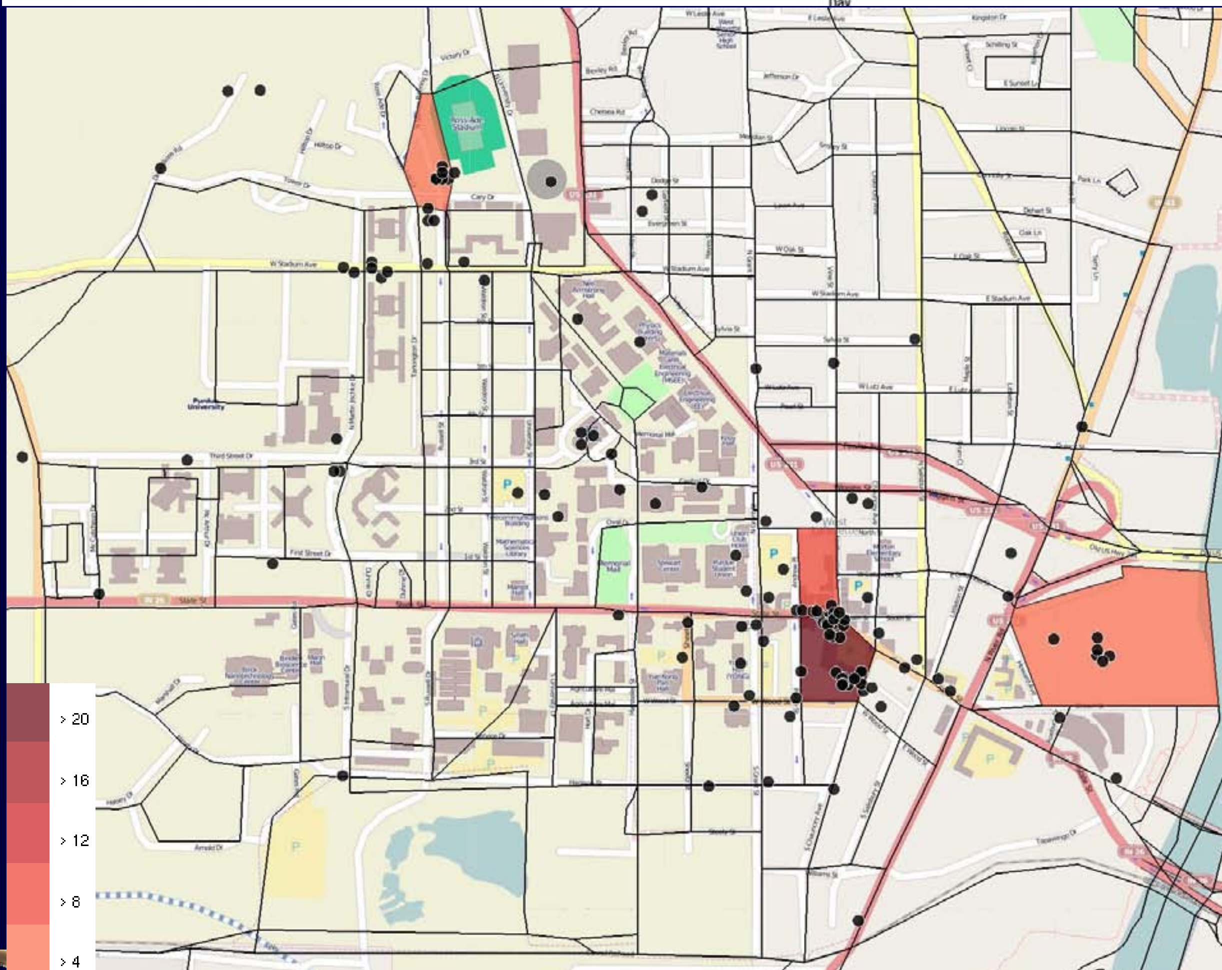


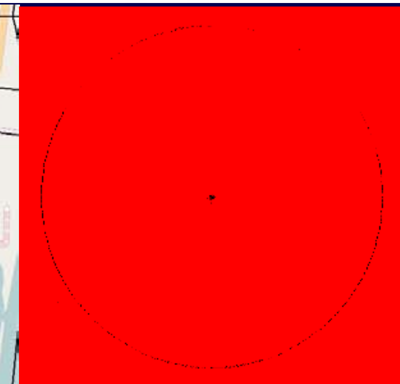


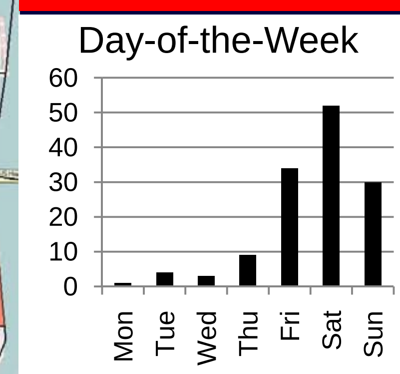


Drunkenness / Public Intoxication

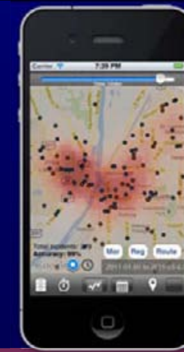






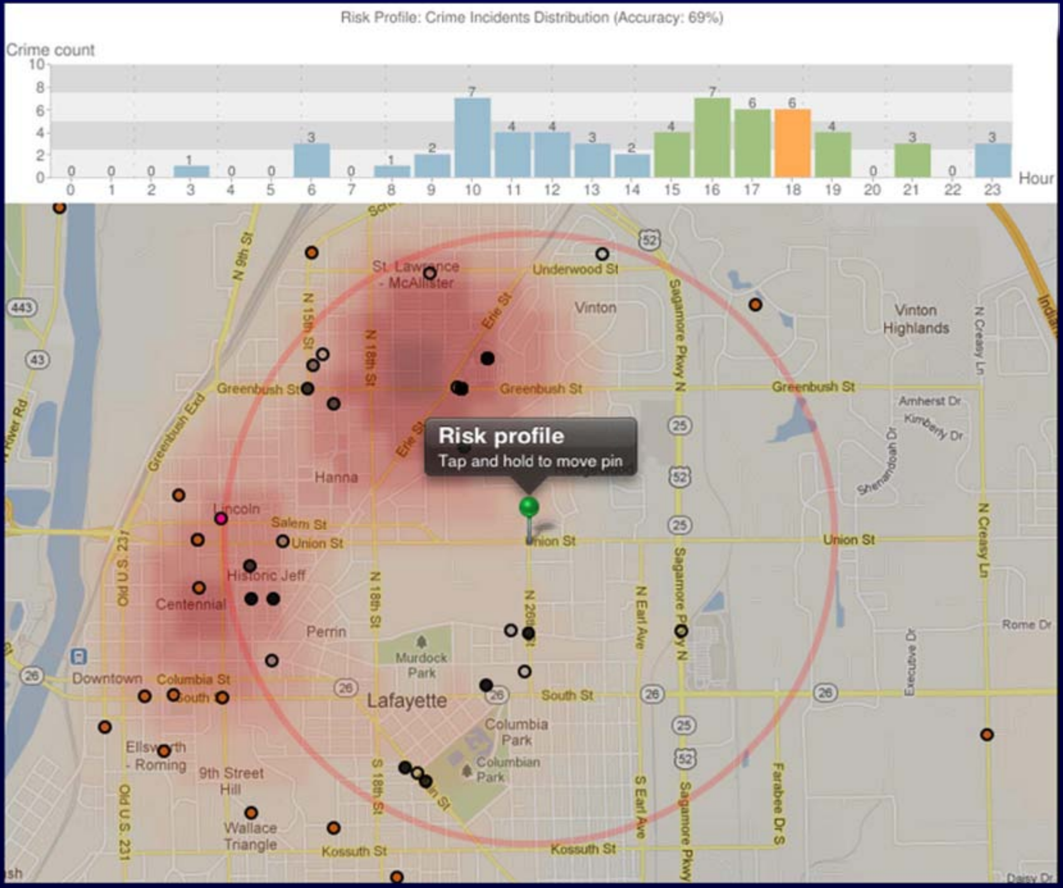


iVALET



- Explore criminal, traffic and civil data on-the-go
- Risk assessment
- Use current spatial + temporal context into analysis

iVALET



- Linked views to explore multivariate spatiotemporal dataset
- Analytical tools to help explore data

VACCINE

GeoVISTA *CrimeViz*

… an extensible web-based geovisualization application that supports exploration of & sensemaking about criminal activity in space & time

Project Personnel:

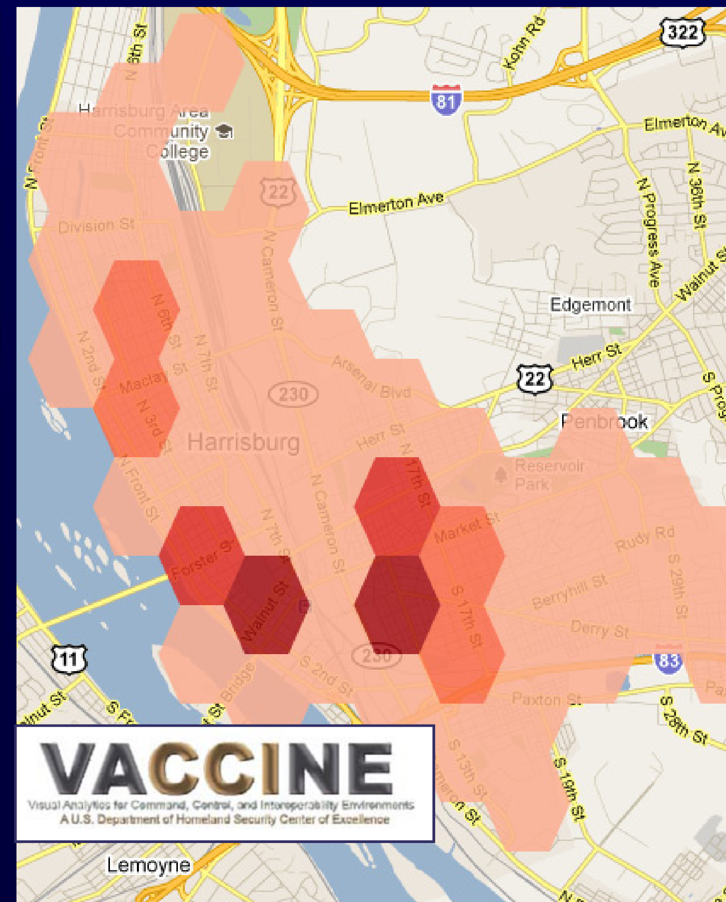
Alan M. MacEachren

Robert E. Roth

Kevin S. Ross

Scott Pezanowski

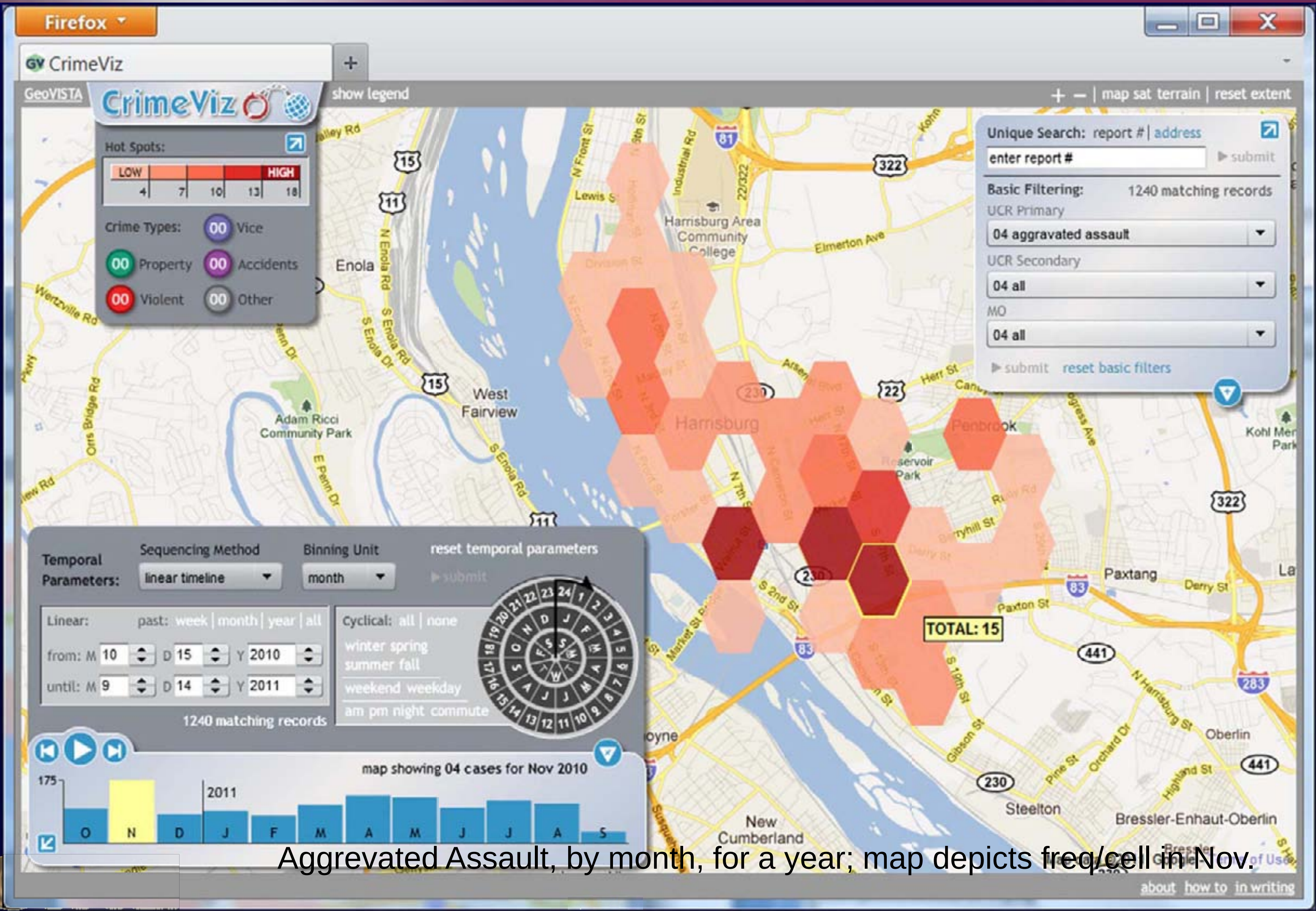
GeoVISTA Center
Department of Geography
Penn State University
maceachren@psu.edu



VACCINE

http://www.geovista.psu.edu/CrimeViz

GeoVISTA CrimeViz: Year/Month View



Aggrevated Assault, by month, for a year; map depicts freq/cell in Nov.

Crime Analytics & SA Issues and Techniques

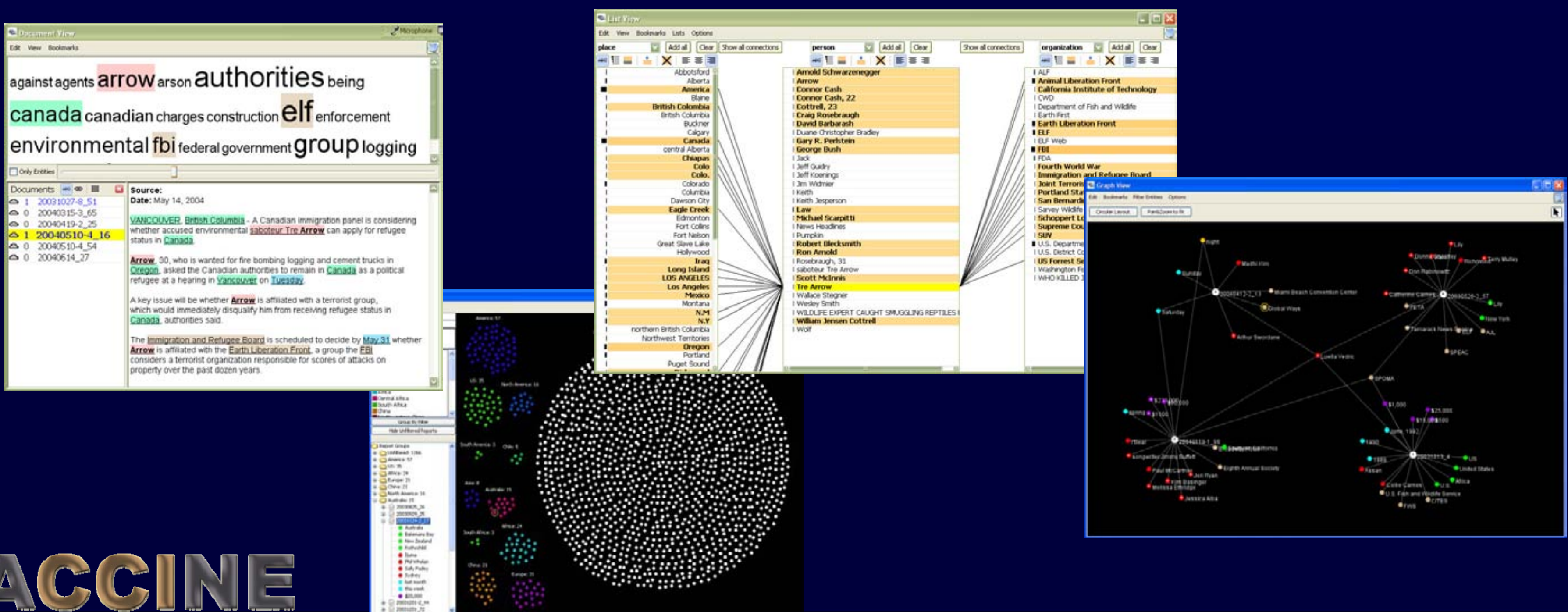
- Fuse data from a variety of sources
  - Law enforcement records management
  - Weather and phases of the moon
  - Street light locations, bus routes
  - Tracking release data of offenders
  - Civil court data
  - Social Media
  - Local event calendar
- Reliable predictive models
- Understandability and trust of predictions
- Main Question: What helps officers, detectives, chief do their variety of jobs?

Jigsaw:
Visual Analytics for Investigative Analysis and Exploration of Document Collections

"Putting the pieces together"

Goal: Assist investigators with understanding, sense-making, and analysis of large, unstructured and structured document collections

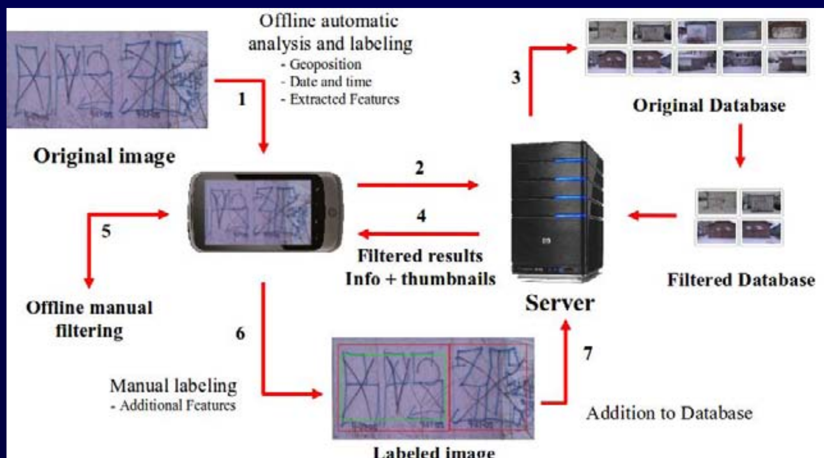
Approach: Provide multiple visual perspectives on the documents and entities within them, highlighting connections between entities



VACCINE

Gang Graffiti Recognition and Analysis Using a Mobile Telephone (GARI)

Edward J. Delp, Mireille Boutin
Collaborating Institution(s): Purdue University
End-User(s): Gary Coons, Chief/Indianapolis Department of Public Safety Division of Homeland Security and IMPD, Indiana Fusion Center Gang Task Force



IMPACT:

- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be “interpreted”
- More than 40 users and 450 graffiti images acquired

GARI delivered:

- Summer 2011:
  - IMPD gang detectives
- August 2011:
  - IMPD at large
  - Ind Fusion Gang Task Force
- October 2011:
  - Gang detectives across Indiana

Path

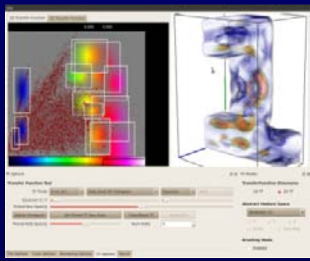
- Introduction to our Center - VACCINE
- Motivation and introduction to visual analytics
- Example solutions we've developed
  - Health visual analytics
  - Public safety
  - Resource allocation

Visual analytic challenges and future

~~Five~~ Six Challenges for Proactive & Predictive Visual Analytics

1. Creating computer-human visual cognition environments
2. Integrating interactive simulation and analytics
3. Solving specific scale issues and cross-scale issues
4. Managing uncertainty and time
5. Enabling risk-based decision making environments
6. Developing the Science of Interaction for Visual Analytics

Visual Analytics At Real-World Scale



- Utilize advanced HPC techniques to enable interactive spatiotemporal analysis (spatiotemporal clustering, prediction)
  - Cluster-based and cloud-based solutions
  - GPGPU solutions
- Develop easily usable HPC visual analytic environments
- Example: Longhorn Exascale Visual Analytic Platform
  - 2048 compute cores (Nehalem quad-core)
  - 512 GPUs (128 NVIDIA Quadro Plex S4s, each containing 4 NVIDIA FX 5800s)
  - 13.5 TB of distributed memory
  - 210 TB global file system

VACCINE

The Science of Interaction

- Definition: The study of methods by which humans create knowledge through the manipulation of an interface.
  - …interaction and inquiry are inextricable. It is through the interactive manipulation of a visual interface – the analytic discourse – that knowledge is constructed, tested, refined and shared." Stasko et al., IV2009

5 Challenges:

1. **Visual Discourse -** Interactive visual thinking tools for the exploration, understanding, collaboration, description, explanation, decision, dissemination, persuasion of concepts & data.
2. **Multi-modal Sensemaking In The Large -** Enable large-scale, distributed teams to interactively make sense of big multi-modal data and problems.
3. **Fluid Interaction -** Designing fluid, high-bandwidth, and powerful interaction models and paradigms for the purpose of individuals reaching their full potential in viewing, analyzing, and understanding large and complex data.
4. **Collaboration –** beyond space and time- synergizing technologies and human users
5. **Mixed Initiative Data Discovery and Manipulation -** Balancing active user input with systematic guidance to enable visual data manipulation and analysis.

Visual Analytics: Remember…

- We need to be cognizant of parameters for visual representations
- Appropriate analysis can guide users to interesting features in the data
- Refined analysis through user interaction and their domain knowledge can help discover hidden problems
- There is no single catch-all visual representation or analysis

Keys for Success

- User and problem driven
- Balance human cognition and automated analysis and modeling
  - Often applied on-the-fly for specific components identified by the user
- Interactivity and easy interaction
  - Utilizing HPC and novel analysis approaches
- Understandability of why predicted value is what it is
- Intuitive visual cognition
- Not overloaded with features

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu

ebertd@purdue.edu

VACCINE

Table of Contents

VACCINE

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

Visual Analytics
≠
Big Data Analytics

David S. Ebert
ebertd@purdue.edu

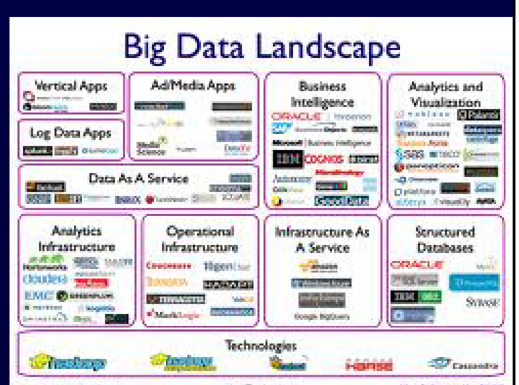
Path

- Introduction
- Motivation and introduction to visual analytics
- Example solutions we've developed
  - Science visual analytics
  - Health visual analytics
  - Maritime public safety and law enforcement
  - Public safety
  - Resource allocation
- Visual analytic challenges and future

Big Data Is The Rage

The Promise of Big Data:

- Commercial Zoo:



Premise of Talk:

- Visual Analytics Is More Than Big Data Analytics

VACCINE

What We Do With Visual Analytics

- Enable effective decision making through interactive visual analytic environments
- Enable effective communication of information
- Provide quantitative, reliable, reproducible evidence
- Enable user to be more effective from planning to detection to response to recovery
- Enable proactive and predictive visual analytics
- Enable effective situational awareness (perception, comprehension, projection)

VACCINE

Path

Introduction

Motivation and introduction to visual analytics

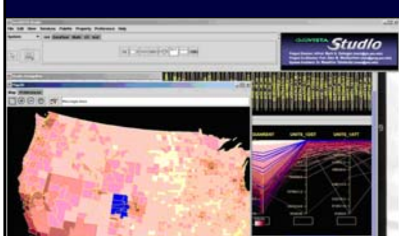
Example solutions we've developed

- Science visual analytics
- Health visual analytics
- Maritime public safety and law enforcement
- Public safety
- Resource allocation

Visual analytic challenges and future

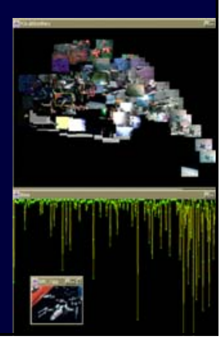
Motivation

To solve today's and tomorrow's problems requires exploring, analyzing, and reasoning with massive, multisource, multiscale, heterogeneous, streaming data

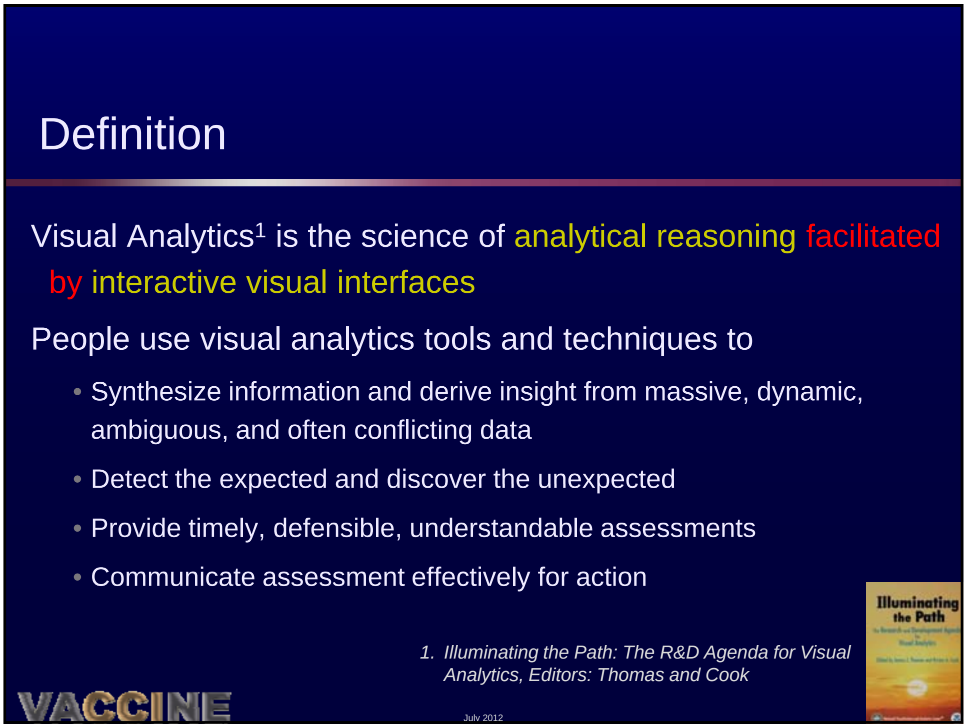


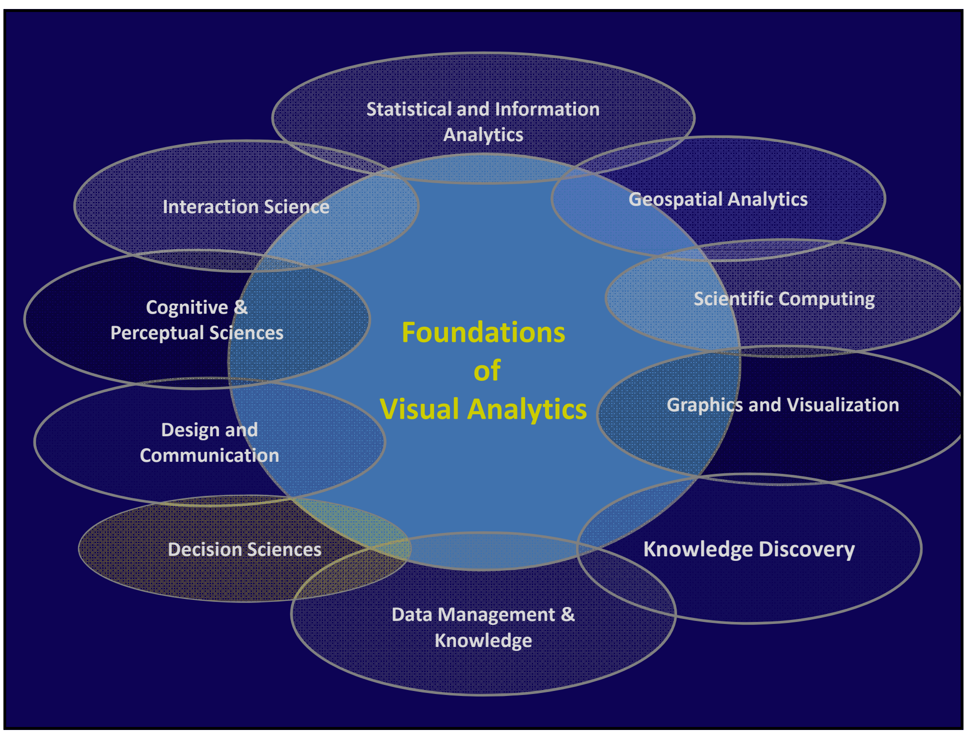
VACCINE

Image of Analyst's Notebook



July 2012


Definition
Visual Analytics1 is the science of analytical reasoning facilitated by interactive visual interfaces
People use visual analytics tools and techniques to
• Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
• Detect the expected and discover the unexpected
• Provide timely, defensible, understandable assessments
• Communicate assessment effectively for action
1. Illuminating the Path: The R&D Agenda for Visual Analytics, Editors: Thomas and Cook
Illuminating the Path
VACCINE
July 2012


Statistical and Information Analytics
Interaction Science
Geospatial Analytics
Cognitive & Perceptual Sciences
Scientific Computing
Foundations of Visual Analytics
Graphics and Visualization
Design and Communication
Decision Sciences
Knowledge Discovery
Data Management & Knowledge

Research Motivation:

- Solving these real-world problems requires
  - Novel theories, techniques, approaches, and adaptations of algorithms
  - Integration of cross-disciplinary expertise
- Solving these real-world problems provides
  - Compelling, publicly understandable value for your research
  - Advances in CS and in other disciplines
  - New publication opportunities
  - Great collaboration partners and proponents
  - Opportunities for new adventures

VACCINE July 2012

Example Challenging Problem - Atmospheric Science: Multi-scale Interactions (in the words of a cloud physicist)

1 mm
1 kHz

1km
5min

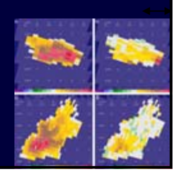
No observing platform can measure the quantities of interest over all needed spatial and temporal scales needed

No numerical model can simulate the quantities of interest over all needed spatial and temporal scales

→We observe/simulate over a subset of the pertinent scales, using different instruments/models, and must assimilate these results to understand the "big picture"

Visual analytics is crucial for this task

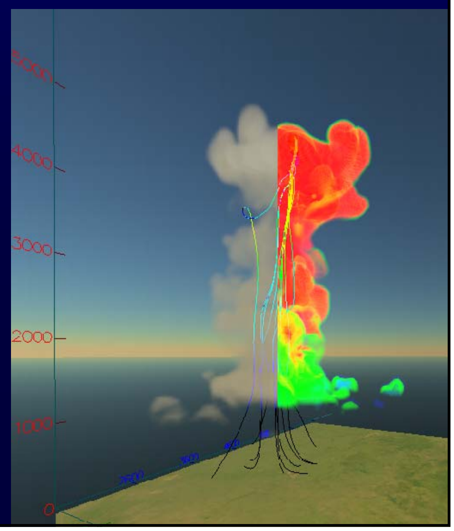
Issues: Multi-scale, multi-system, multisource, massive, data & simulations



VACCINE July 2012

One Solution in Use: Our Atmospheric Visual Analytic Environment

- Utilize multiple rendering styles
- Provide interactive data exploration and user directed analysis
- Allow user specified analysis and queries on the fly
- Allow interactive correlative analysis of multisource data



VACCINE

July 2012

What's Needed for Proactive and Predictive Visual Analytics?

- Reliable and reproducible models and simulation
- Understanding of the data
  - Distribution and skewness, errors, appropriate analysis techniques
- Understanding of the sources and types of data
- Comparable or Correlative sources of data
  - Appropriate transformations applied to enable meaningful comparison and correlation
- Understanding of the use and problem to be solved!

VACCINE

July 2012

How to Harness Visual Representations

- Visual representations translate data into a visible form, highlighting important features, such as commonalities and anomalies
- Visual representations make it easy for users to perceive salient aspects of their data quickly
- These visual representations augment the cognitive reasoning process with perceptual reasoning which enhances the analytical reasoning process

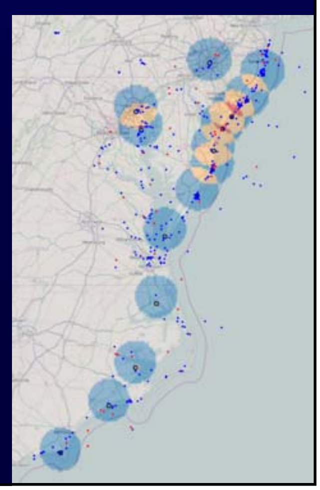
1 – J. J. Thomas and K. A. Cook (Ed.). *Illuminating the Path: The R&D Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.

VACCINE

July 2012

Visualization, Data, and Decision Making

- At some level most decision making is driven by data
- What role does visual representation play?
- Good or bad – people believe charts and graphics – “seeing is believing”



VACCINE

July 2012

Path

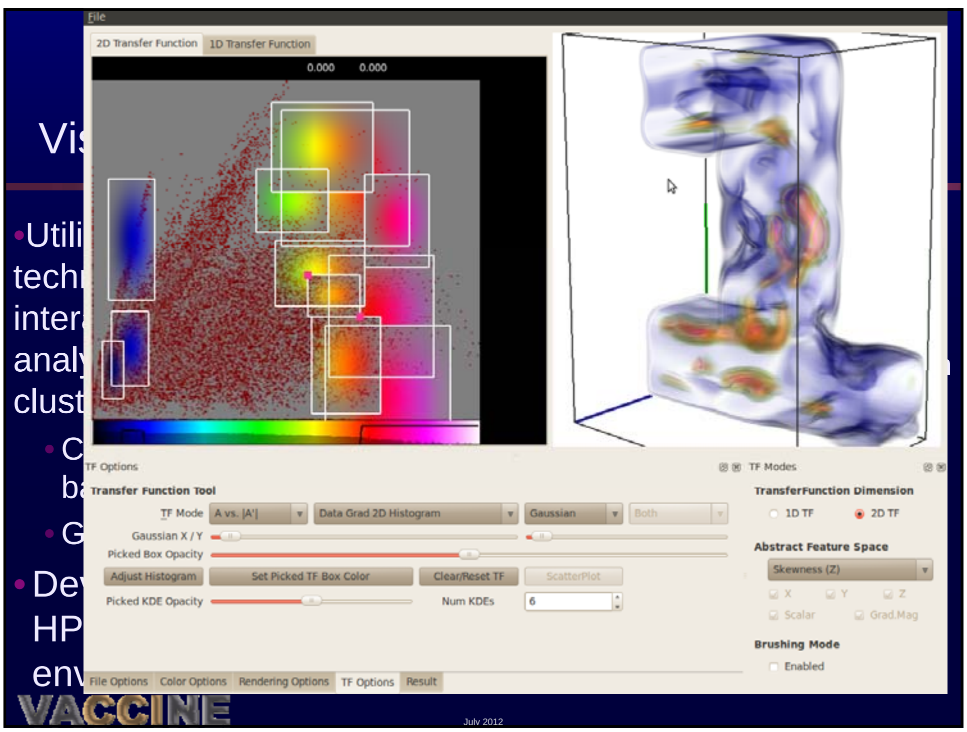
Introduction to our Center - VACCINE

Motivation and introduction to visual analytics

Example solutions we've developed

- Science visual analytics
- Health visual analytics
- Maritime public safety and law enforcement
- Public safety
- Resource allocation and business intelligence

Visual analytic challenges and future

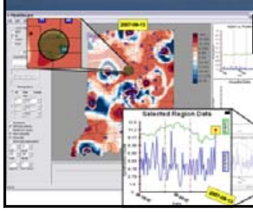


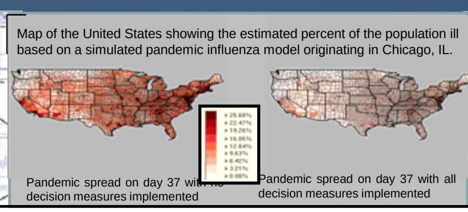
Example: Public Health Visual Analytics

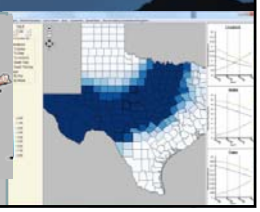
Health Surveillance

Pandemic Simulation and Planning

Zoonoses Planning and Response

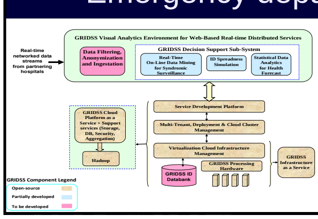


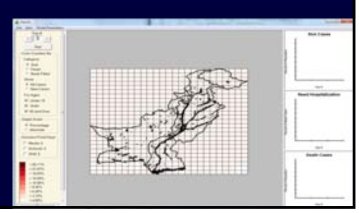




Visual Analytics For Syndromic Surveillance

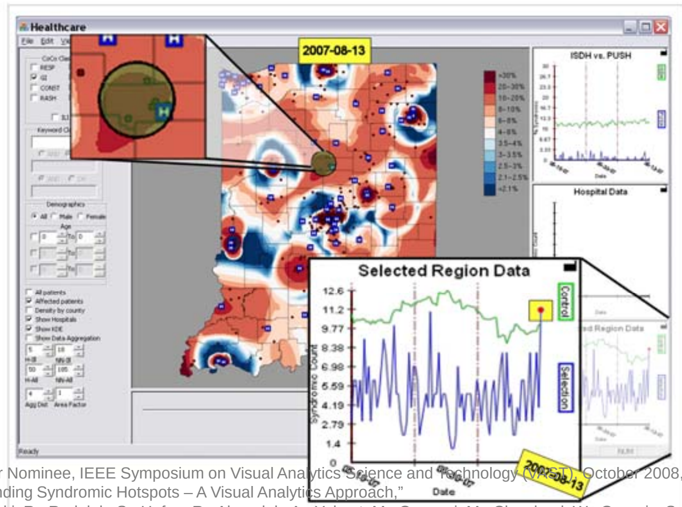
- Syndromic surveillance: detection of adverse health events focusing on pre-diagnosis information to improve response time
- Pre-diagnosis information can consist of multiple data sources:
  - Over the counter medicine sales
  - News reports on emerging diseases
  - Pro-med news feeds
  - Emergency department chief complaints





July 2012

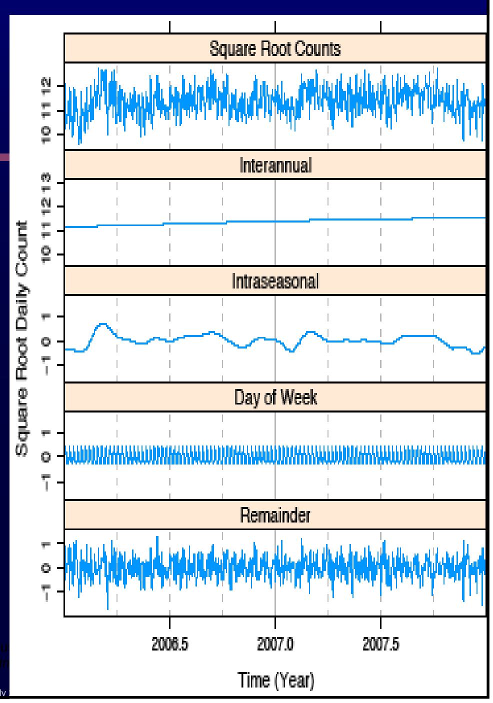
Hypothesis Generation and Exploration



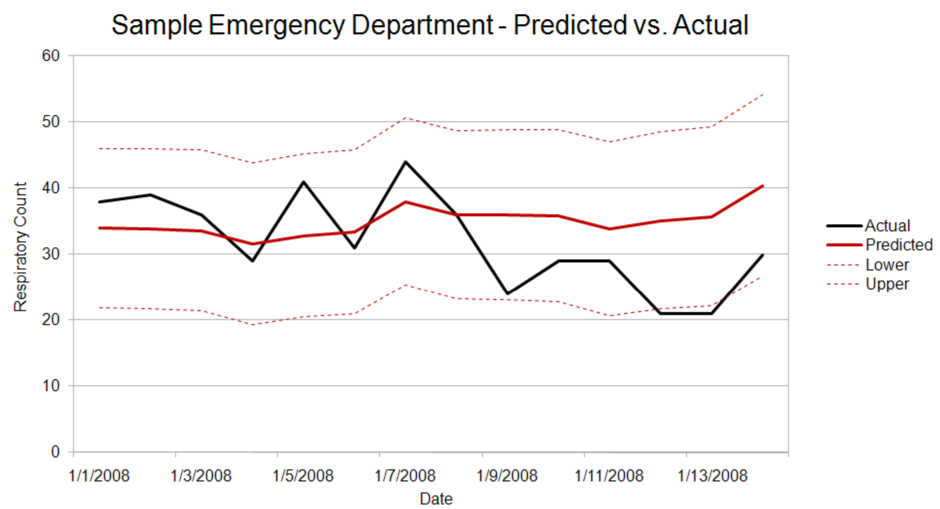
Best Paper Nominee, IEEE Symposium on Visual Analytics Science and Technology (VAST), October 2008, for "Understanding Syndromic Hotspots – A Visual Analytics Approach," (Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W., Grannis, S., Wade, M., Ebert, D.)

Time Series Modeling

- Seasonal-Trend Decomposition Based on Loess
  - Time series can be viewed as the sum of multiple trend components
  - For each data signal, components are extracted
  - Can then analyze correlation between components

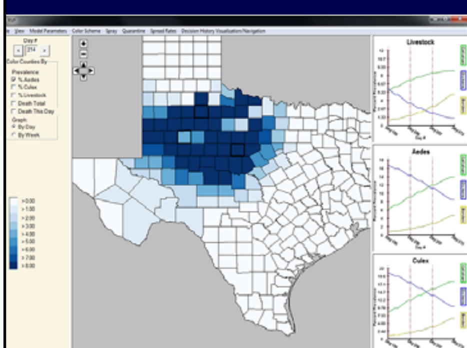


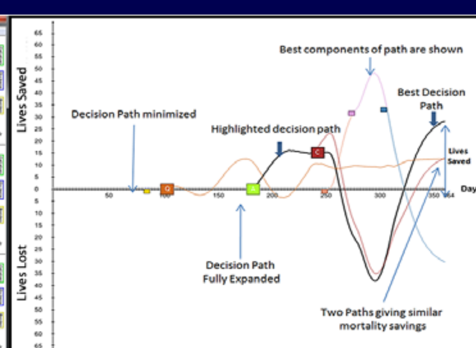
Predictive Visual Analytics



July 2012

Advanced Decision Support Tools: Rift Valley Fever

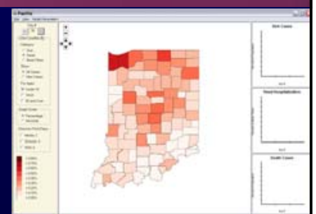




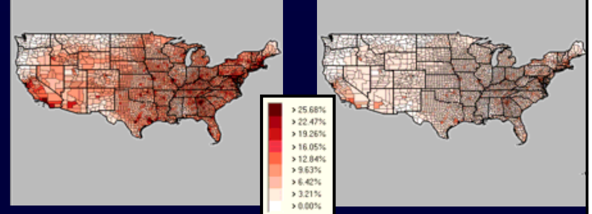
July 2012

Modeling a Pandemic

- Pandemic Influenza Planning Tool
- Models user specified:
  - Pandemic influenza characteristics
  - County population, demographics, hospital beds
- Decision measures
  - Strategic National Stockpile deployment
  - School Closures
  - Media Alerts



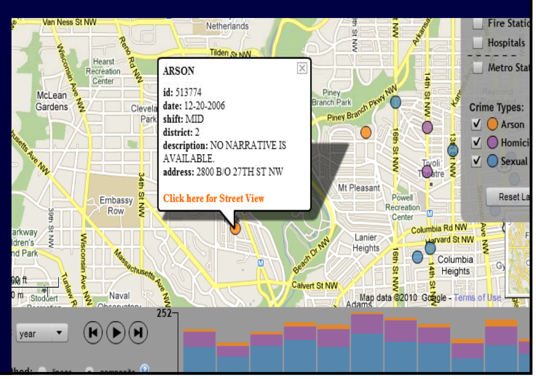
Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.



VACCINE

Visual Analytics Uses for Crisis Management and Response

- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis
- Purpose:
  - Planning for resiliency
  - Long-term analysis
  - Predictive analytics
  - Training
  - Detection
  - Investigation
  - Response
  - Recovery, remediation

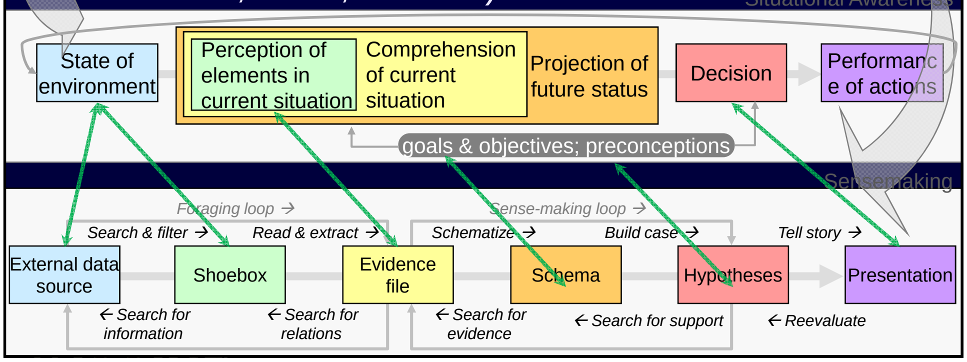


VACCINE

Visual Analytics vs. Situational Awareness
Situational Awareness ←→ Sensemaking

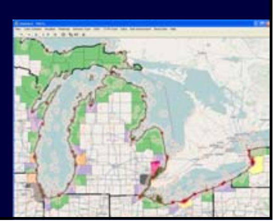
situational awareness as a sensemaking process (courtesy Alan MacEachren)

Situational awareness model (after Endsley, 1995; top) compared to model of sense making (after Pirolli & Card, 2005; bottom)



Visual Analytics Uses for Public Safety

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



VACCINE

July 2012

Example Maritime Public Safety Projects

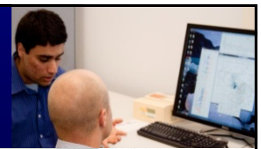
- Search and rescue resource allocation
- Swimmer death analysis
- PWCS analysis
- Economic impact analysis
- Resource allocation and risk-based decision making

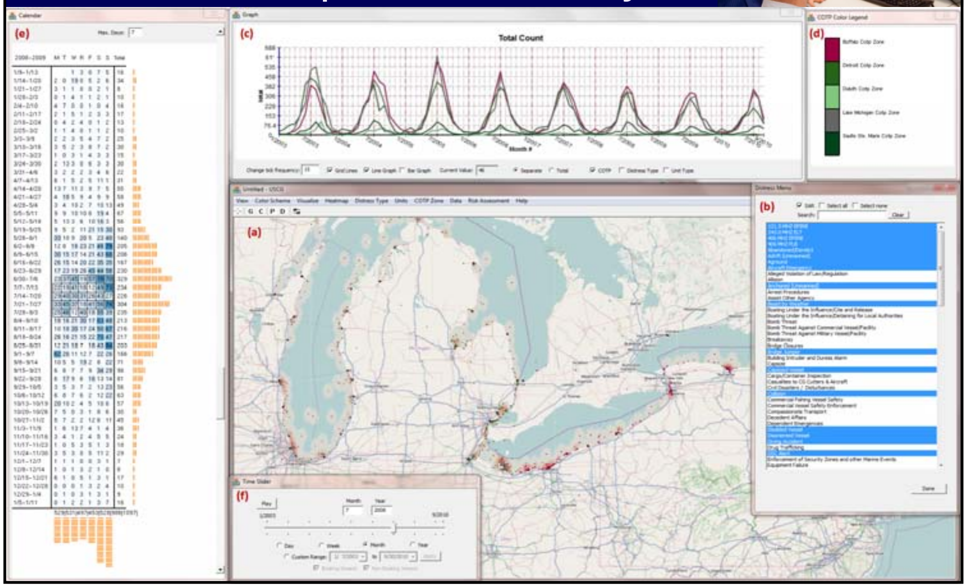
Example: Visual Analytics Environment

- Supports decision making and risk assessment
- Interactive exploration and analysis of trends, patterns and anomalies
- Allows analysis of risks associated with
  - Closing one or more Coast Guard stations
    - Find optimal stations that absorb work load of the closing station
  - Allocating new resources
    - Impact on safety and efficiency of operation
- Currently being used by analysts at the U.S. Ninth District, HQ, and Atlantic Commands

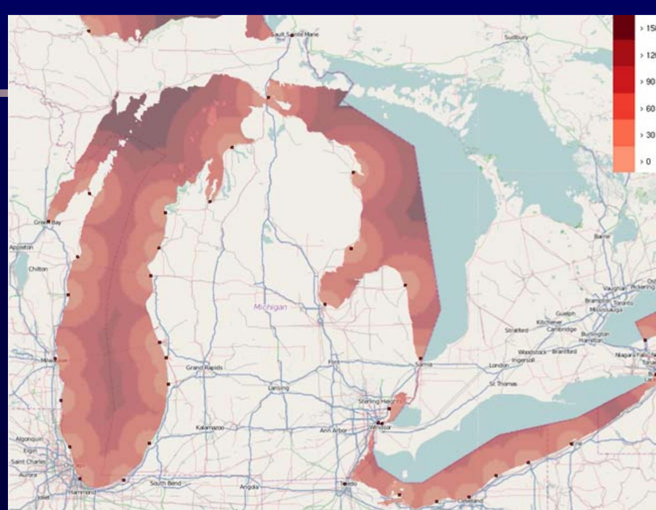
VACCINE July 2012

Example: USCG D9 Search And Rescue Operational Analysis





System Features: SAR Risk Profile



Time taken by CG stations to deploy an asset to the Great Lakes to respond to a SAR incident.

VACCINE

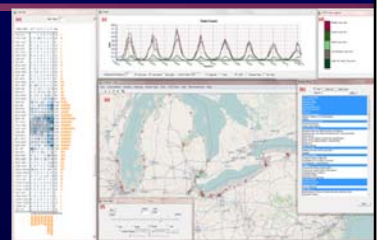
July 2012

U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7 (Operational Analysis) , USCG D9, USCG D5

IMPACTS:

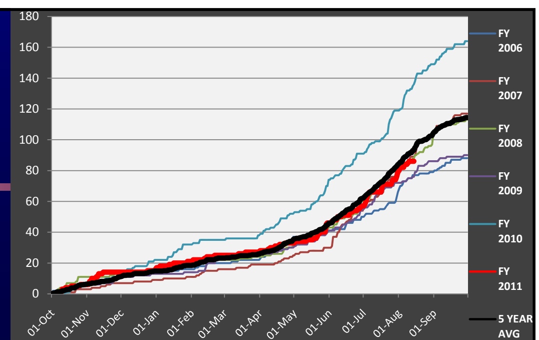
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided evidence of temporal and spatial patterns used in planning – new insights to SAR mission
- Hurricane Irene resource allocation decision based on cgSARva analysis and visualization
    - Highest SAR workload that weekend for D9



VACCINE

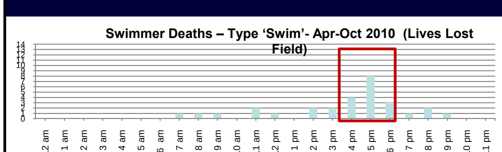
July 2012

U.S. Coast Guard Swimmer Death Analysis



Impact:

- Analyzed spatial and temporal patterns of shore-based and boat-based swimmer deaths to understand death dramatic increase in D9 in Summer 2010
- Provided information and visualizations used for public information campaign 2011 and for patrols 2011
- Significant decrease in deaths in 2011



Findings:

- Swimmer deaths
    - August highest frequency
    - Late afternoon highest frequency
    - Lake Michigan (south and west shore) have high concentration
- Boating deaths
    - Fri, Sat, Sun account for almost all deaths
    - Mid July to Mid August have highest frequency (only 1 week significantly high)
- 2009-2010 from MISLE Data
    - Large increase on Mon, Thu, Fri, Sun
    - Early and late season increase

VACCINE

Resource Allocation and Risk-Based Decision Making

- Explore risk-based decision making and utilize historical data for analysis and prediction
  - Total Risk, Mitigated Risk, Residual Risk
  - Explore 11 different USCG missions
  - Explore allocation of assets with different capabilities
  - Explore staffing, utilization, assets vs. risk measures
  - Perform What-If scenarios

VACCINE

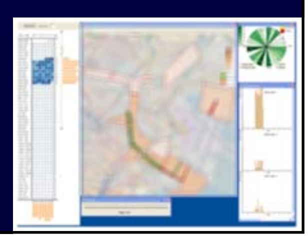
July 2012

USCG PWCS PROTECT Project

Partners: USC CREATE, USCG RDC, USCG D1, USCG LANT, USCG PAC

Impact:

- Provided insight and analysis into historical security patrols in Boston
- Analyzed developing PROTECT model and provided pattern analysis to improve model during operational deployment in Boston and for improvements for deployment in NYC
- Developed end-user tool for evaluation and validation of patrol routes



VACCINE

July 2012

USCG Port Closure Economic Impact VA

Partners: USC CREATE, USCG RDC, USCG D7, USCG LANT

IMPACT:

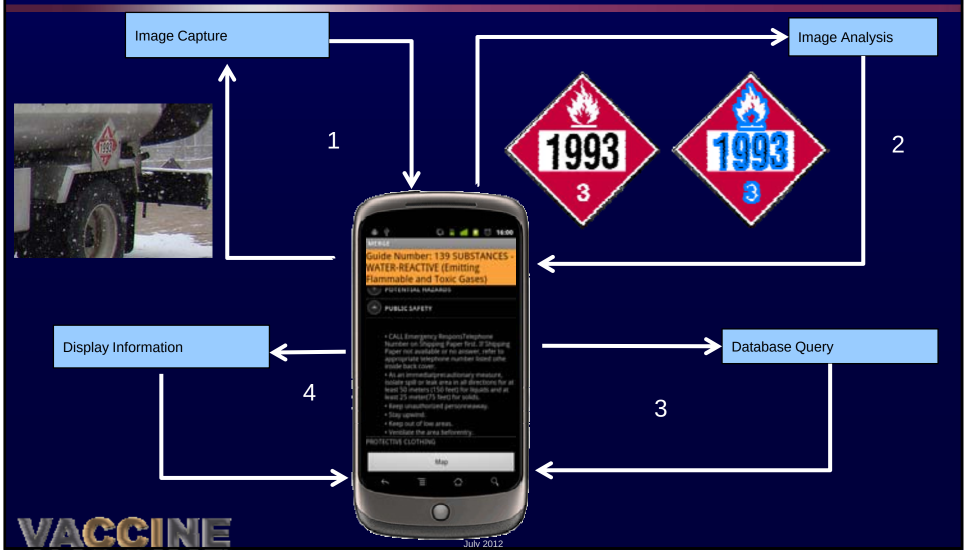
- Provided tool for use analysis and planning for impact of port closure in Port Arthur, Tx
- Economic sector impact, local and national impact
- Impact and effectiveness of alternative mitigation strategies

VACCINE

Example Public Safety Projects:

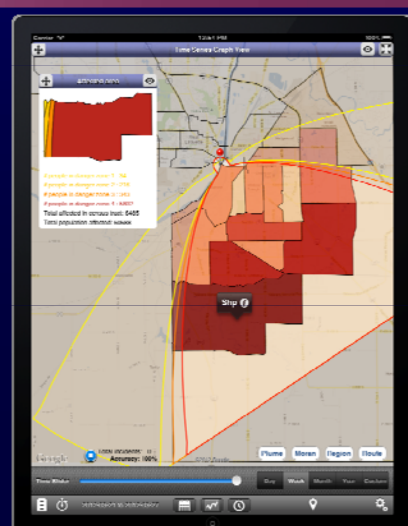
- Evacuation planning and management
- Hazardous Materials - mobile imaging
- Law enforcement visual analytics
- Additional projects
  - GARI – gang graffiti
  - Jigsaw

MERGE: Mobile Emergency Response Guide – System Overview

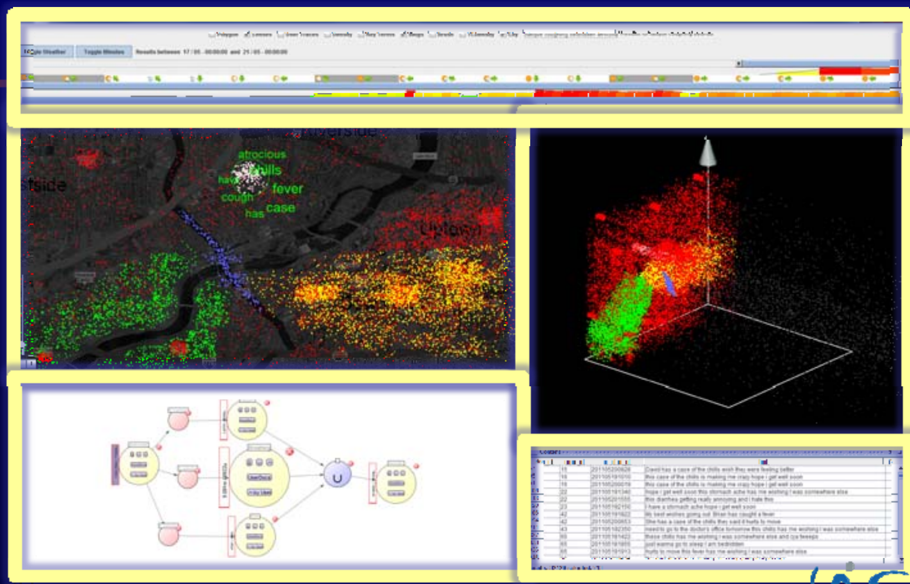


MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

- Chemical release plume modeling identifies census tracts with the highest number of expected people affected



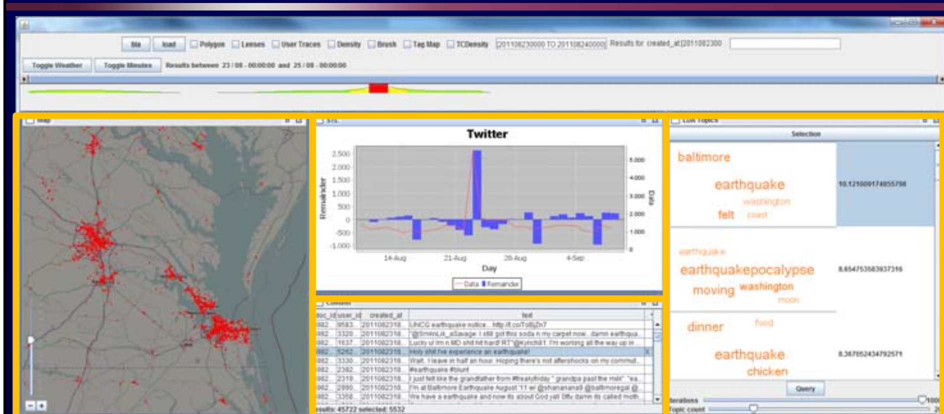
Scatterblogs: Geo-Spatial Document Analysis



VACCINE

July 2012

Spatiotemporal Social Media Analytics for Abnormal Event Detection



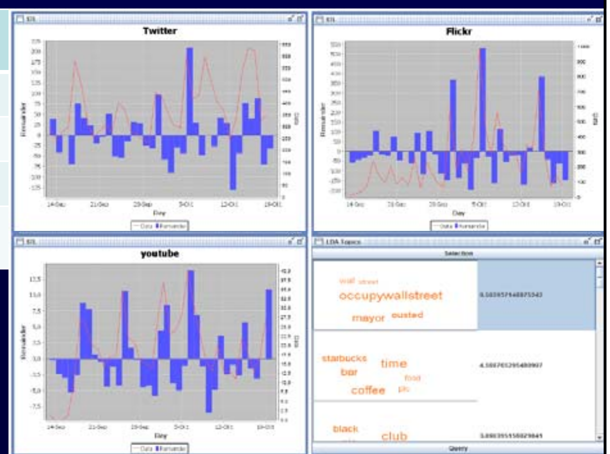
VACCINE

July 2012

Correlation Between Multiple Social Media

- Occupy Wall Street event

Date	Twitter (z-score)	Flickr (z-score)	Date	YouTube (z-score)
9/17, 2011	2.53	0.97	9/18, 2011	5.98
9/30, 2011	3.20	3.50	10/2, 2011	1.38
10/5, 2011	3.17	3.36	10/6, 2011	2.75
10/15, 2011	4.90	4.53	10/15, 2011	2.19



VACCINE

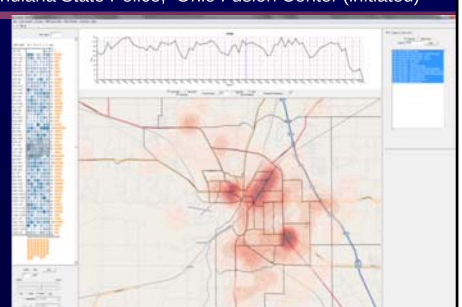
July 2012

Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

End-Users: Lafayette, WL Police, Indiana Fusion Center, Indiana State Police, Ohio Fusion Center (initiated)

Impacts:

- In use to analyze crime patterns in Lafayette, Indiana and connect strings of activities
- Mobile version being released to public (November 2011) for community-based policing
- Investigating correlation of bus routes and crime, street lights and crime
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning

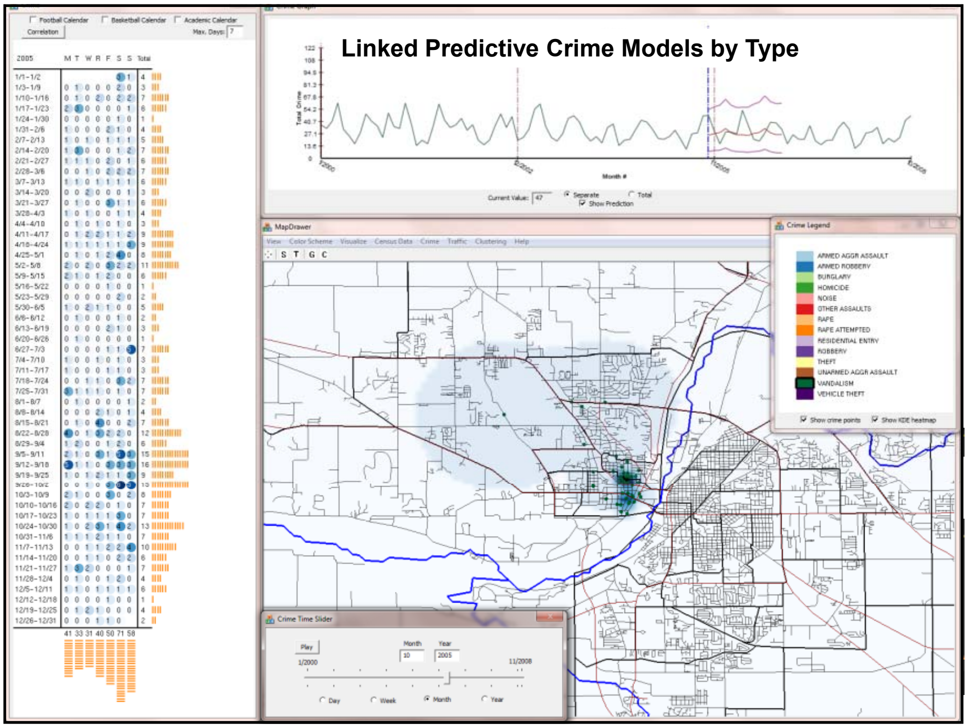


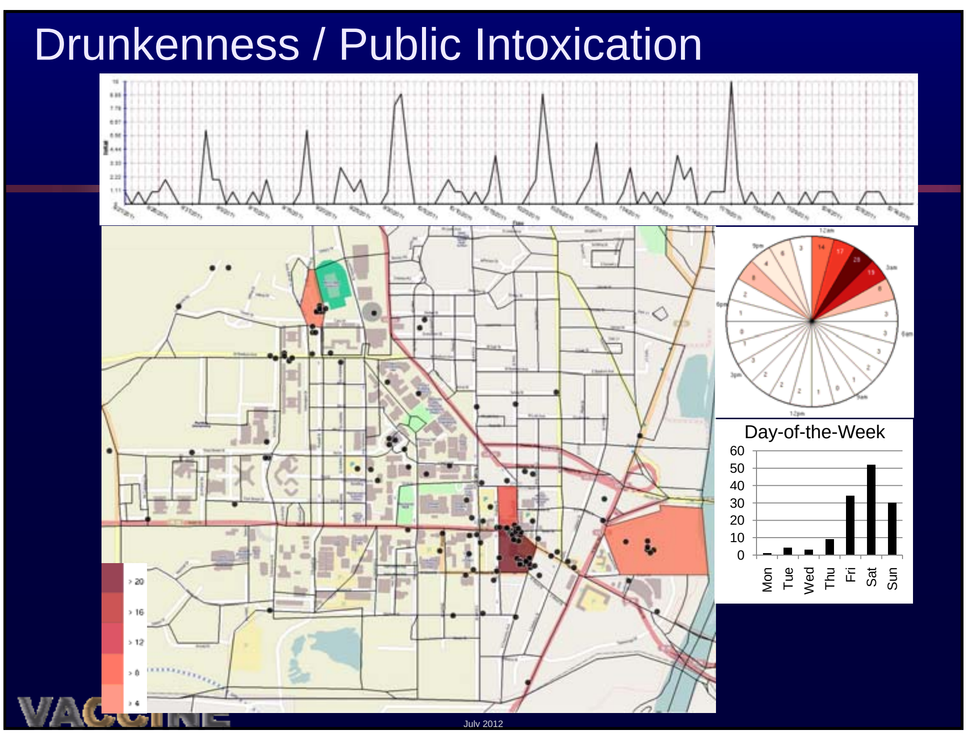
VALET delivered:
- Spring 2011: WL, Lafayette Police

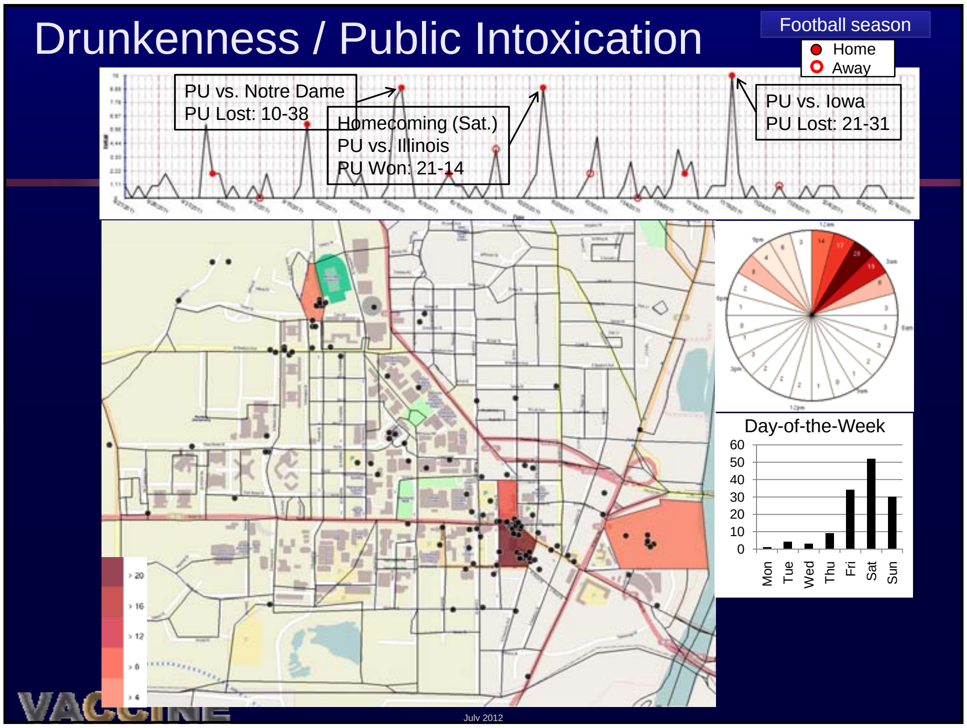
iVALET delivered:
- October 2011: Purdue, WL Police

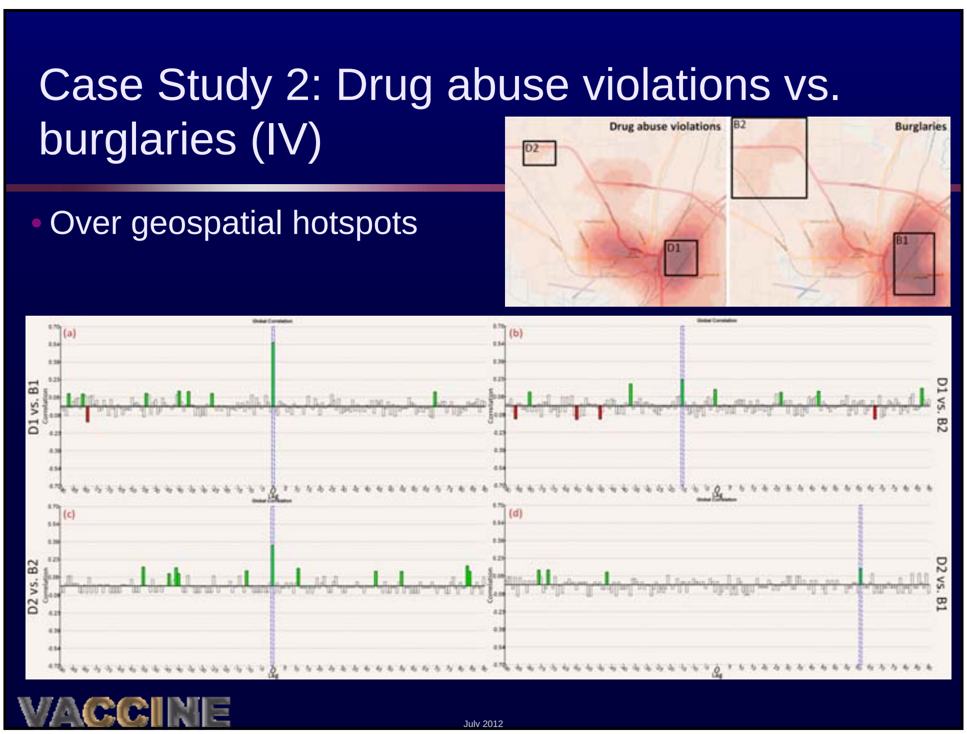
VACCINE

July 2012

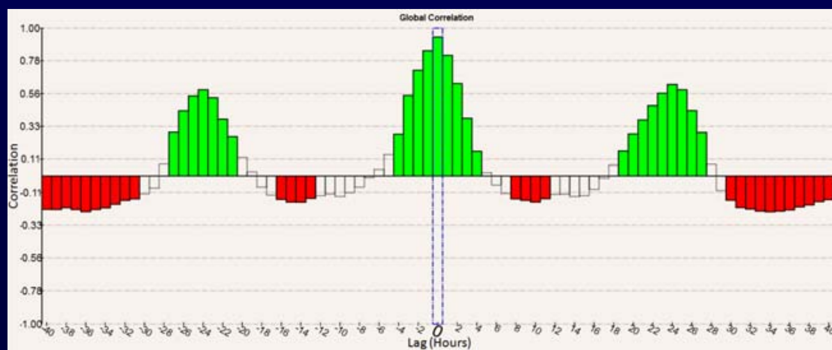

Linked Predictive Crime Models by Type
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
MapDrawer
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Crime Time Slider
Play
Month
Year
Day
Week
Month
Year


Drunkenness / Public Intoxication
Day-of-the-Week
60
50
40
30
20
10
0
Mon
Tue
Wed
Thu
Fri
Sat
Sun
> 20
> 16
> 12
> 8
> 4
VACCINE
July 2012


Drunkenness / Public Intoxication
Football season
Home
Away
PU vs. Notre Dame
PU Lost: 10-38
Homecoming (Sat.)
PU vs. Illinois
PU Won: 21-14
PU vs. Iowa
PU Lost: 21-31
Day-of-the-Week
60
50
40
30
20
10
0
Mon
Tue
Wed
Thu
Fri
Sat
Sun
VACCINE
July 2012

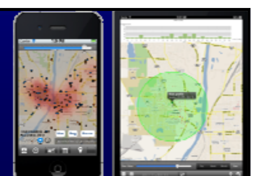

Case Study 2: Drug abuse violations vs. burglaries (IV)
• Over geospatial hotspots
Drug abuse violations
Burglaries
D2
D1
B2
B1
D1 vs. B1
D1 vs. B2
D2 vs. B2
D2 vs. B1
(a)
(b)
(c)
(d)
VACCINE
July 2012

Case Study 3: Drunkenness/public intoxication and noise complaints vs. traffic accidents (II)



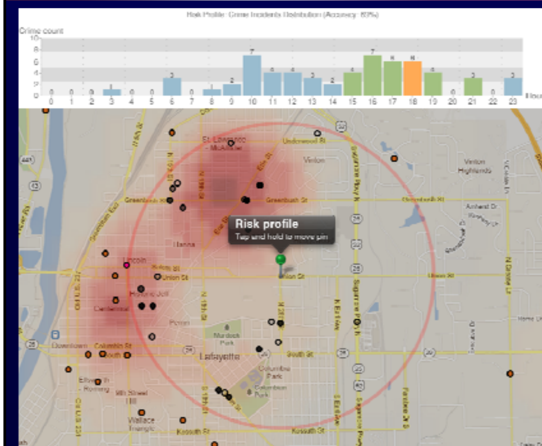
By Hour

iVALET



- Explore criminal, traffic and civil data on-the-go
- Risk assessment
- Use current spatial + temporal context into analysis

iVALET



- Linked views to explore multivariate spatiotemporal dataset
- Analytical tools to help explore data

VACCINE

Crime Analytics & SA Issues and Techniques

- Fuse data from a variety of sources
  - Law enforcement records management
  - Weather and phases of the moon
  - Street light locations, bus routes
  - Tracking release data of offenders
  - Civil court data
  - Social Media
  - Local event calendar
- Reliable predictive models
- Understandability and trust of predictions
- Main Question: What helps officers, detectives, chief do their variety of jobs?

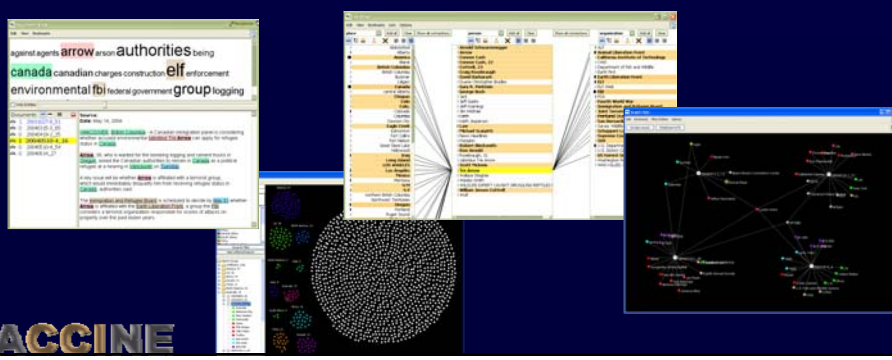
VACCINE

Jigsaw:
Visual Analytics for Investigative Analysis and Exploration of Document Collections

"Putting the pieces together"

Goal: Assist investigators with understanding, sense-making, and analysis of large, unstructured and structured document collections

Approach: Provide multiple visual perspectives on the documents and entities within them, highlighting connections between entities



VACCINE

Gang Graffiti Recognition and Analysis Using a Mobile Telephone (GARI)

Edward J. Delp, Mireille Boutin
Collaborating Institution(s): Purdue University
End-User(s): Gary Coons, Chief/Indianapolis Department of Public Safety Division of Homeland Security and IMPD, Indiana Fusion Center Gang Task Force

IMPACT:

- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be "interpreted"
- More than 40 users and 450 graffiti images acquired

GARI delivered:

- Summer 2011:
  - IMPD gang detectives
- August 2011:
  - IMPD at large
  - Ind Fusion Gang Task Force
- October 2011:
  - Gang detectives across Indiana

VACCINE

July 2012

Path

Introduction to our Center - VACCINE

Motivation and introduction to visual analytics

Example solutions we've developed

- Science visual analytics
- Health visual analytics
- Public safety
- Resource allocation

Visual analytic challenges and future

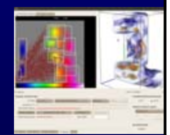
Six Challenges for Proactive & Predictive Visual Analytics

1. Creating computer-human visual cognition environments
2. Integrating interactive simulation and analytics
3. Solving specific scale issues and cross-scale issues
4. Managing uncertainty and time
5. Enabling risk-based decision making environments
6. Developing the Science of Interaction for Visual Analytics

VACCINE

July 2012

Visual Analytics At Real-World Scale



- Utilize advanced HPC techniques to enable interactive spatiotemporal analysis (spatiotemporal clustering, prediction)
- Cluster-based and cloud-based solutions
- GPGPU solutions
- Develop easily usable HPC visual analytic environments
- Example: Longhorn Exascale Visual Analytic Platform
  - 2048 compute cores (Nehalem quad-core)
  - 512 GPUs (128 NVIDIA Quadro Plex S4s, each containing 4 NVIDIA FX 5800s)
  - 13.5 TB of distributed memory
  - 210 TB global file system

The Science of Interaction

- Definition: The study of methods by which humans create knowledge through the manipulation of an interface.
  - …interaction and inquiry are inextricable. It is through the interactive manipulation of a visual interface – the analytic discourse – that knowledge is constructed, tested, refined and shared." Stasko et al., IV2009

5 Challenges:

1. **Visual Discourse -** Interactive visual thinking tools for the exploration, understanding, collaboration, description, explanation, decision, dissemination, persuasion of concepts & data.
2. **Multi-modal Sensemaking In The Large -** Enable large-scale, distributed teams to interactively make sense of big multi-modal data and problems.
3. **Fluid Interaction -** Designing fluid, high-bandwidth, and powerful interaction models and paradigms for the purpose of individuals reaching their full potential in viewing, analyzing, and understanding large and complex data.
4. **Collaboration –** beyond space and time- synergizing technologies and human users
5. **Mixed Initiative Data Discovery and Manipulation -** Balancing active user input with systematic guidance to enable visual data manipulation and analysis.

Visual Analytics: Remember…

- We need to be cognizant of parameters for visual representations
- Appropriate analysis can guide users to interesting features in the data
- Refined analysis through user interaction and their domain knowledge can help discover hidden problems
- There is no single catch-all visual representation or analysis

VACCINE July 2012

Keys for Success

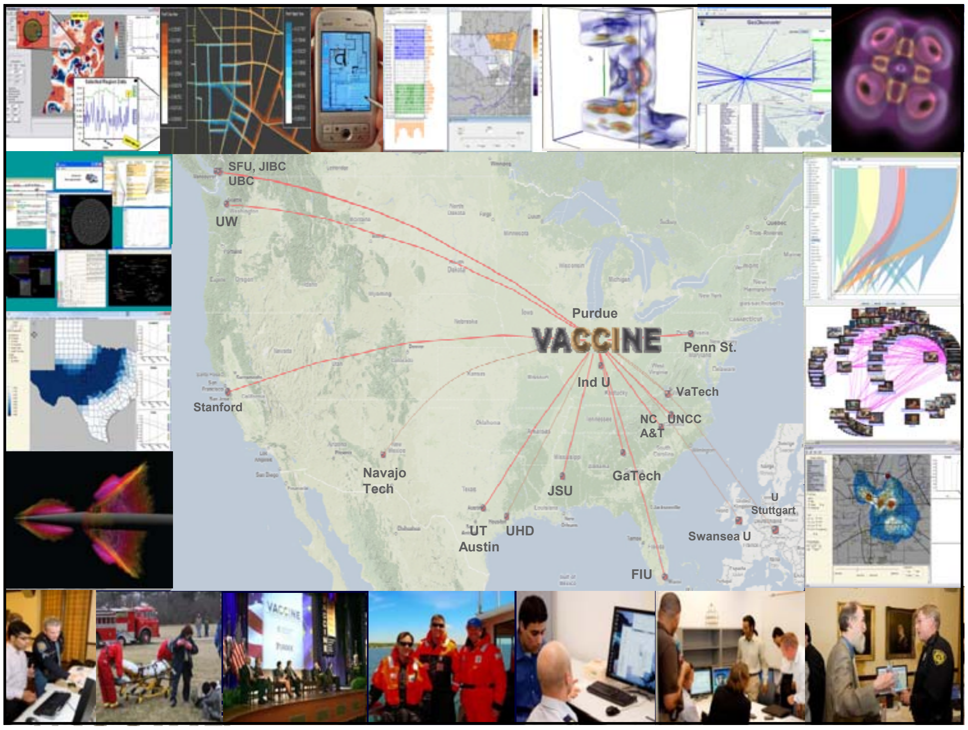
- User and problem driven
- Balance human cognition and automated analysis and modeling
  - Often applied on-the-fly for specific components identified by the user
- Interactivity and easy interaction
  - Utilizing HPC and novel analysis approaches
- Understandability of why predicted value is what it is
- Intuitive visual cognition
- Not overloaded with features

VACCINE July 2012

For Further Information

www.VisualAnalytics-CCI.org

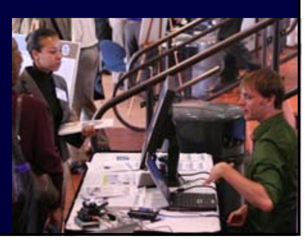
vaccine@purdue.edu
ebertd@purdue.edu



VACCINE Mission

- Provide visual analytic and scalable solutions for all 2.3 million extended homeland security personnel
  - 185,000 DHS personnel, 350,000 law enforcement personnel, 750,000 homeland security practitioners, 2.3 million extended personnel

- Achieve excellence in visual analytics and visualization sciences
- Educate homeland security stakeholders and the next generation of talent



July 2012

Who We Are: International Team of Experts - 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

VACCINE

July 2012

Some VACCINE Regional and Corporate Partners

- Indiana Department of Homeland Security
- Indiana State Department of Health
- Indianapolis Public Safety
- Charlotte Mecklenburg Police Department
- Lafayette, West Lafayette, Purdue Police Departments
- Tippecanoe County EMA & Sherriff's Department
- Harrisburg, PA Police
- Coast Guard Sector Boston, Seattle, LA
- Joint Harbor Operations Command Center
- Port of Seattle
- Indianapolis Public Safety
- USCG District 9, D5, D1
- Motorola
- Harris Corporation
- Boeing
- Next Wave Systems, LLC
- Banfield, The Pet Hospital
- Raytheon
- MacDonald, Dettwiler and Associates
- Oculus Info Inc.
- Kx Systems
- Bank of America
- Duke Energy
- World Vision International
- Gates Foundation
- Kimberly Clark
- Hallmark

VACCINE

July 2012

Table of Contents

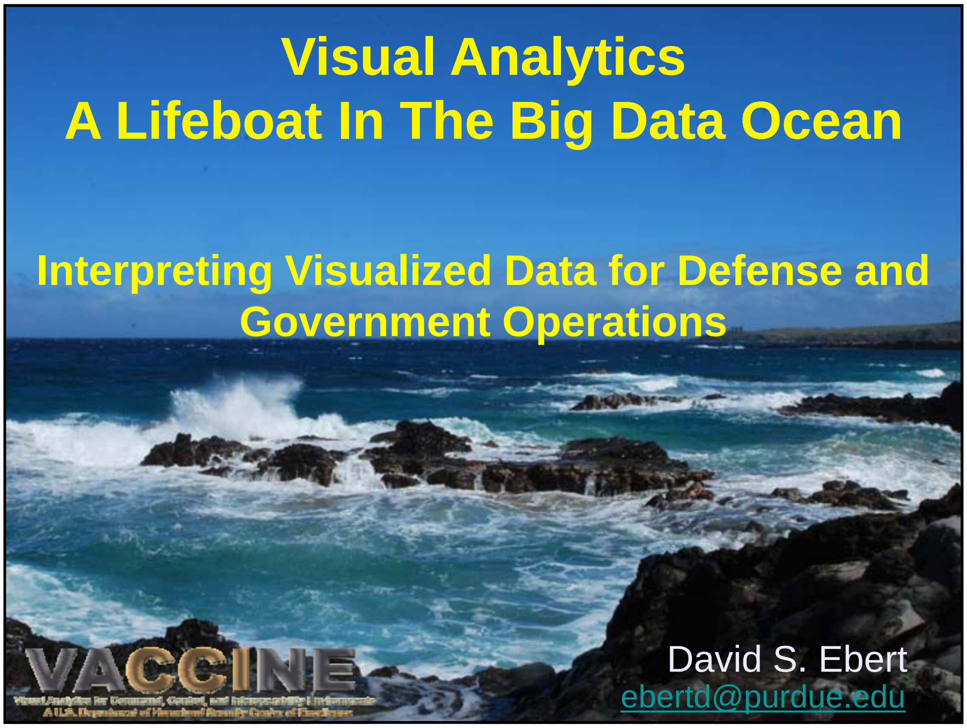
Some VACCINE Government Partners

- U.S. Coast Guard
- Federal Emergency Management Agency
- Customs and Border Patrol
- National Geospatial Intelligence Agency
- National Science Foundation
- Army Research Office
- Department of Defense
- Department of Health and Human Services
- U.S. Dept. of State
- National Institute of Justice
- National Maritime Intelligence Center
- Defence Research & Development Canada
- Foreign Broadcast Information Service
- US-CERT
- Oak Ridge National Laboratories
- US Army Corps of Engineers-ERDC
- Pacific Northwest National Laboratory
- ARL CERDEC & ARDEC
- US Attorney's Office (David Capp)

VACCINE

July 2012

Table of Contents



Big Data

- **Big Data Definition:** A phenomenon defined by the rapid acceleration in the expanding volume of high velocity, complex and diverse types of data. (TechAmerica Foundation)
  Big Data is often defined along three dimensions:
  - **Volume** – size
  - **Velocity** – rate of input, update, change
  - **Variety** – different types, sources, variables
- Need:
  - Advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of information (TechAmerica Foundation)
  - **Enable effective, efficient analysis, decision making, planning, and action**

VACCINE

February 2013

Big Data Challenges

- Heterogeneity
- Unprecedented volume and/or complexity
- Multivariate, Multidimensional
- Streaming
- Qualitative, Quantitative, Categorical, Temporal,…
- Distributed Sources: sensors, text, images, maps, …
- Inconsistent and Uncertain
- Different Scales and Cross Scale

Big Data Analytics – The Solution?

- Advanced data analytics for big data can help manage big data
  - Machine learning
  - Advanced statistical analysis
  - Hadoop-deployed analytics
- However, it falls short in several aspects

What Big Data Analytics Can't Do
(inspired and adapted from David Brooks, New York Times, 2/18/2013)

- Qualitative, fuzzy, and social data
  - Preferences, significance of one relationship over another
  - Trust
- Context
  - Data is rarely complete, nor does it incorporate all relevant information for decision making
  - Humans have extensive information and experience that never make it into the collected data

VACCINE February 2013

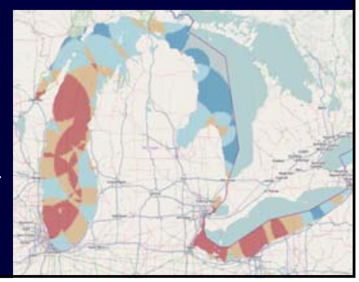
What Big Data Analytics Can't Do
(inspired and adapted from David Brooks, New York Times, 2/18/2013)

- Spurious vs. Significant
  - Big data means more statistically significant events and correlations, but they may not have any relevance
  - Increases noise to signal ratio
- Big problems
  - Complex, multifaceted, multiparameter big challenges with unquantified dependencies

VACCINE February 2013

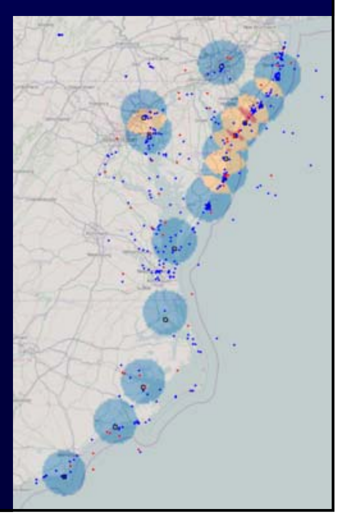
The Power of Visualization

- Turn the abstract into the concrete
- Enable rapid perception and visual cognition of information to increase efficient understanding and decision making
- Goal: Transform *data* into *information and knowledge*
- Data size doesn't matter, size of insights matters
- Visualization is a key method to increase understandability of data and analytical results



Visualization, Data, and Decision Making

- At some level most decision making is driven by data
- What role does visual representation play?
- Good or bad – people believe charts and graphics – "seeing is believing"



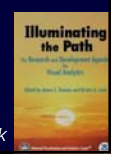
Beyond Visualization to Visual Analytics

Visual Analytics[1] is the science of analytical reasoning facilitated by interactive visual interfaces

Interactive visualization, data analysis, exploration, and decision making with human in the loop!

People use visual analytics tools and techniques to

- Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
- Detect the expected and discover the unexpected
- Provide timely, defensible, understandable assessments
- Communicate assessment effectively for action



1. Illuminating the Path: The R&D Agenda for Visual Analytics, Editors: Thomas and Cook

VACCINE

February 2013

VA Approaches to Big Data

- Don't display all the data!
- Extract relevant information at the appropriate natural scales using analytics and statistical models to reduce data space so that it is mappable to the visual space
- User interaction to guide above process
- Adapt to available hardware for the problem
  - Should we use big iron or cell phones?

VACCINE

February 2013

Visual Analytics' Value

Enable users to be more effective through innovative interactive visualization, analysis, and decision making tools

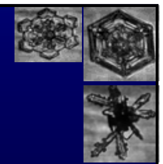
- Provide the right information, in the right format within the right time to solve the problem
- Turn data deluge into a pool of relevant, actionable knowledge
- Enable user to be more effective from planning to detection to response to recovery
- Enable effective communication of information

Approach: End User-driven solutions and research

Examples And Visual Analytic Solutions

- Example 1: Basic question of understanding why and when a cumulus cloud will produce precipitation
  - Data: Multiscale, multisource, incomplete data
- Example 2: Predictive crime analytics – when and where will crimes occur today
  - Data: A few million crime records with uncertain, missing data, dozens of fields, associated risks, related social media data, images
- Example 3: US Coast Guard decision making for efficiency and effectiveness
  - Data: Mission record, risk models, costs, weather data

Atmospheric Science: Multi-scale Interactions (in the words of a cloud physicist)



1 mm
1 kHz

No observing platform can measure the quantities of interest over all needed spatial and temporal scales needed

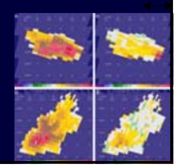
No numerical model can simulate the quantities of interest over all needed spatial and temporal scales

→We observe/simulate over a subset of the pertinent scales, using different instruments/models, and must assimilate these results to understand the "big picture"

Visual analytics is crucial for this task

1km
5min

Issues: Multi-scale, multi-system, multisource, massive, data & simulations



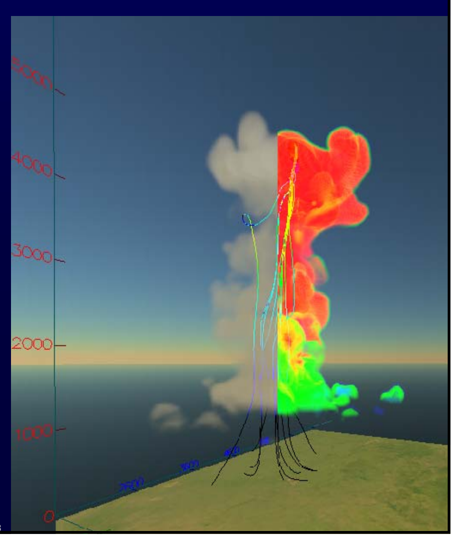
Simulation grid
90 billion cells

VACCINE

February 2013

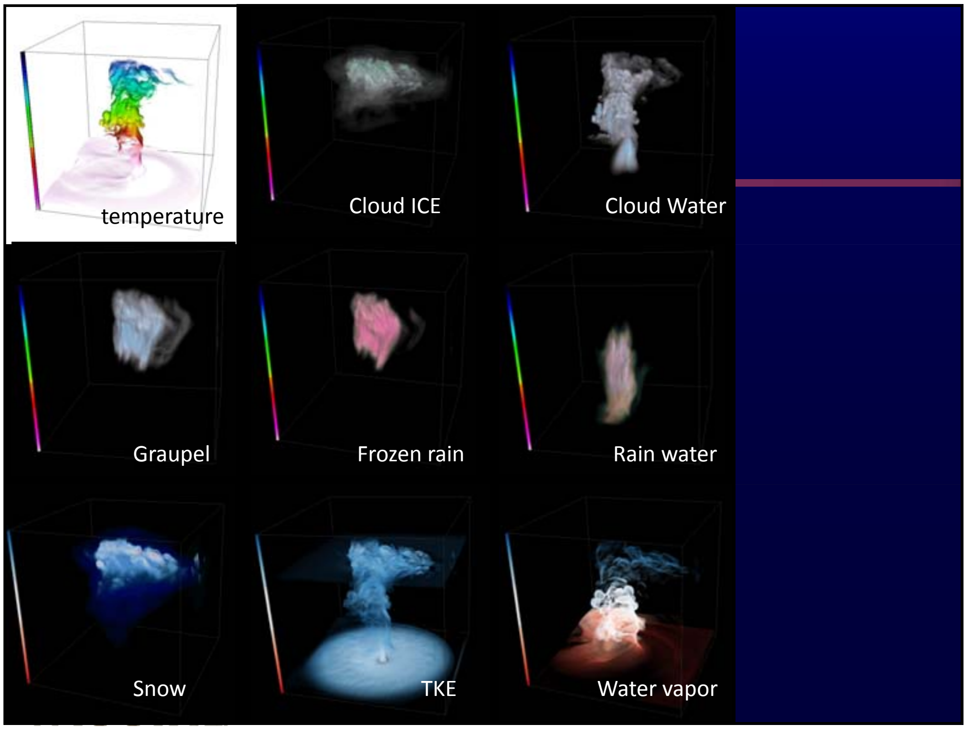
One Solution in Use: Our Atmospheric Visual Analytic Environment

- Utilize multiple rendering styles
- Provide interactive data exploration and user directed analysis
- Allow user specified analysis and queries on the fly
- Allow interactive correlative analysis of multisource data



VACCINE

February 2013



Crime Analytics Issues And Techniques

- Requirement: Fuse data from a variety of sources
  - Law enforcement records management
  - Weather and phases of the moon
  - Street light locations, bus routes
  - Tracking release data of offenders
  - Civil court data
  - Social Media
  - Local event calendar
- Need:
  - Reliable predictive models
  - Understandability and trust of predictions
- Main Question: What helps officers, detectives, chief do their variety of jobs?

VACCINE February 2013

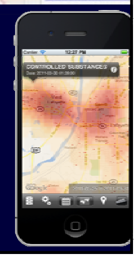
Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

Impacts:

- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (February 2013) for community-based policing
- Investigating correlation factors
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning
- Incorporating predictive alerts

VALET delivered:
- Spring 2011: WL, Lafayette Police

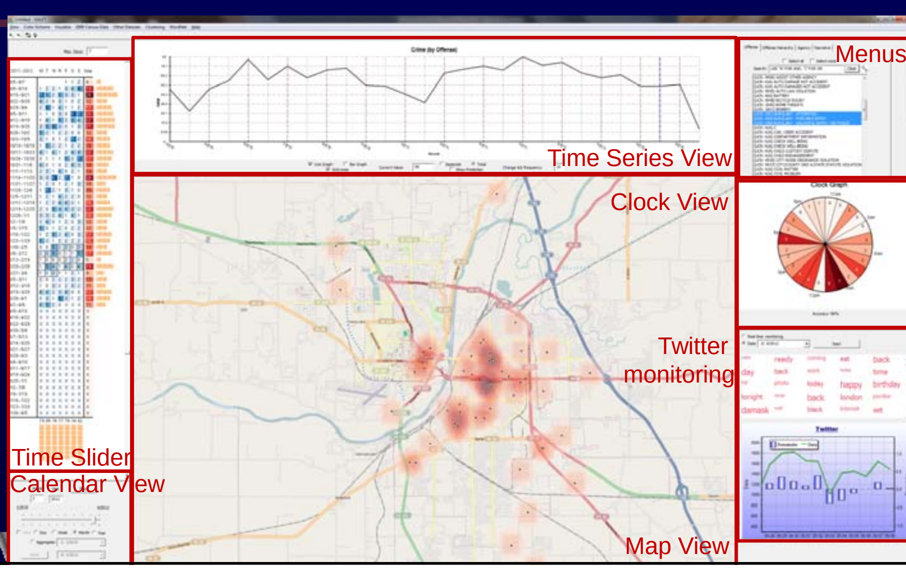
iVALET delivered:
- October 2011: Purdue, WL Police



VACCINE

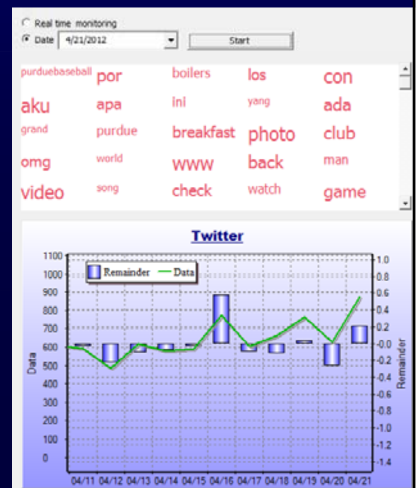
February 2013

VALET Overview



Social Media: Real-time Twitter Monitoring and Integration into Tools
(Purdue, Stuttgart, Penn St.)

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent Twitter feeds
- Topics extracted shown as a dynamic word cloud

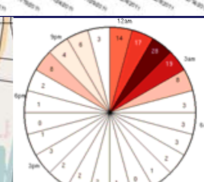


Grand Prix Weekend, Purdue University

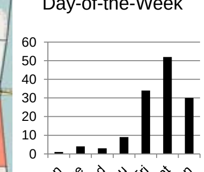
VACCINE

February 2013

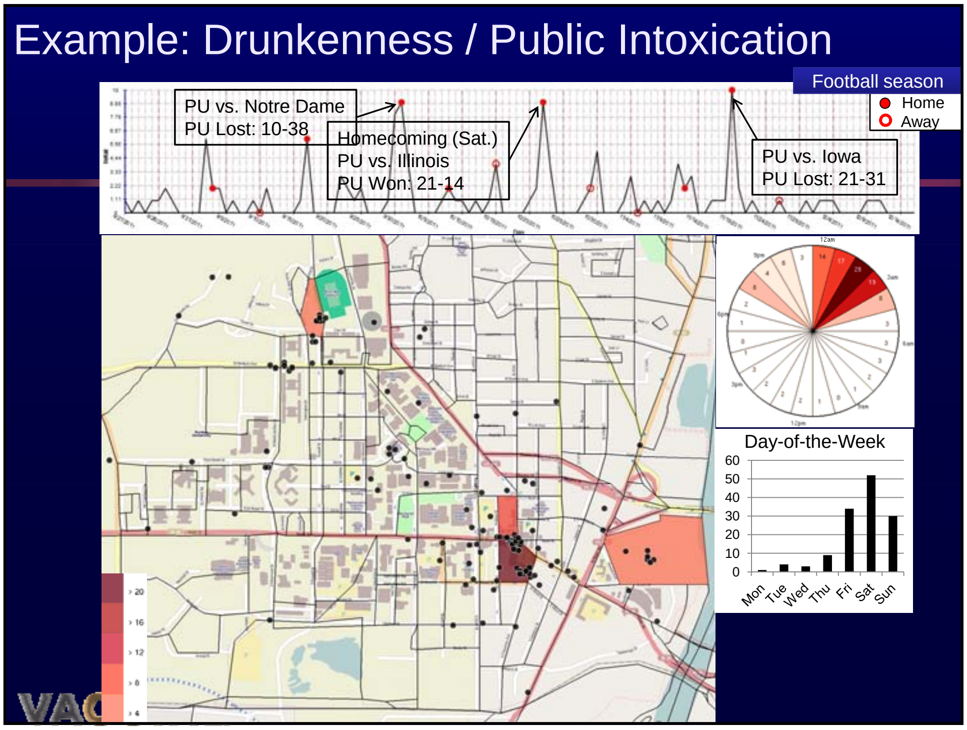
Example: Drunkenness / Public Intoxication

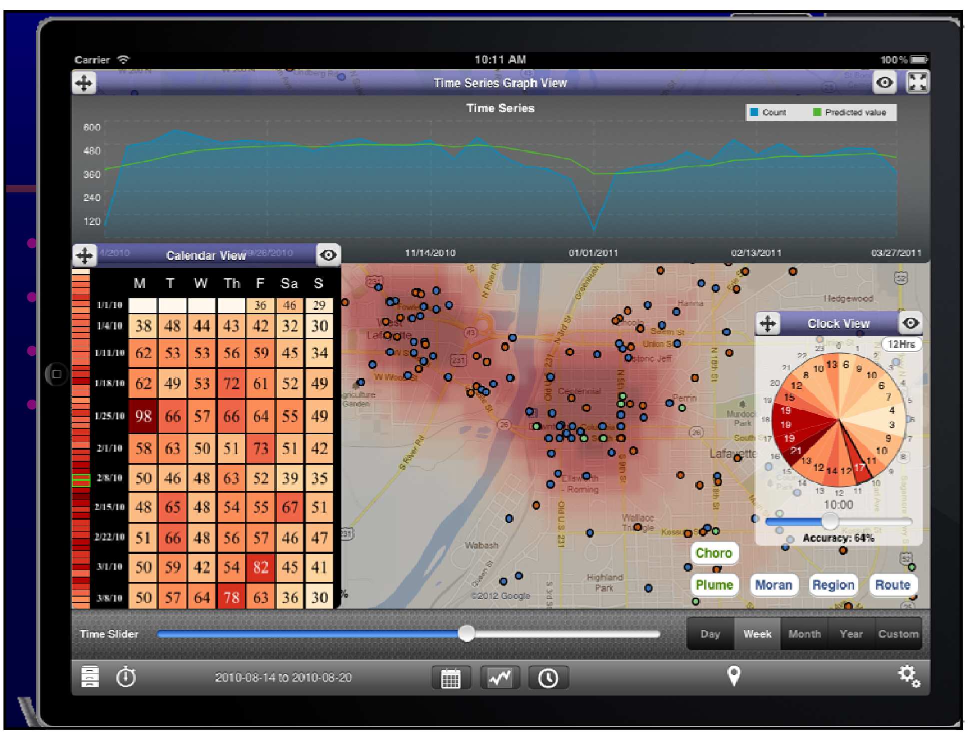


Day-of-the-Week



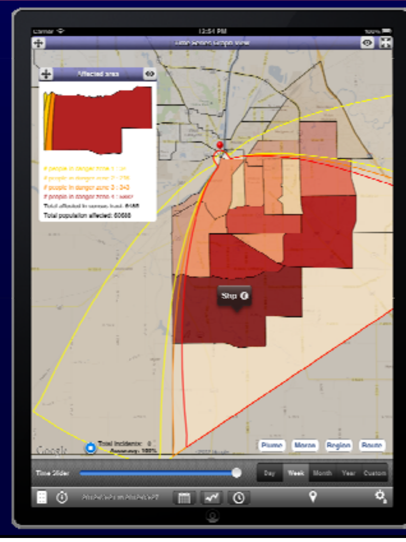
VAC


Example: Drunkenness / Public Intoxication
Football season
Home
Away
PU vs. Notre Dame
PU Lost: 10-38
Homecoming (Sat.)
PU vs. Illinois
PU Won: 21-14
PU vs. Iowa
PU Lost: 21-31
Day-of-the-Week
60
50
40
30
20
10
0
Mon
Tue
Wed
Thu
Fri
Sat
Sun


Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
Calendar View
11/14/2010
01/01/2011
02/13/2011
03/27/2011
M T W Th F Sa S
Clock View
12Hrs
10:00
Accuracy: 64%
Choro
Plume
Moran
Region
Route
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20

MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

- Chemical release plume modeling identifies census tracts with the highest number of expected people affected



VACCINE

February 2013

Gang Graffiti Analysis and Recognition Using a Mobile Telephone (GARI)

IMPACT:

- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be "interpreted"
- More than 60 users and 500 graffiti images acquired

GARI delivered:

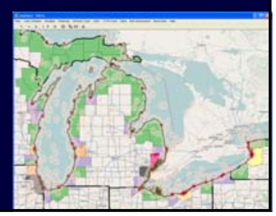
- Summer 2011:
  - IPD gang detectives
- August 2011:
  - IPD at large
  - Ind Fusion Gang Task Force
- September 2011:
  - Gang detectives across Indiana

VACCINE

February 2013

Visual Analytics Uses for Risk-Based Decision Making

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



VACCINE February 2013

USCG: Effective Risk-based Decision Making and Resource Allocation Visual Analytics

- Evaluate current and historical mission area:
  - Demands
  - Risks (total, mitigated, residual)
  - Resource allocation
  - Return on investment
- Evaluate courses of action
- Evaluate at both Strategic and Tactical/Operational level

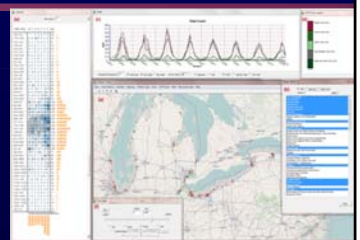
VACCINE February 2013

U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7, USCG D9, USCG D5, USCG HQ 771

IMPACTS:

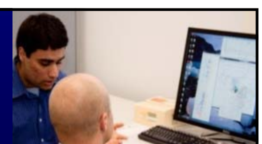
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Superstorm Sandy: Used for resource allocation in response and in rebuilding
- Used for Hurricane Irene resource allocation decisions
- Informed Commandant's budget testimony and recommendations to Congress
- Key component of USCG D9 reallocation plan for 2011-2012 based on decreased budget
- Key component of Coastal Operations Allocation Suite of Tools (COAST) – USCG HQ

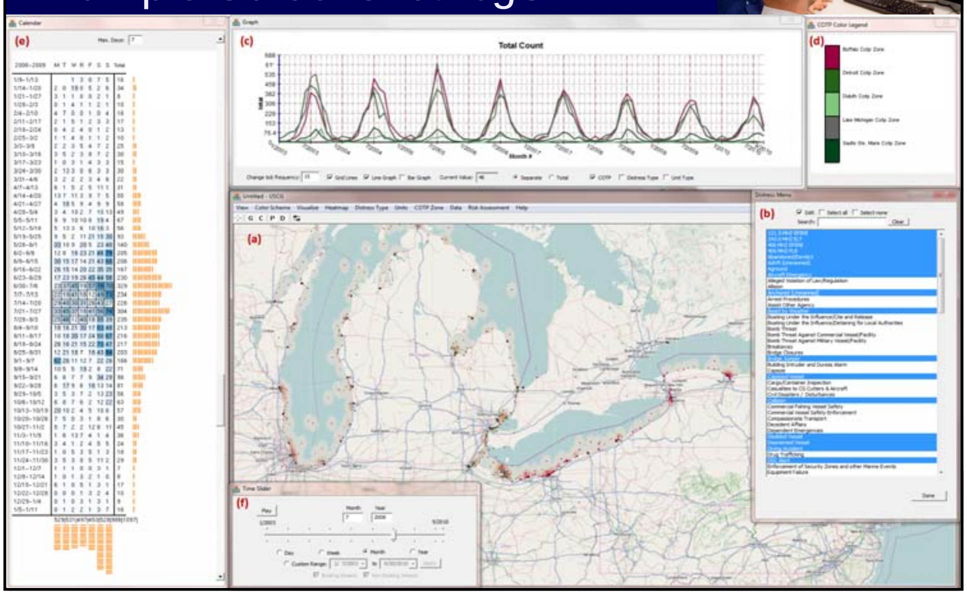


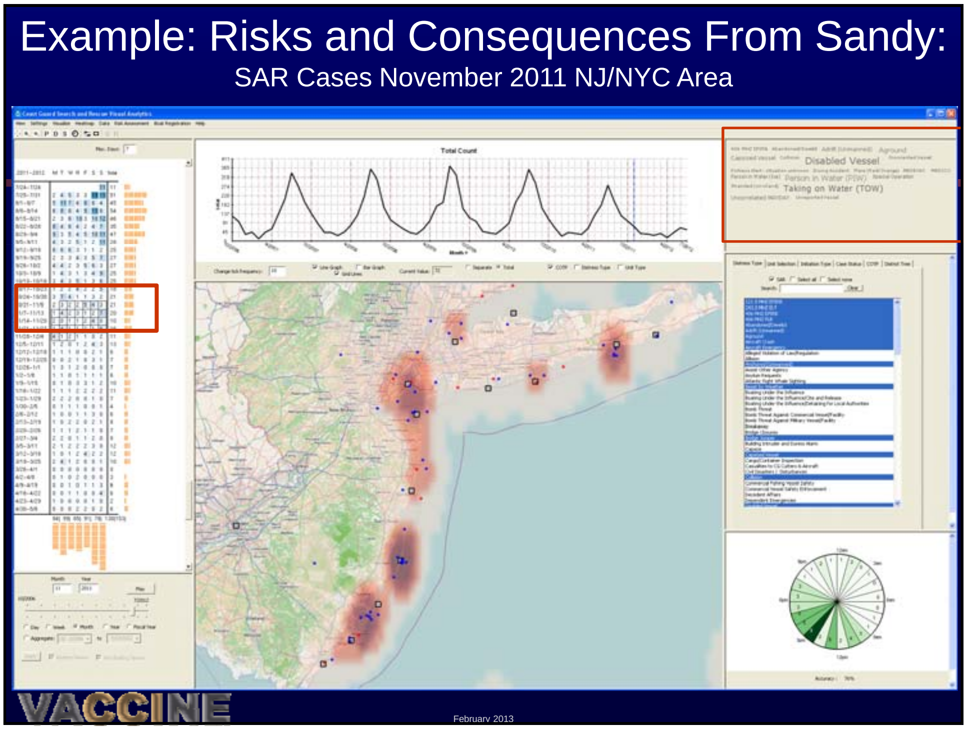
VACCINE

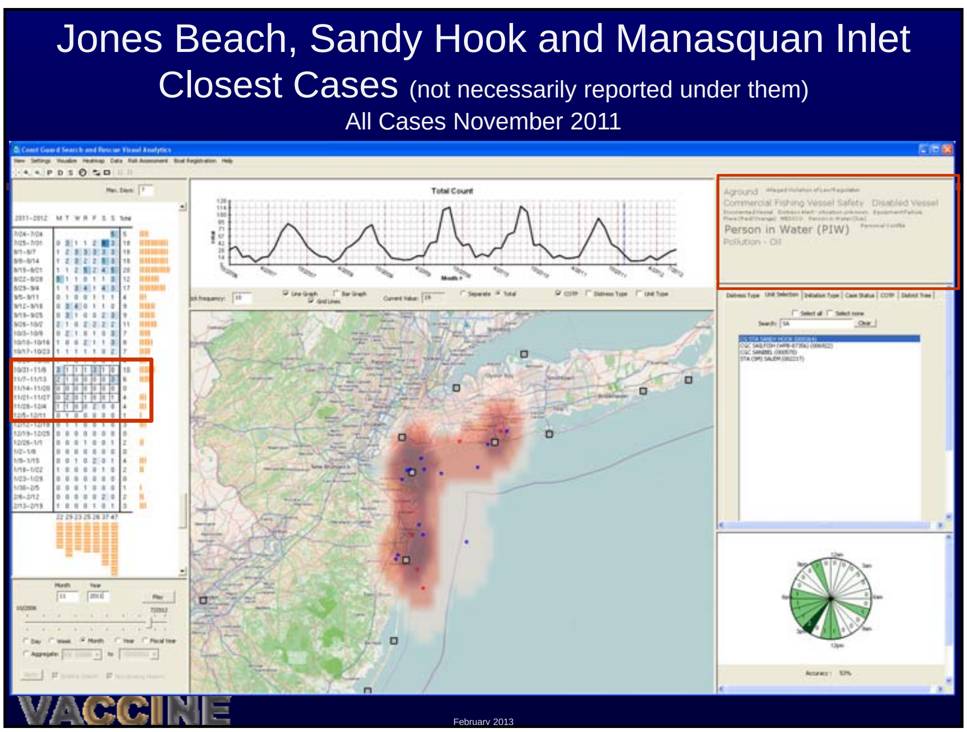
February 2013

Example Screenshot: cgSARVA

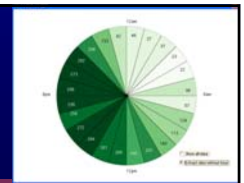





Example: Risks and Consequences From Sandy:
SAR Cases November 2011 NJ/NYC Area
Total Count
Disabled Vessel
Person in Water (PIW)
Taking on Water (TOW)
VACCINE
February 2013

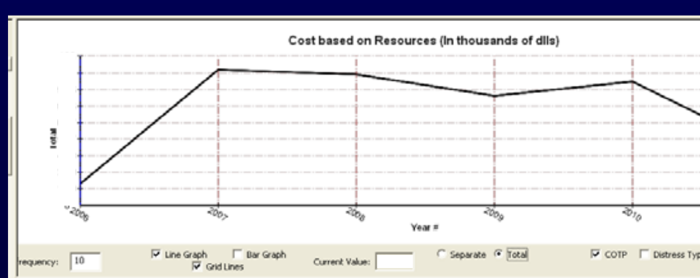

Jones Beach, Sandy Hook and Manasquan Inlet
Closest Cases (not necessarily reported under them)
All Cases November 2011
Total Count
Aground
Commercial Fishing Vessel Safety
Disabled Vessel
Person in Water (PIW)
Pollution - Oil
VACCINE
February 2013

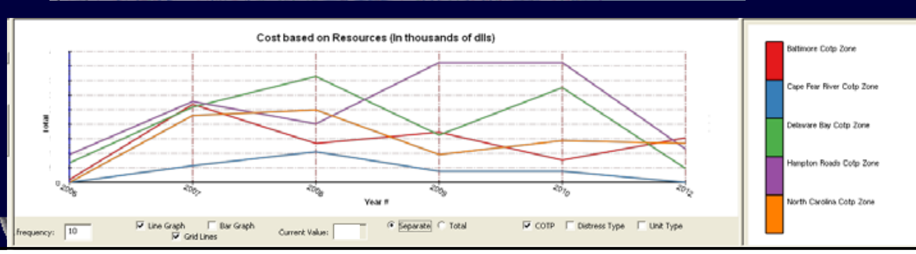
Risk-Based Allocations



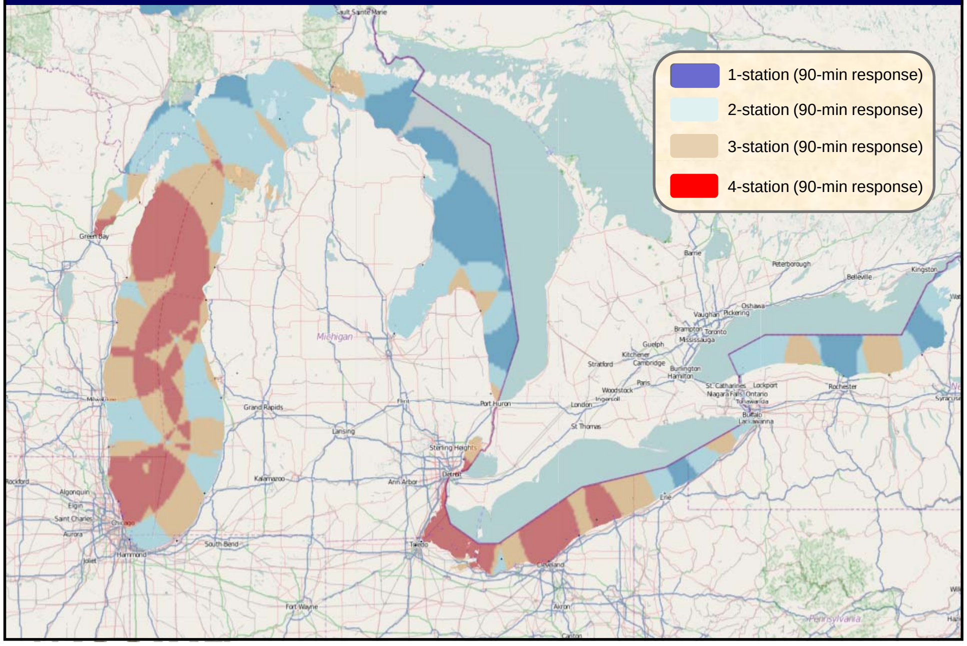
- Comparative visual analysis of mission cases/hours vs. staffing hours
- Comparative visualization of resources vs. risk
- Trend visual analytics
  - Increase/decrease in resource allocation
  - Increase/decrease in risk (total, mitigated, residual)
  - Increase/decrease in incidents
- Exploration of alternatives and effect on risk
- Predictive analytics based on historical data
- Effectiveness and efficiency measures, cost measures

Cost Based on Resources Used for the Open-Suspended Sorties Over 5 years (2006-2011)

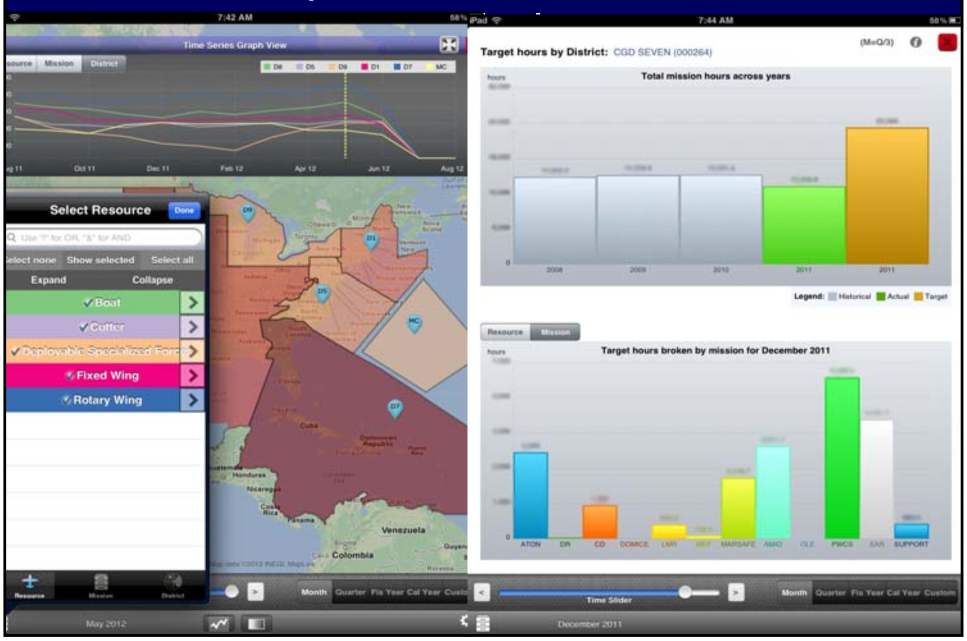




Response Efficiency – Notional Future Assets



Interactive Operational Performance



Visual Analytics: Remember…

- We need to be cognizant of parameters for visual representations
- Appropriate analysis can guide users to interesting features in the data
- Refined analysis through user interaction and their domain knowledge can help discover hidden problems
- There is no single catch-all visual representation or analysis

VACCINE February 2013

Keys for Success

- User and problem driven
- Balance human cognition and automated analysis and modeling
  - Often applied on-the-fly for specific components identified by the user
- Interactivity and easy interaction
  - Utilizing HPC and novel analysis approaches
- Understandability of why predicted value is what it is
- Intuitive visual cognition
- Not overloaded with features

VACCINE February 2013

Table of Contents

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu
ebertd@purdue.edu

VACCINE

VACCINE

February 2013

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

Table of Contents

Visual Analytics and Decision Making

David S. Ebert

ebertd@purdue.edu

What We Do

- *Enable effective decision making through interactive visual analytic environments*
- *Enable effective communication of information*
- *Provide quantitative, reliable, reproducible evidence*
- *Enable user to be more effective from planning to detection to response to recovery*
- *Enable proactive and predictive visual analytics*

Some Example Projects

- ***Science Visual Analytics** (weather, nanotechnology, cancer care)*
- ***Scalable Law Enforcement Projects***
- ***USCG Projects***
  - Search and rescue and swimmer deaths
  - PWCS analysis and economic impact analysis
  - Resource allocation and risk-based decision making
- ***Predictive spatiotemporal analytics – Indiana Intelligence Fusion Center***
- ***Multivariate visual analytics (business analytics)***
- ***Decision support with linked simulations***
- ***Investigative analytics (text and streaming data)***

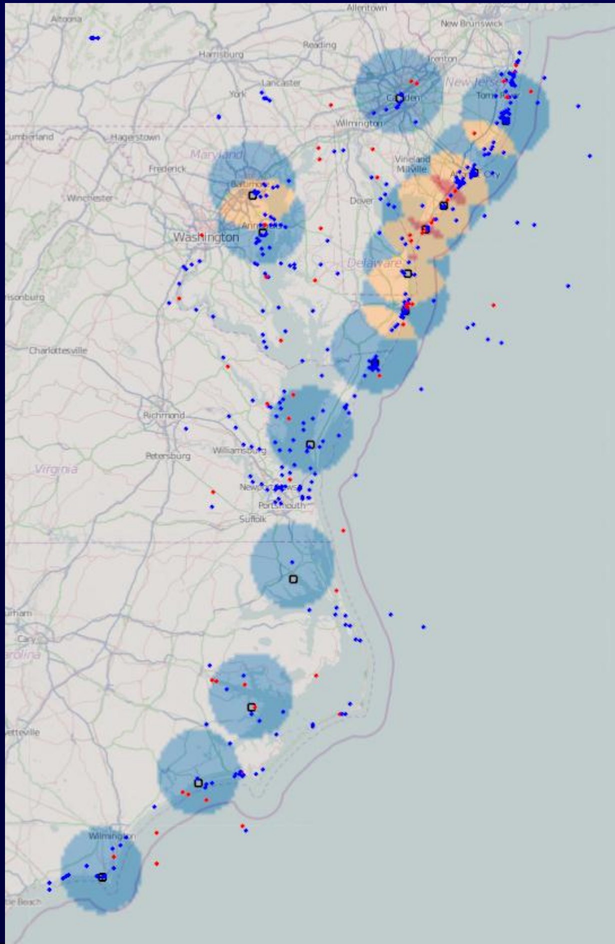
Visual Representations[1]

- *Visual representations translate data into a visible form, highlighting important features, such as commonalities and anomalies*
- *Visual representations make it easy for users to perceive salient aspects of their data quickly*
- *These visual representations augment the cognitive reasoning process with perceptual reasoning which enhances the analytical reasoning process*

1 – J. J. Thomas and K. A. Cook (Ed.). *Illuminating the Path: The R&D Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.

Visualization, Data, and Decision Making

- ***At some level most decision making is driven by data***
- ***What role does visual representation play?***
- ***Good or bad – people believe charts and graphics – "seeing is believing"***



Visualization, Data, and Decision Making

- *Unfortunately, it is easy to mislead with visual representations, even unintentionally!*
- *Fortunately, visual perception is the largest bandwidth into the human brain*
- *Visualization is an **effective, efficient communication mechanism** and a **common language***
  - *"Even a simple sailor like me can understand this" RDM Mike Parks, USCG D9*

Current: Utilization and Effectiveness

- ***At some level, most policies and decisions are data driven***
- ***Beginning to explore conveyance of decisions***
- ***Ease and effectiveness across scales of***
  - Decision makers
  - Affected population
  - Time for consideration decreases exponentially with increase in management level

Need: Interactive VA with integrated consequence modeling

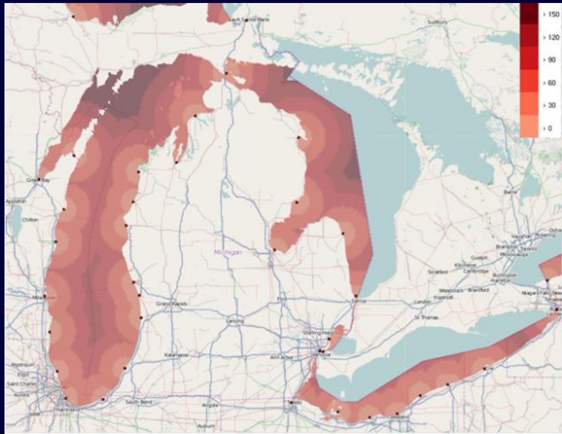


Table of Contents

Caveats

- ***Reliability and certainty must be conveyed***
- ***Someone must understand and verify the data interpretation***

Table of Contents

CVADA

Center for Visual Analytics and Data Analytics

A U.S. Department of Homeland Security Center of Excellence

Visual Analytics for Decision Making

David S. Ebert

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

April 2013

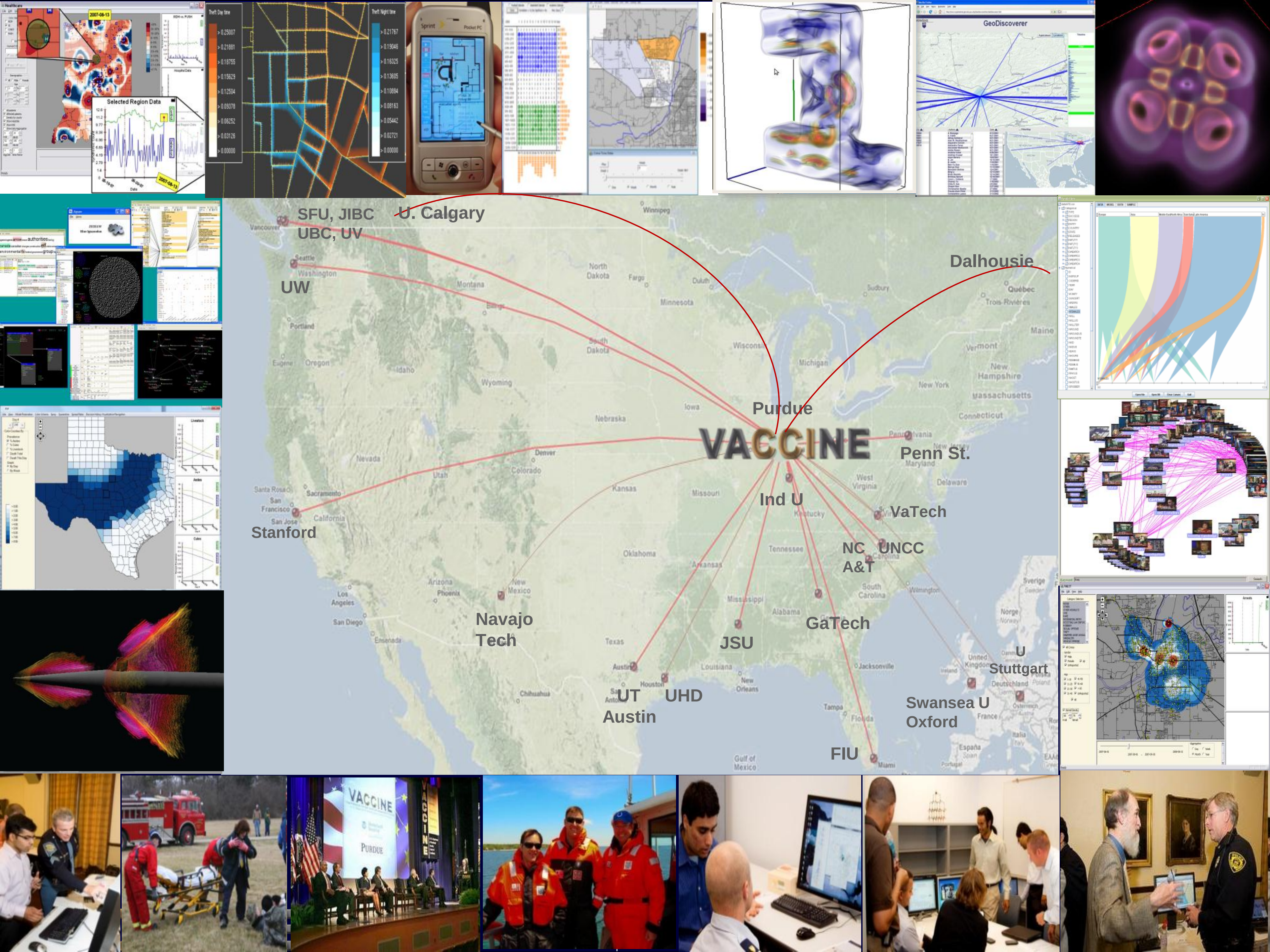
VACCINE:

Who We Are
&
What We Do

Who We Are: International Team of Experts - 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

VACCINE


SFU, JIBC
UBC, UV
U. Calgary
UW
Dalhousie
Purdue
VACCINE
Penn St.
Ind U
VaTech
NC A&T
UNCC
Stanford
Navajo Tech
JSU
GaTech
U Stuttgart
UT Austin
UHD
Swansea U
Oxford
FIU
GeoDiscoverer
Selected Region Data

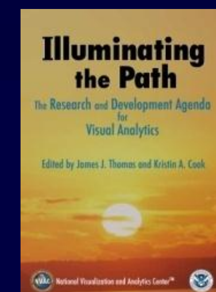
What We Do: Visual Analytics

Visual Analytics[1] is the science of analytical reasoning facilitated by interactive visual interfaces

People use visual analytics tools and techniques to

- Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
- Detect the expected and discover the unexpected
- Provide timely, defensible, understandable assessments
- Communicate assessment effectively for action

1. Illuminating the Path: The R&D Agenda
Analytics, Editors: Thomas and Cook



VACCINE

VACCINE Value to DHS/ HSE

- **Innovative Fielded Solutions:** We provide innovative visual analytic and scalable solutions to the extended homeland security community
- **Improved Effectiveness**: We enable users to be more effective through innovative, interactive visualization, analysis, and decision making tools
  - Provide the right information, in the right format, within the right time to solve the problem
  - Enable user to be more effective from planning to detection to response to recovery
  - Enable effective communication of information
- **People and Partnerships**
  - Interdisciplinary world-leading team of researchers and students
  - Actively Engaged Partners –
    We define and extend the new science of visual analytics driven by real-world, real-scale problems of engaged partners (local, state, federal)

VADM Robert Parker with VACCINE student researchers (cgSARVA, COAST, iOPAR)

"cgSARVA has proven its worth time and again, providing key analytic information for decision makers for large scale projects…"
VADM Robert Parker, 2012 MRS Keynote Address

Engaged End-Users

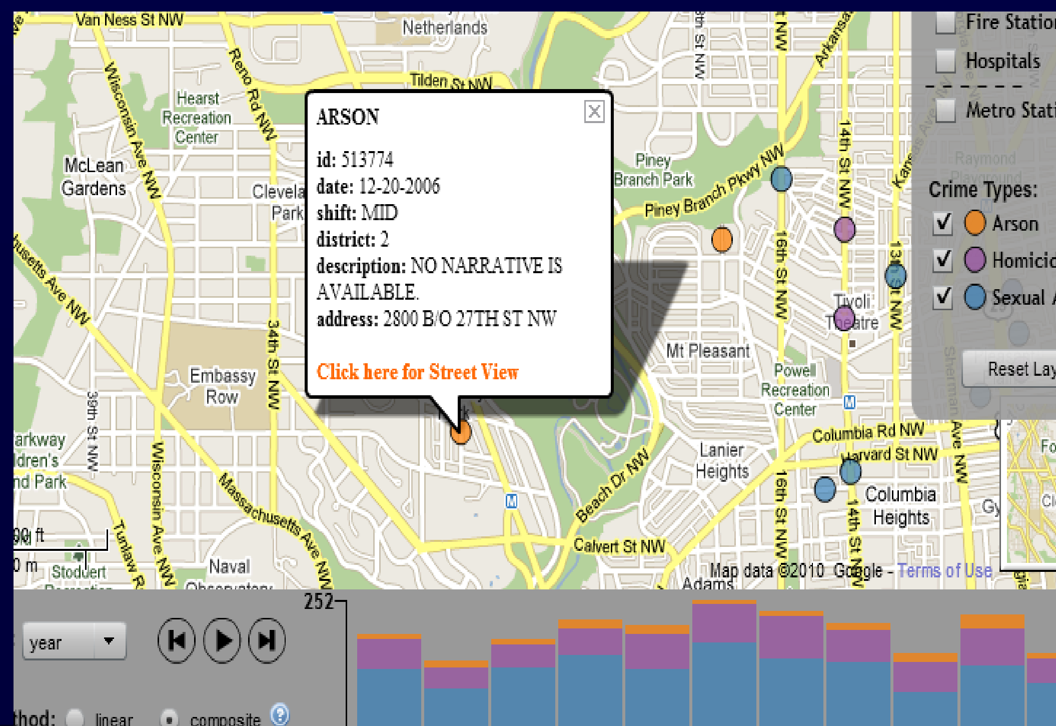
- Federal Operating Components:
  - US Coast Guard
  - US Transportation Security Agency
  - US Customs and Immigration Service
  - US Federal Emergency Management Agency
  - US Customs and Border Patrol
  - US CERT
  - US ICE (in progress)
- Law Enforcement
  - Over 40 local and state agencies
- Fusion Centers
  - Ohio (SAIC)
  - Indiana (IIFC)

VACCINE:

Our Work

Visual Analytics Uses for Crisis Management and Response

- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis
- Purpose:
  - Planning for resiliency
  - Long-term analysis
  - Predictive analytics
  - Training
  - Detection
  - Investigation
  - Response
  - Recovery, remediation



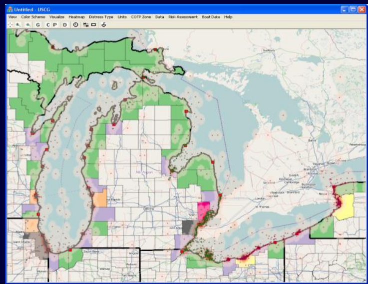
VACCINE

Situational Awareness, Planning, Investigation, Response Visual Analytics

- Situational Awareness and Assessment from massive data
  - In-field and desktop historical and current incident/crime information
  - Innovative spatiotemporal visualization, analysis, prediction
  - Image analysis tools
  - Integrated statistical forecasting and simulation models
  - Advanced social media analytics
- Tools for Intelligence analysis
- Planning and response support
  - Integrated simulation, resource allocation, census, weather data

Visual Analytics Uses for Public Safety

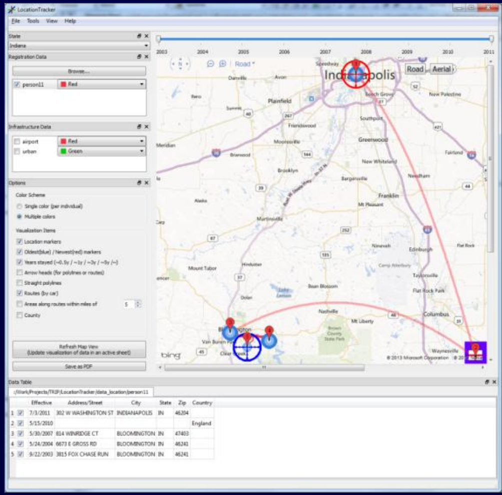
- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia massive data integration & analysis
- Visual standardization and interoperability



Current Public Safety Coalition Projects: Law Enforcement

Desktop crime analytics and intelligence analytics

- VALET & CrimeViz
  - Delivered to Lafayette PD 2011, WL PD Spring 2012
  - Delivered to Harrisburg Bureau of Police Spring 2011
- TRIP
  - Delivered to Indiana Intelligence Fusion Center May 2012
- Jigsaw
  - Delivered to West Lafayette PD, IMPD 09/2010, TSA 06/2012



In-field crime analytics

- Gang Graffiti Mobile Application (Gari)
  - Delivered to IMPD 09/2011, in use by 15 different police agencies in Indiana
- iValet
  - Delivered to WLPD, LPD, Purdue PD Fall 2011

Current Public Safety Coalition Projects: Emergency Management

Desktop

- Miami-Dade EOC Image-based Disaster Assessment
  - Delivered to EOC Spring 2011, students deployed 2011, 2012
- Flooding damage assessment (cooperation with CHC COE)
  - Used by CHC for assessment and by FEMA/DHS
- Visual Analytics in EOCs (partnership with DRDC Canada)
  - Used in Richmond, BC EOC and for exercises

In-field

- HazMat Visual Analytic Application (MERGE)
  - Delivered to Carroll County EMA, Purdue Police, Ind. Fire, 08/2011
  - Delivered to TSA July 2012

A Few Examples of Working with Stakeholders

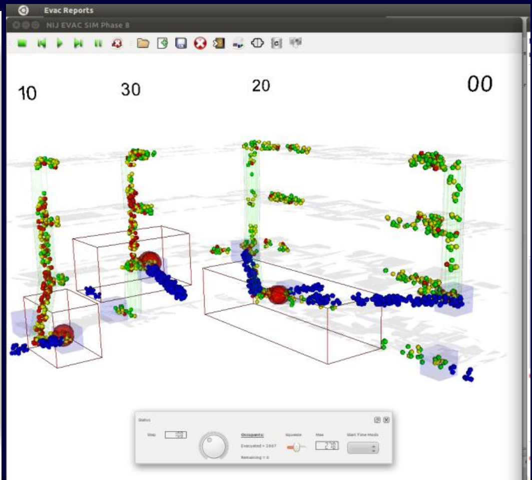
- GARI – initiated by Chief John Cox, Purdue Police
- MERGE – initiated by FLETC meeting
- VALET – initiated by Sheriff Tracy Brown, Tippecanoe County
- Jigsaw – initiated by briefs with intelligence analysts
- TRIP – initiated by Indiana Intelligence Fusion Center
- cgSARVA – initiated by Dr. Joe DiRenzo, Chief Operations Analysis, USCG LANT

VACCINE:

Our Projects

Example Projects

- VASA
- TSA
- Public Safety and Law Enforcement
- USCG
- Financial Fraud
- Cybersecurity



VASA: Visual Analytics for Security Applications

Cascading Critical Infrastructure Resiliency Modeling and Analytics

- Purpose: Apply visual analytics to the problem of monitoring and understanding cyber networks and critical infrastructures during detrimental cascading effects, and to the management of the ensuing crisis response.

- Collaborating Institution(s):
  Purdue, UNC Charlotte, U. Minn. (NCFPD), U. Konstanz, U. Stuttgart, Fraunhofer IGD, Siemens, German utilities

- End-User(s): Power Suppliers (e.g., Duke Energy),
  Cyber Community (e.g., Cisco),
  Quick Service Restaurants and suppliers

VASA

Image by Pak Wong

CI/KR's: Transportation, Food Safety, Cyber Networks and Power Grids at the national levels.

Key Outcome: Visual analytics prototype to analyze large ensembles of simulations (massive, dependant data) for a range of situations to fully understand areas of stability and key failure points.

Year 1:

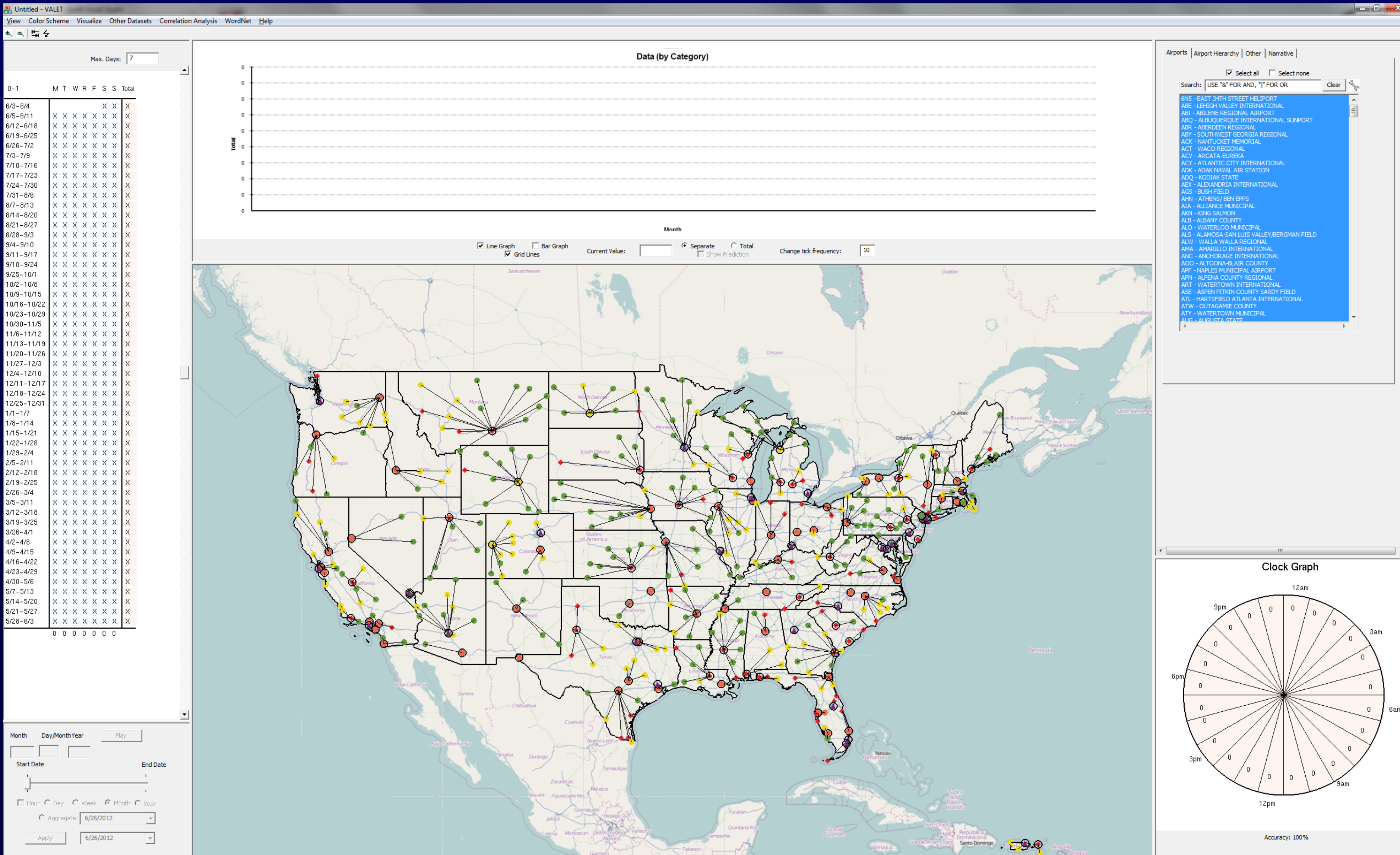
- Develop QSR model based on measurements and link in basic VA system prototype
- Develop utility model and powergrid interactions and link in basic VA system prototype

VACCINE

TSA Projects

- Visual Analytics for multifaceted operational analysis
  - Analysis and correlative analysis of
    - Operational performance (throughput per lane, use of AIT, targets)
    - Security performance (incidents)
    - Financial performance (expenses vs. performance metrics)
  - Spatial, temporal, spoke, hub, and regional analysis
- Customer satisfaction visual analytics
  - Jigsaw use to understand and find patterns in compliment/complaint database
  - Potential linkage to above data
- MERGE for rail inspectors

TSA Visual Analysis Prototype



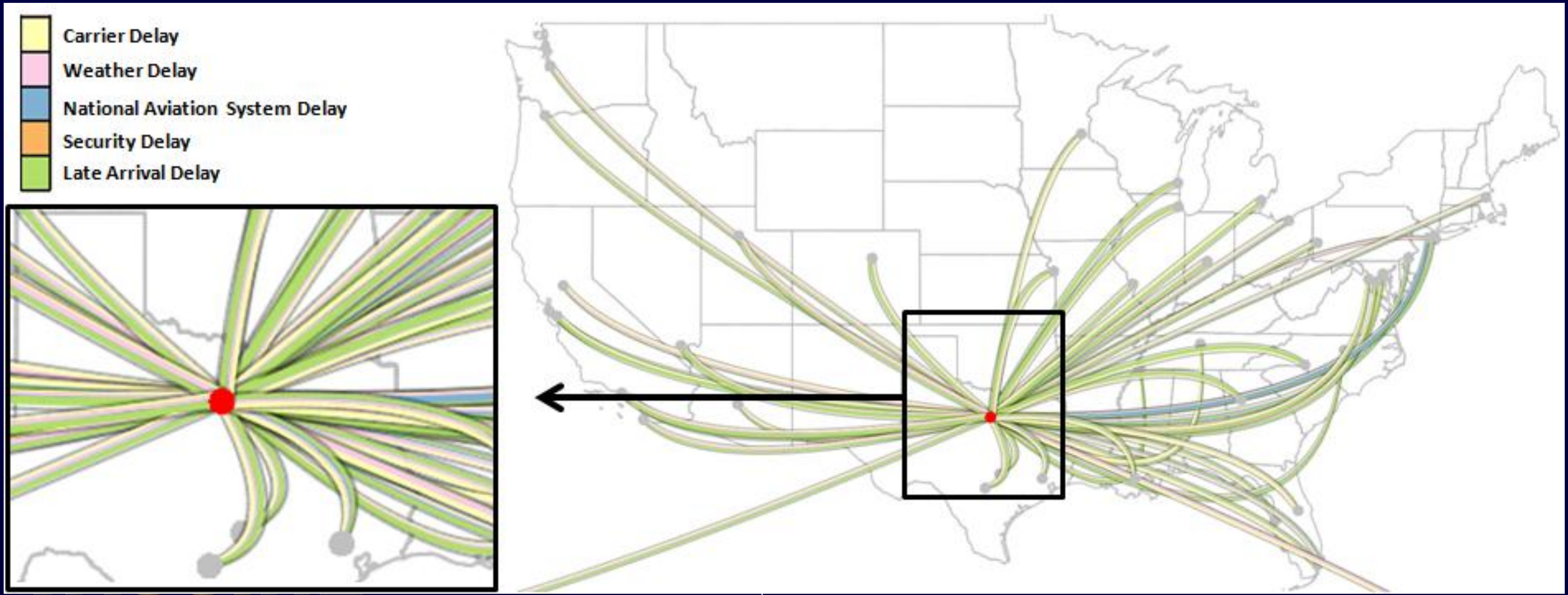
Category X (28) | Category I (57) | Category II (77) | Category III (127) | Category IV (157) | Hub (90) | Infrequent Public Charter | Coordination Centers (121) | Total (446)

? ?

Flight Delay Visual Analysis Prototype

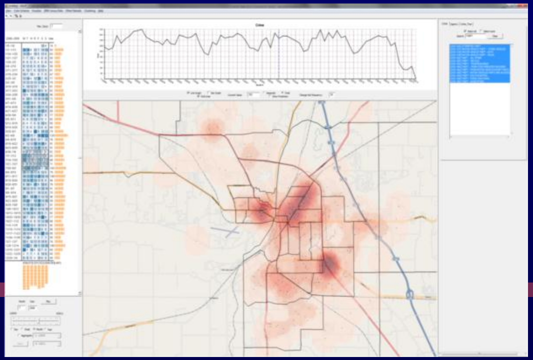
Flight Delay Visual Analysis Prototype

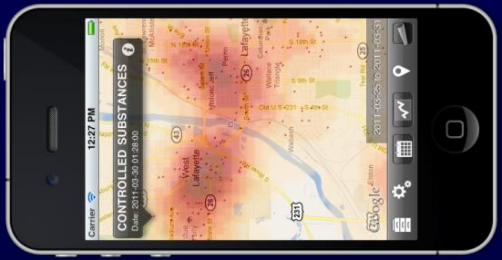
- Delay at DFW causing delays at destinations by delay type

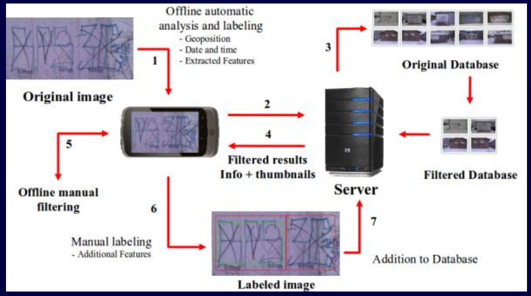


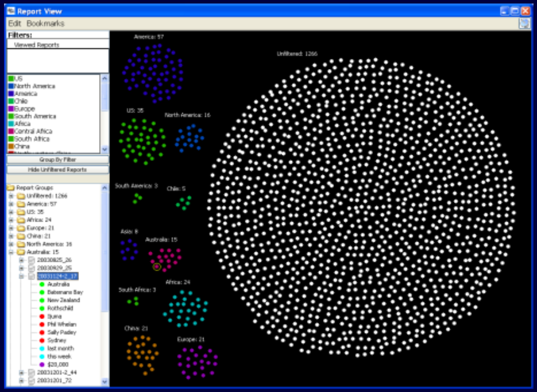
Public Safety & Law Enforcement

- Law enforcement visual analytics
  - VALET, iVALET (iPhone/iPad), CrimeViz
    - Visual Analytics Law Enforcement Toolkit
    - Analyzing crime patterns and time of day problems
- Gang activity analytics
  - GARI
    - Gang Graffiti Recognition and Interpretation using a mobile telephone
    - Allowing police to catalog and analyze gang graffiti images, better track and determine gang activities
- Document visual analytics
  - JIGSAW
    - Visualization for investigative analysis
    - Discovery of hidden relationship and threats across documents









Jigsaw

Visualization for Investigative Analysis across Document Collections

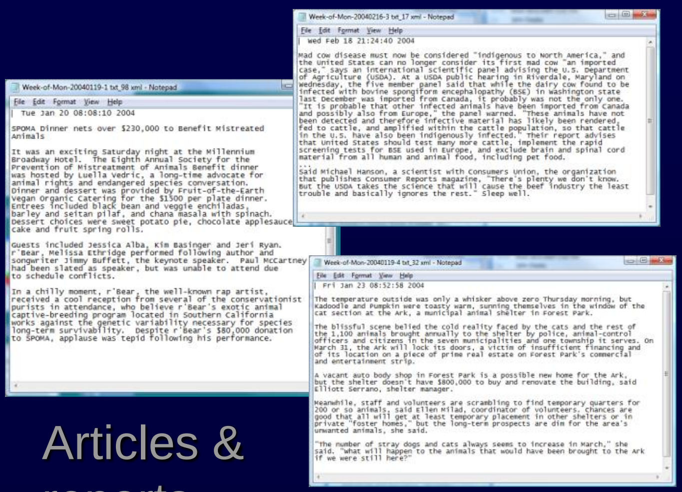
- Law enforcement & intelligence community
- Fraud (finance, accounting, banking)
- Academic research
- Journalism & reporting
- Consumer research

"Putting the pieces together"

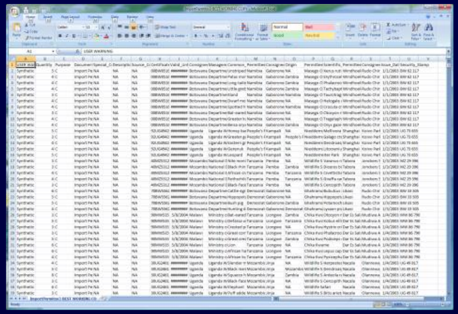
VACCINE

Jigsaw: Problem Addressed

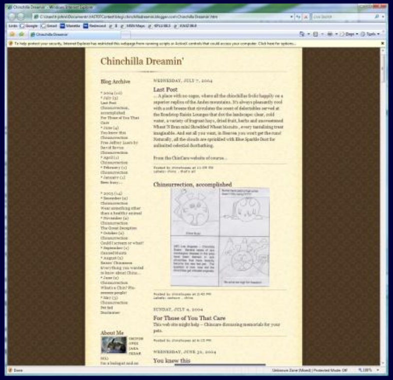
Help “investigators” explore, analyze and understand large document collections



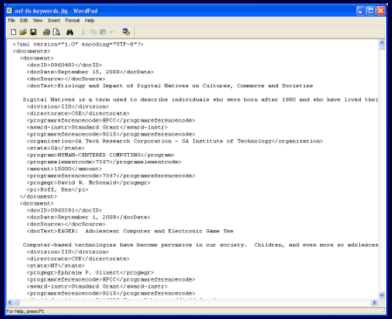
Articles & reports



Spreadsheets

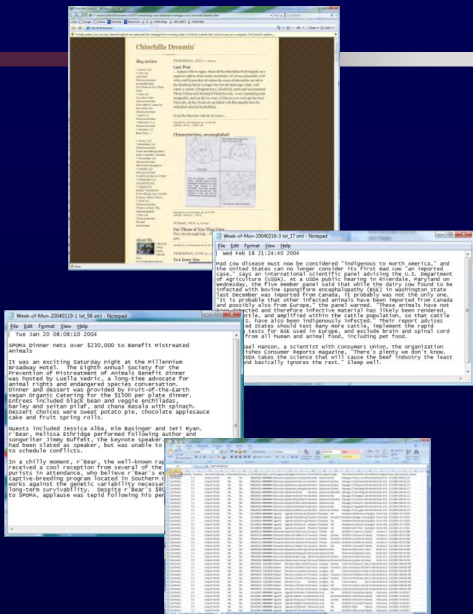


Blogs



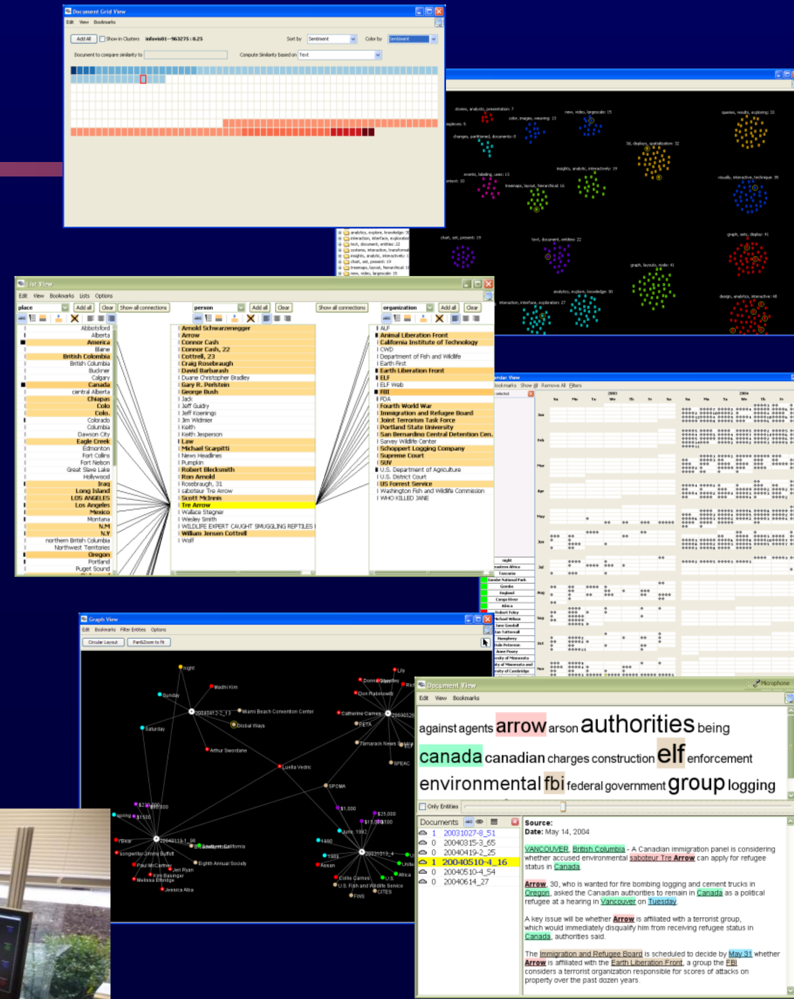
XML documents

Jigsaw: Visual Analytics for Investigative Analysis



Documents in

Knowledge out



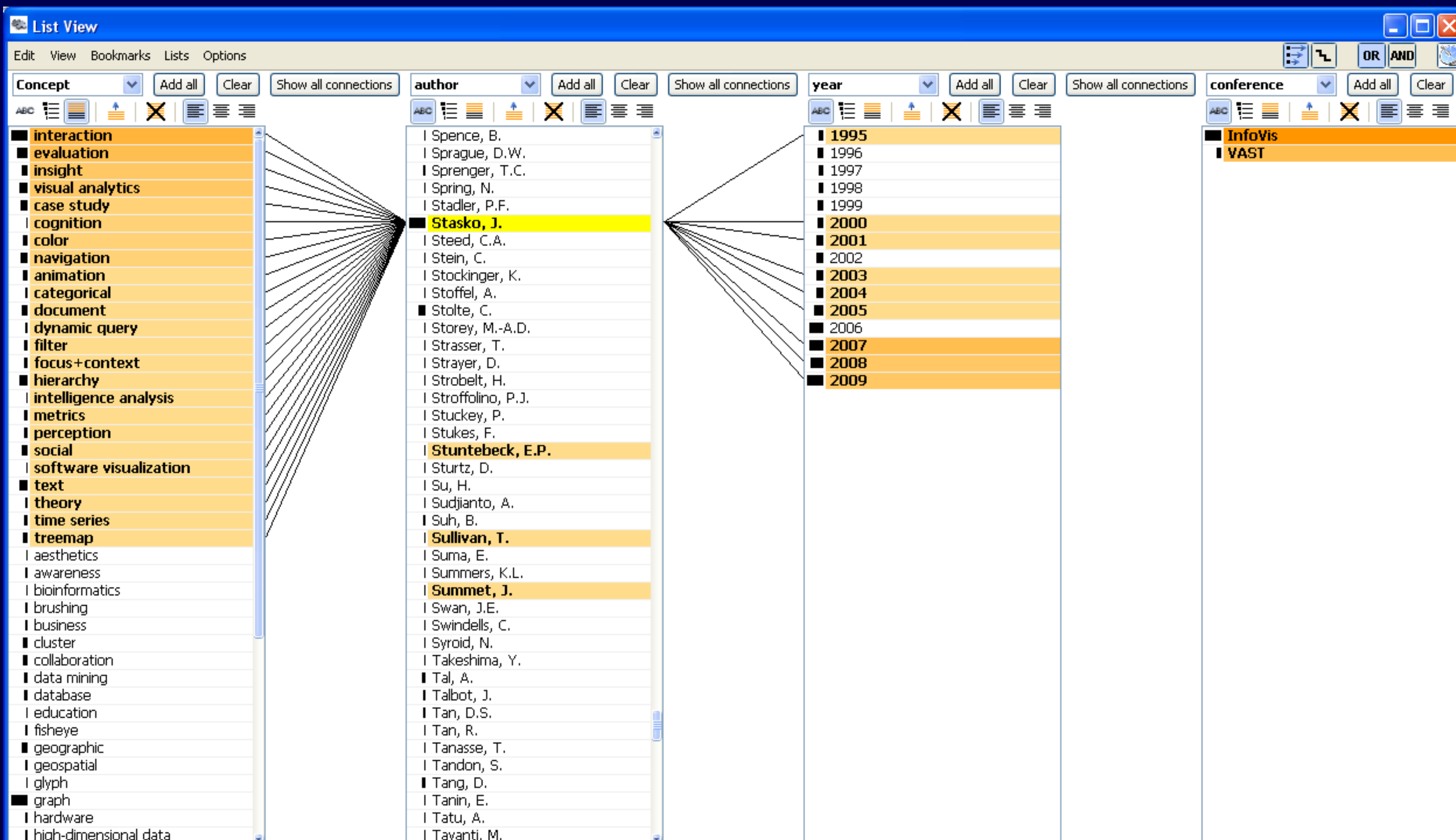
John Stasko

VACCINE

Georgia Tech College of Computing
School of Interactive Computing

Jigsaw List View

Lists of entities by type
Connections highlighted



VACCINE

WAJAC, Seattle WA

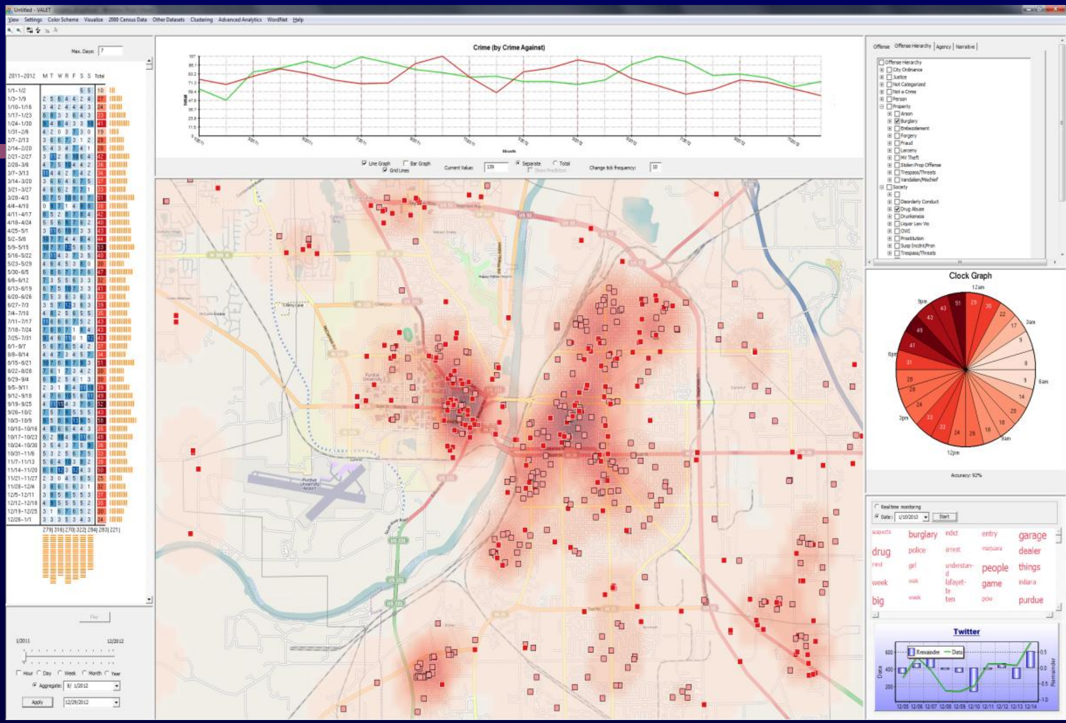
"Recently, the Washington Joint Analytical Center (WAJAC) assisted a local police department in analyzing 345 incident reports pertaining to graffiti. Jigsaw was an invaluable tool when it came to examining reports. Without the use of Jigsaw, examining reports would have been tedious and time consuming, and connections between entities may not have been seen.

The local police department was very impressed with Jigsaw and stated that the program will become very useful in tracking graffiti activity, and possibly other criminal activity.

WAJAC uses Jigsaw on a continuous basis to examine information in our intelligence database. Jigsaw is also used to analyze incidents relating to the security of the Washington State Ferry System. It provides a searchable database and quick analytical tool for tracking potential threats to the largest ferry transportation system in the U.S."

CHAD R MELTON
Criminal Intelligence Analyst
WAJAC/FBI-FIG

Visual Analytics Law Enforcement Toolkit (VALET, iVALET)



Impacts:

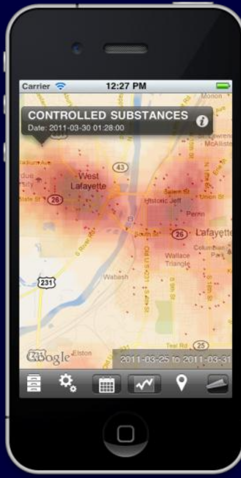
- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (February 2013) for community-based policing
- Investigating correlation factors
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning
- Incorporating predictive alerts

VALET delivered:

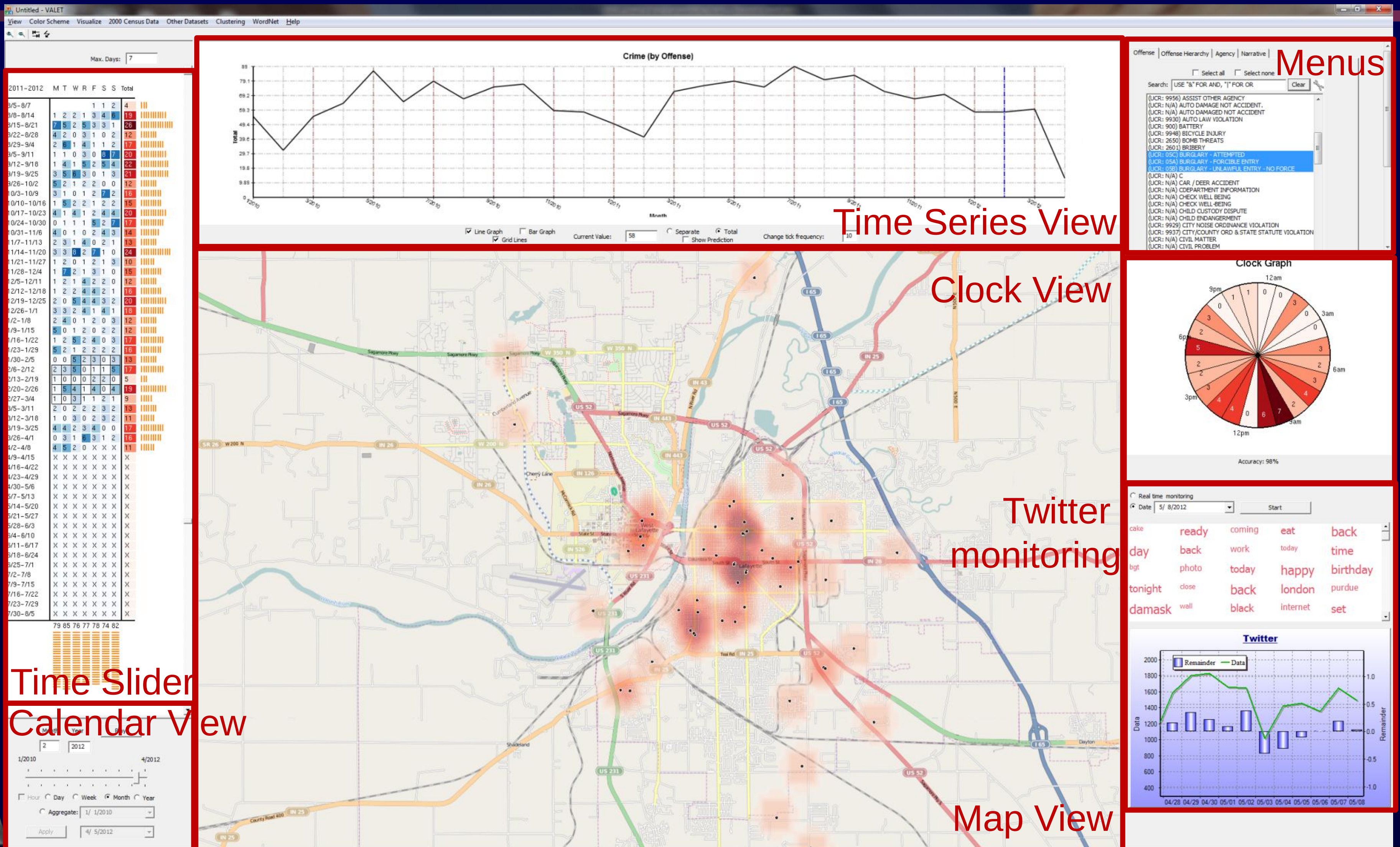
- Spring 2011: WL, Lafayette Police

iVALET delivered:

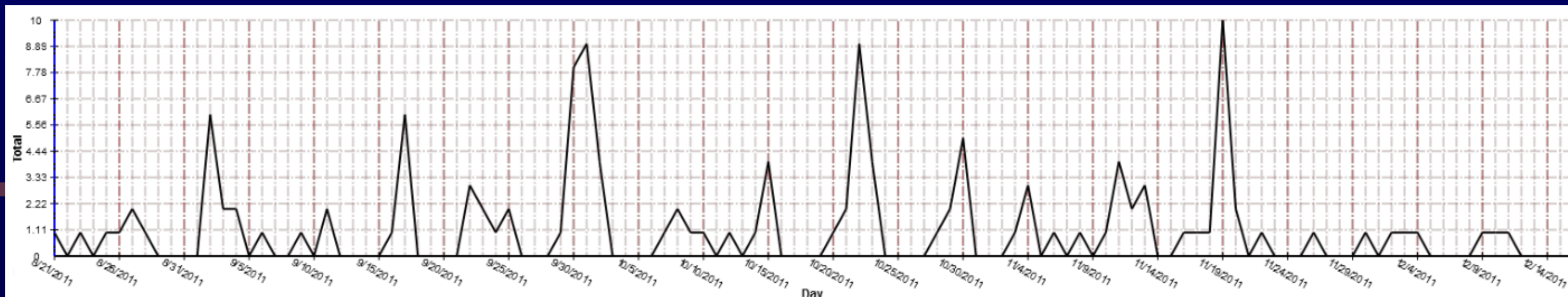
- October 2011: Purdue, WL Police

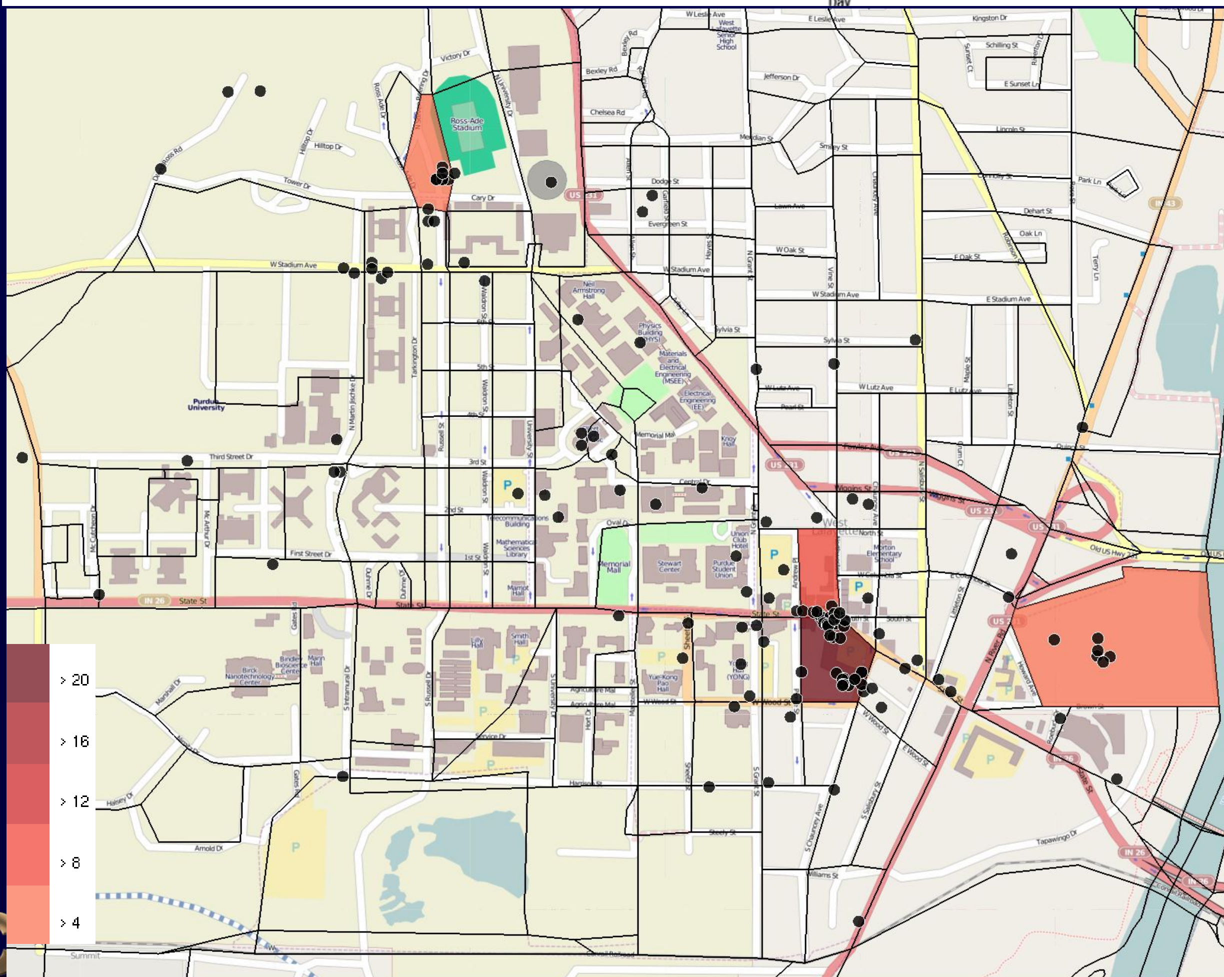


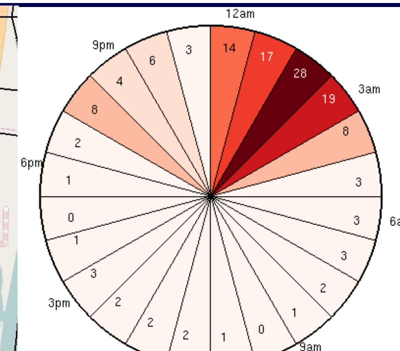
VALET Overview



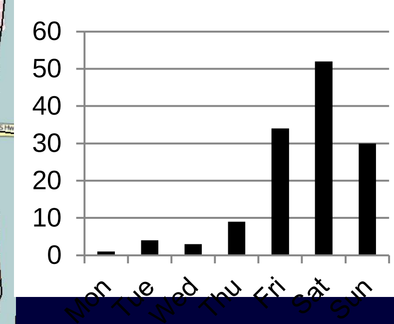
Example: Drunkenness / Public Intoxication



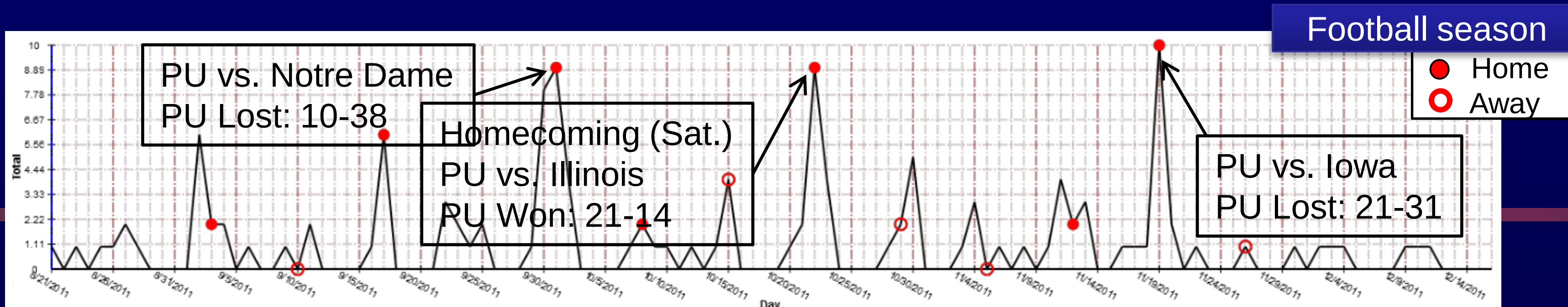


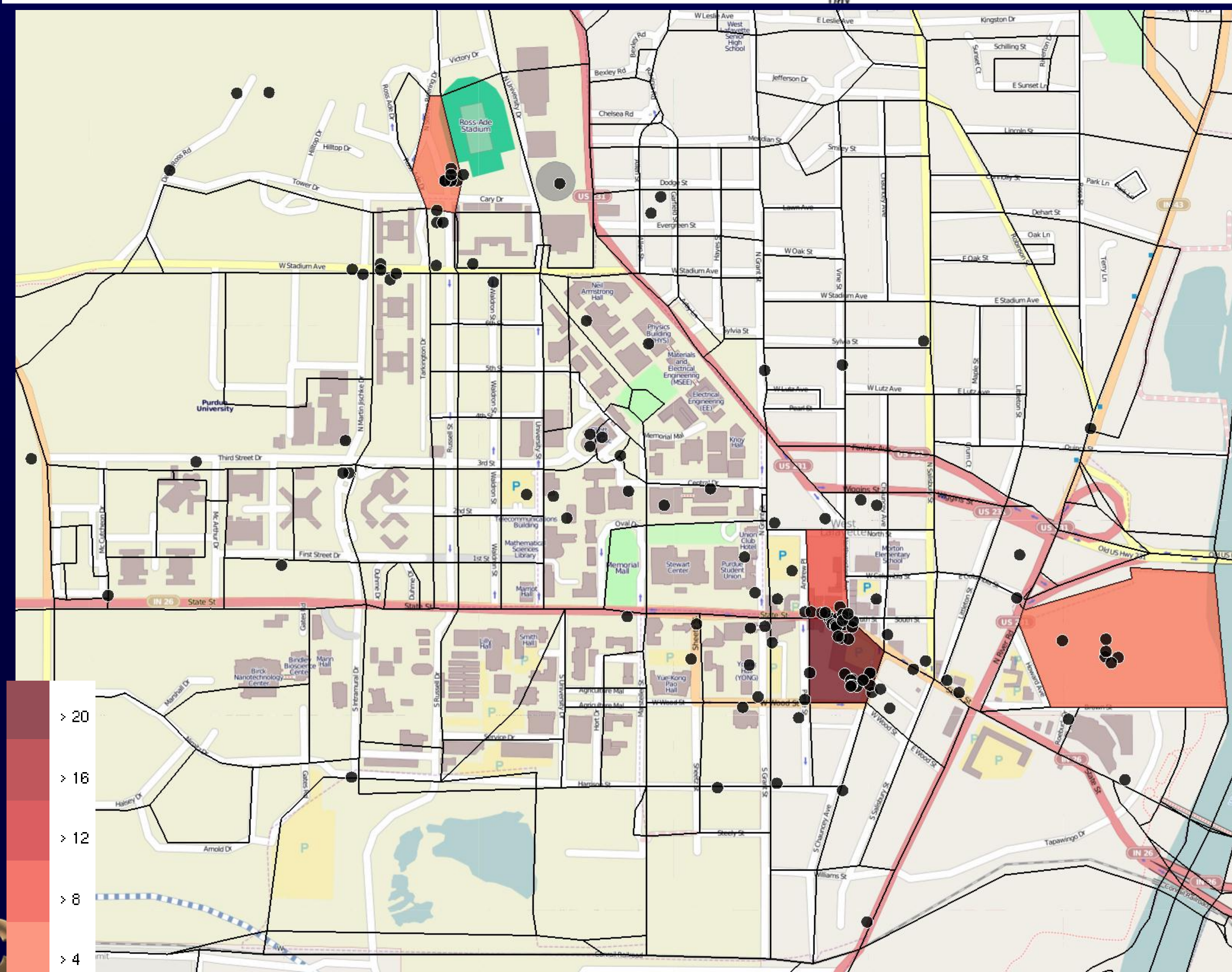


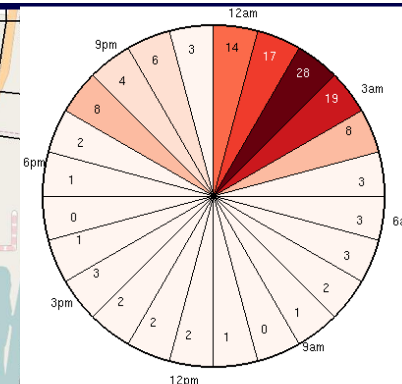
Day-of-the-Week

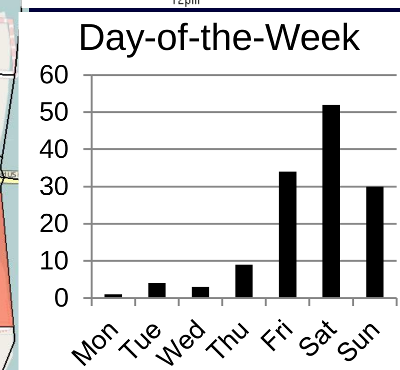


Example: Drunkenness / Public Intoxication





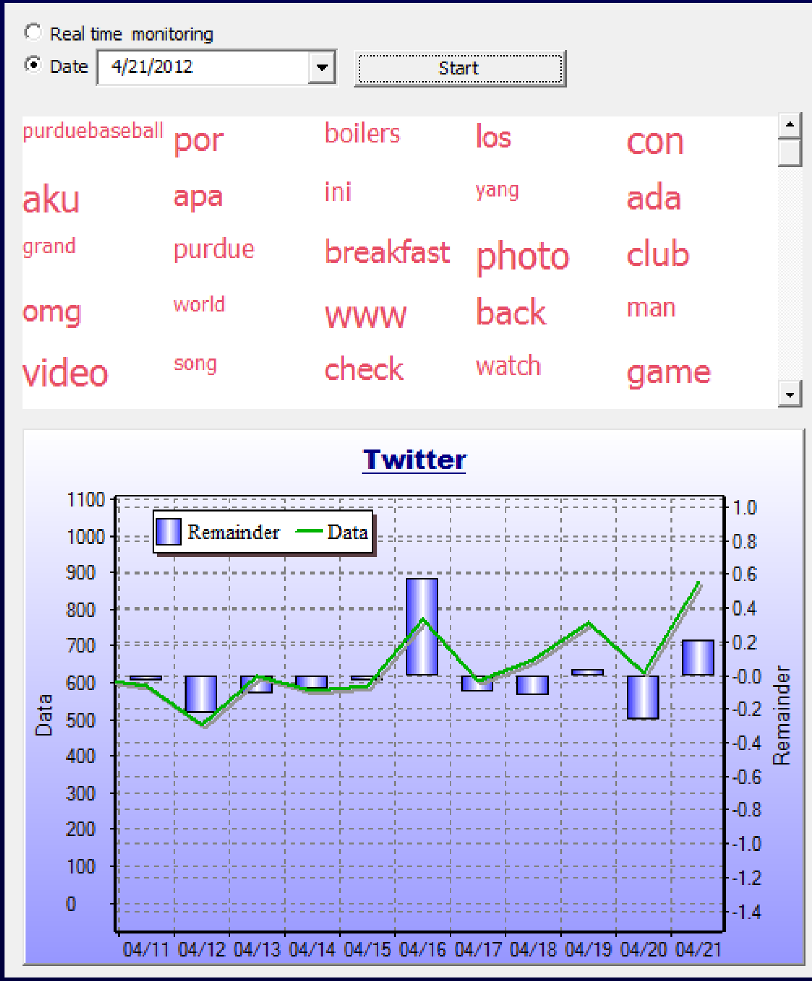




Social Media:Real-time Twitter Monitoring and Integration into Tools

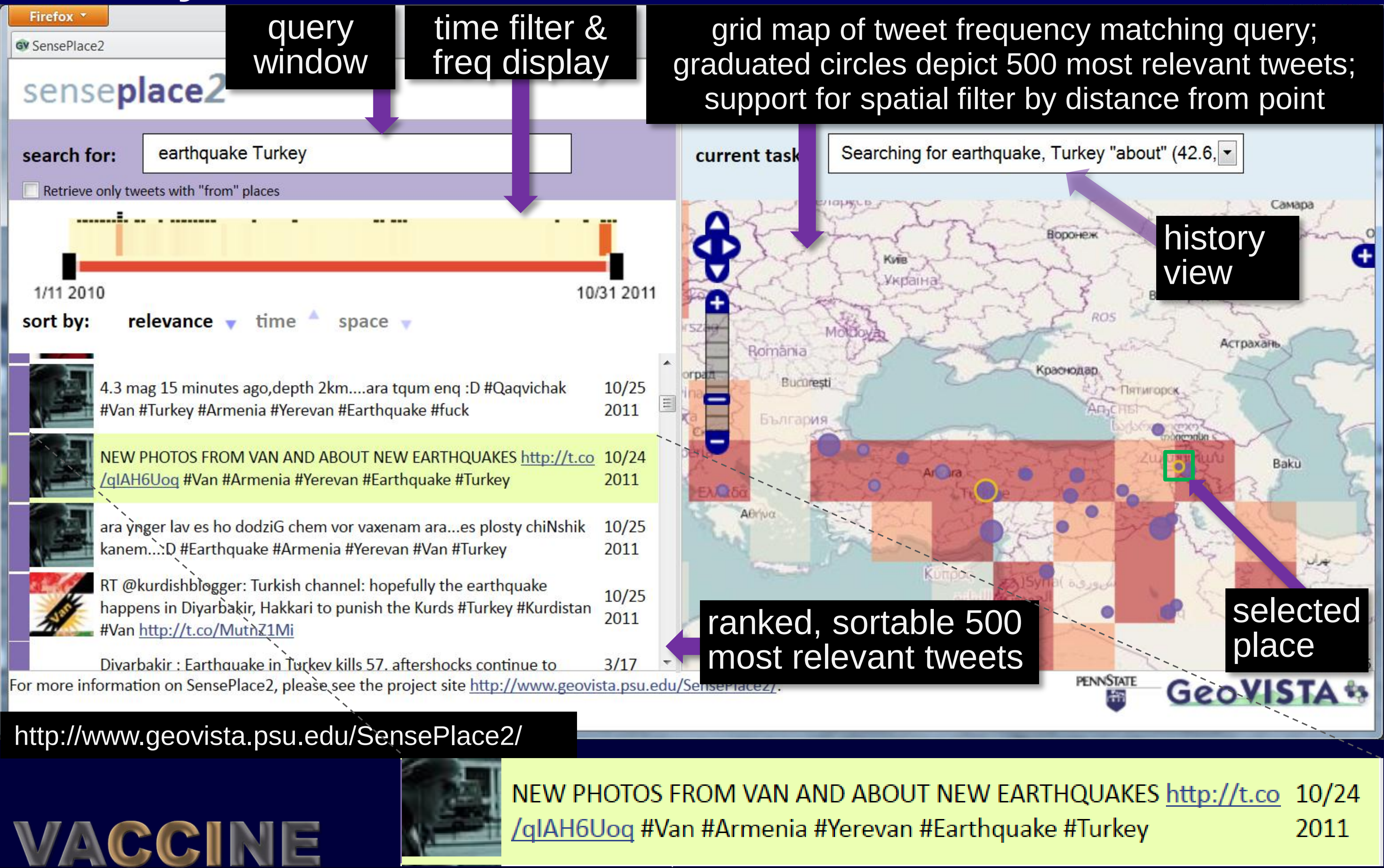
(Purdue, Stuttgart, Penn St.)

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud



Grand Prix Weekend, Purdue University

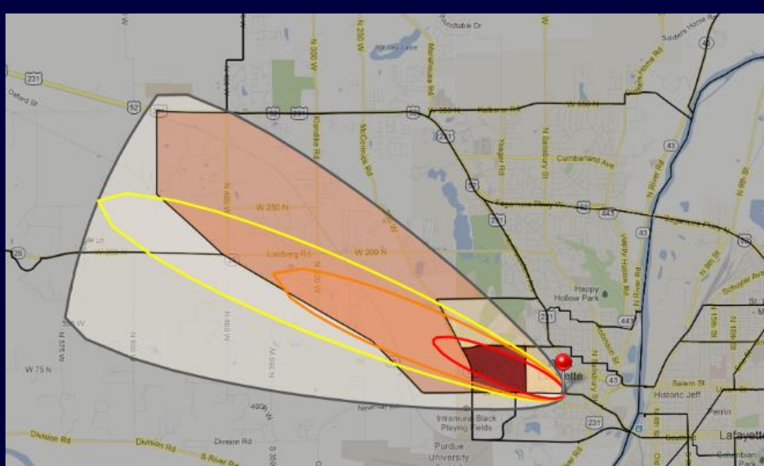
SensePlace2: Place-Time-Concept Analytics

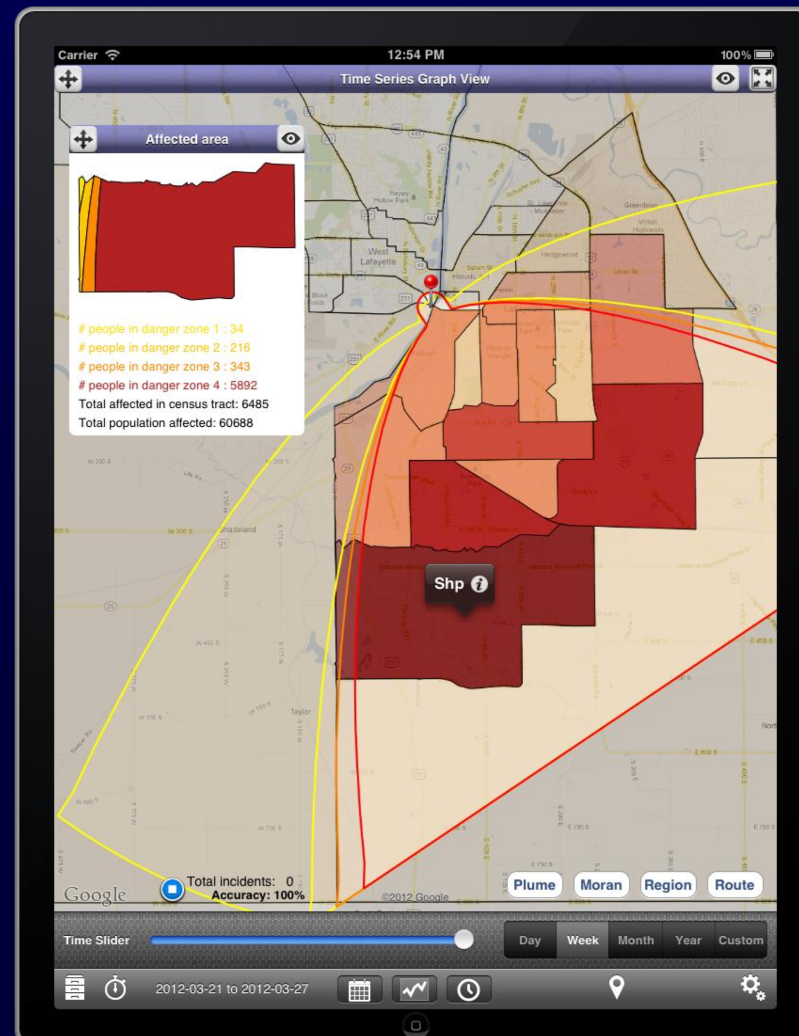



Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
11/14/2010
01/01/2011
02/13/2011
03/27/2011
Calendar View
M T W Th F Sa S
1/1/10 36 46 29
1/4/10 38 48 44 43 42 32 30
1/11/10 62 53 53 56 59 45 34
1/18/10 62 49 53 72 61 52 49
1/25/10 98 66 57 66 64 55 49
2/1/10 58 63 50 51 73 51 42
2/8/10 50 46 48 63 52 39 35
2/15/10 48 65 48 54 55 67 51
2/22/10 51 66 48 56 57 46 47
3/1/10 50 59 42 54 82 45 41
3/8/10 50 57 64 78 63 36 30
Clock View
12Hrs
10:00
Accuracy: 64%
Hedgewood
Hanna
Centennial
Perrin
Lafayette
Wabash
Highland Park
Wallace Triangle
©2012 Google
Choro
Plume
Moran
Region
Route
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20

MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

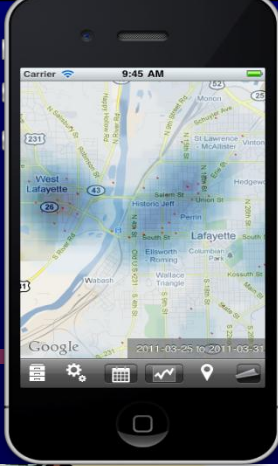
- Chemical release plume modeling identifies census tracts with the highest number of expected people affected





We Have an App For That!

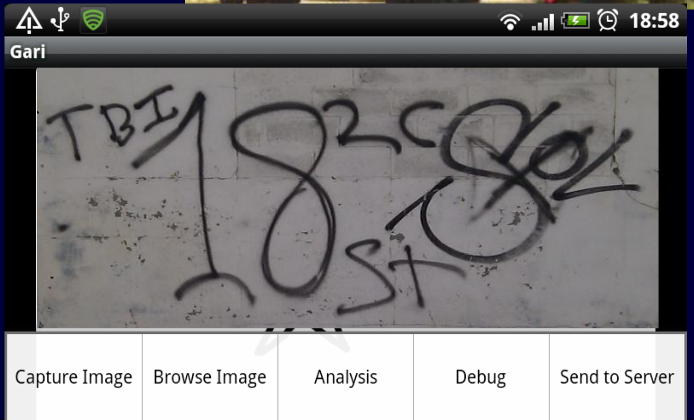
Our Mobile Tools



- VALET

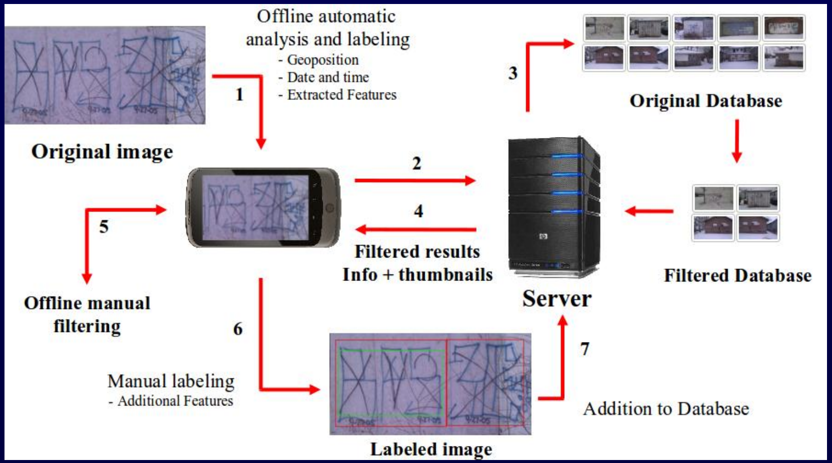
- Evacuation Planning
- Rosetta Phone

- Hazmat app
- Gang graffiti app



- Tatoo app

Gang Graffiti Analysis and Recognition Using a Mobile Telephone (GARI)



IMPACT:

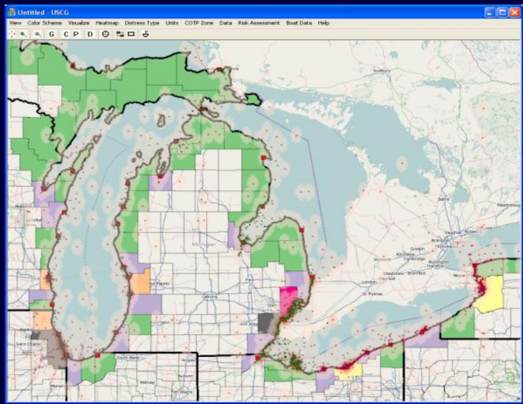
- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be "interpreted"
- More than 75 users and 800 graffiti images acquired

GARI delivered:

- Summer 2011:
  - IPD gang detectives
- August 2011:
  - IPD at large
  - Ind Fusion Gang Task Force
- September 2011:
  - Gang detectives across Indiana

Visual Analytics Uses for Risk-Based Decision Making

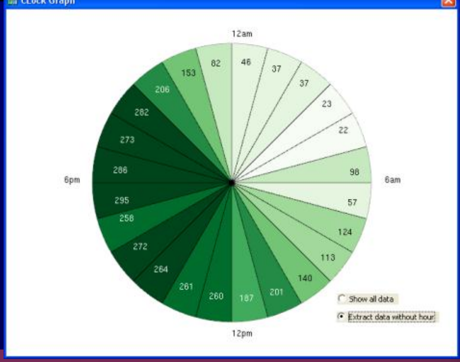
- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



USCG: Effective Risk-based Decision Making and Resource Allocation Visual Analytics

- Evaluate current and historical mission area:
  - Demands
  - Risks (total, mitigated, residual)
  - Resource allocation
  - Return on investment
- Evaluate courses of action
- Evaluate above at both Strategic and Tactical/Operational level

Risk-Based Allocations



- Comparative visual analysis of mission cases/hours vs. staffing hours
- Comparative visualization of resources vs. risk
- Trend visual analytics
  - Increase/decrease in resource allocation
  - Increase/decrease in risk (total, mitigated, residual)
  - Increase/decrease in incidents
- Exploration of alternatives and effect on risk
- Predictive analytics based on historical data (STL and EWMA)

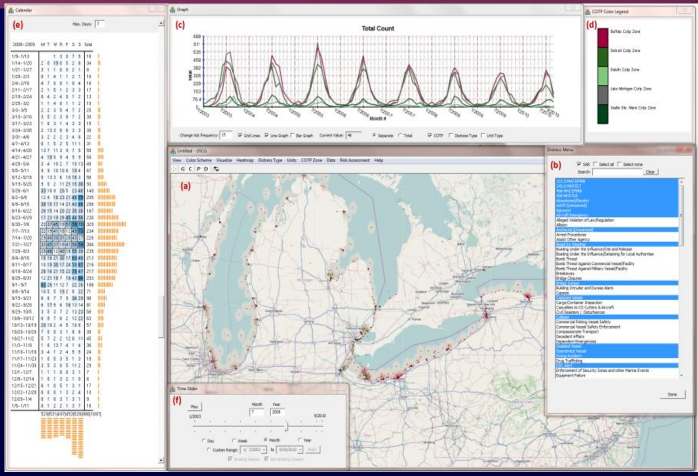
VACCINE

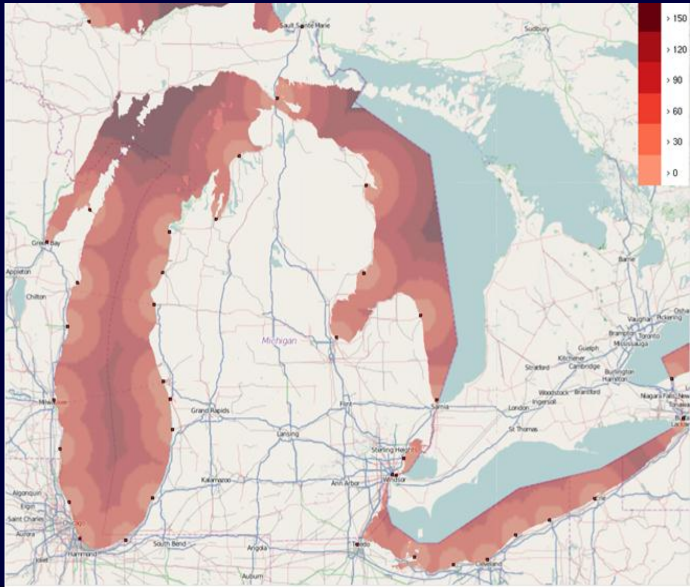
U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7, USCG D9, USCG D5, USCG HQ 771

IMPACTS:

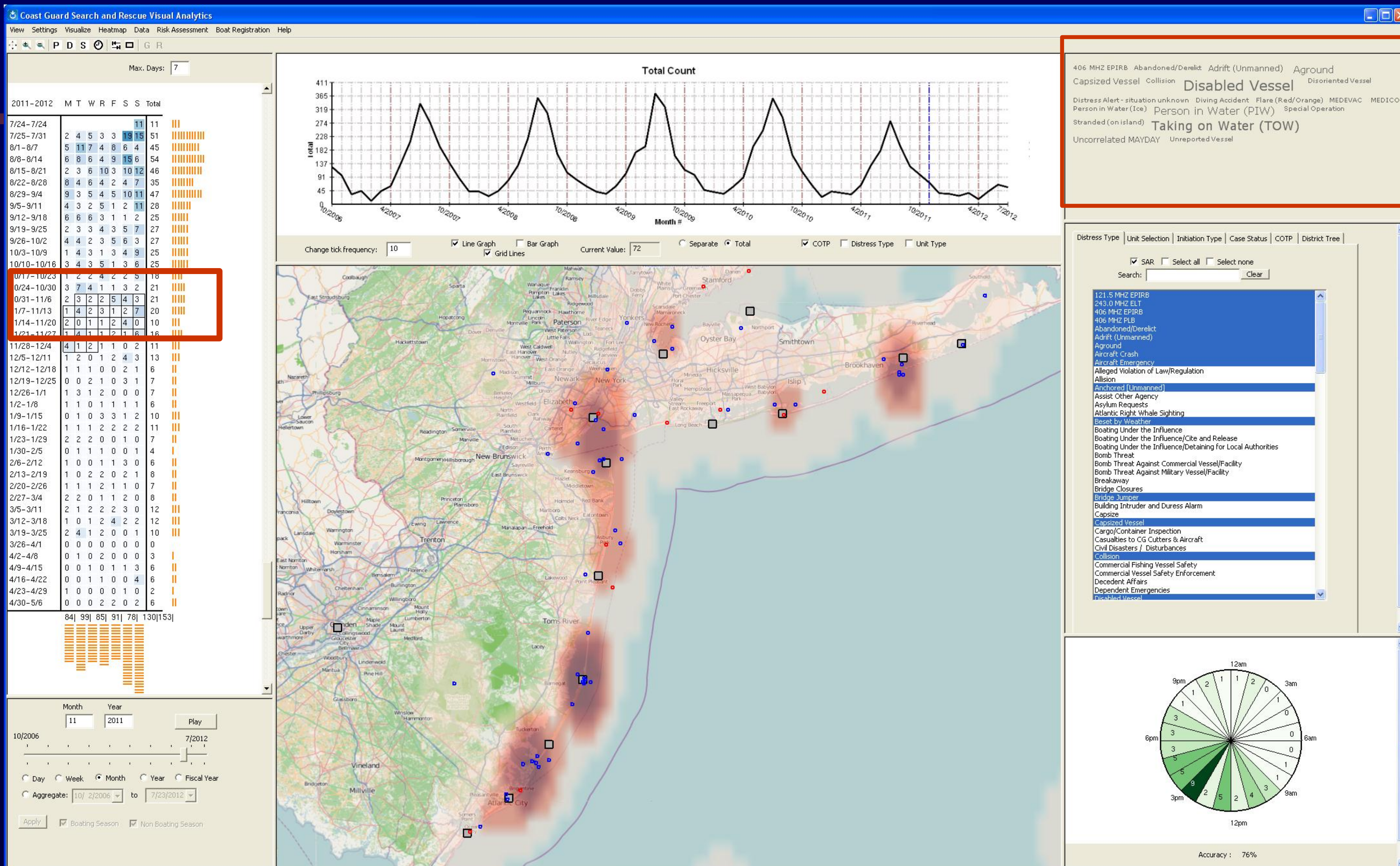
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Hurricanes Sandy and Irene resource allocation decisions based on cgSARVA analysis and visualization
- Informed Commandant's budget testimony and recommendations to Congress
- Key component of USCG D9 reallocation plan for 2011-2012 based on decreased budget
- Key component of Coastal Operations Allocation Suite of Tools (COAST) – USCG HQ



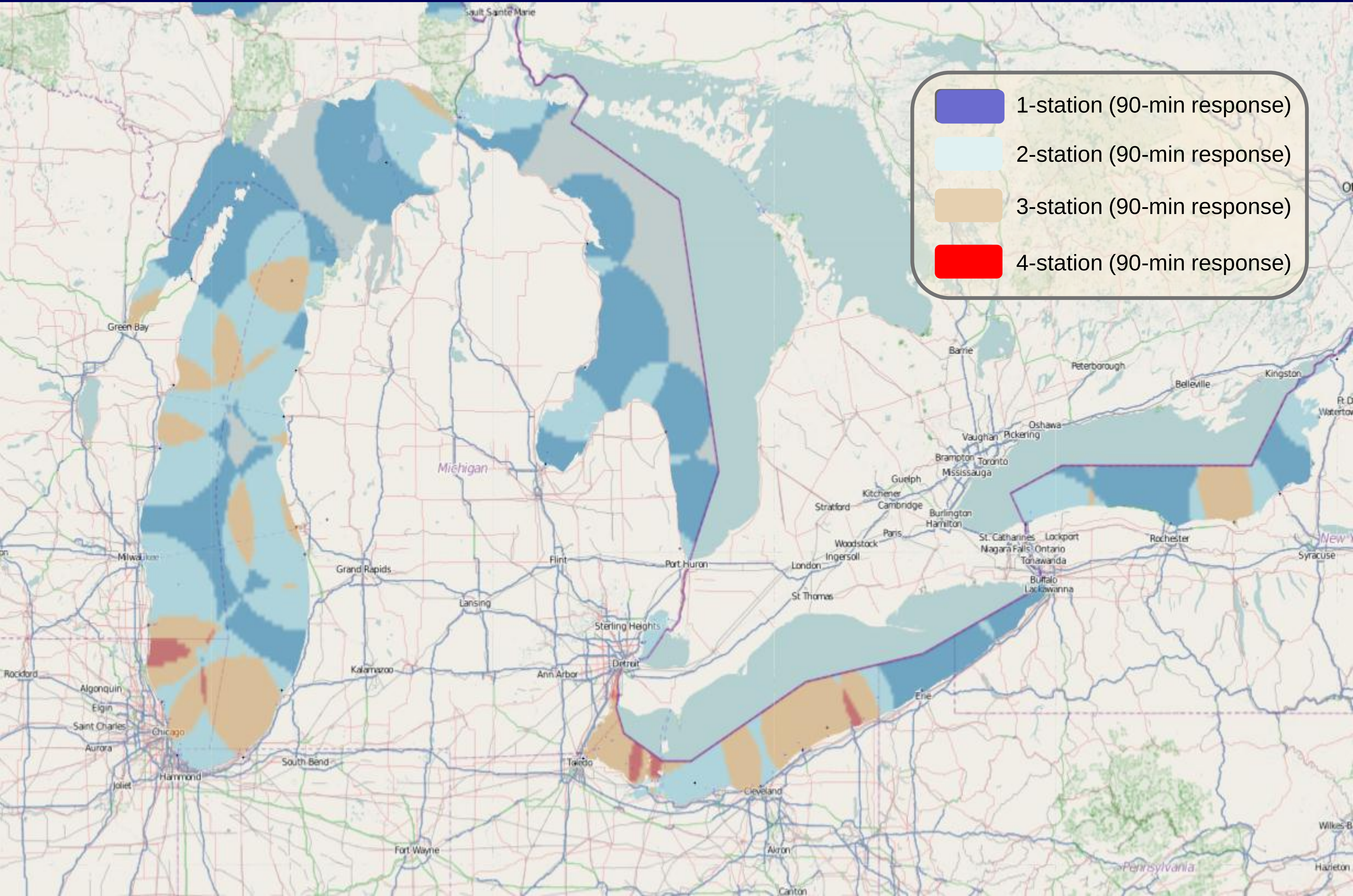


Example: Risks and Consequences From Sandy:

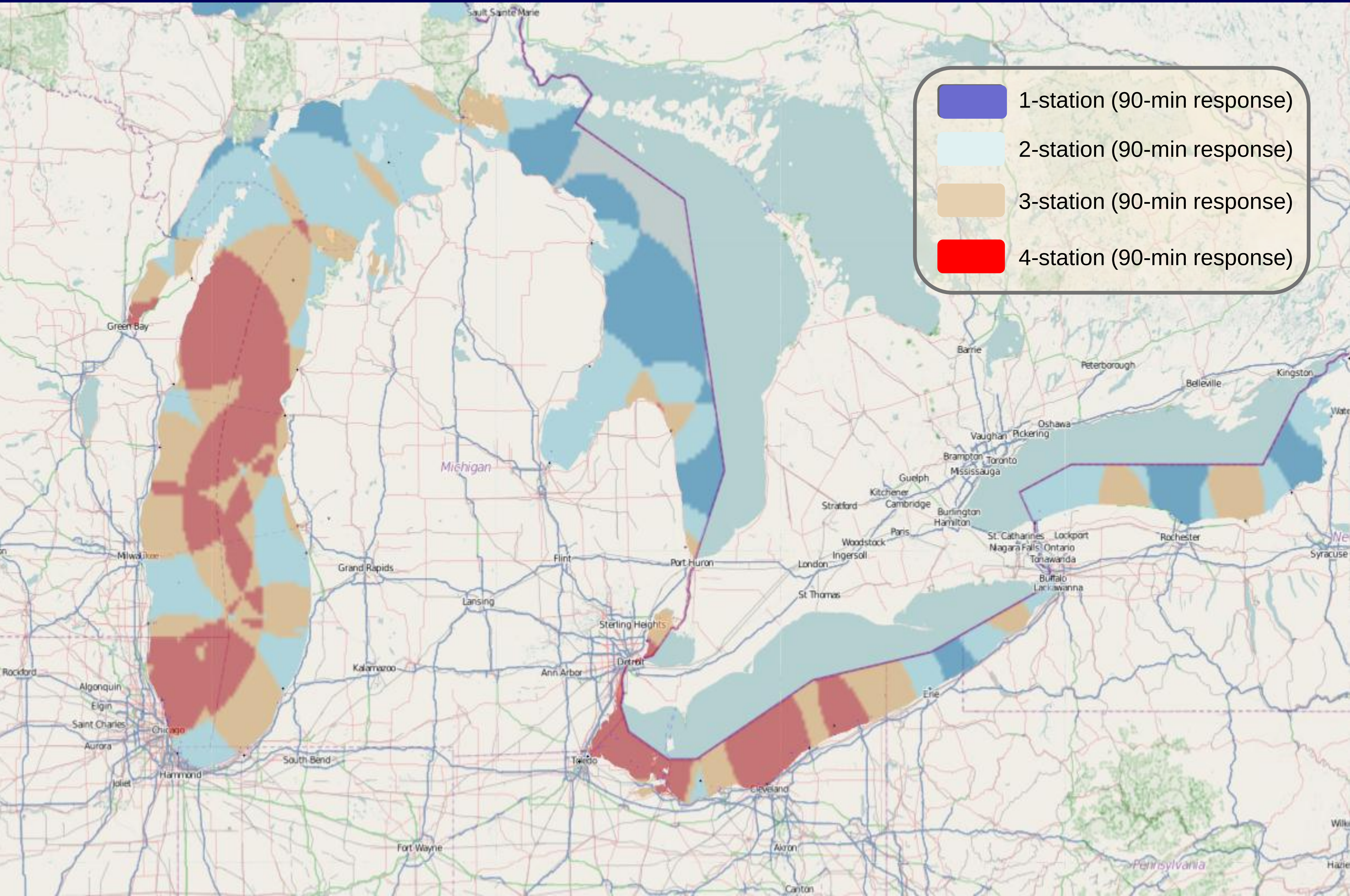
SAR Cases November 2011 NJ/NYC Area



Response Efficiency – Current Assets



Response Efficiency – Potential Future Assets

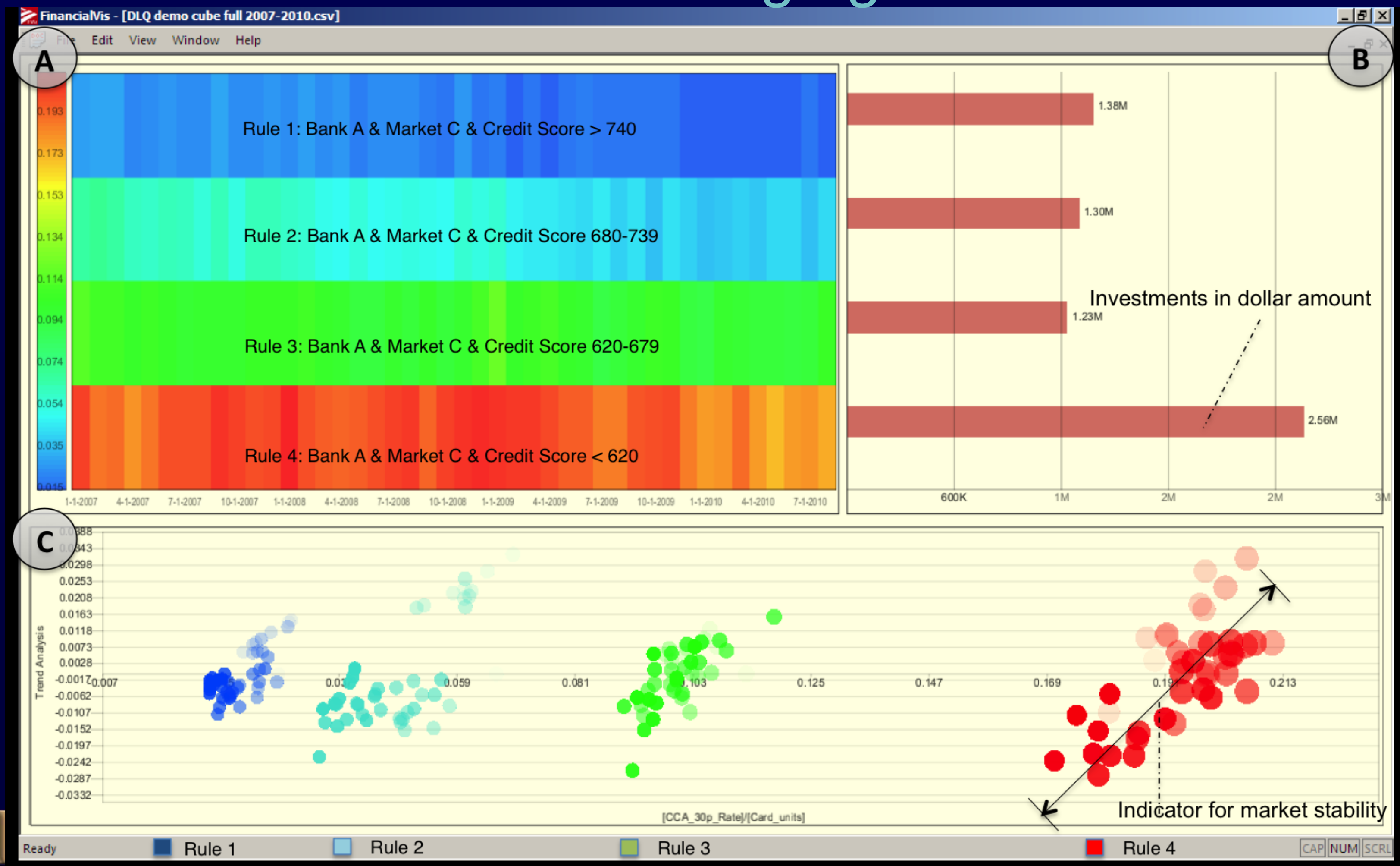


RiskVA: Key Credit Risk Analytics Challenges

- Consumer credit data is large, temporal, and related across multiple investments and financial markets.
- The data are heterogeneous, not clean, have missing values, may be misleading and inefficient to explore.
- The data contains important behaviors and relations/groupings that change over time.
- Analysts need to have customer-centric interactive systems to achieve full analysis of risks

RiskVA Overview: Interactive Exploratory Visualizations

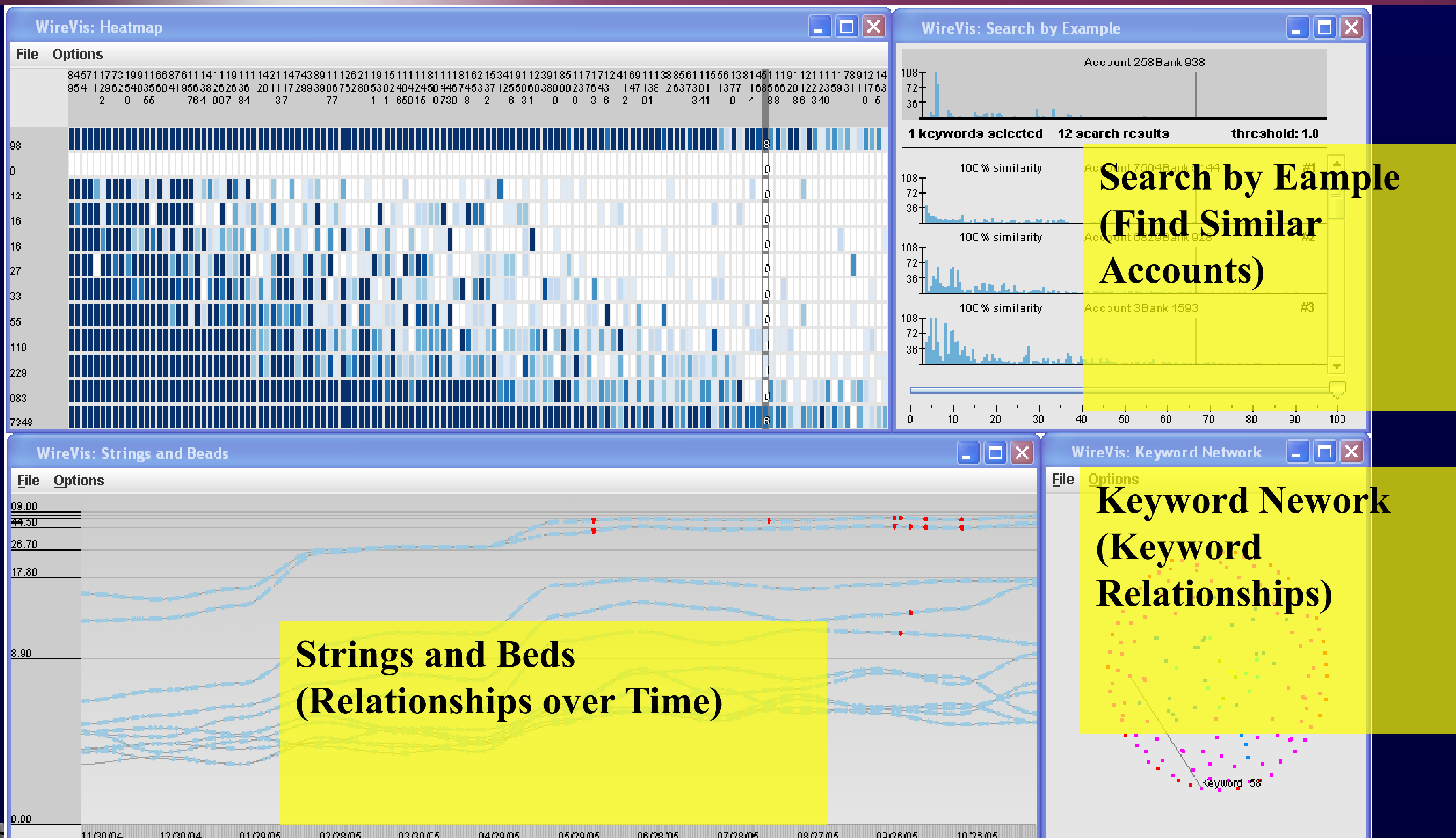
Identification of Emerging Risks



VACCI

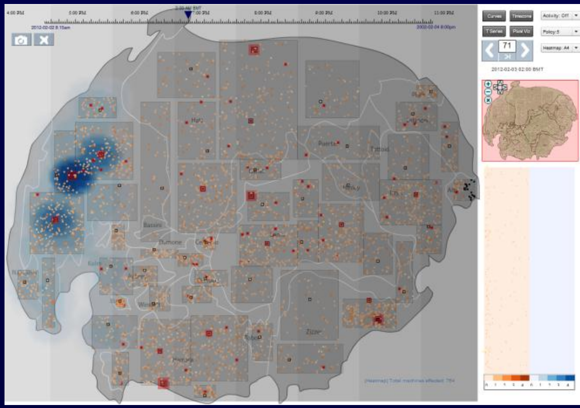
WireVis – Streaming Data - Multiple Linked Views

- Temporal, geospatial, theme, cluster, list views with association linkages between views



Example Work in Cybersecurity

- Corporate Insider Threat Detection (Oxford, Leicester, Cardiff)
- Sensor Forensics (Purdue)
- SemanticPrism (Purdue)
- Multiscreen, Multiview Interactive Cyber Investigation (VaTech, PNNL)
- Log Visualization (Purdue)



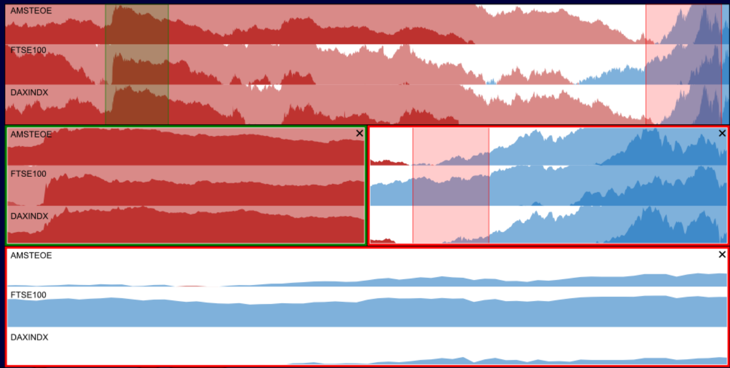


Table of Contents

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu

ebertd@purdue.edu

VACCINE

Table of Contents

CVADA

Center for Visual Analytics and Data Analytics

A U.S. Department of Homeland Security Center of Excellence

Visual Analytics for Effective Planning, Analysis, and Decision Making

David S. Ebert

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

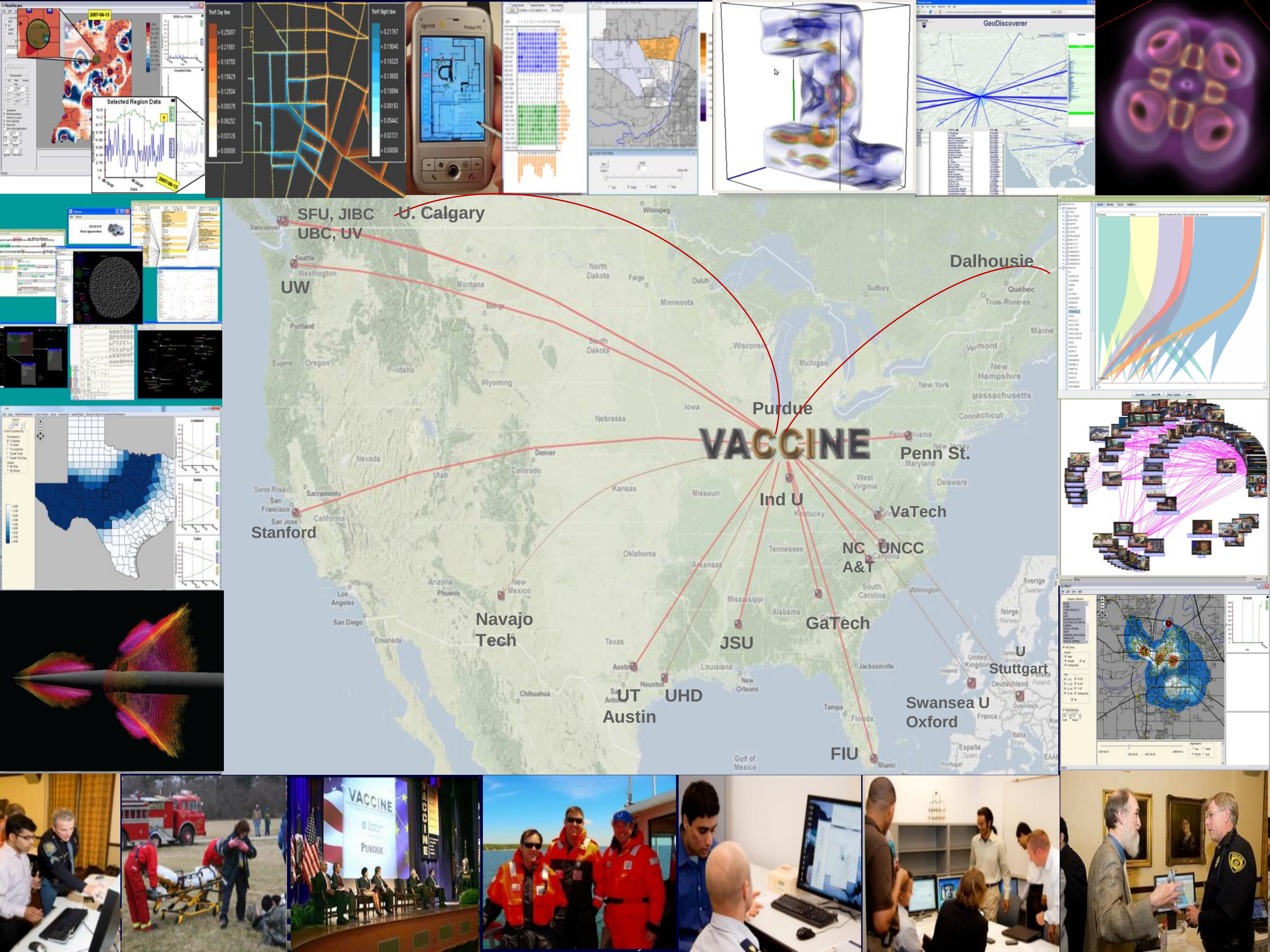
February 2013

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

Who We Are
&
What We Do


SFU, JIBC
UBC, UV
U. Calgary
UW
Dalhousie
Purdue
VACCINE
Penn St.
Ind U
VaTech
NC A&T
UNCC
Stanford
Navajo Tech
GaTech
JSU
U Stuttgart
UT Austin
UHD
Swansea U
Oxford
FIU
Selected Region Data
GeoDiscoverer

Who We Are:
International Team of Experts
- 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

VACCINE

VACCINE Multidisciplinary Team

Principal Partners	Areas of Expertise/Core Capabilities
Purdue University	Visualization of Structured, Unstructured and Streaming Data; Scalable Filtering and Dissemination; Video and Image Processing; Mobile, Lightweight Information Analytics; Public Safety Operations; Undergraduate/Graduate Education; Professional Training
Georgia Institute of Technology	Information Visualization; Visual Analytics; Peripheral Awareness; Undergraduate/Graduate Education;
Pennsylvania State University	Spatial cognition; Geo-information Representation, Spatial Analysis, Cartography; Undergraduate/Graduate Education; Professional Training
University of NC-Charlotte	3D Multimodal Interaction; Bioinformatics Visualization; Virtual Environments; Visual Reasoning; Interactive Visualization of Large-scale Information Spaces; Undergraduate/Graduate Education; Professional Training

Extended Partner Network
Florida International University*; University of Houston-Downtown*; University of Texas, Austin; California State University Dominguez Hills*; Morgan State University*; Virginia Polytechnic Institute and State University; Navajo Technical College*; Simon Fraser University, Canada; University of British Columbia, Canada; Oxford University, UK; Swansea University, UK; Stuttgart University, Germany; Stanford University; University of Washington; Tennessee State University; Middlesex University, UK; Dalhousie University, Canada; University of Manitoba; University of Victoria; Indiana University; Jackson State University*; University of Calgary, Canada; Carleton University; Justice Institute of British Columbia, Canada; Ontario Institute of Technology, Canada.
* Indicates MSI

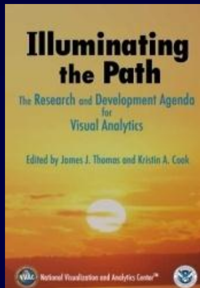
What We Do: Visual Analytics

Visual Analytics[1] is the science of analytical reasoning facilitated by interactive visual interfaces

People use visual analytics tools and techniques to

- Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
- Detect the expected and discover the unexpected
- Provide timely, defensible, understandable assessments
- Communicate assessment effectively for action

1. *Illuminating the Path: The R&D Agenda Analytics, Editors: Thomas and Cook*



VACCINE

VACCINE's Value

We enable users to be more effective through innovative interactive visualization, analysis, and decision making tools

- Provide the right information, in the right format within the right time to solve the problem
- Turn data deluge into a pool of relevant, actionable knowledge
- Enable user to be more effective from planning to detection to response to recovery
- Enable effective communication of information

Approach: Partner-driven solutions and research

Table of Contents

VACCINE Research Themes To Solve Real World Problems

- **Visual Analytics –** integrated, interactive visual exploration, analysis, and decision making environments to enable effective decisions from massive, time-evolving multimedia data
- **Interactive scalable analytics techniques –** statistical, spatiotemporal, image, video, signal, machine learning
- **Science of Interaction for visual analytics**
- **Evaluation** in real world environments

VACCINE Research Topics/Projects

1. Big Data visual analytics
   - Both in size and complexity
2. Event/anomaly detection
   - Health and crime STL, spatial anomalies
3. Predictive analytics
   - Incorporating statistical analytical models (STL)
4. Social media analysis
   - Scatterblogs, SensePlaceII, …
5. Combining complex data (Structured, unstructured, streaming, variable)
   - Multi-source police data (Valet, iValet)
   - Multiple text sources (Jigsaw)

VACCINE Research Topics/Projects

6. Visual analytics solutions for risk-based decision making
   - US Coast Guard projects
7. Making "data science" available to non-data scientists
   - iVALET, iOPAR
8. Active Data
   - Alert lists based on unusual recent activity
9. Video analytics
10. Infield visual analytics and image analytics
    - MERGE, GARI, RosettaPhone

Topic 1: Big Data Challenges

- Heterogeneity
- Unprecedented volume and/or complexity
- Multivariate, Multidimensional
- Streaming
- Qualitative, Quantitative, Categorical, Temporal,…
- Distributed Sources: sensors, text, images, maps, …
- Inconsistent and Uncertain
- Different Scales and Cross Scale

The Curse of Dimensionality (Bellman 1961)

- Problems associated with multivariate data analysis as the dimensionality increases the available data becomes sparse
- Sparsity is a problem for any method that requires statistical significance
- In visualization, we are also limited with screen space and the number of available visual variables, so choosing the most appropriate dimensions is key

Richard Ernest Bellman; Rand Corporation (1957). *Dynamic programming*. Princeton University Press.

Big Data Isn't Just Large Data!

- All of the above characteristics lead to big, complex data
- Example : A few million crime records with uncertain, missing data, dozens of fields, associated risks, related social media data, text reports, images, video…

VA Approaches to Big Data

- Don't display all the data!!
- Extract relevant information at the appropriate natural scales using analytics and statistical models to reduce data space so that it is mappable to the visual space
- User interaction to guide above process
- Should we use big iron or cell phones?
  - Solution should adapt to the available hardware

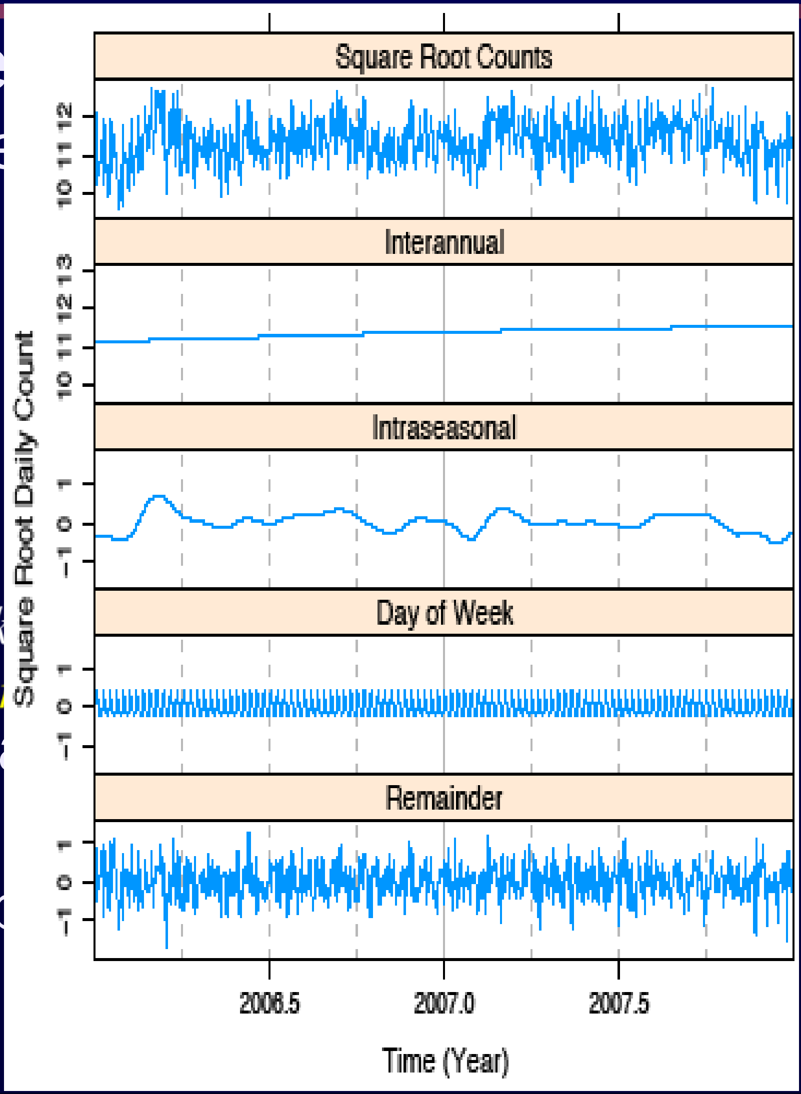
Topic 2: Event/Anomaly Detection - STL

- Temporal trend analysis and pre
  - Seasonal Trend decomposition base weighted regression) (STL)
  - Time series data viewed as the sum

$$Y = T$$

Y : *original time series*
T : *trend component*
S : *seasonal (daily/week*
R : *remainder compone*

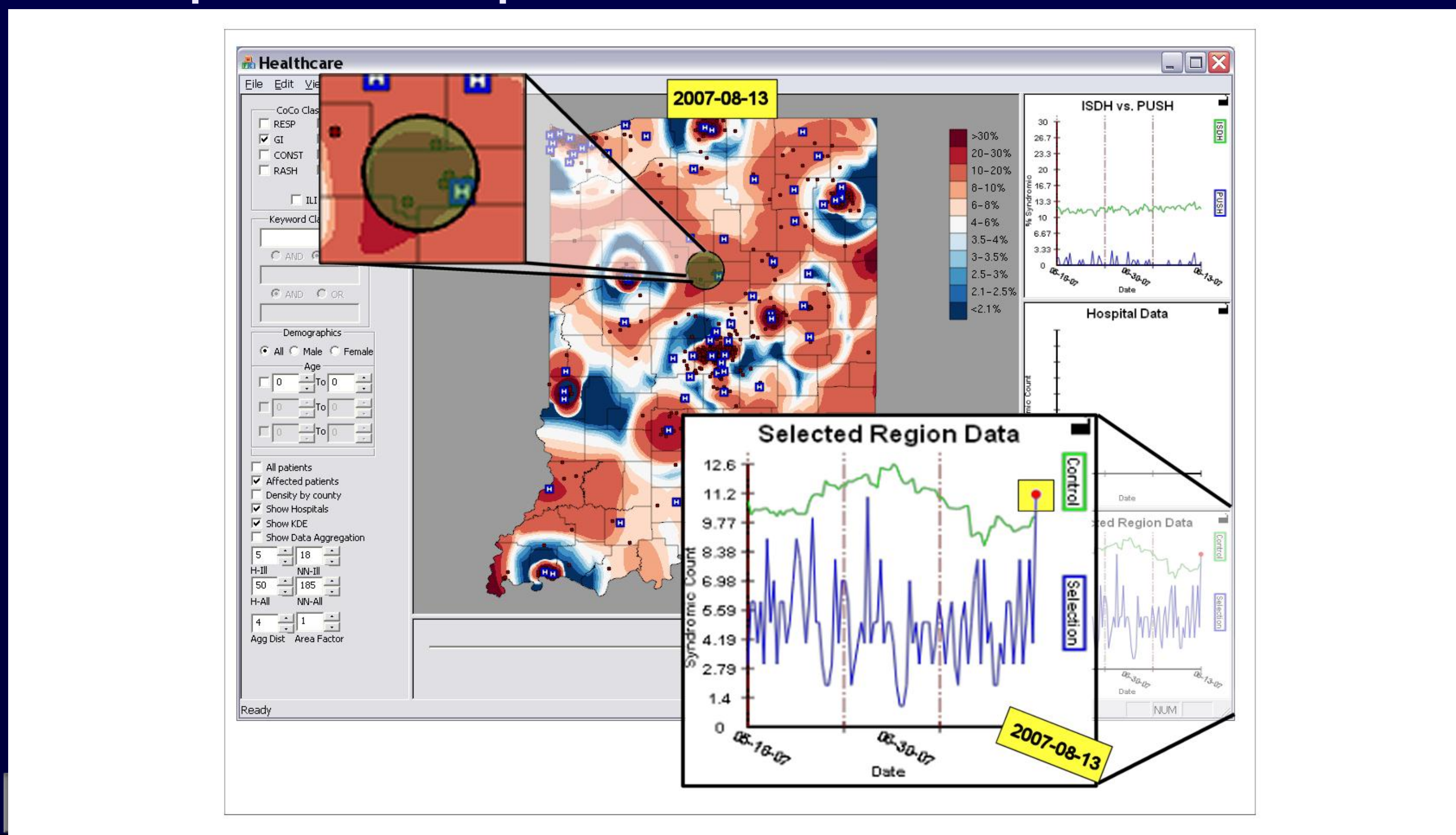
  - Large value of R indicates substanti
  - Filter out the noise component
  - If value varies by 2 SD from STL exp



VACCINE

Topics 2: Spatial Anomaly/Hot Spot Detection

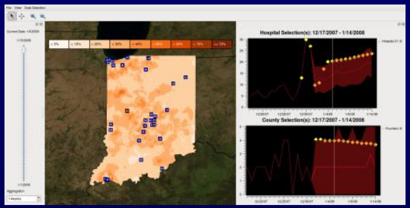
- Use Kernel Density Estimation to generate population adaptive hotspot detection

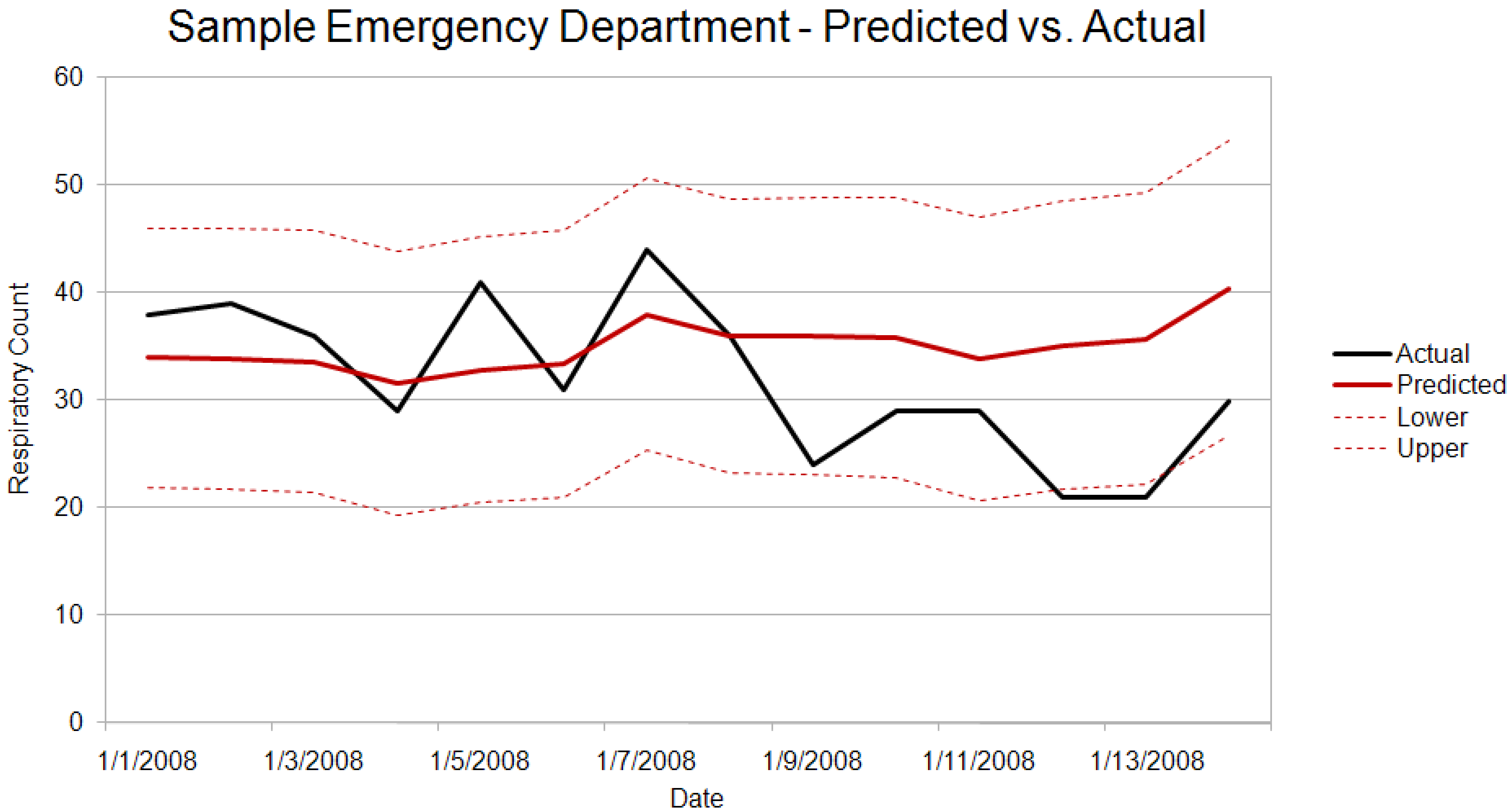


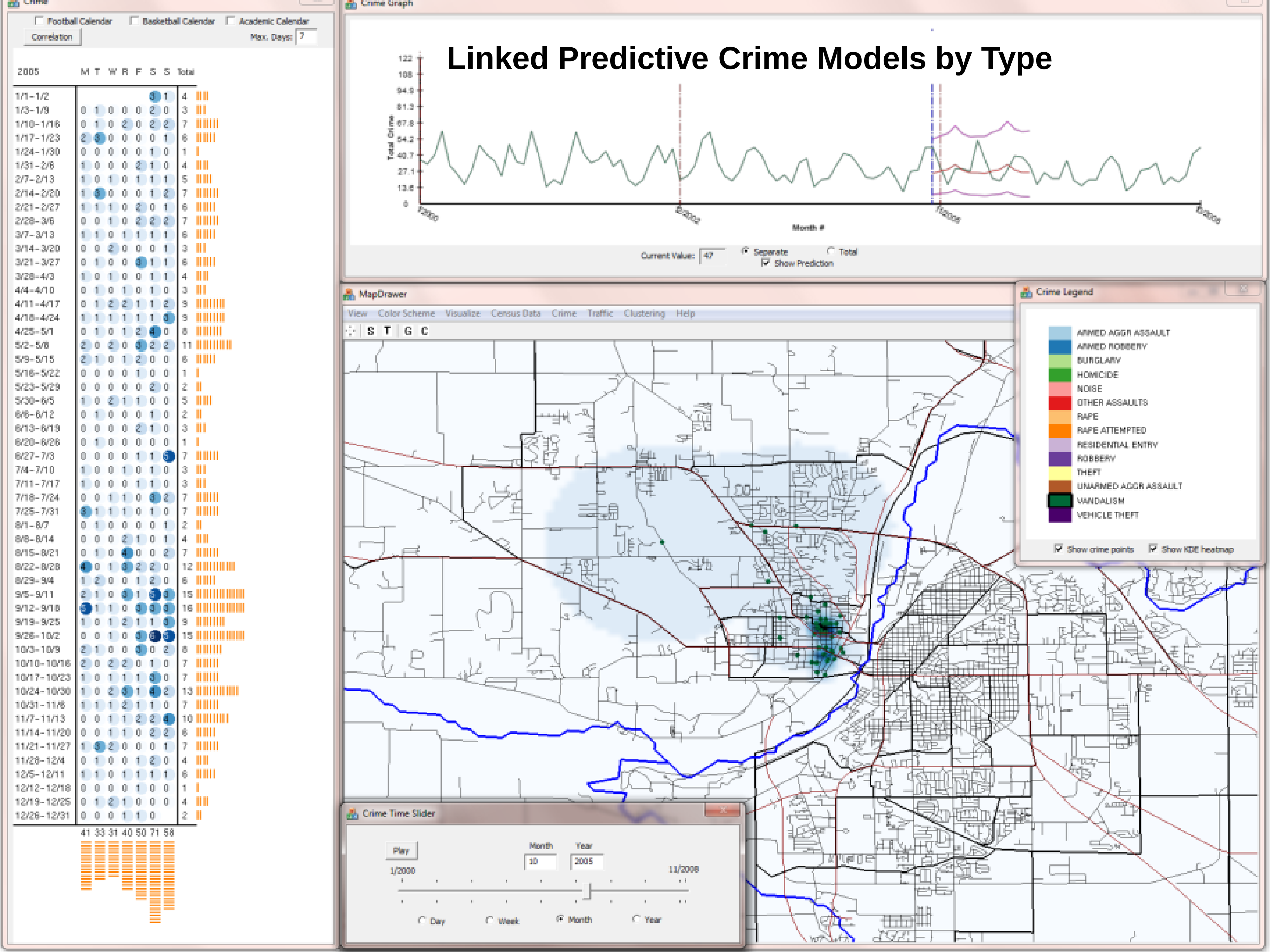
Topic 3: Predictive Analytics

- Use STL per data item to develop predicted values with confidence intervals
- Applied to health data and crime data
- Adapting to appropriate spatial resolution for most accurate predictions

Predictive Visual Analytics

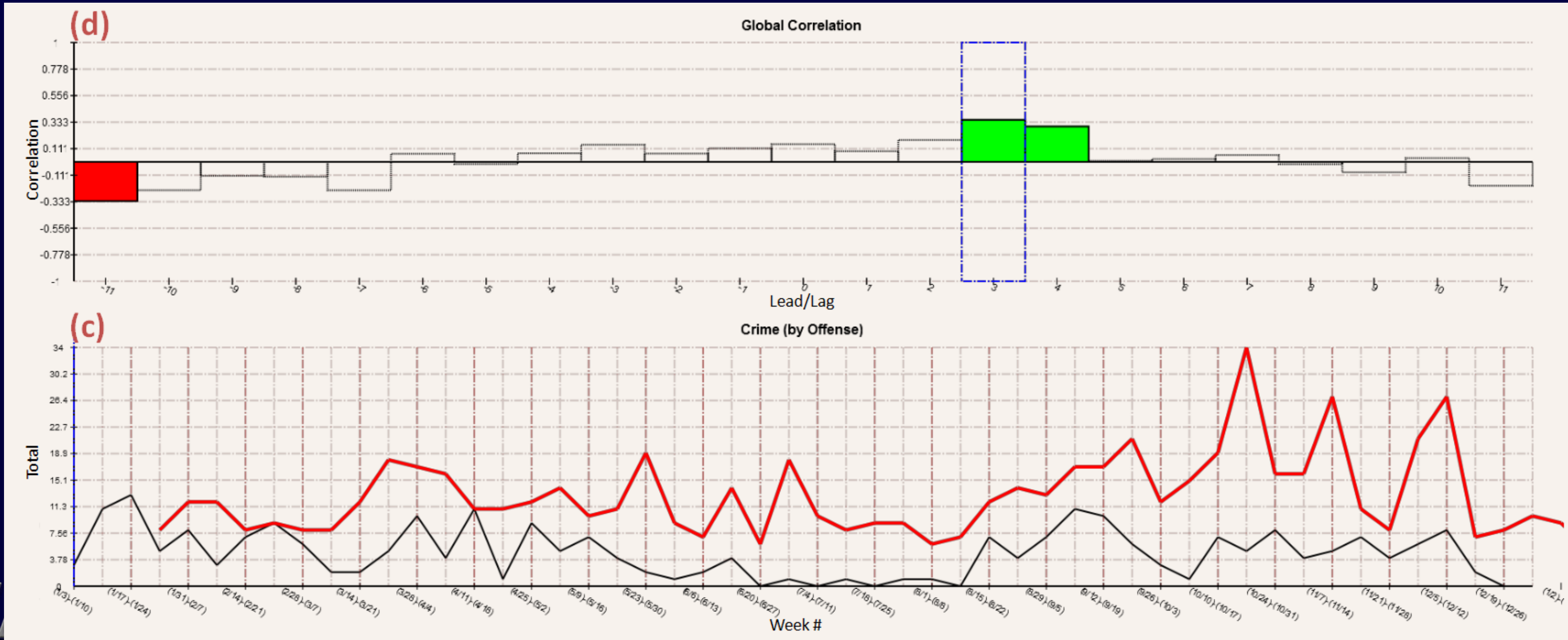




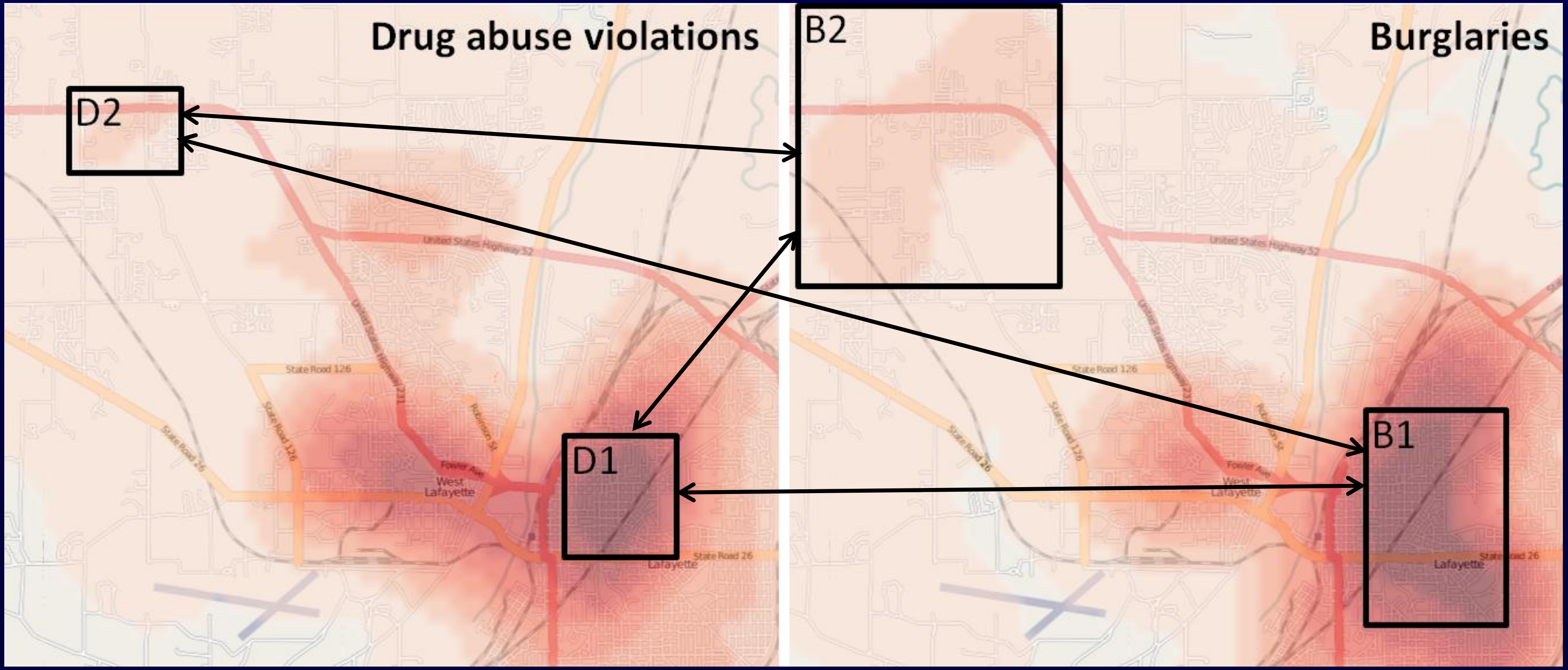

Linked Predictive Crime Models by Type
Crime
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
Max. Days: 7
2005
M T W R F S S Total
41 33 31 40 50 71 58
Crime Graph
Total Crime
Month #
Current Value: 47
Separate
Total
Show Prediction
MapDrawer
View Color Scheme Visualize Census Data Crime Traffic Clustering Help
S T G C
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Show crime points
Show KDE heatmap
Crime Time Slider
Play
Month
Year
10
2005
1/2000
11/2008
Day
Week
Month
Year

Multivariate Correlative Predictive Analytics: Three Approaches

- Automatic correlation computation against lead/lags
- Temporal and spatial windowing
- Data category parameter space



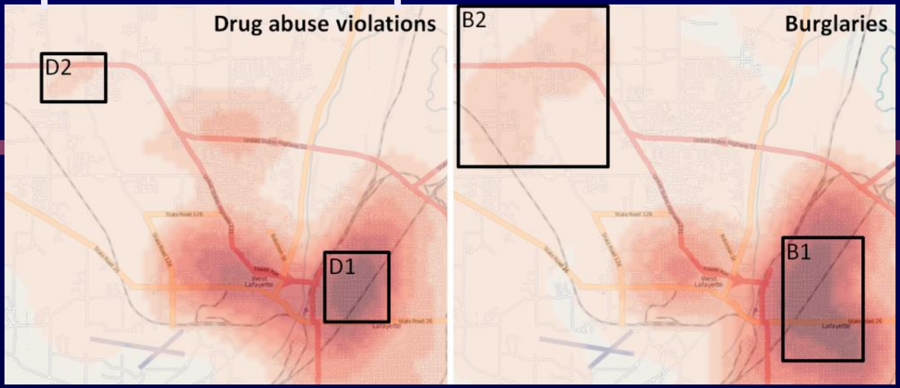
Example: Drug Abuse Violations Vs. Burglaries – Focus On Geospatial Hotspots

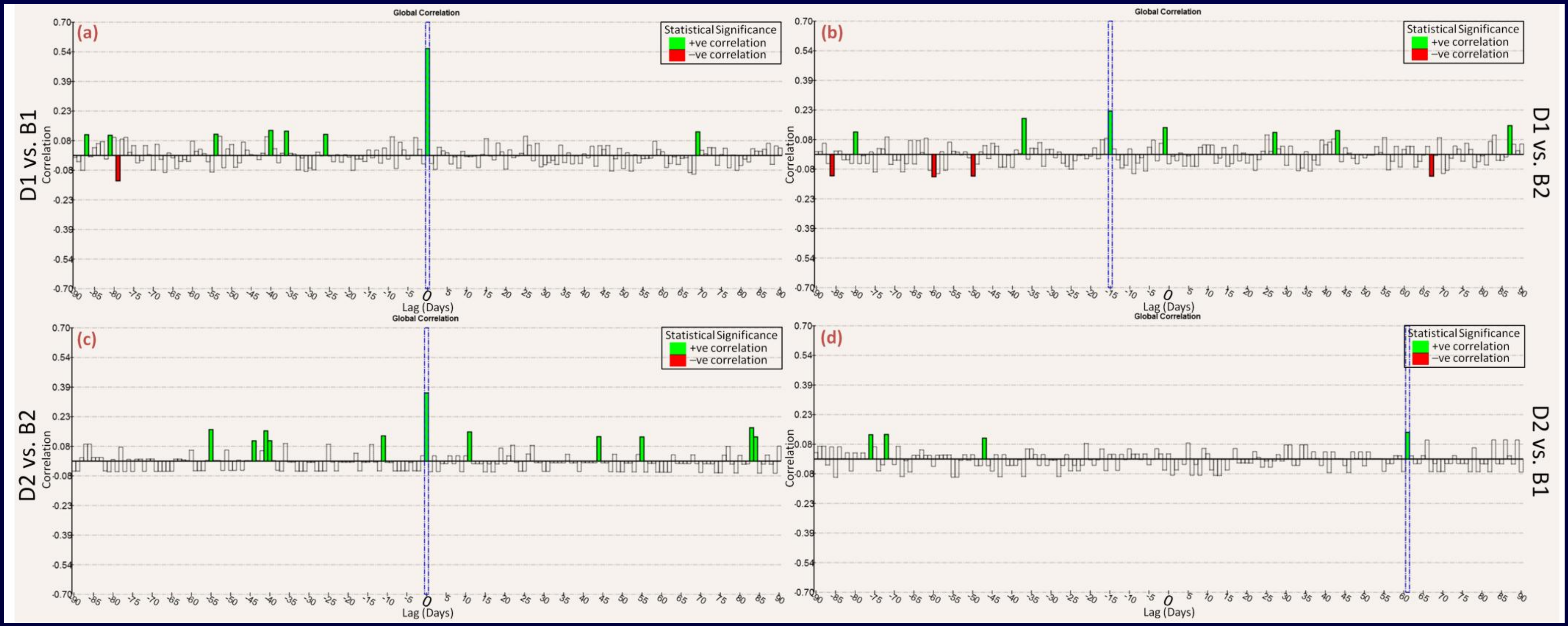


VACCINE

Example: Drug Abuse Violations Vs. Burglaries – Focus on Geospatial Hotspots

- High positive correlation in same neighborhood with zero lag

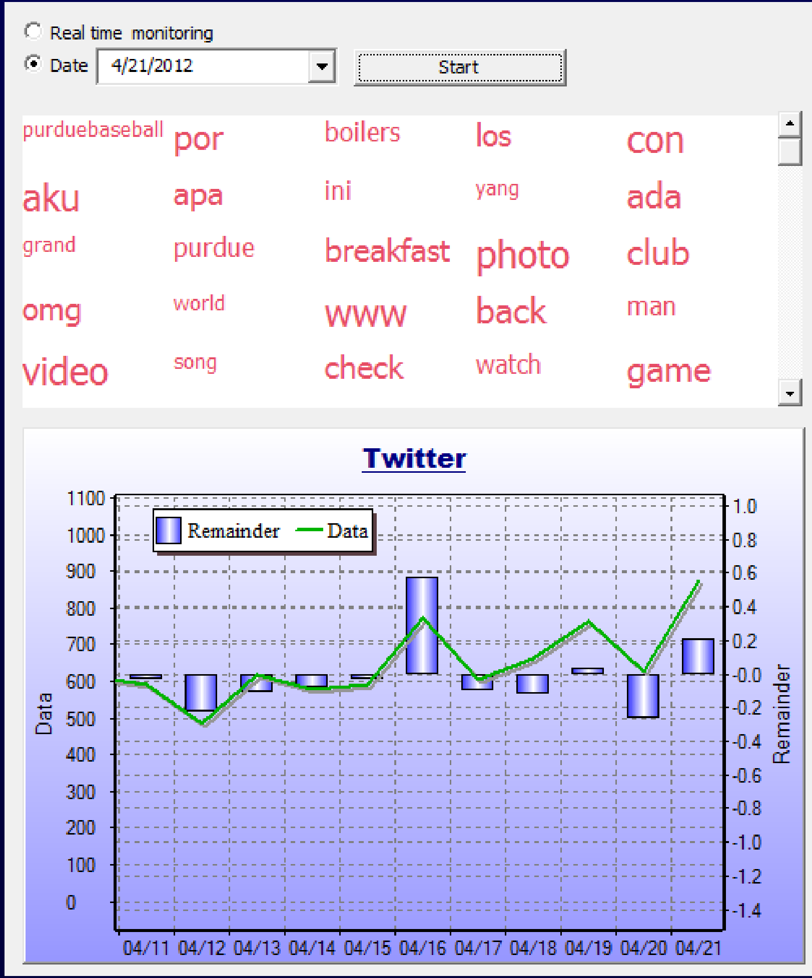




VACCINE

Topic 4: Social Media Analytics :Real-time Twitter Monitoring and Integration into Tools
(Purdue, Stuttgart, Penn St.)

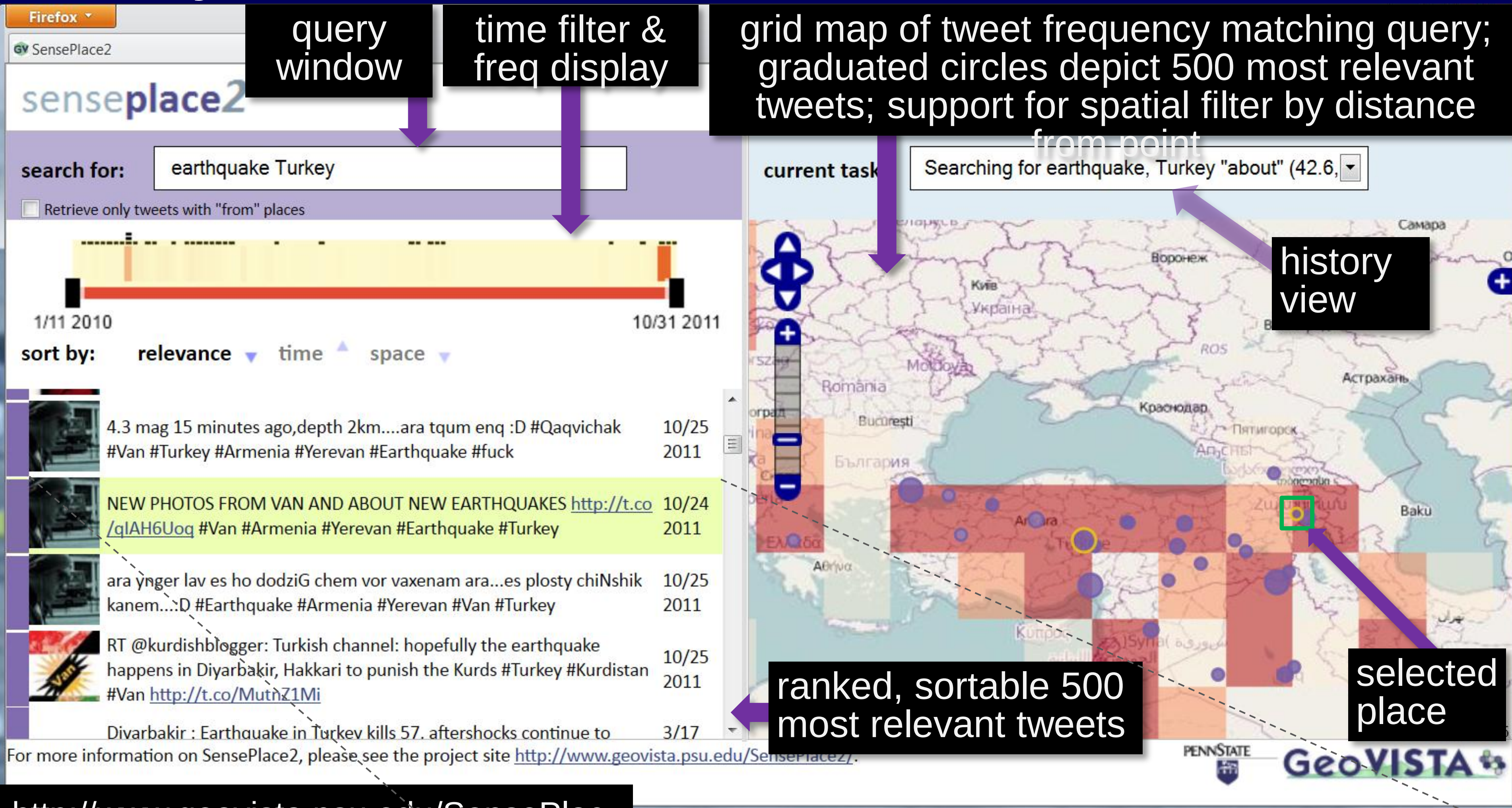
- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud



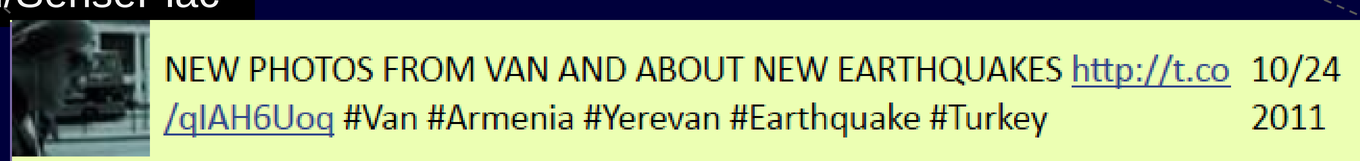
Grand Prix Weekend, Purdue University

VACCINE

SensePlace2: Place-Time-Concept Analytics (PSU)



http://www.geovista.psu.edu/SensePlace2/

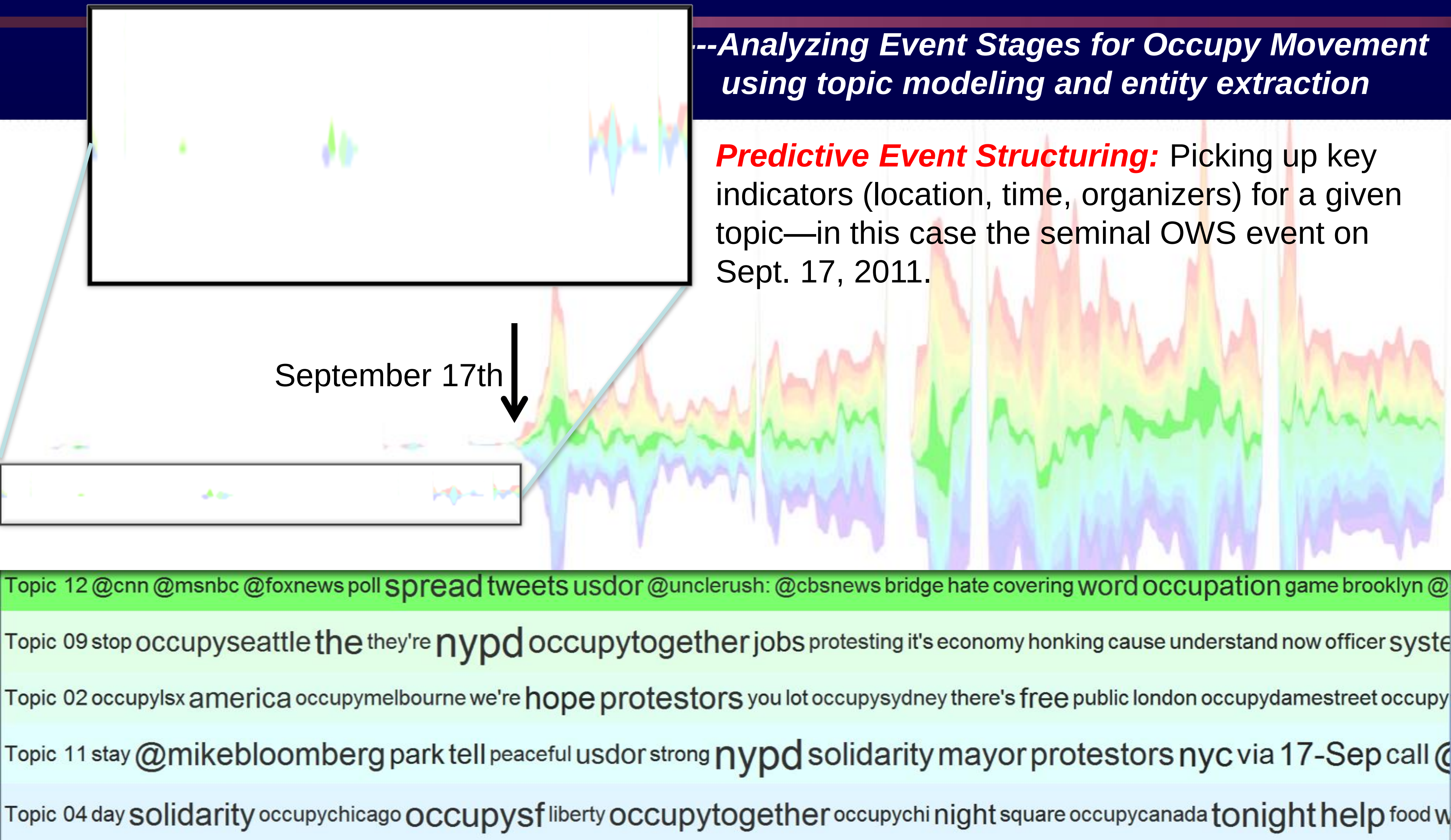


VACCINE

Deep Data Analytics: Temporal Event Structuring from Streaming Media (UNCC)

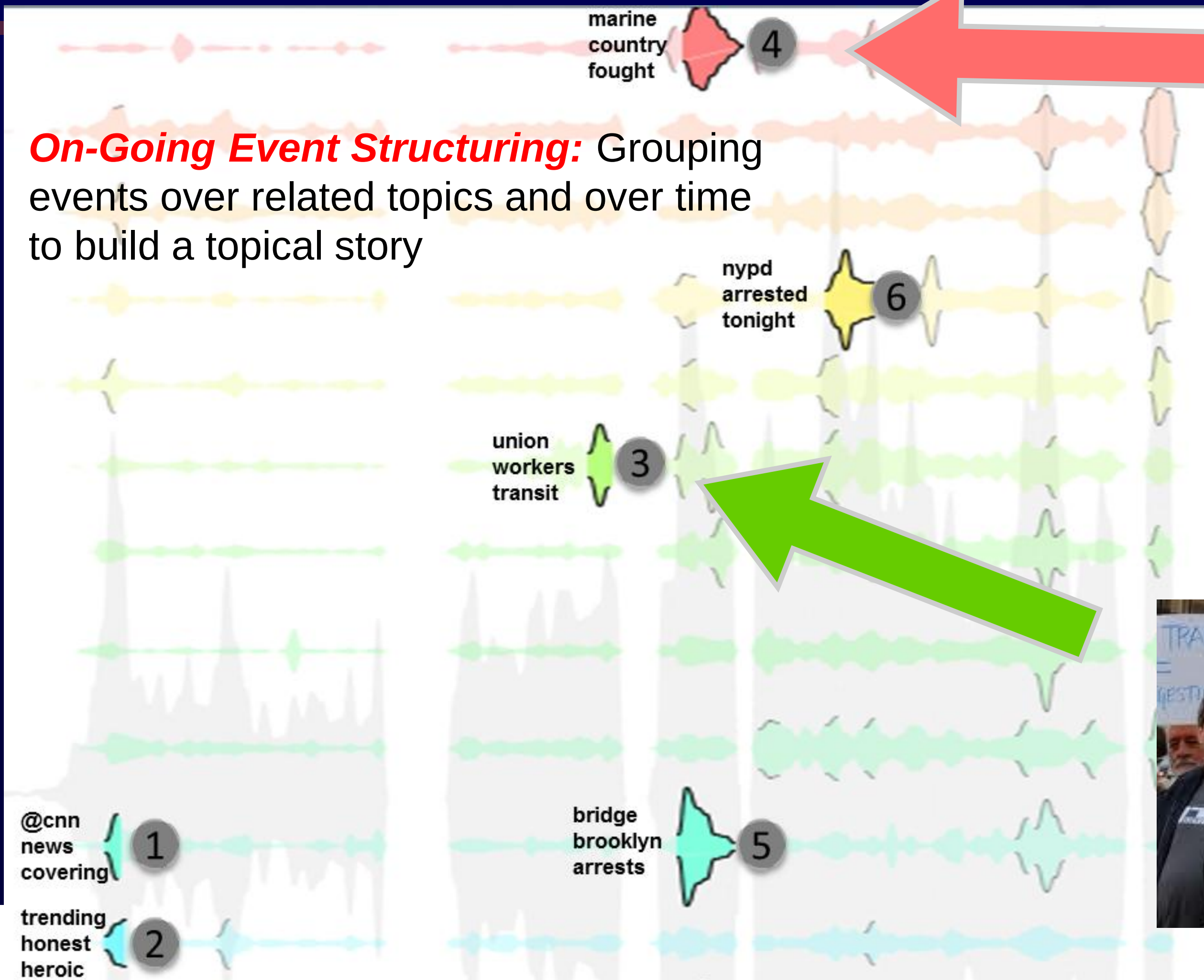
---Analyzing Event Stages for Occupy Movement using topic modeling and entity extraction

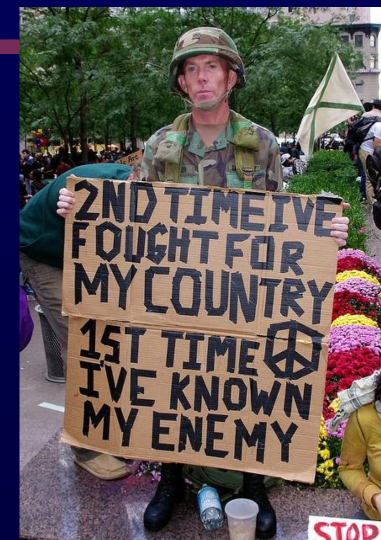
Predictive Event Structuring: Picking up key indicators (location, time, organizers) for a given topic—in this case the seminal OWS event on Sept. 17, 2011.



Events (bursts of activity for a given topic)
Example: Story of Occupy Wall Street

On-Going Event Structuring: Grouping events over related topics and over time to build a topical story



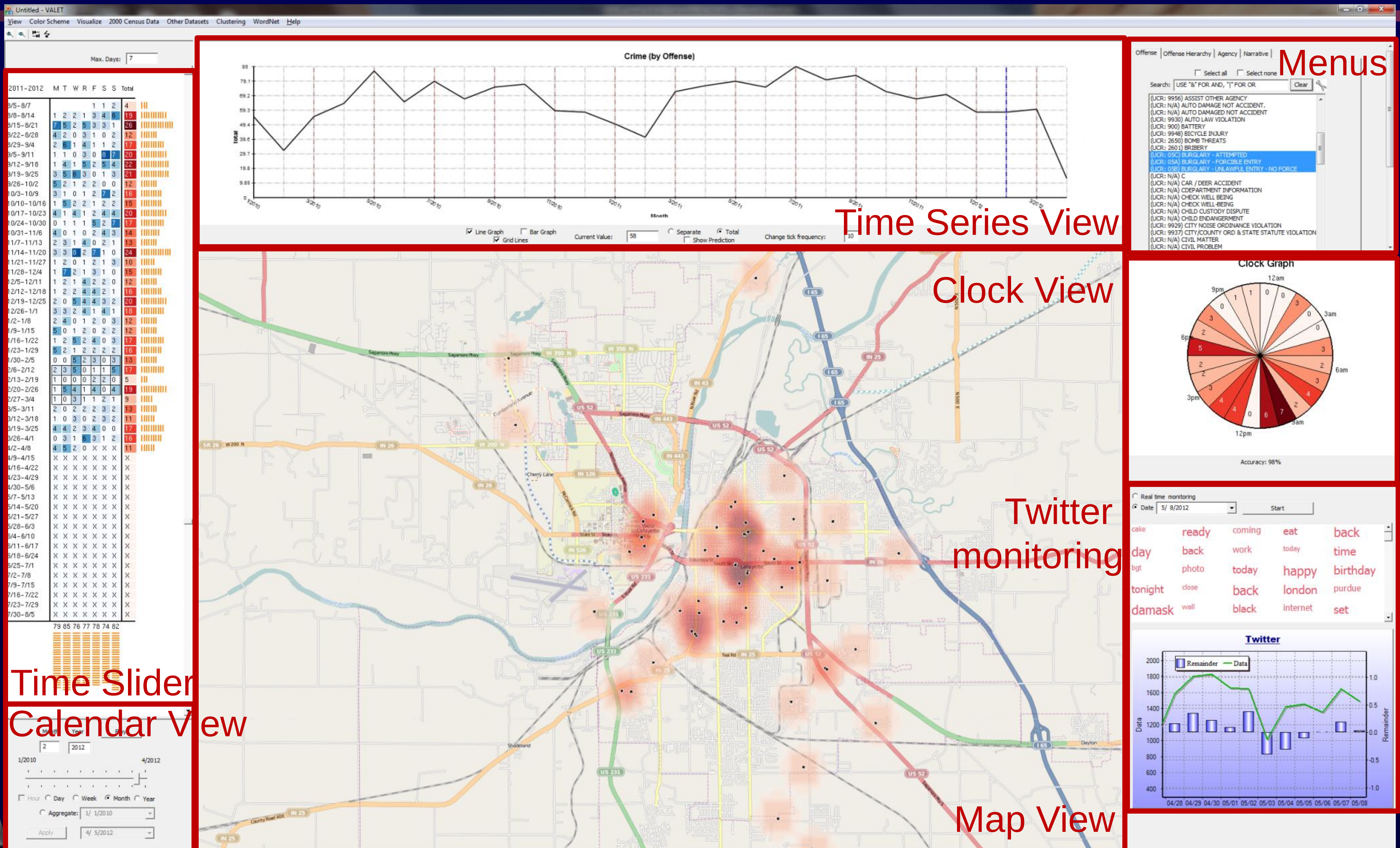


Sept. 17

Topic 5: Complex Data

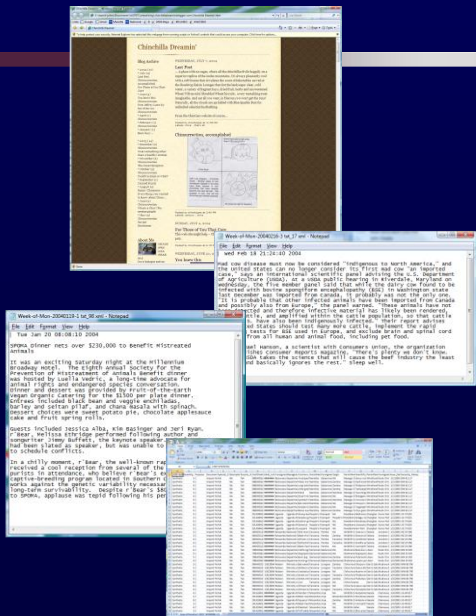
- Deep Data Analytics tool
  - Multisource news feeds and twitter feeds
- Scatterblogs
  - Twitter, Flickr, YouTube with anomaly detection and prediction
  - Developed predictive capability of each vs. event characteristics
  - VALET – multisource data
    - Crime case data, dispatch data, EMS data
    - Weather data
    - Social media
    - Gang graffiti
    - Public civil/legal court data to be added
  - Jigsaw
  - Cybersecurity VA

VALET Overview



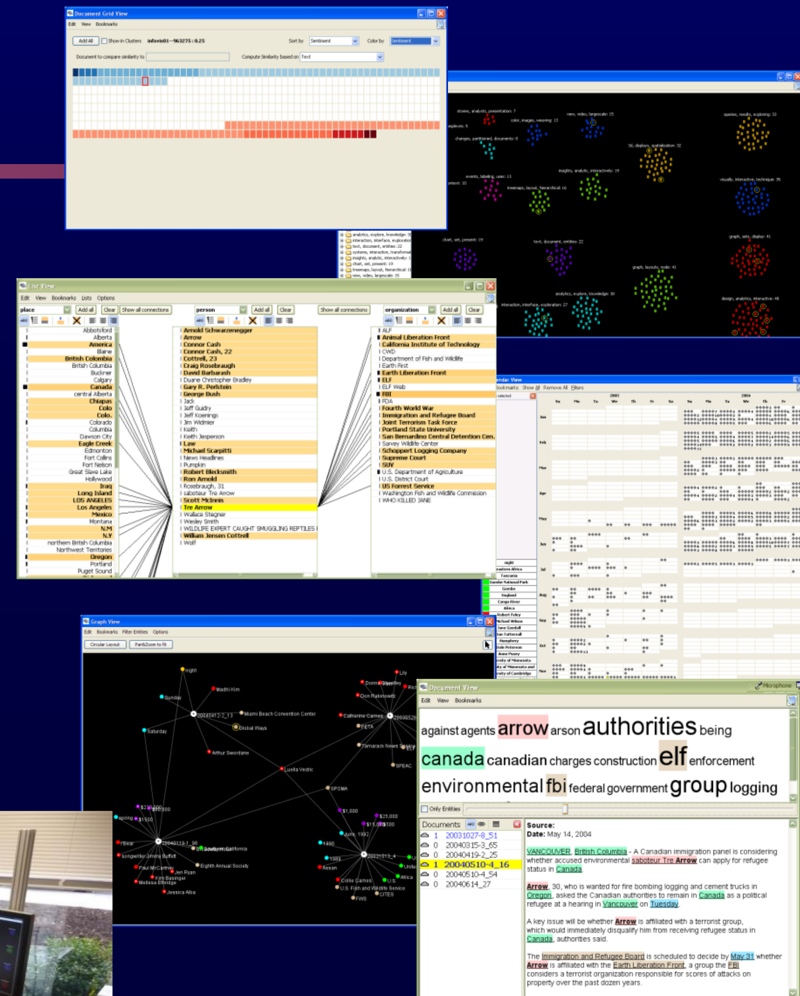
Jigsaw: Visual Analytics for Investigative Analysis

John Stasko



Documents in

Knowledge out

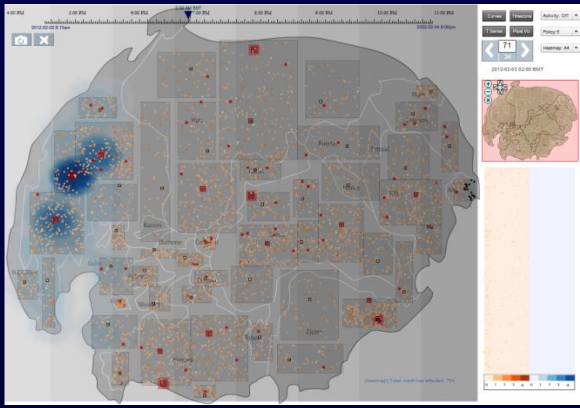


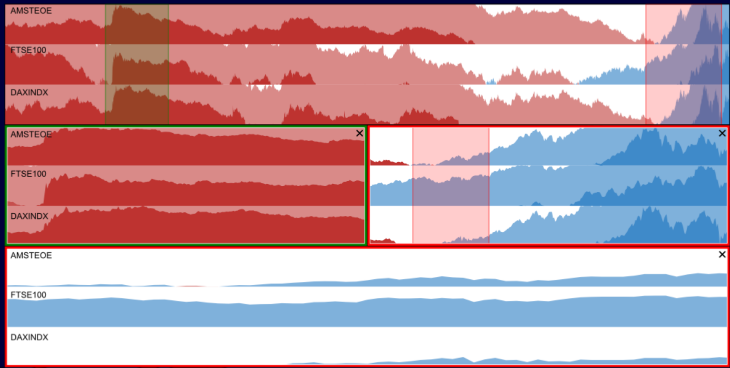
VACCINE

Georgia Tech College of Computing
School of Interactive Computing

Example Work in Cybersecurity

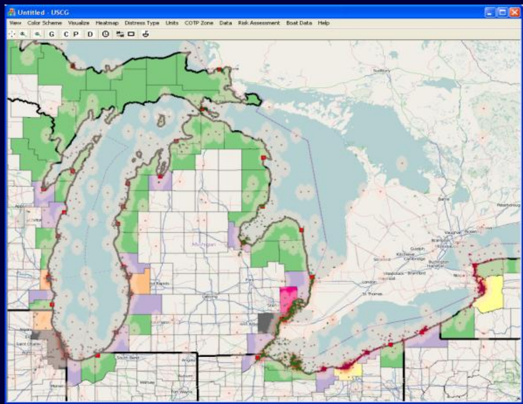
- Corporate Insider Threat Detection (Oxford, Leicester, Cardiff)
- Sensor Forensics (Purdue)
- SemanticPrism (Purdue)
- Multiscreen, Multiview Interactive Cyber Investigation (VaTech, PNNL)
- Log Visualization (Purdue)





Topic 6: Visual Analytics Uses for Risk-Based Decision Making

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



VACCINE

USCG: Effective Risk-based Decision Making and Resource Allocation Visual Analytics

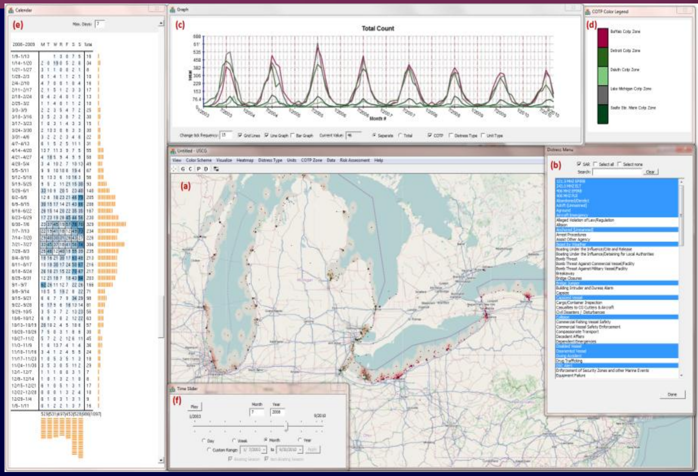
- Evaluate current and historical mission area:
  - Demands
  - Risks (total, mitigated, residual)
  - Resource allocation
  - Return on investment
- Evaluate courses of action
- Evaluate above at both Strategic and Tactical/Operational level

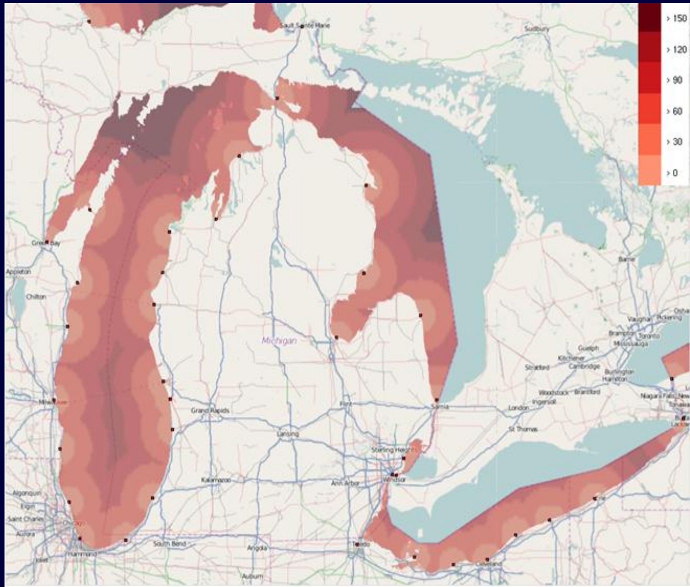
U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7, USCG D9, USCG D5, USCG HQ 771

IMPACTS:

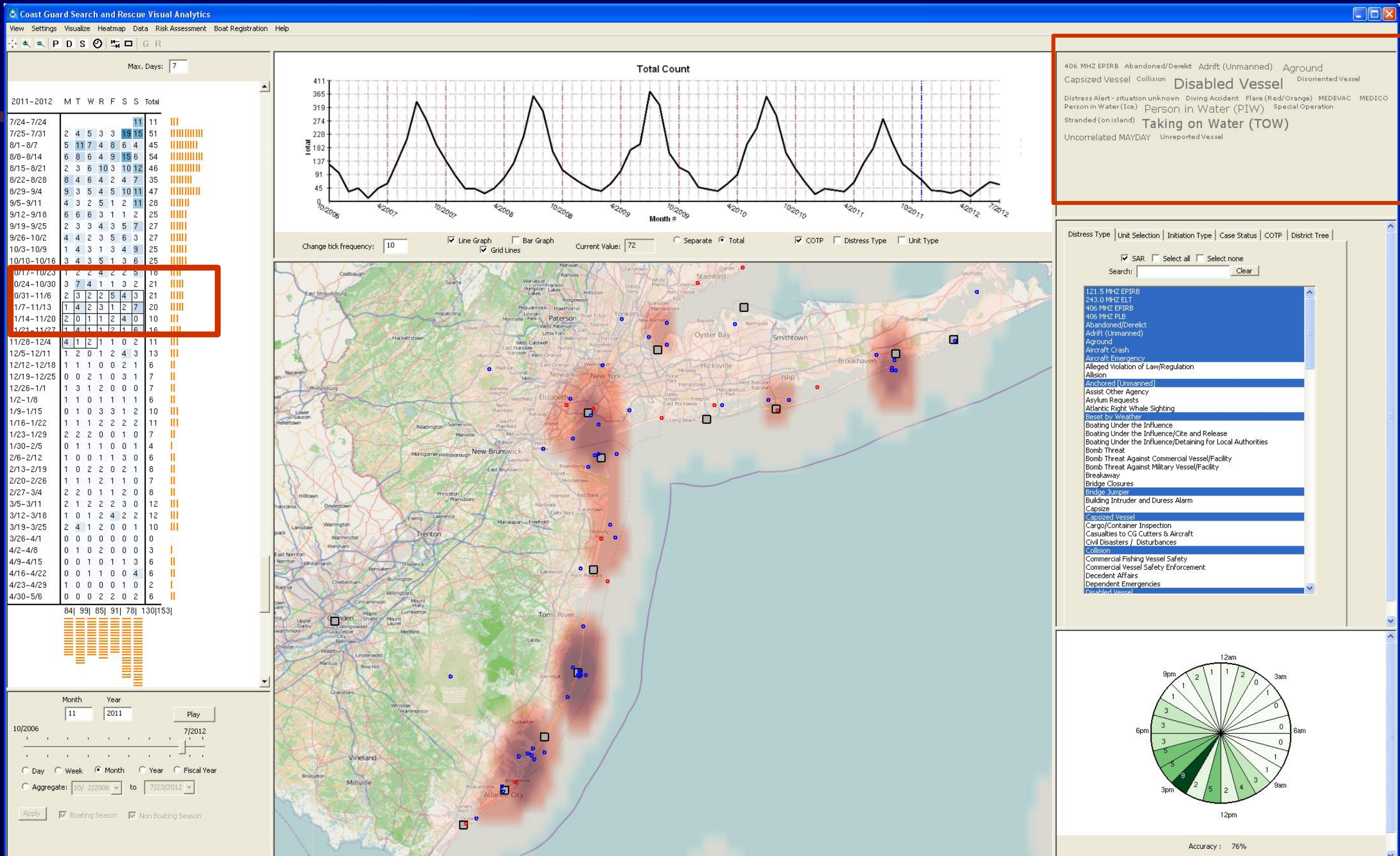
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Used for response to SuperStorm Sandy
- Hurricane Irene resource allocation decision based on cgSARVA analysis and visualization
- Informed Commandant's budget testimony and recommendations to Congress
- Key component of USCG D9 reallocation plan for 2011-2012 based on decreased budget
- Key component of Coastal Operations Allocation Suite of Tools (COAST) – USCG HQ





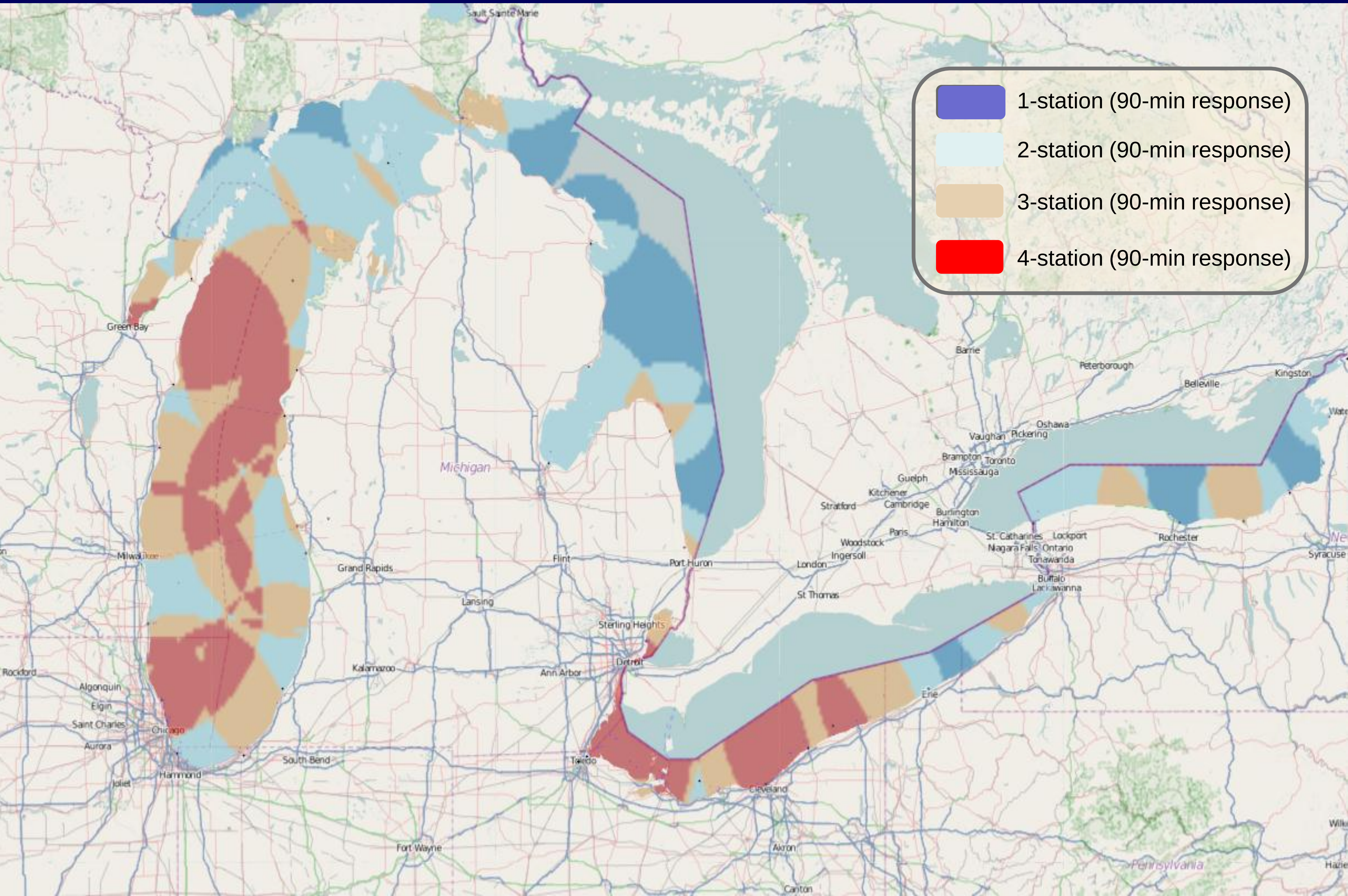
Example: Risks and Consequences From Sandy:

SAR Cases November 2011 NJ/NYC Area



VACCINE

Response Efficiency – Potential Future Assets



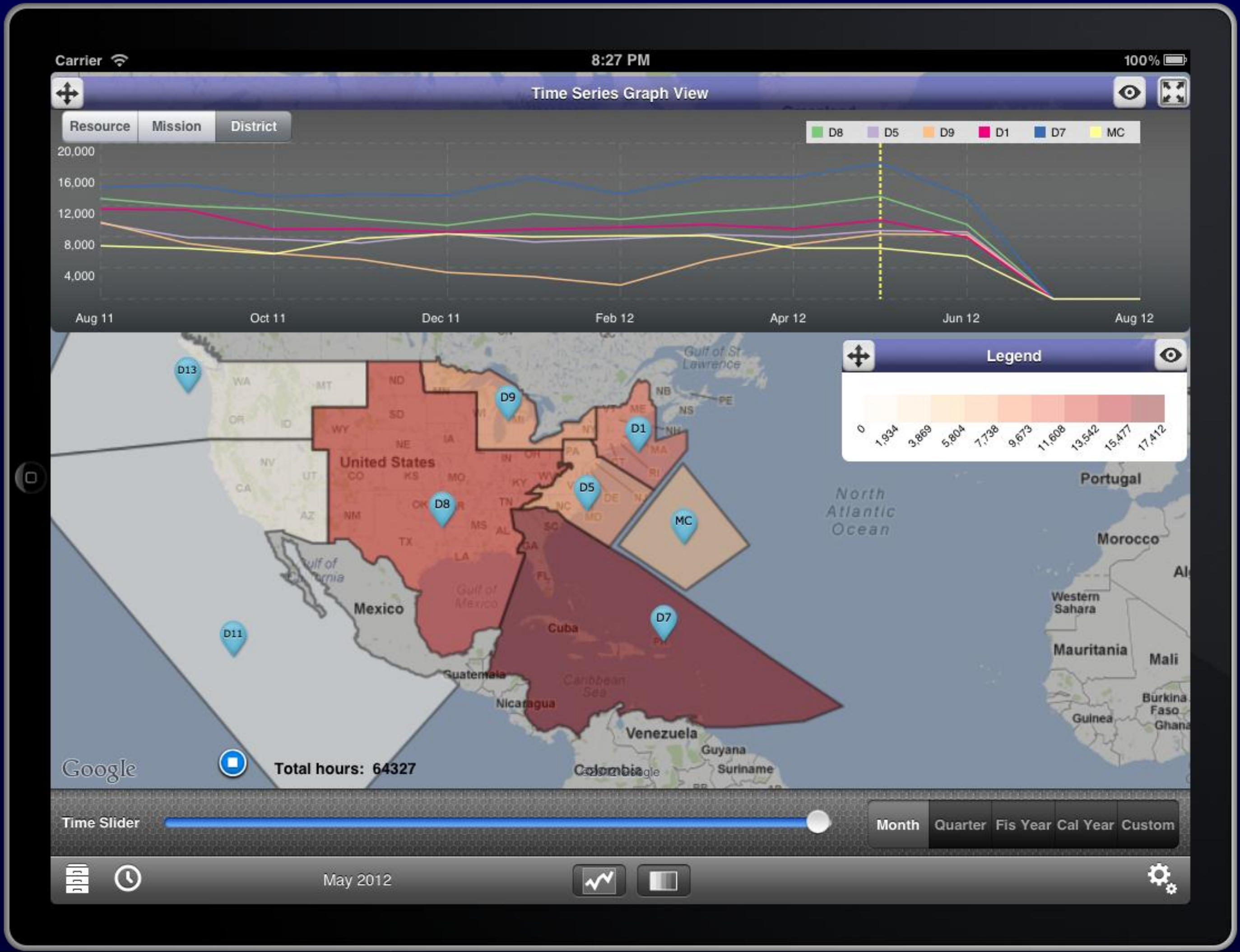
Topic 7: Non Data Scientists

Interactive Operational Performance Assessment Report (iOPAR)

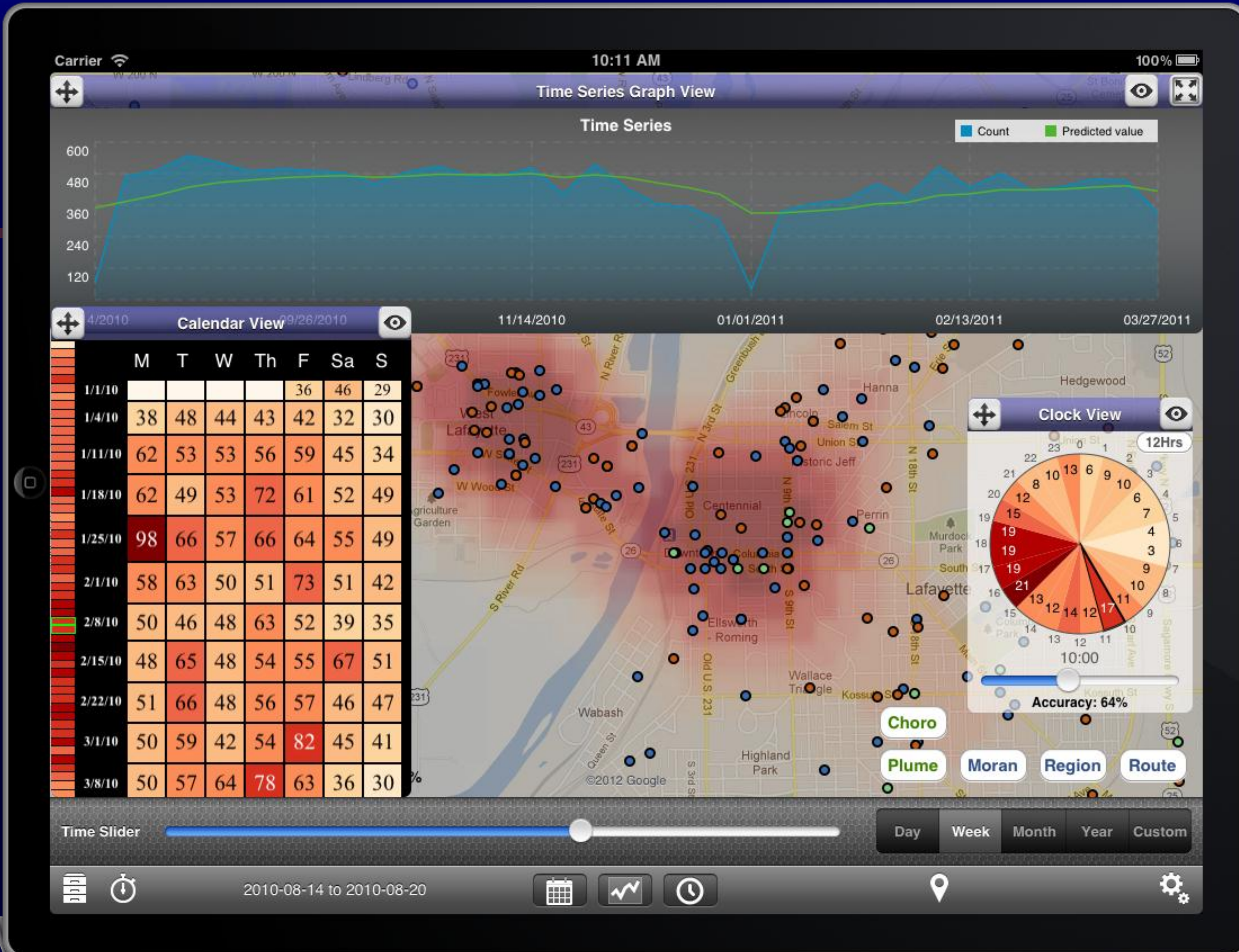
- OPAR
  - Standard report within the Coast Guard Business Intelligence (CGBI) system
  - Displays resource use and performance by core CG mission areas
- iOPAR – interactive iPad version to allow interactive visualization and analysis and inform decision making
  - Anaylze/visualize performance, targets, seasonal trends, predictions by boat, aircraft, cutter type, region levels

iOPAR Alpha Version: Spatiotemporal View

- May 2012



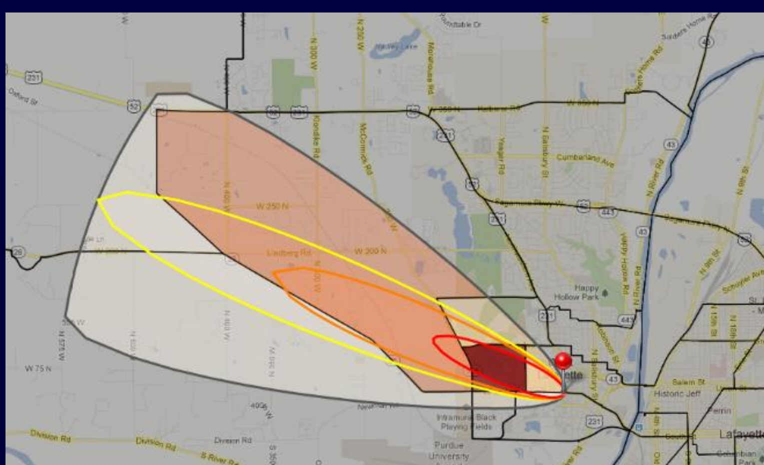
VACCINE

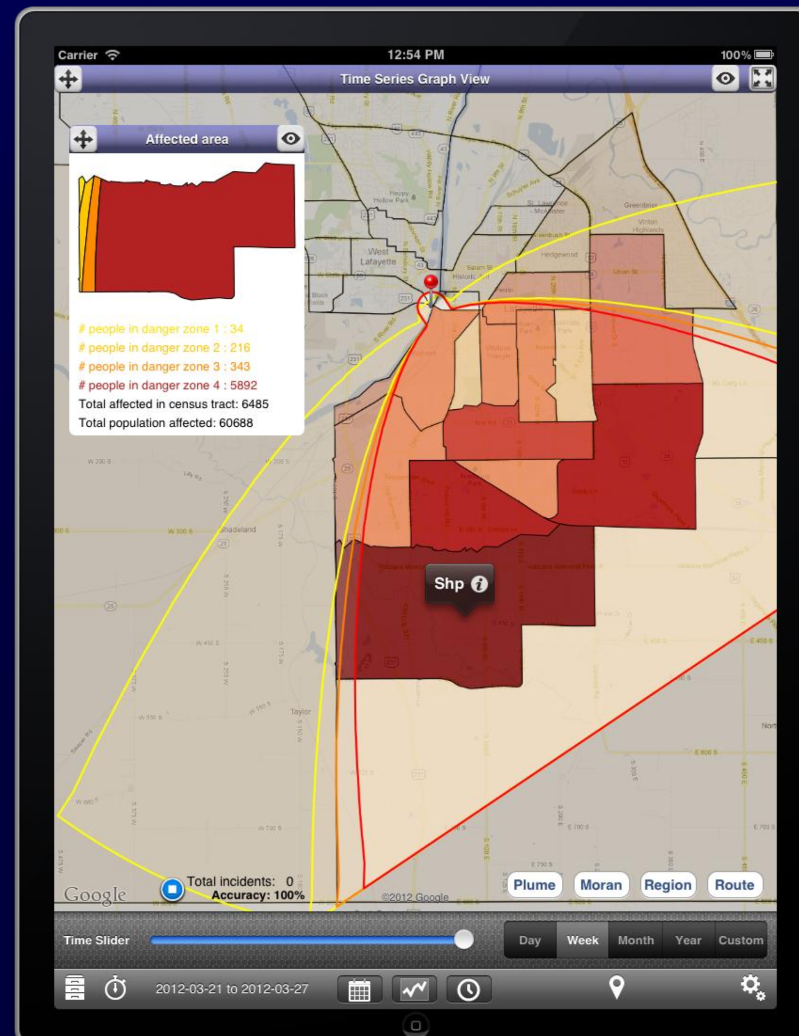


Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
11/14/2010
01/01/2011
02/13/2011
03/27/2011
Calendar View
M T W Th F Sa S
1/1/10 36 46 29
1/4/10 38 48 44 43 42 32 30
1/11/10 62 53 53 56 59 45 34
1/18/10 62 49 53 72 61 52 49
1/25/10 98 66 57 66 64 55 49
2/1/10 58 63 50 51 73 51 42
2/8/10 50 46 48 63 52 39 35
2/15/10 48 65 48 54 55 67 51
2/22/10 51 66 48 56 57 46 47
3/1/10 50 59 42 54 82 45 41
3/8/10 50 57 64 78 63 36 30
Clock View
12Hrs
10:00
Accuracy: 64%
Choro
Plume
Moran
Region
Route
Hedgewood
Lafayette
Wabash
Highland Park
©2012 Google
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20

MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

- Chemical release plume modeling identifies census tracts with the highest number of expected people affected





Topic 8: Active Data

- Incorporating automatic predication and relation algorithms to incoming data
- Exploring methods for visualizing/alerting user:
  - Passive visualizations on screen (e.g., related tweets, related incidents and how related)
  - Periodic alerts based on some criteria
    - Top 10 crimes that occurred in areas with unusual volume
    - All crimes that occurred in anomalous (>2SD) areas

Topic 9: Video and Sensor Analytics

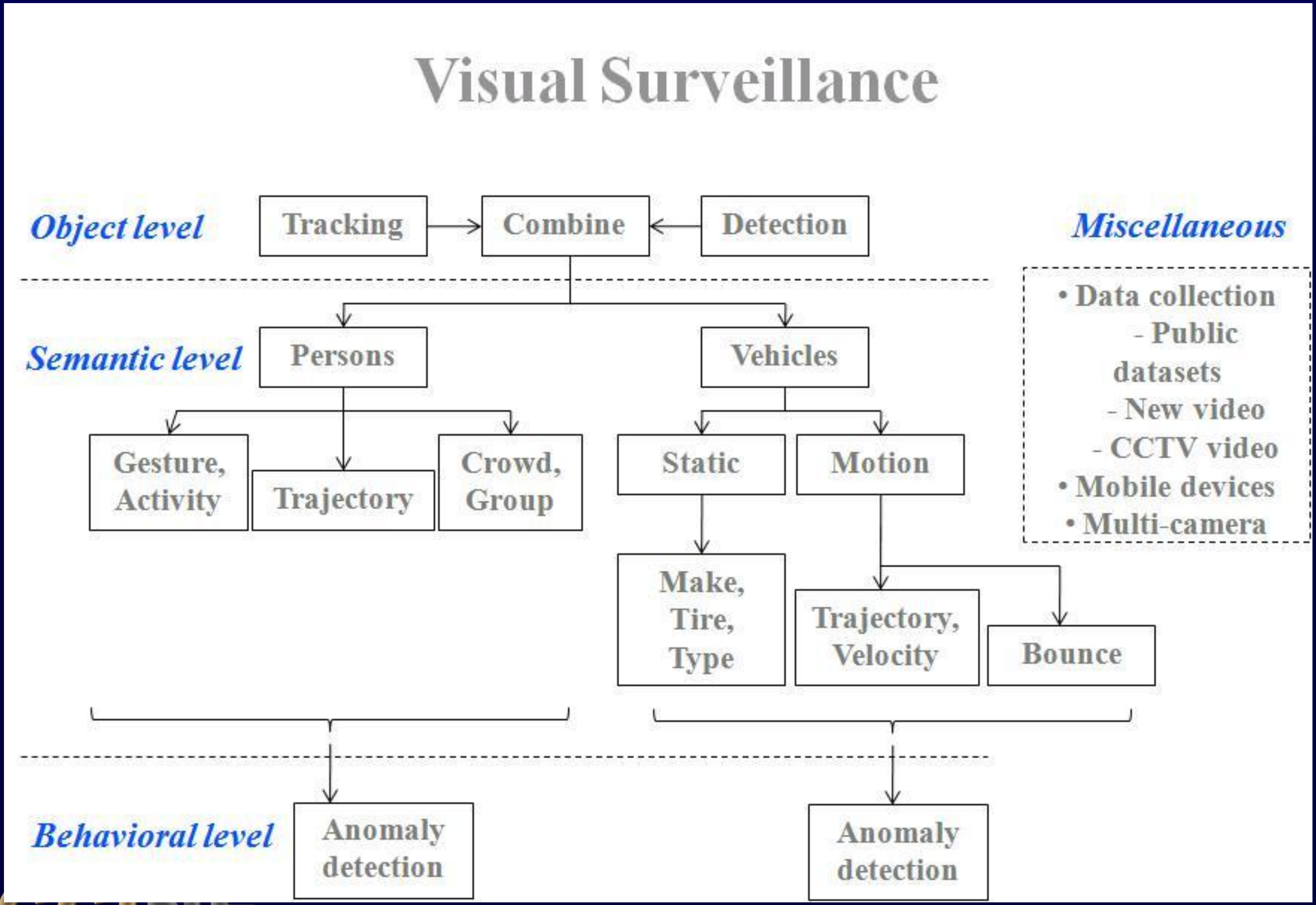
Sensor Forensics – Based on *Signatures*

- Forensic characterization
  - Observe device output ➔ which device produced it?
  - Exploit how the device “makes” its output
- Device authentication
  - Performed using forensic characterization
  - Identify device type, make, model, configuration
  - Can the sensor be trusted?
- Detection of data forgery or alteration
- Fingerprint and trace
- Devices: Printers, Cameras, Scanners, Sensors Nodes, RF Devices

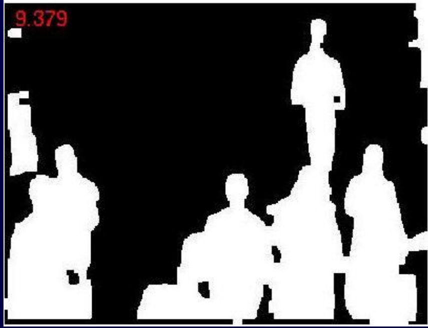
Visual Surveillance

- Video contains enormous information
- Systems “easy” to obtain/install
- Allows monitoring from a remote location
- Passive operation, traditionally a forensic tool
- Manned operations prone to human error
  - Fatigue, distraction, information overload
- Automated surveillance systems:
  - Model/understand/summarize the scene
  - Assist human operator by filtering information

Visual Surveillance



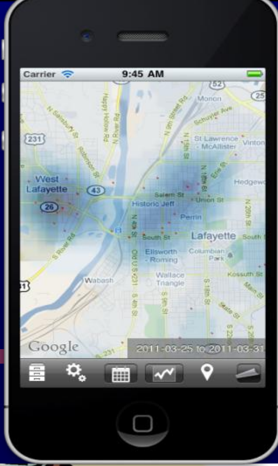
Visual Surveillance Topics:

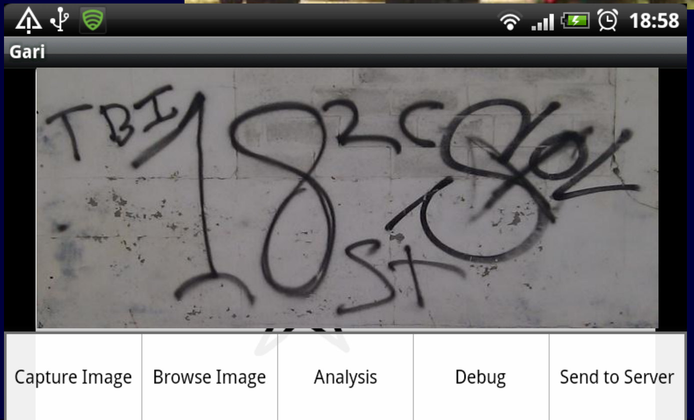


- Moving object detection and tracking
- Video-analysis based video summarization
- Crowd Density and Flow Estimation
  - Crowd flow, congestion, unusual patterns

Topic 10: In-Field Analytics: We Have an App for That!

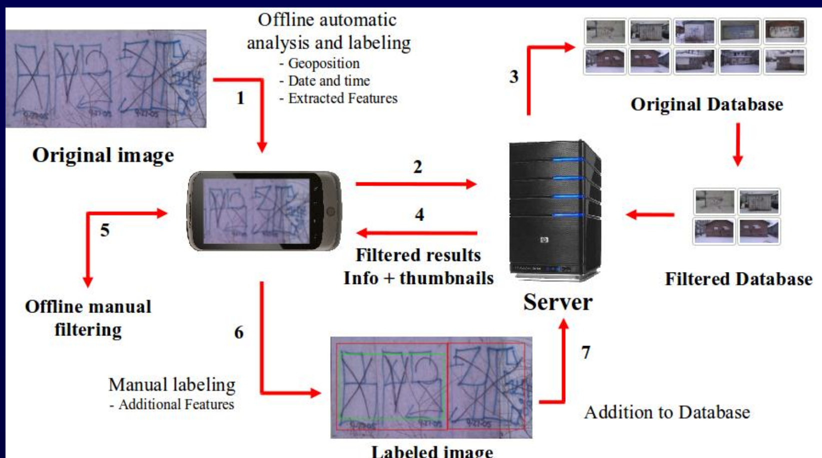
- iVALET
- Evacuation Planning
- Rosetta Phone
- Hazmat app
- Gang graffiti app
- Tatoo app





VACCINE

Gang Graffiti Analysis and Recognition Using a Mobile Telephone (GARI)



IMPACT:

- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be "interpreted"
- More than 60 users and 500 graffiti images acquired

GARI delivered:

- Summer 2011:
  - IPD gang detectives
- August 2011:
  - IPD at large
  - Ind Fusion Gang Task Force
- September 2011:
  - Gang detectives across Indiana

VACCINE

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu
ebertd@purdue.edu

VACCINE

Table of Contents

Table of Contents

CVADA

Center for Visual Analytics and Data Analytics

A U.S. Department of Homeland Security Center of Excellence

Visual Analytics for Financial Analytics, Fraud, and Decision Making

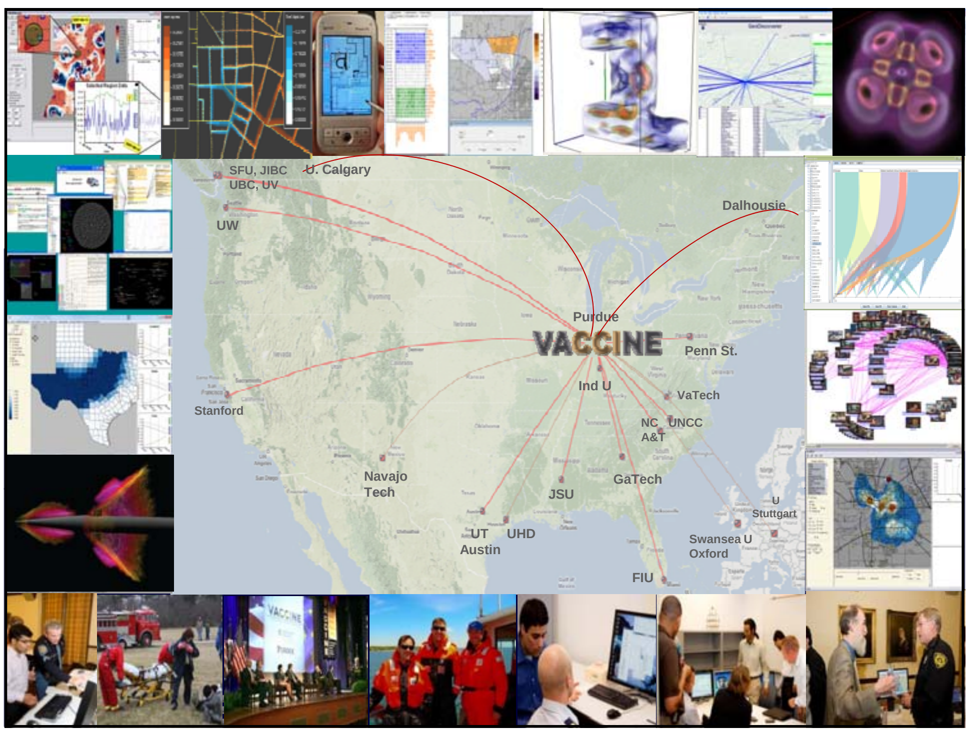
David S. Ebert

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

February 2013



Who We Are: International Team of Experts - 75+ Faculty, 25 institutions

- Purdue University
- Georgia Institute of Technology
- Pennsylvania State University
- Stanford University
- University of North Carolina at Charlotte
- University of Washington
- Simon Fraser University
- University of British Columbia
- Justice Institute of British Columbia
- Ontario Institute of Technology
- Dalhousie University
- University of Houston, Downtown
- Virginia Tech
- Indiana University
- Florida International University
- University of Texas at Austin
- Morgan State University
- Navajo Technical College
- University of Stuttgart
- University of Swansea, UK
- Oxford
- University of Calgary
- University of Manitoba
- Carleton University
- University of Victoria

VACCINE Multidisciplinary Team

- Experts in computer science, image analysis, signal processing, cognitive science, decision theory, management science, geography and GIS, HCI, visualization, HPC, statistics, political science …
- Partners in nursing, public health, economics, databases, command and control, emergency management

VACCINE's Value

We enable users to be more effective through innovative interactive visualization, analysis, and decision making tools

- Provide the right information, in the right format within the right time to solve the problem
- Turn data deluge into a pool of relevant, actionable knowledge
- Enable user to be more effective from planning to detection to response to recovery
- Enable effective communication of information

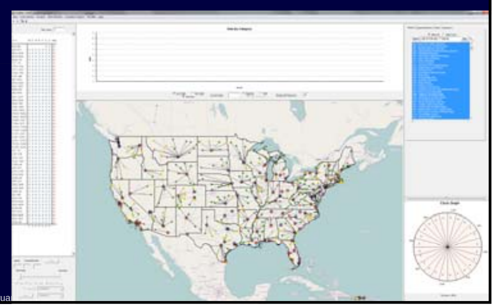
Approach: Partner-driven solutions and research

Engaged End-Users

- Federal Operating Compoents:
  - US Coast Guard
  - US Transportation Security Agency
  - US Federal Emergency Management Agency
  - US Customs and Border Patrol
  - US CERT
- Law Enforcement
  - Over 40 local and state agencies
- Fusion Centers
  - Ohio and Indiana

Example Projects

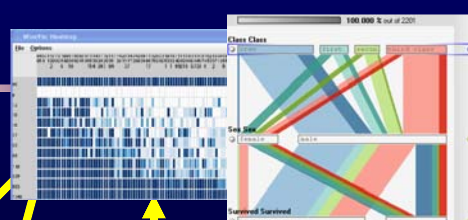
- Financial Visual Analytics
- TSA
  - Integrating staffing, financial, performance metric, legal data into a unified VA system
- Law Enforcement
- USCG – Risk-based decision making



TSA System

The Pipeline for Financial Anomaly Analysis

All transaction activity



Identify

Prioritize

Interactive Visualization

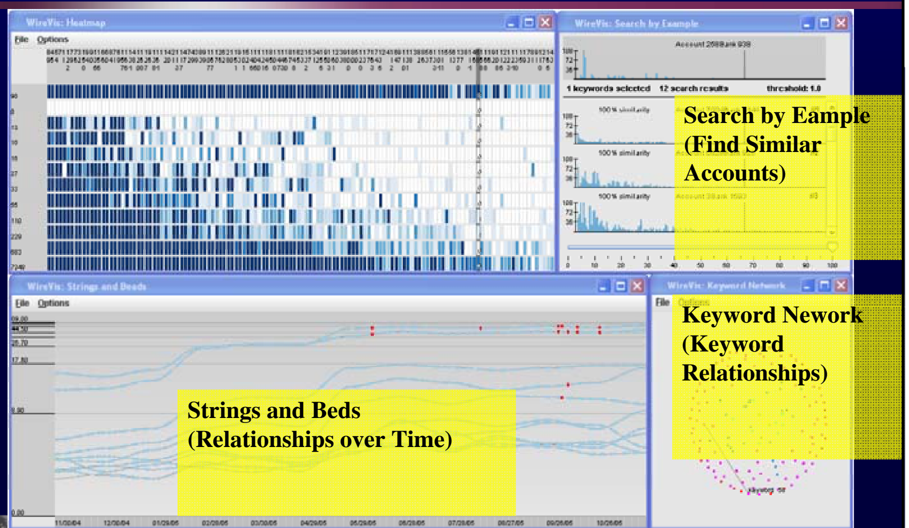
Google

Investigate

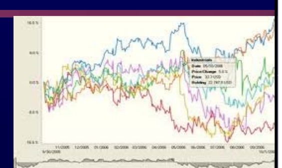
Report

WireVis – Streaming Data - Multiple Linked Views

- Temporal, geospatial, theme, cluster, list views with association linkages between views

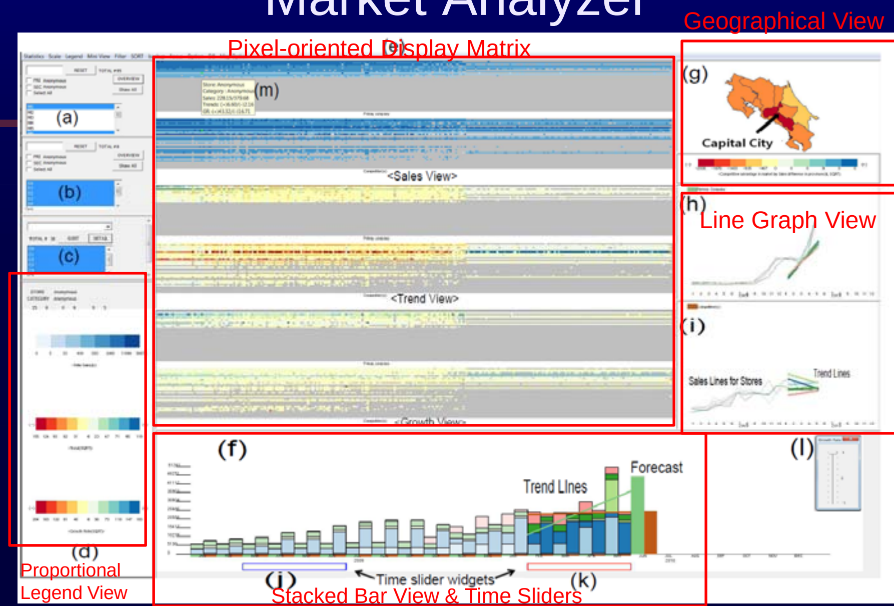


Competitive Business Intelligence Based On Point of Sales Data



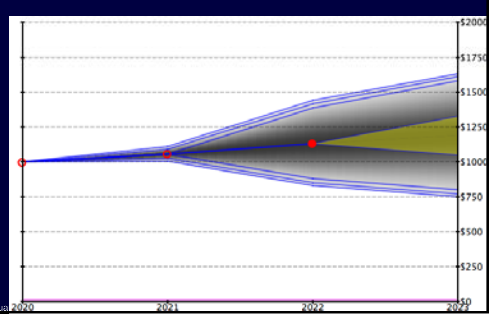
- Characteristics of point of sale data
  - Multivariate
    - Large # of dimensions
      – E.g., 38 categories(products) in 288 stores
  - Temporal : 18 months
  - Spatial
    - Stores located all over Costa Rica
  - Requirements
    - Supporting easy comparison among companies
      – E.g., visualize all data at once, sorting by importance
      – Enable geographical comparison
    - Easy recognition of any change in sales
      – E.g., proportional legends
    - Forecast for decision-making

Market Analyzer



FinVis: Applied Visual Analytics for Personal Financial Planning

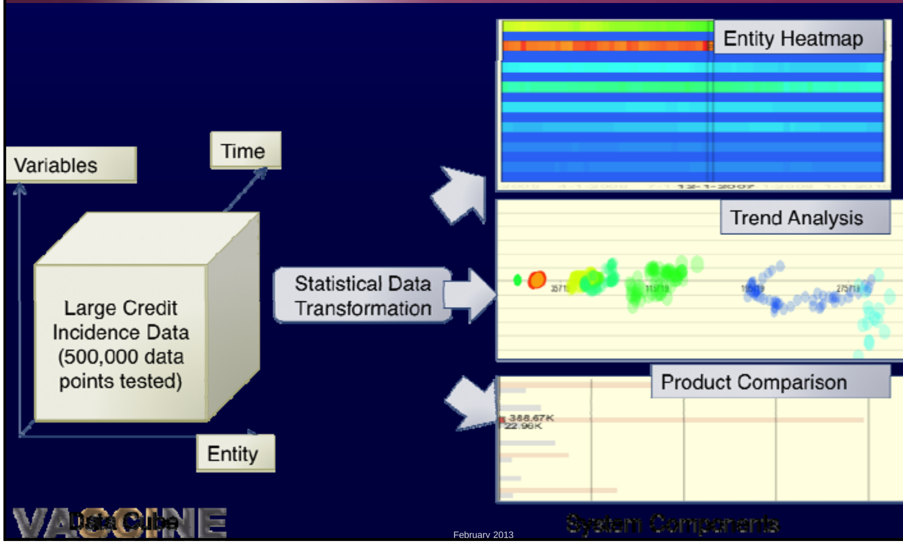
- First visual analytics work to address portfolio planning for casual users
  - *General public could benefit from this tool*
- Decisions are better when using FinVis
  - *Quantified using experimental economics*
  - *Improved*
    - *Decisions*
    - *Learning*
    - *exploration,*
    - *confidence*



RiskVA: Key Credit Risk Analytics Challenges

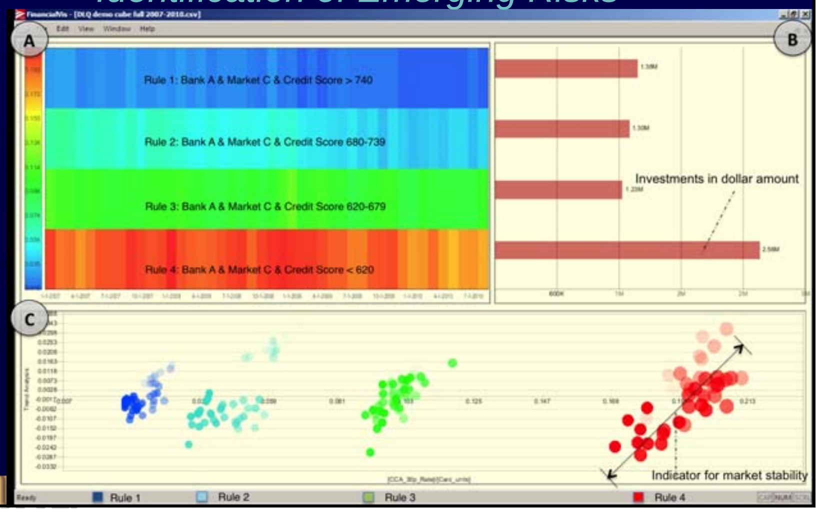
- Consumer credit data is large, temporal, and related across multiple investments and financial markets.
- The data are heterogeneous, not clean, have missing values, may be misleading and inefficient to explore.
- The data contains important behaviors and relations/groupings that change over time.
- Analysts need to have customer-centric interactive systems to achieve full analysis of risks

RiskVA Data Integration



RiskVA Overview: Interactive Exploratory Visualizations

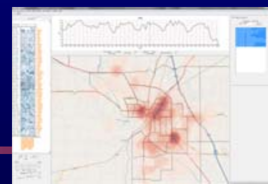
Identification of Emerging Risks

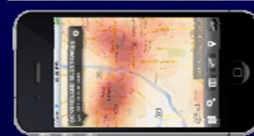


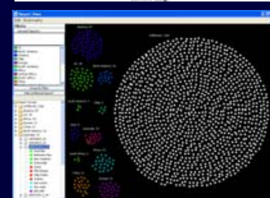
Sensor Forensics – Based on *Signatures*

- Forensic characterization
  - Observe device output ➔ which device produced it?
  - Exploit how the device "makes" its output
- Device authentication
  - Performed using forensic characterization
  - Identify device type, make, model, configuration
  - Can the sensor be trusted?
- Detection of data forgery or alteration
- Fingerprint and trace
- Devices: Printers, Cameras, Scanners, Sensors Nodes, RF Devices

Law Enforcement







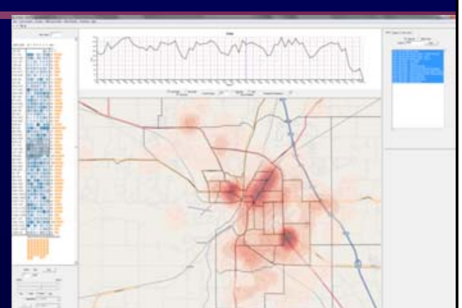
- Law enforcement visual analytics
  - VALET, iVALET (iPhone/iPad), CrimeViz
    - Visual Analytics Law Enforcement Toolkit
    - Analyzing crime patterns and time of day problems
- Gang activity analytics
  - GARI
    - Gang Graffiti Recognition and Interpretation using a mobile telephone
    - Allowing police to catalog and analyze gang graffiti images, better track and determine gang activities
- Document visual analytics
  - JIGSAW
    - Visualization for investigative analysis
    - Discovery of hidden relationship and threats across documents

VACCINE

Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

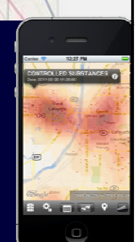
Impacts:

- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (September 2012) for community-based policing
- Investigating correlation of bus routes and crime, street lights and crime
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning



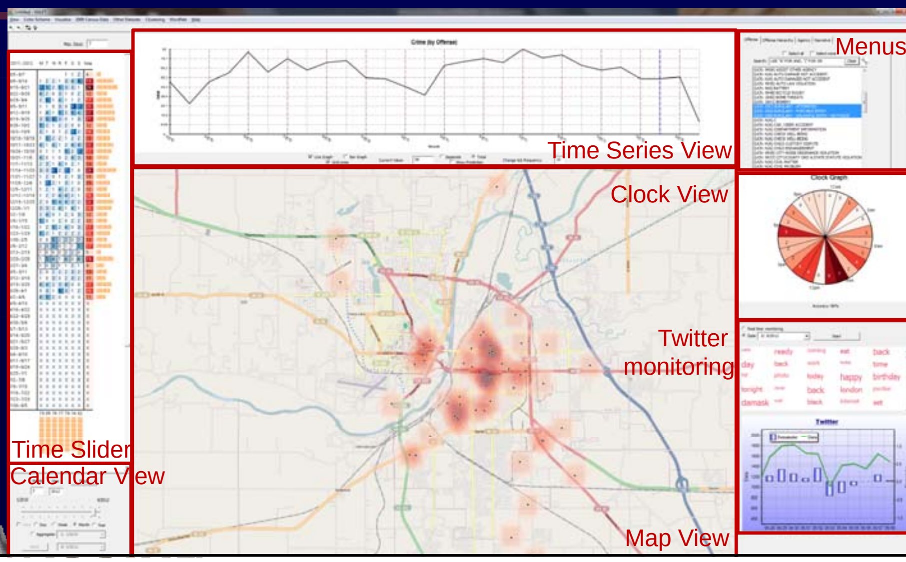
VALET delivered:
- Spring 2011: WL, Lafayette Police

iVALET delivered:
- October 2011: Purdue, WL Police



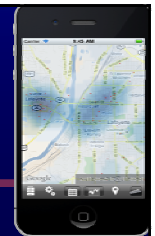
VACCINE

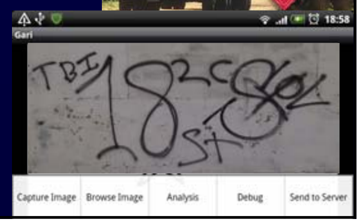
VALET Overview



We Have an App For That!

Our Mobile Tools





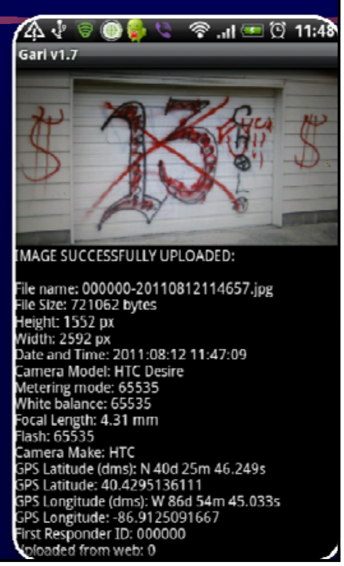
- VALET
- Evacuation Planning
- Rosetta Phone
- Hazmat app
- Gang graffiti app
- Tatoo app

VACCINE

February 2013

Gang Graffiti Recognition and Analysis Using a Mobile Telephone (GARI) - There's An App For That

- What is GARI?
  - A mobile device application that analyzes gang graffiti
- How does GARI work?
  - User takes an image of the graffiti
  - User receives an analysis
    - Geographic locations
    - Graffiti colors, shapes, meaning
  - Database search of similar tags



VACCINE

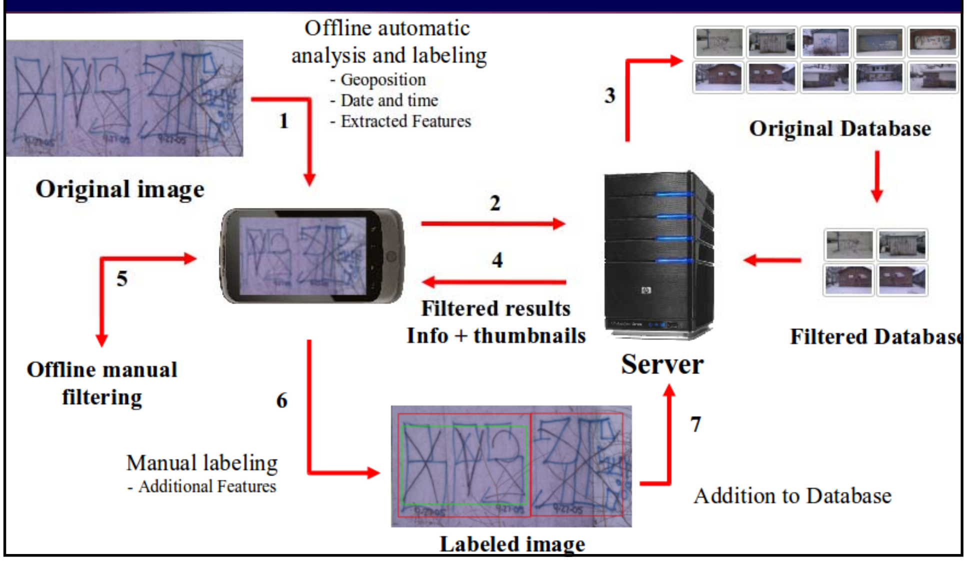
February 2013

GARI On The Street

- Over 50 law enforcement officials are using or field testing GARI
- Delivered to:
  - IMPD gang detectives (Summer 2011)
  - IMPD at large, Indiana Fusion Gang Task Force (August 2011)
  - Gang detectives across Indiana (October 2011)
- Homeland Security applications

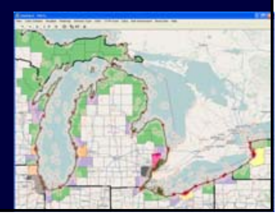
VACCINE

GARI System Overview



Visual Analytics Uses for Risk-Based Decision Making

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



VACCINE

February 2013

USCG: Effective Risk-based Decision Making and Resource Allocation Visual Analytics

- Evaluate current and historical mission area:
  - Demands
  - Risks (total, mitigated, residual)
  - Resource allocation
  - Return on investment
- Evaluate courses of action
- Evaluate above at both Strategic and Tactical/Operational level

VACCINE

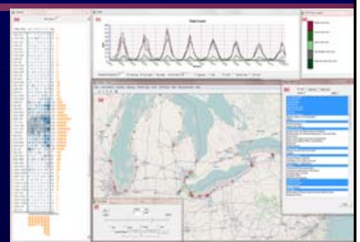
February 2013

U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7, USCG D9, USCG D5, USCG HQ 771

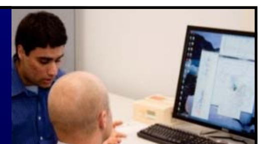
IMPACTS:

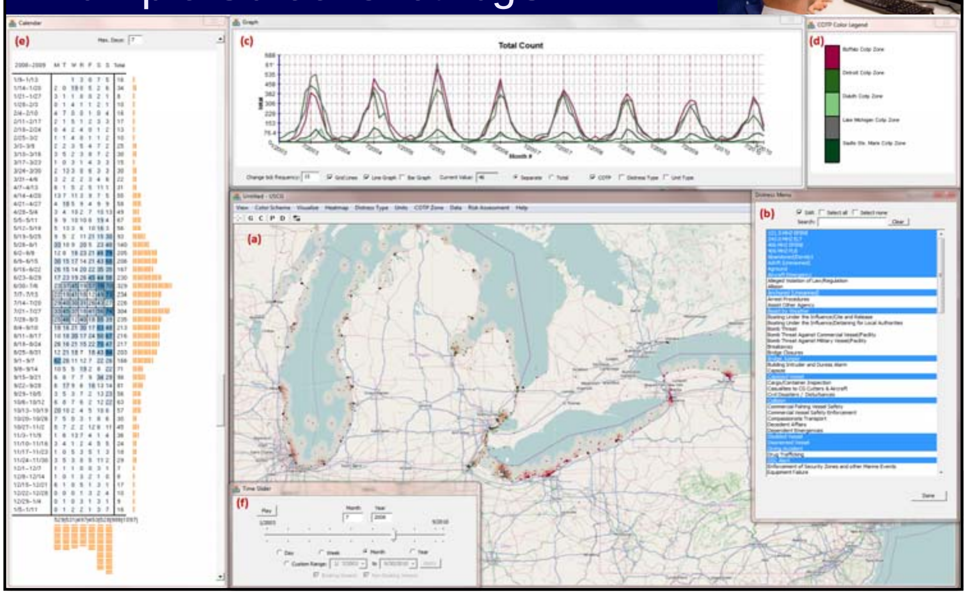
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Superstorm Sandy: Used for resource allocation in response and in rebuilding
- Used for Hurricane Irene resource allocation decisions
- Informed Commandant's budget testimony and recommendations to Congress
- Key component of USCG D9 reallocation plan for 2011-2012 based on decreased budget
- Key component of Coastal Operations Allocation Suite of Tools (COAST) – USCG HQ



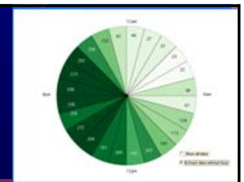
February 2013

Example Screenshot: cgSARVA





Risk-Based Allocations

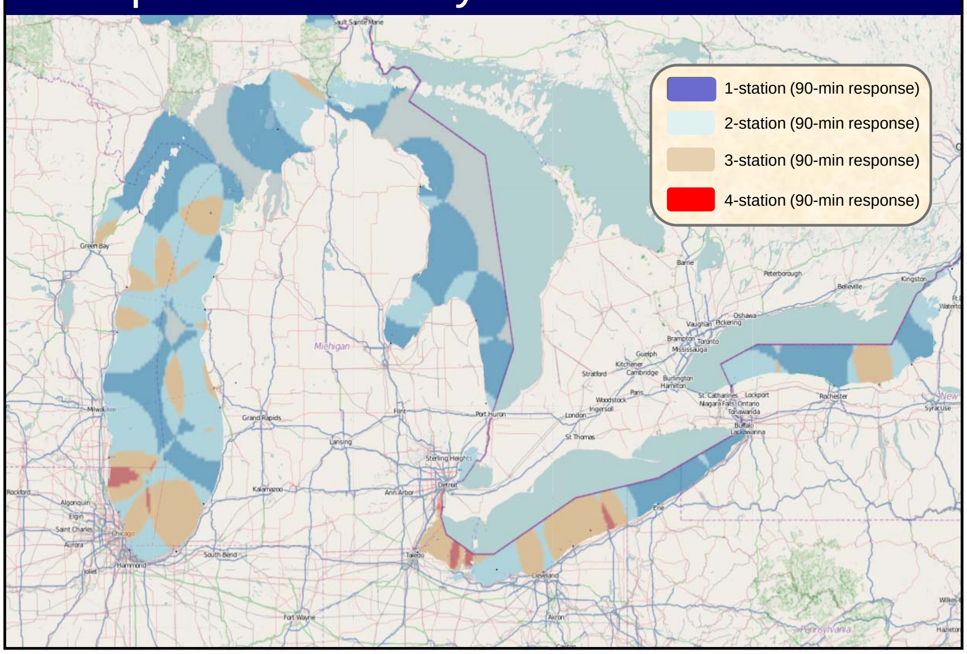


- Comparative visual analysis of mission cases/hours vs. staffing hours
- Comparative visualization of resources vs. risk
- Trend visual analytics
  - Increase/decrease in resource allocation
  - Increase/decrease in risk (total, mitigated, residual)
  - Increase/decrease in incidents
- Exploration of alternatives and effect on risk
- Predictive analytics based on historical data (STL and EWMA)

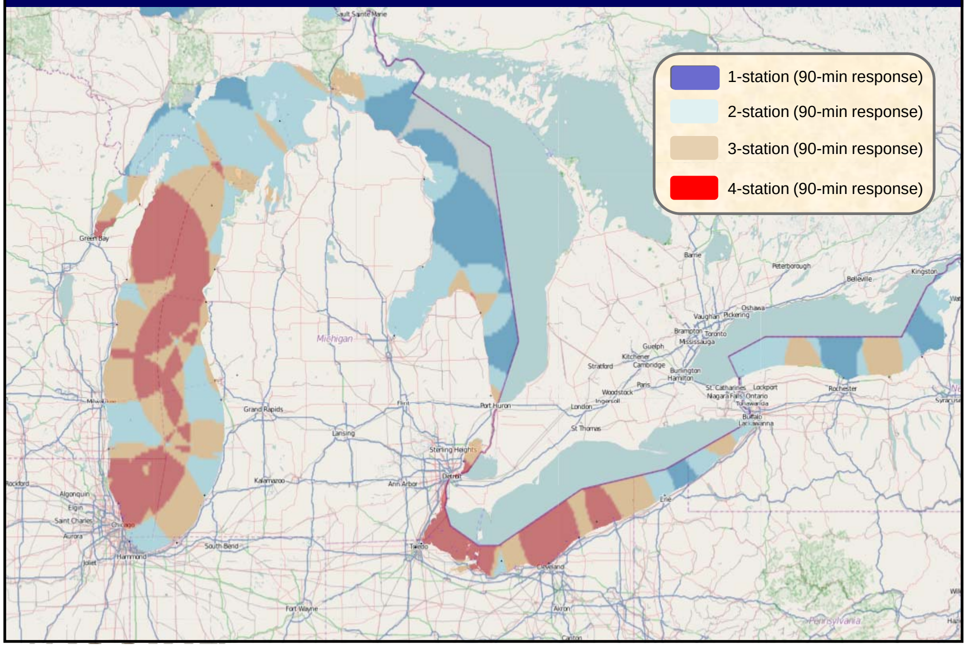
VACCINE

February 2013

Response Efficiency – Current Assets



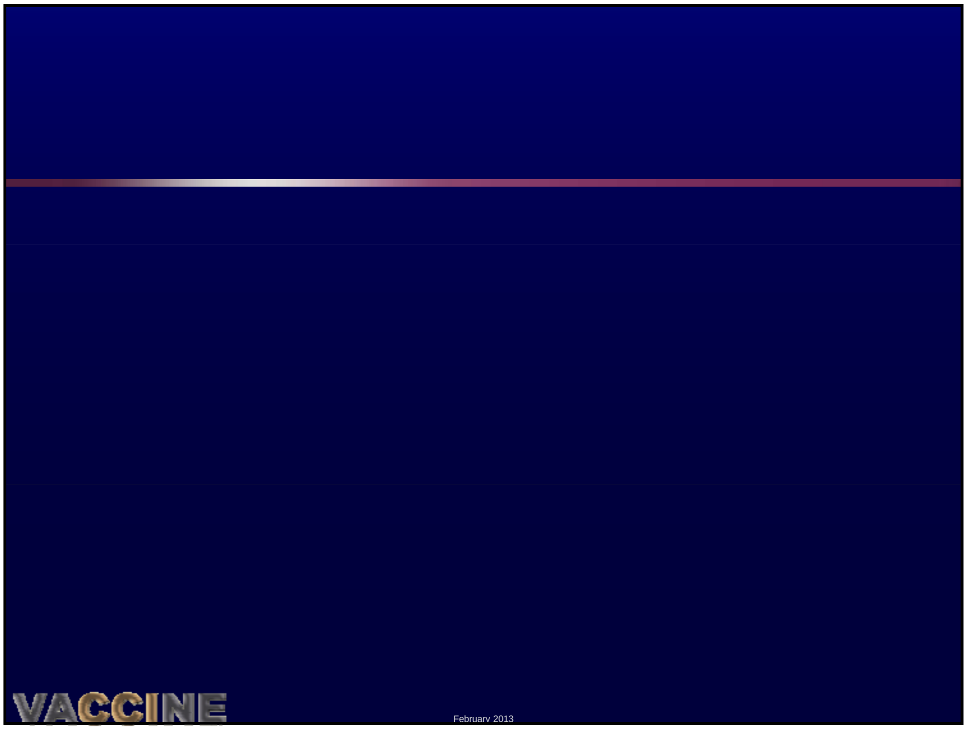
Response Efficiency – Potential Future Assets



For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu
ebertd@purdue.edu


VACCINE
February 2013

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

Table of Contents

Visual Analytics for Risk-Based Decision Making: Anytime and Anywhere

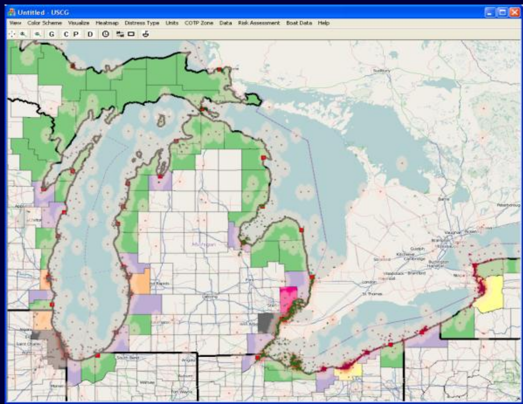
David S. Ebert
Purdue University
Abish Malik, Silvia Oliveros, Ahmad Razip, Sungahn Ko, Yang Yang

CVADA
Center for Visual Analytics and Data Analytics
A U.S. Department of Homeland Security Center of Excellence

November 2012

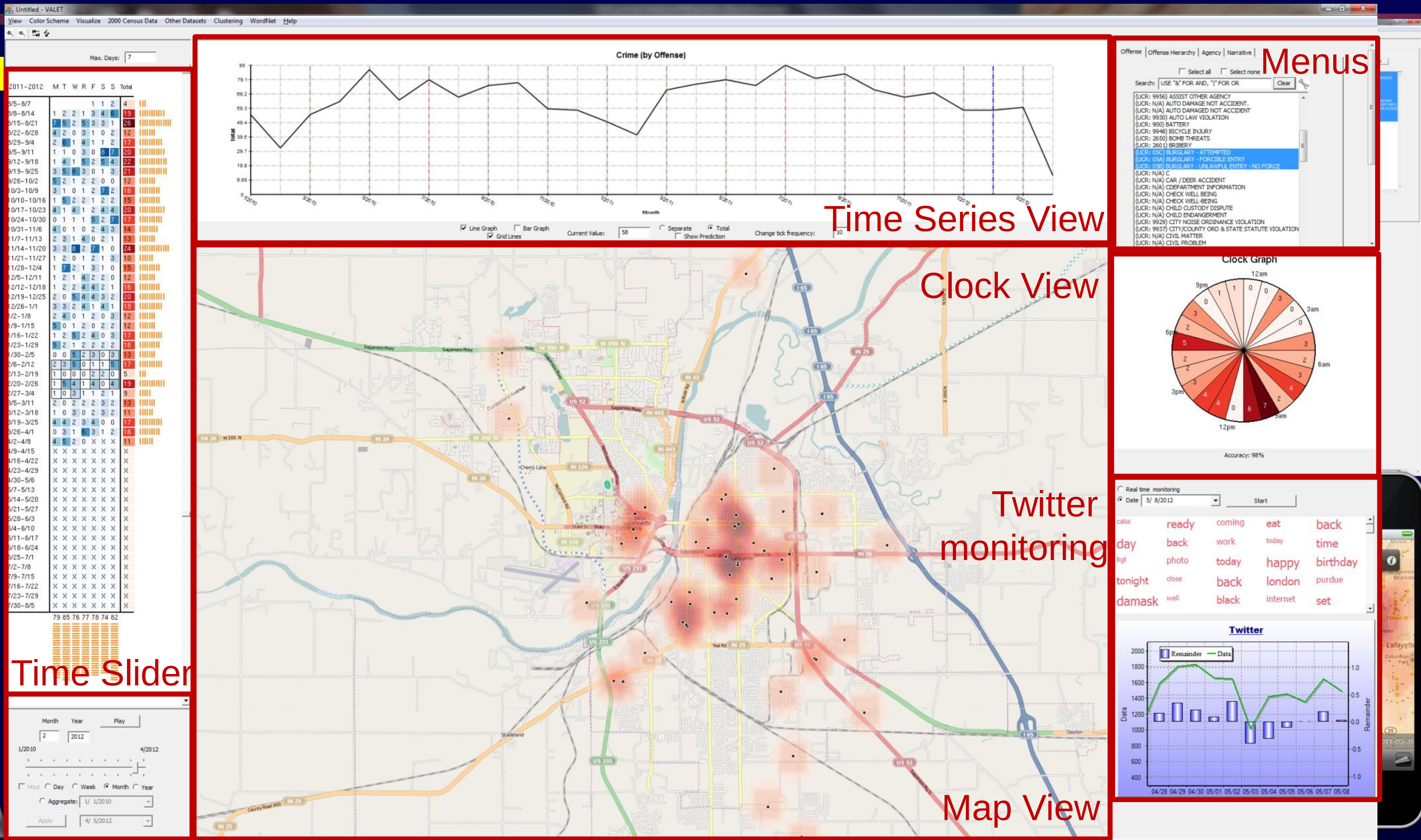
Visual Analytics Uses for Risk-Based Decision Making

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



VACCINE

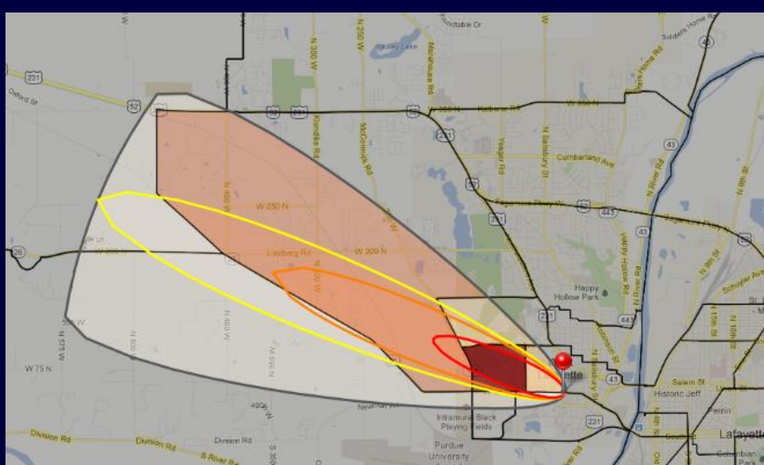
Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

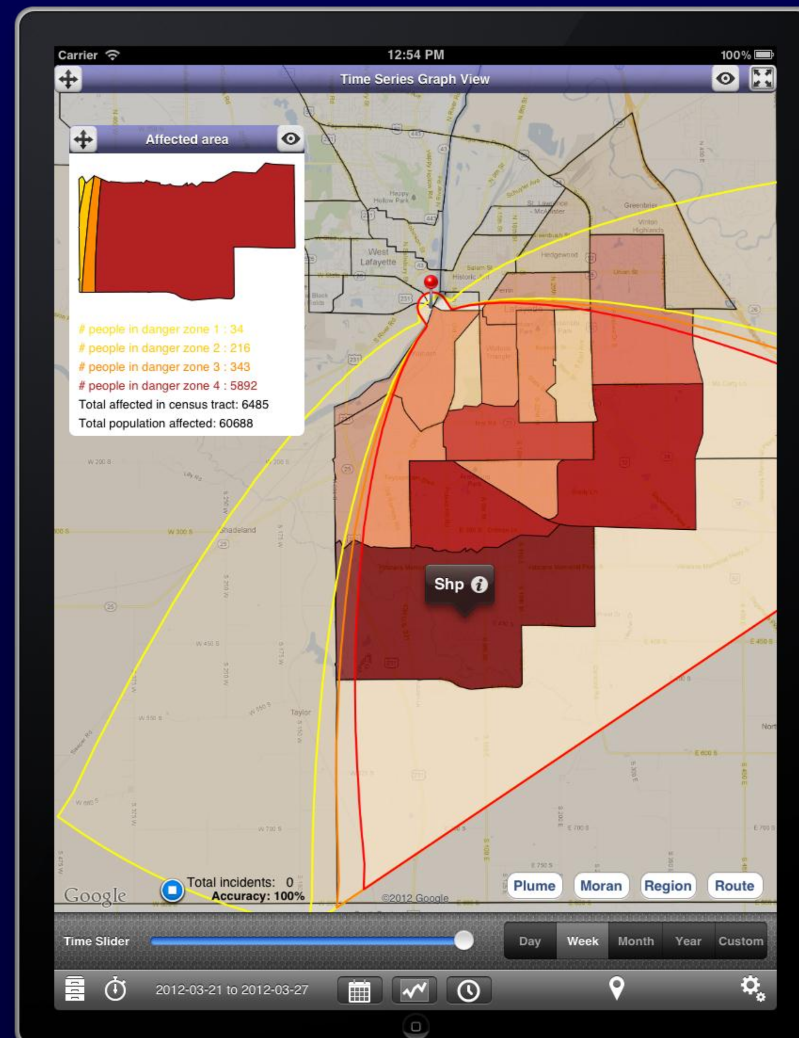



Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
11/14/2010
01/01/2011
02/13/2011
03/27/2011
Calendar View
M T W Th F Sa S
1/1/10 36 46 29
1/4/10 38 48 44 43 42 32 30
1/11/10 62 53 53 56 59 45 34
1/18/10 62 49 53 72 61 52 49
1/25/10 98 66 57 66 64 55 49
2/1/10 58 63 50 51 73 51 42
2/8/10 50 46 48 63 52 39 35
2/15/10 48 65 48 54 55 67 51
2/22/10 51 66 48 56 57 46 47
3/1/10 50 59 42 54 82 45 41
3/8/10 50 57 64 78 63 36 30
Clock View
12Hrs
10:00
Accuracy: 64%
Choro
Plume
Moran
Region
Route
Lafayette
©2012 Google
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20

MERGE – iVALET Interactive Plume Visualization and Evacuation Planning

- Chemical release plume modeling identifies census tracts with the highest number of expected people affected





VACCINE

Summary

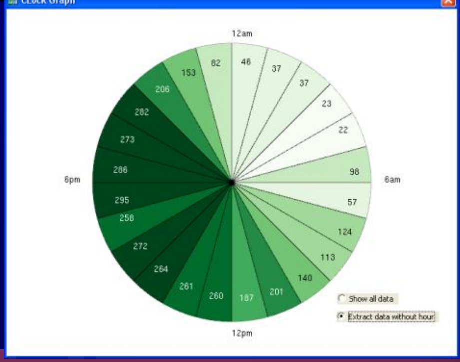
- Advanced analytics running at real-time rates
- Real-time date feeds and analysis
- Scenario and course of action planning
- Integration of multiple mobile tools for decision making

- Now let’s look at maritime applications

USCG: Effective Risk-based Decision Making and Resource Allocation Visual Analytics

- Evaluate current and historical mission area:
  - Demands
  - Risks (total, mitigated, residual)
  - Resource allocation
  - Return on investment
- Evaluate courses of action
- Evaluate above at both Strategic and Tactical/Operational level

Risk-Based Allocations

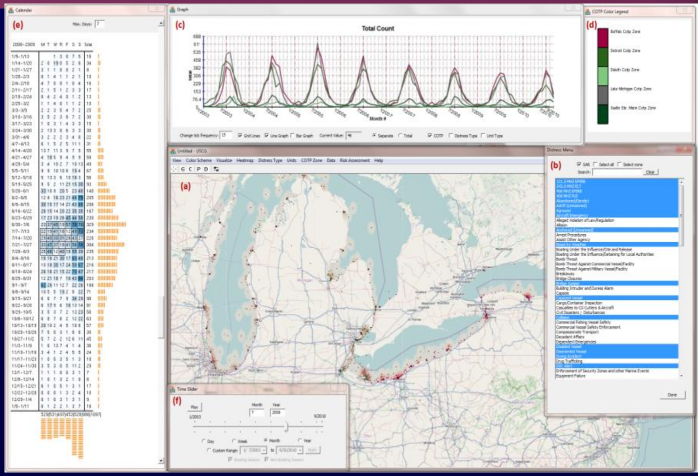


- Comparative visual analysis of mission cases/hours vs. staffing hours
- Comparative visualization of resources vs. risk
- Trend visual analytics
  - Increase/decrease in resource allocation
  - Increase/decrease in risk (total, mitigated, residual)
  - Increase/decrease in incidents
- Exploration of alternatives and effect on risk
- Predictive analytics based on historical data (STL and EWMA)

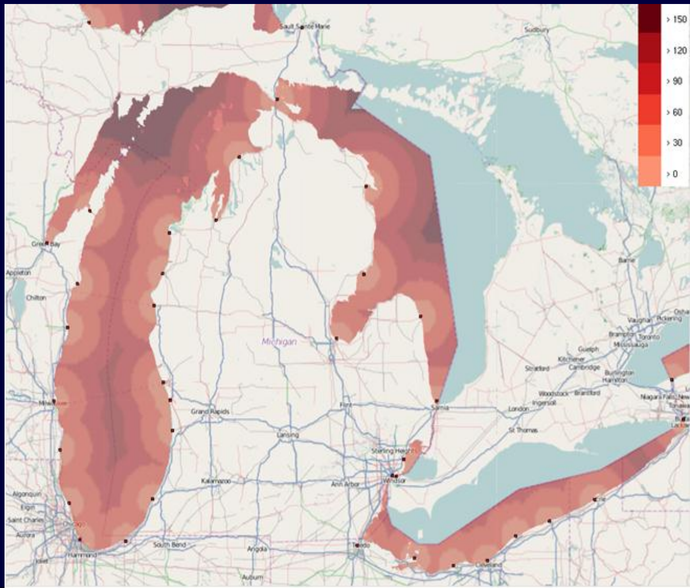
U.S. Coast Guard Missions VA (cgSARVA/COAST)

Partners: USCG LANT 7, USCG D9, USCG D5, USCG HQ 771

- Resource allocation and realignment analysis
- Risk-based analysis
- Economic analysis
- Examples:



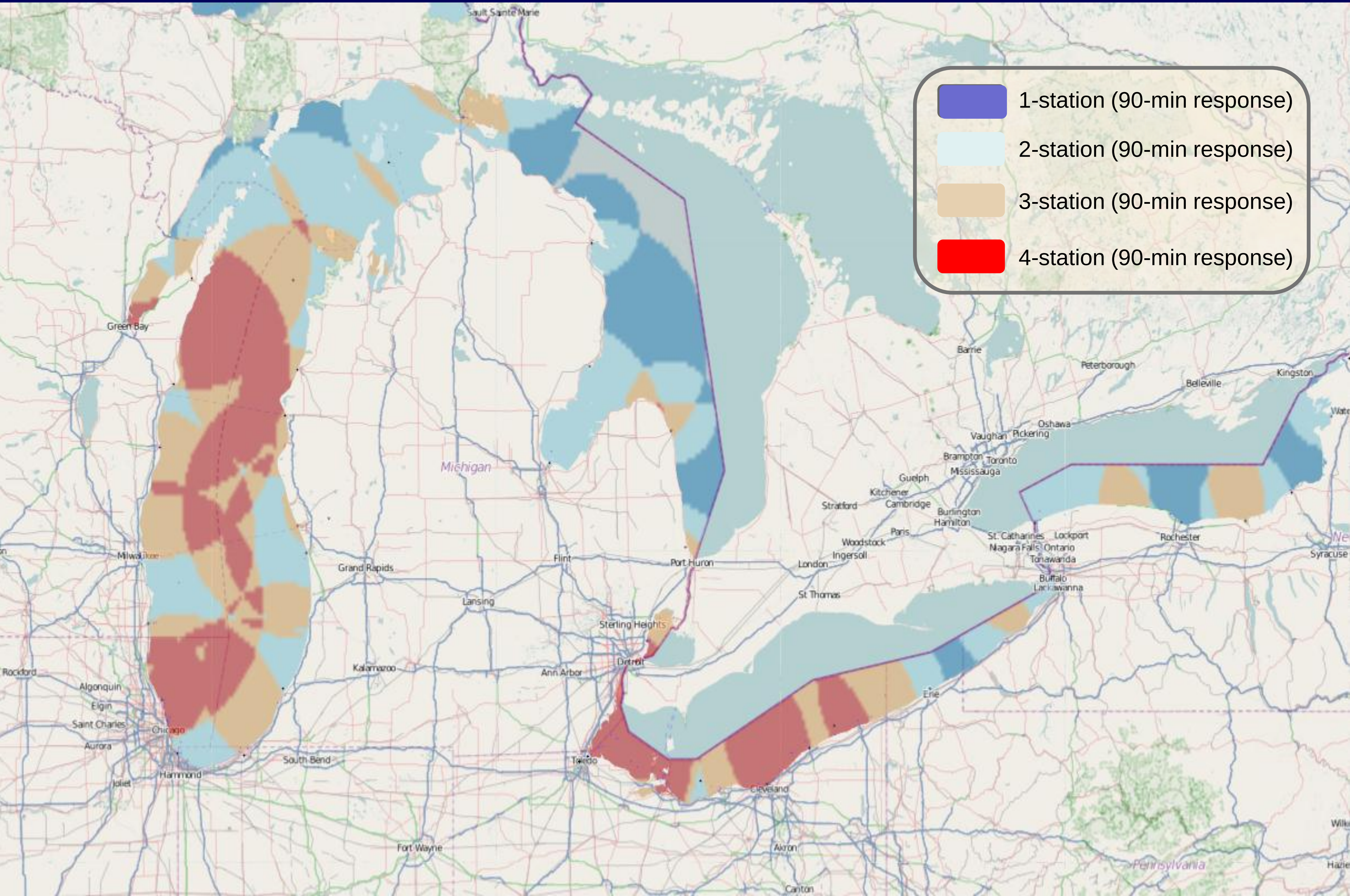
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Hurricane Irene resource allocation decision based on cgSARVA analysis and visualization



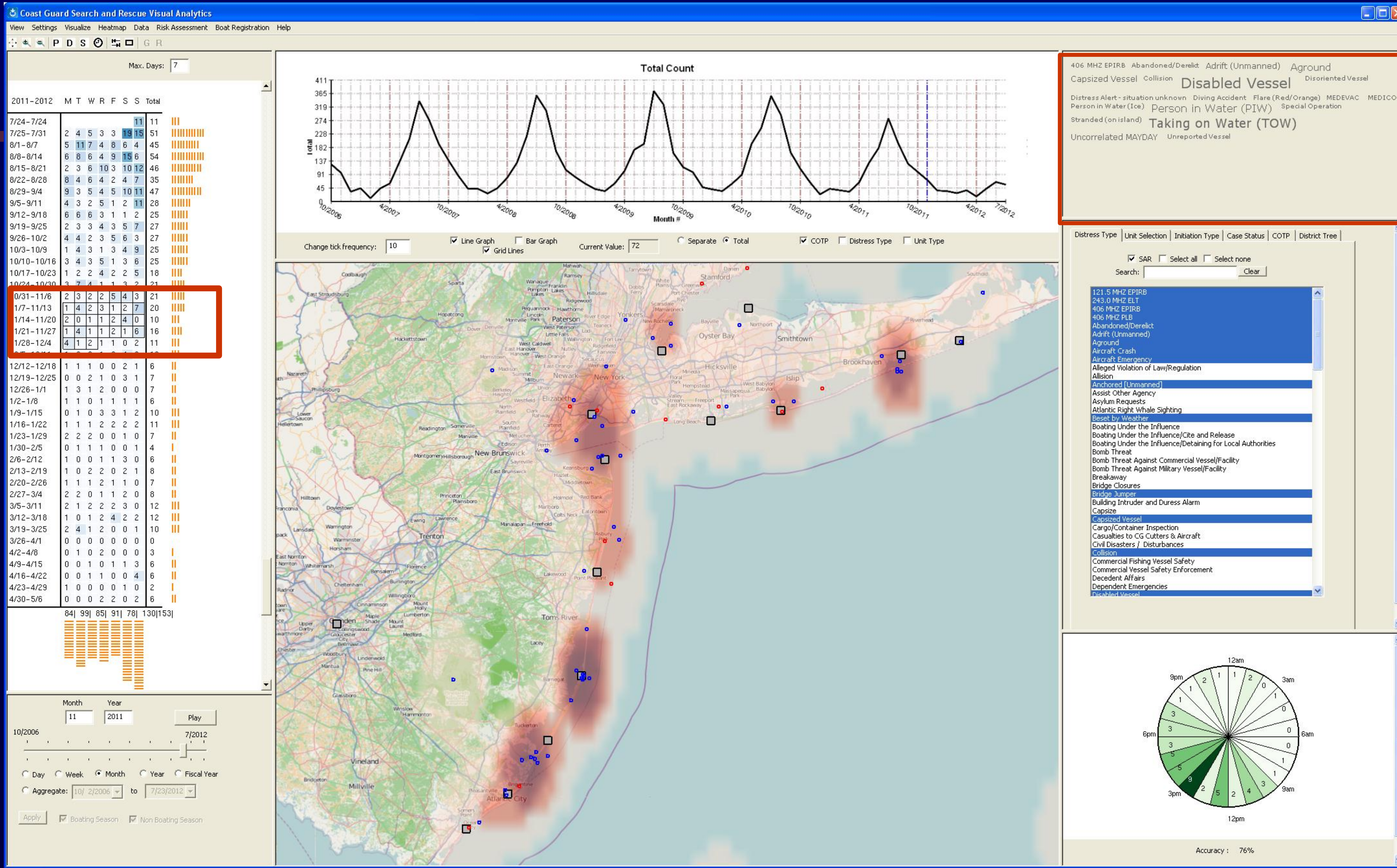
Mission-based Risk Analysis Capabilities

- Efficiencies in CG station coverage
  - Resource allocation for SAR
  - Temporal and spatial patterns used in planning
  - Hypothetical station closure analysis
  - Operational risk analysis and assessment (total risk, mitigated risk, residual risk)
  - Effects of different mitigation strategies
  - Balancing costs of different decisions

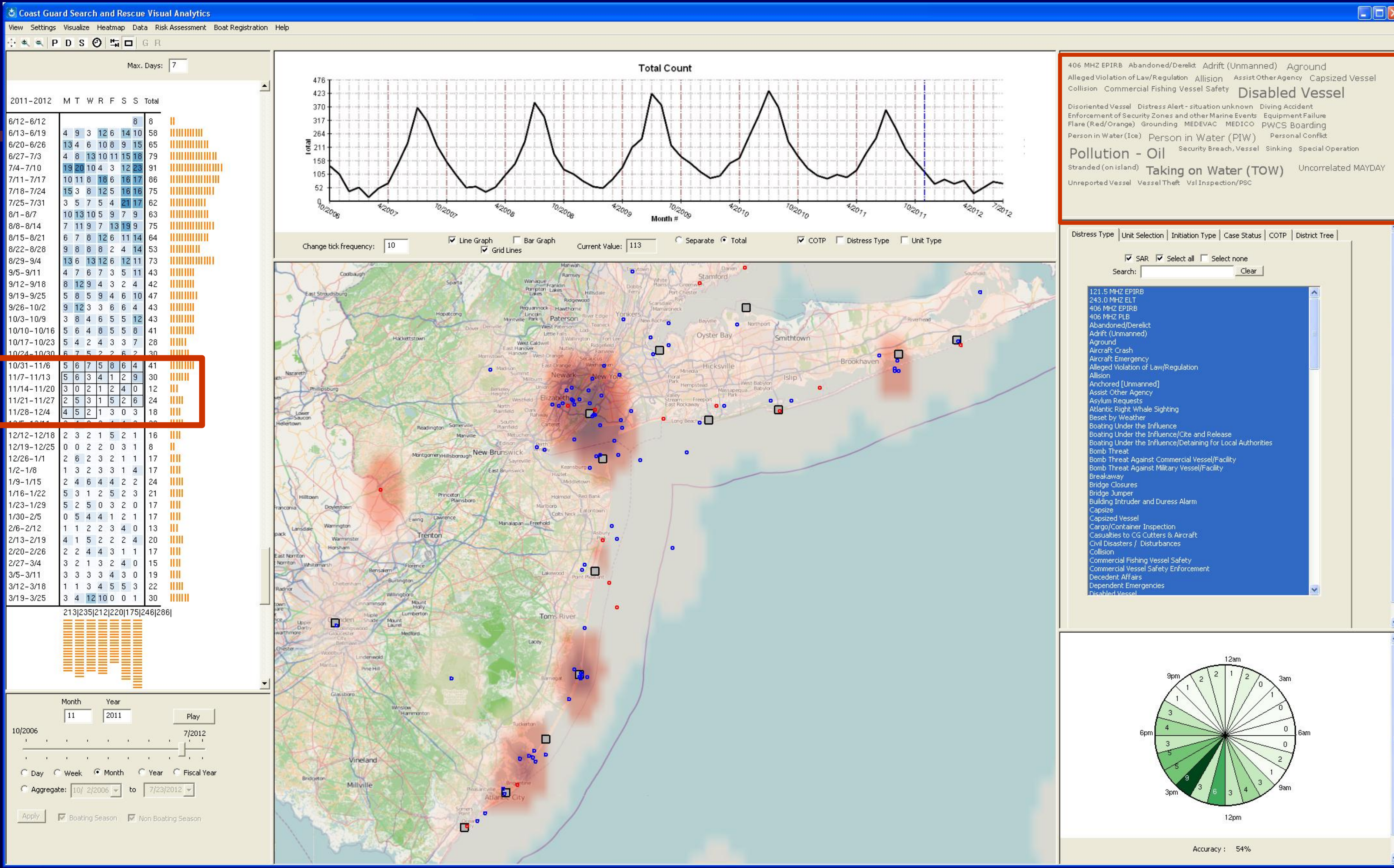
Response Efficiency: Hypothetical Asset Allocation



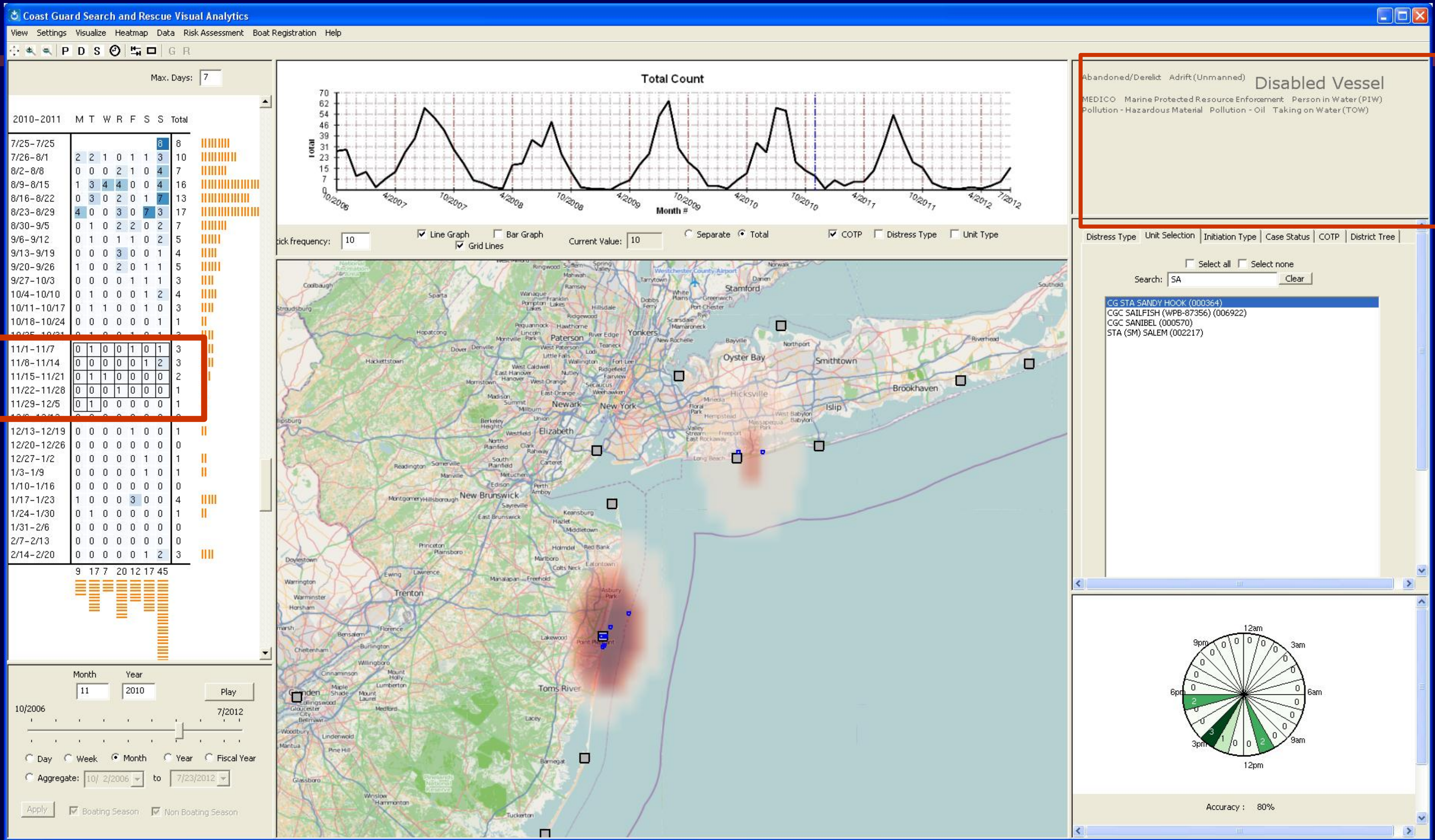
Example: Risks and Consequences From Sandy:
SAR Cases November 2011 NJ/NYC Area



All Cases November 2011

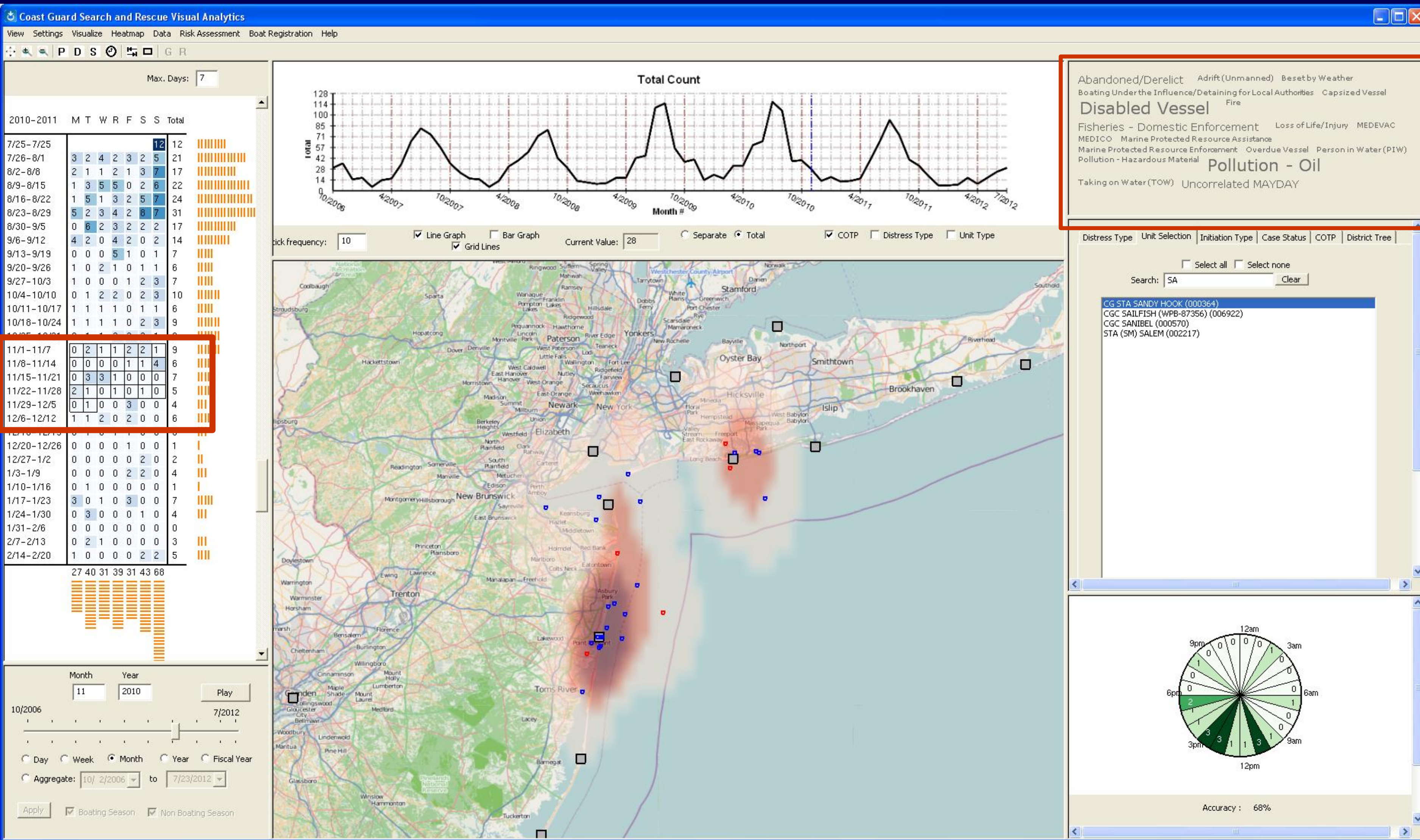


Jones Beach, Sandy Hook, and Manasquan Inlet cases recorded as responded by these stations .
All Cases November 2010

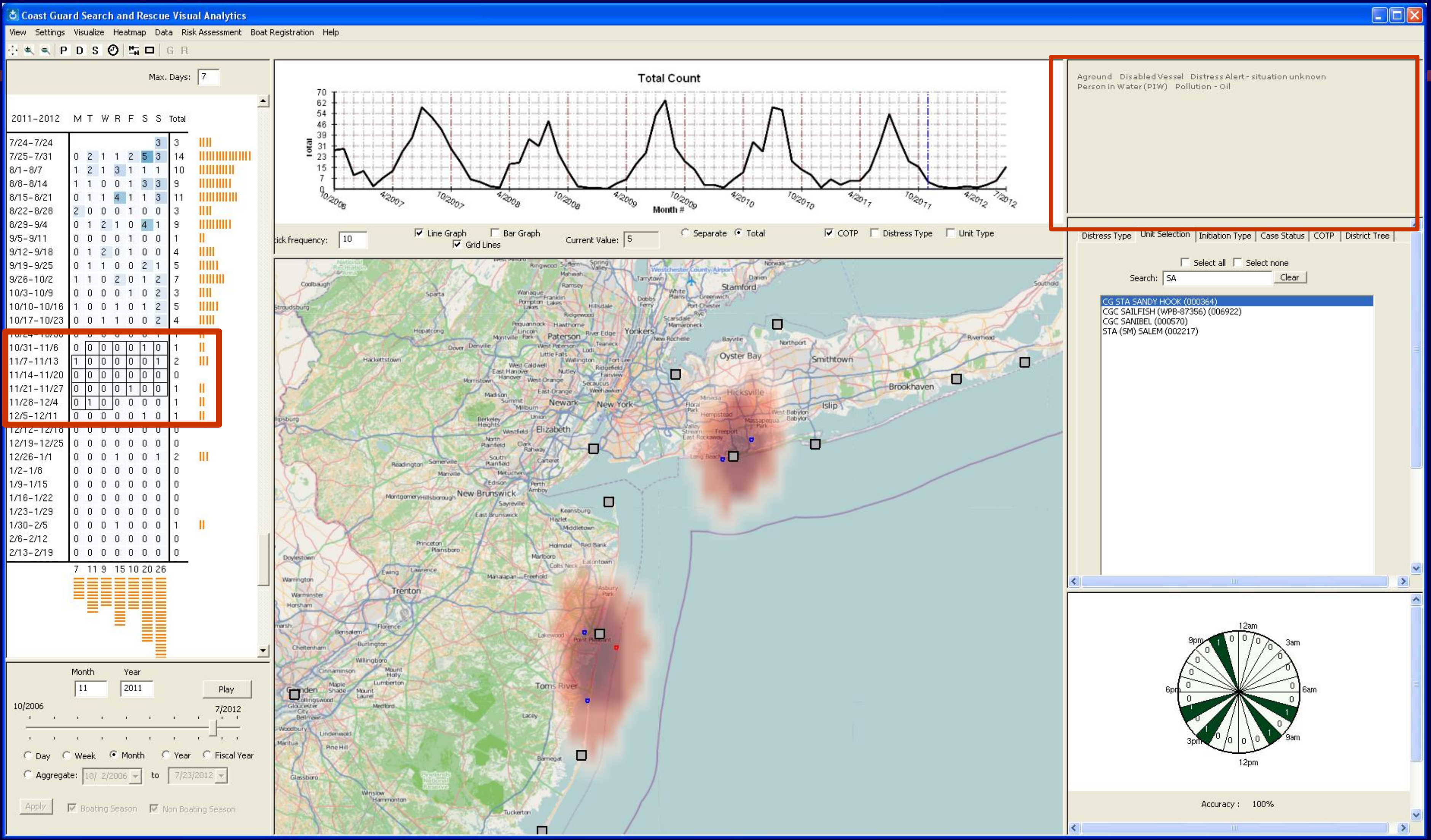


VACCINE

Jones Beach, Sandy Hook and Manasquan Inlet closest cases (not necessarily reported under them).
All Cases November 2010

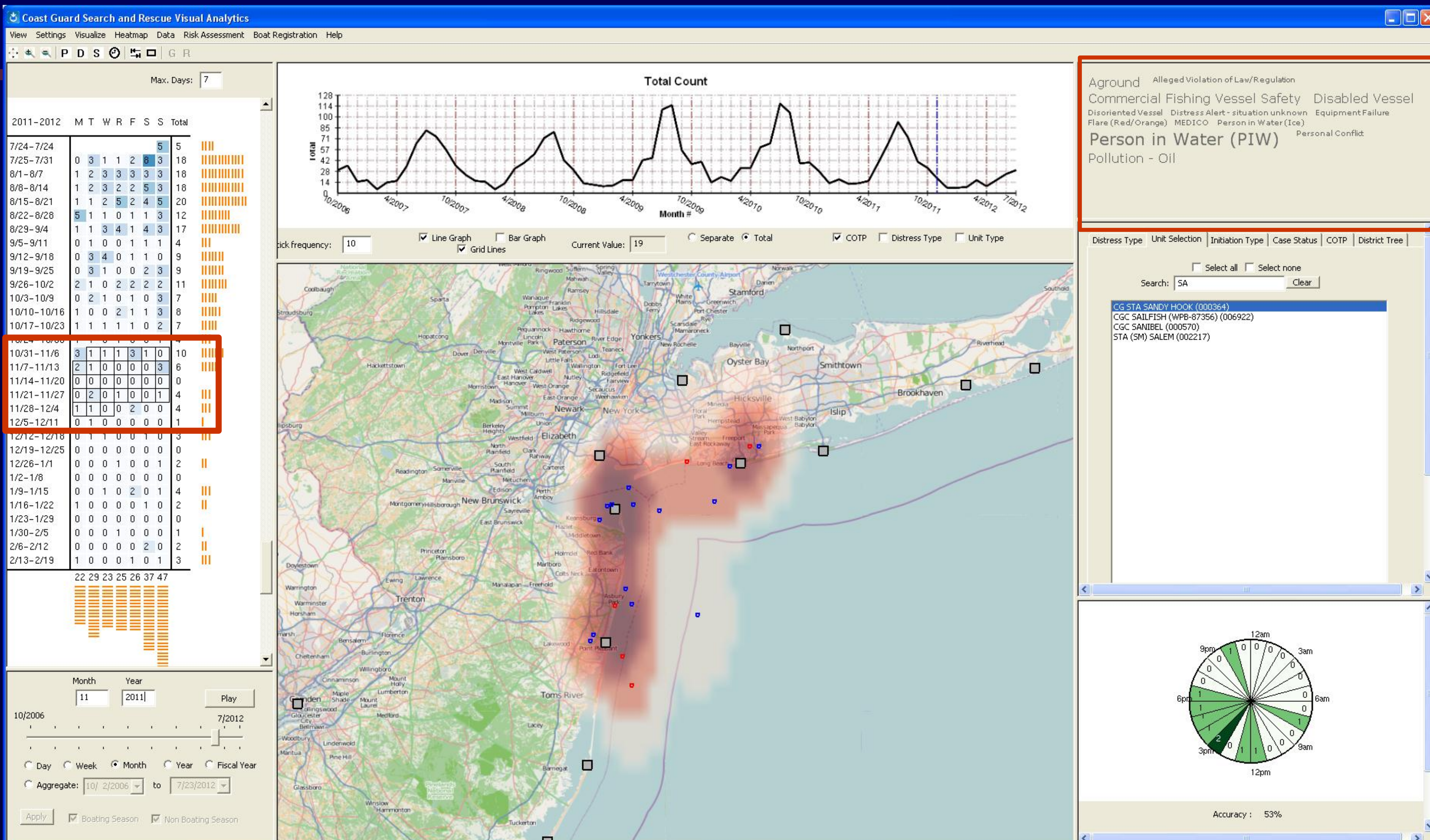


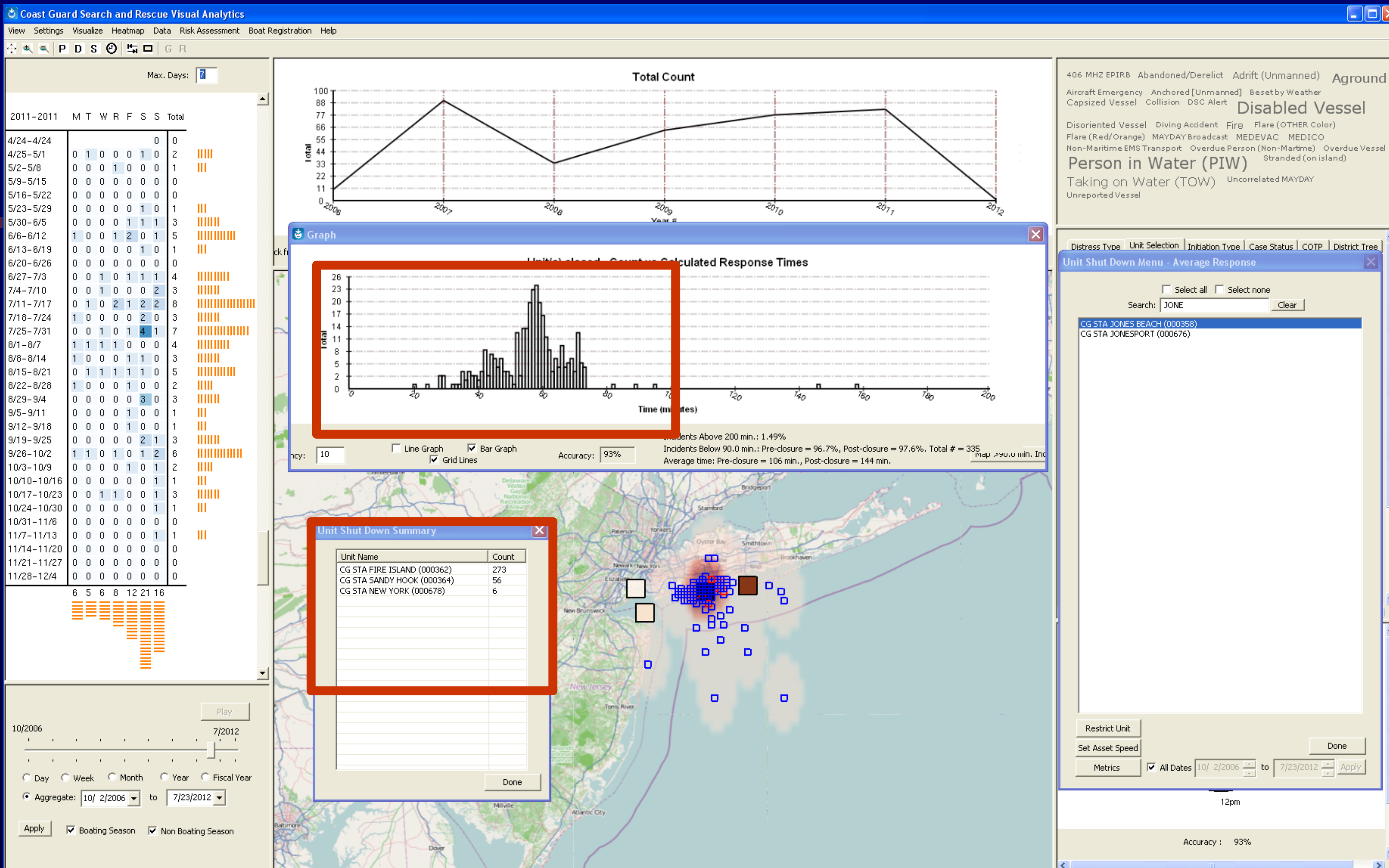
Jones Beach, Sandy Hook, and Manasquan Inlet cases recorded as responded by these stations .
All Cases November 2011



VACCINE

Jones Beach, Sandy Hook and Manasquan Inlet closest cases (not necessarily reported under them).
All Cases November 2011

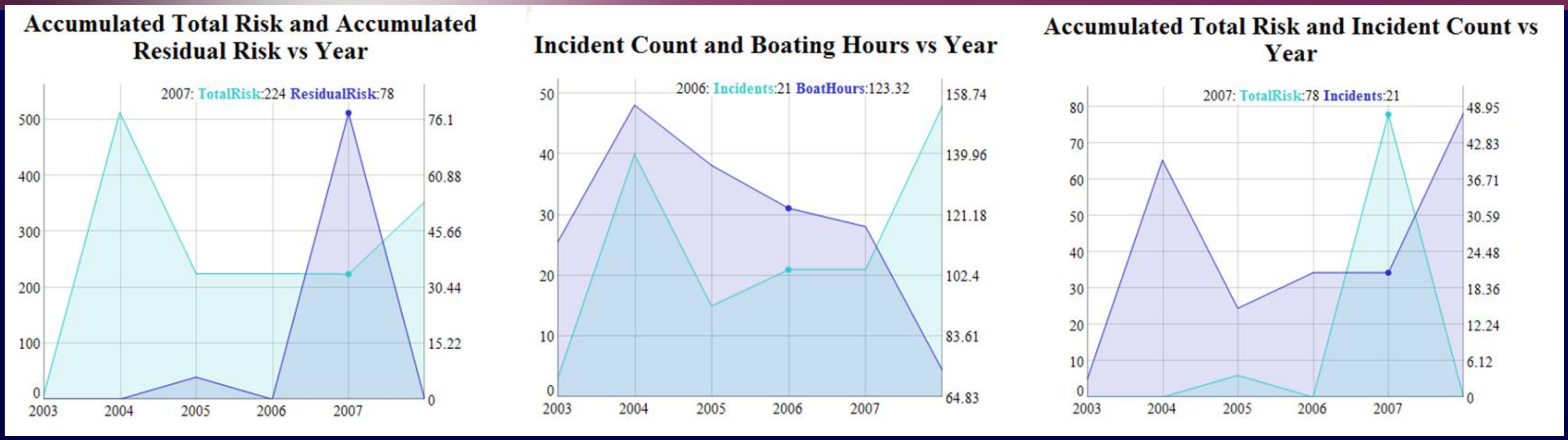




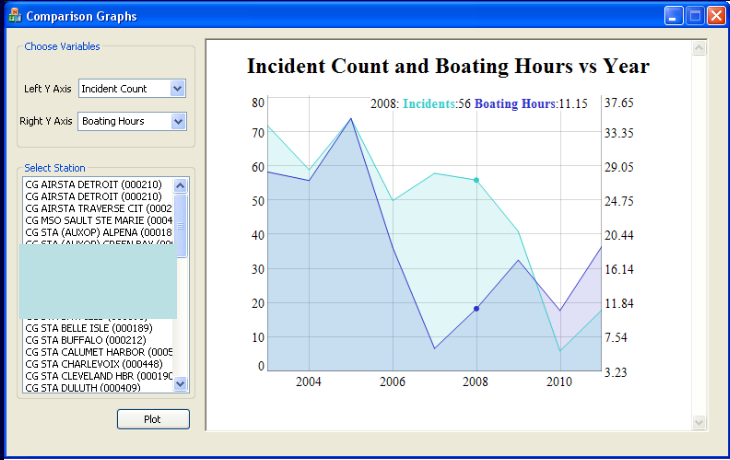
Closure for Jones Beach (with all the regular operations for the past 5 years taken into account).

VACCINE

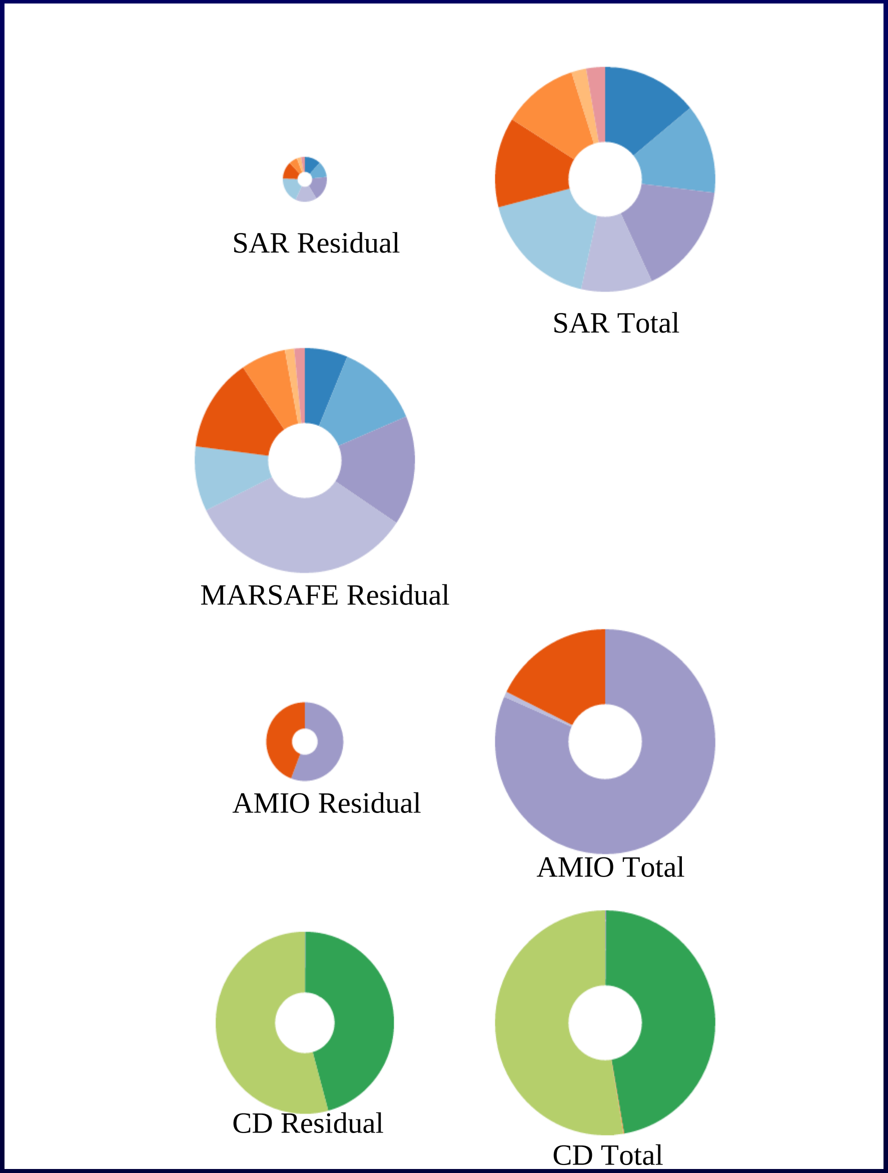
Resource Allocation and Risk-Based Decision Making (e.g., ORAM)

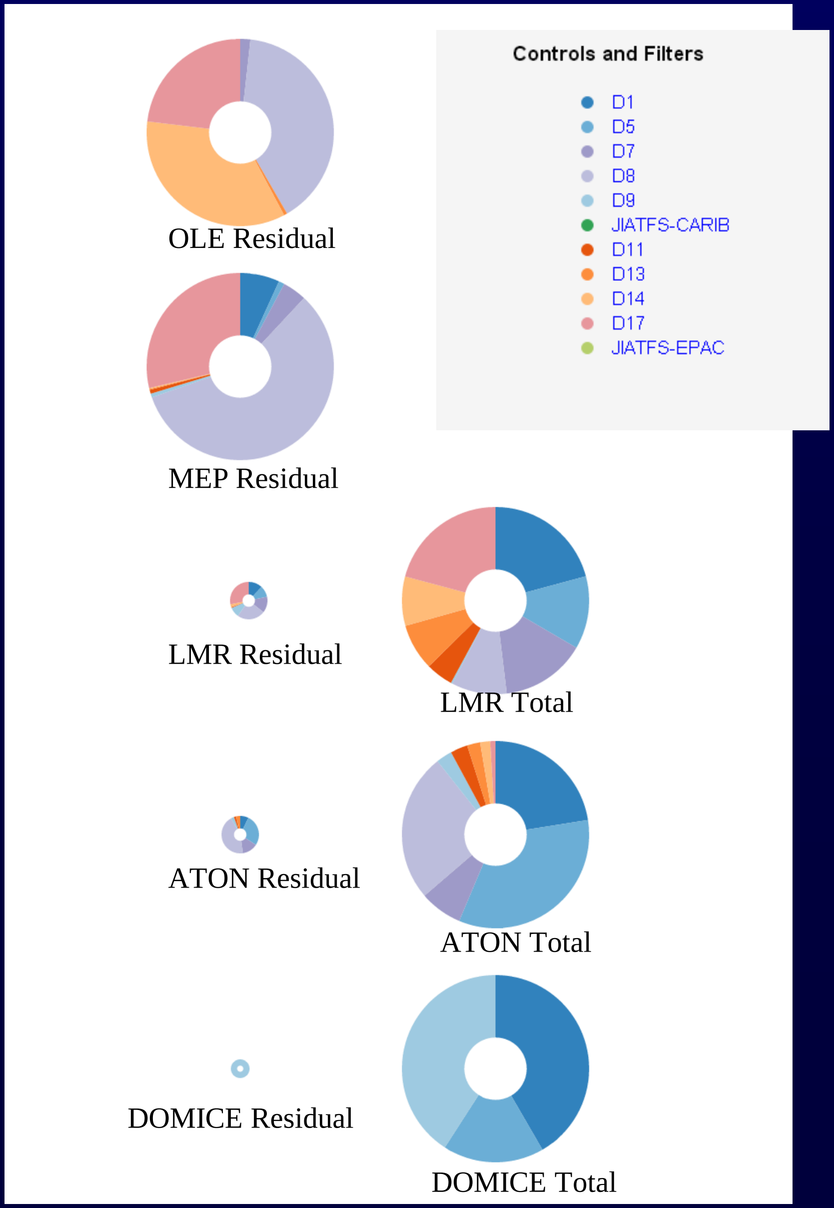


- Explore risk-based decision making and utilize historical data for analysis and prediction
  - Total Risk, Mitigated Risk, Residual Risk
  - Explore 11 different USCG missions
  - Explore allocation of assets w/ different capabilities
  - Explore staffing, utilization, assets vs. risk
  - Perform What-If scenarios

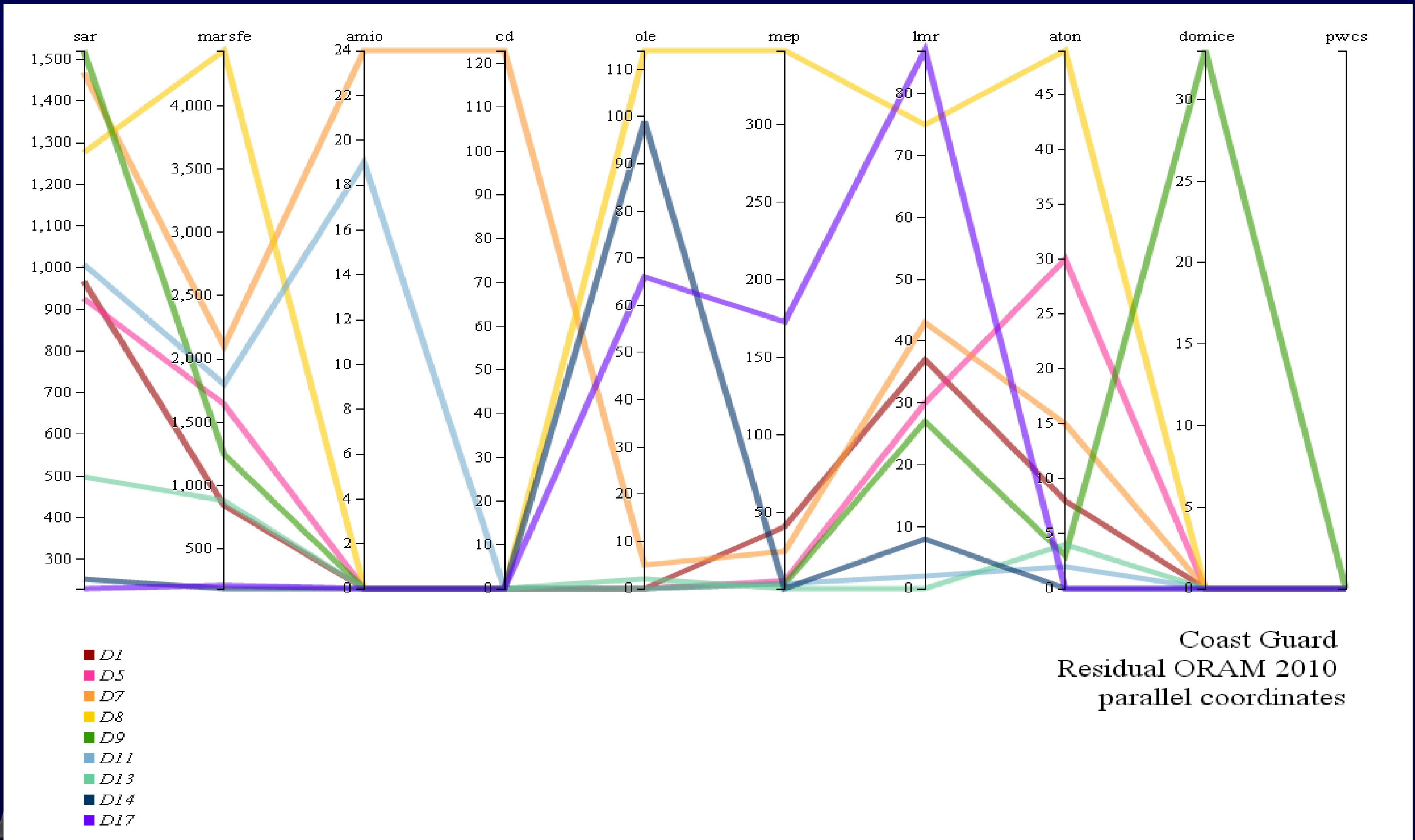


ORAM Risk Pie Graph by Mission Type

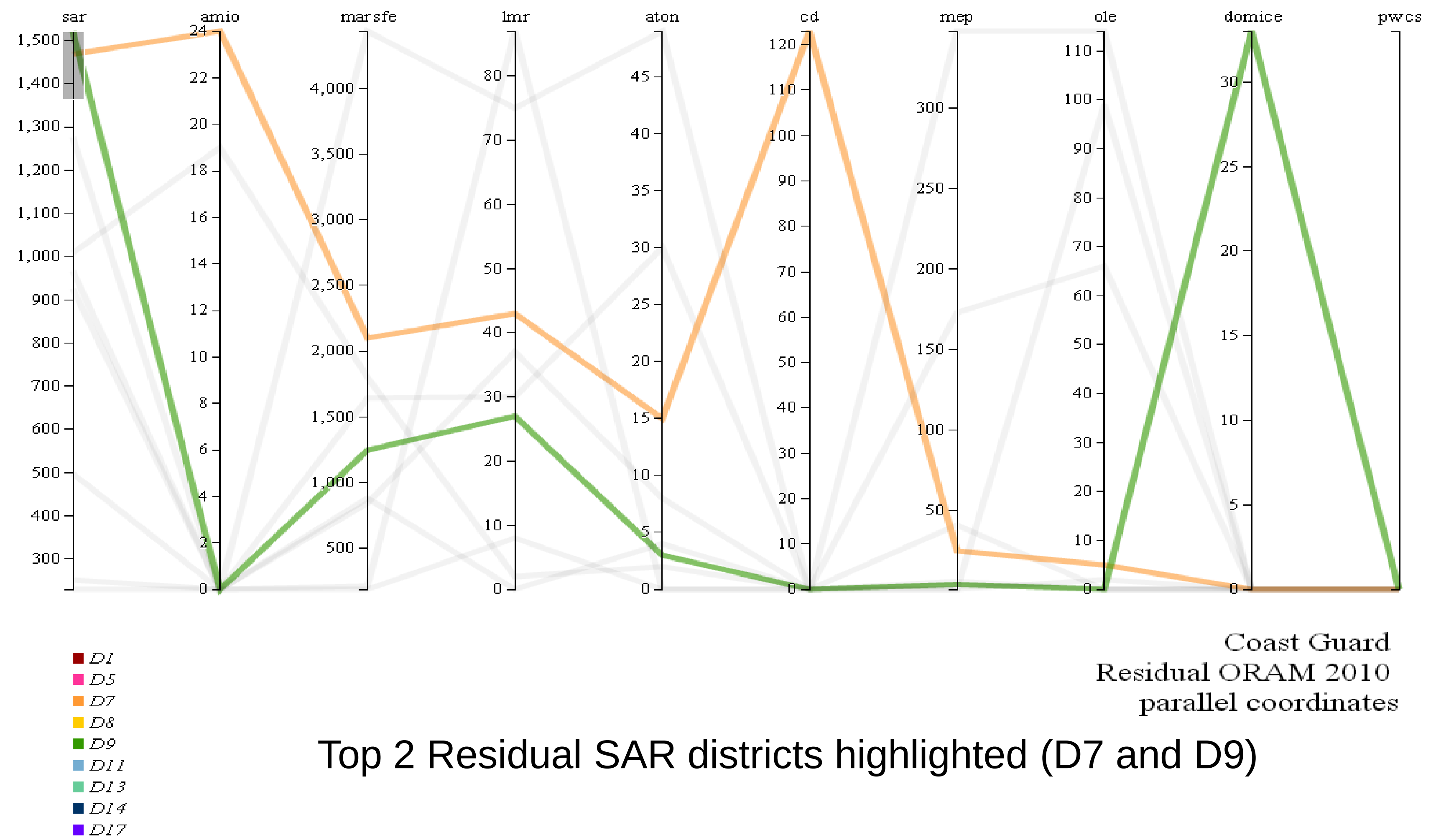




ORAM Parallel Coordinate Analysis



ORAM Top 2 Districts based on Residual SAR Risk



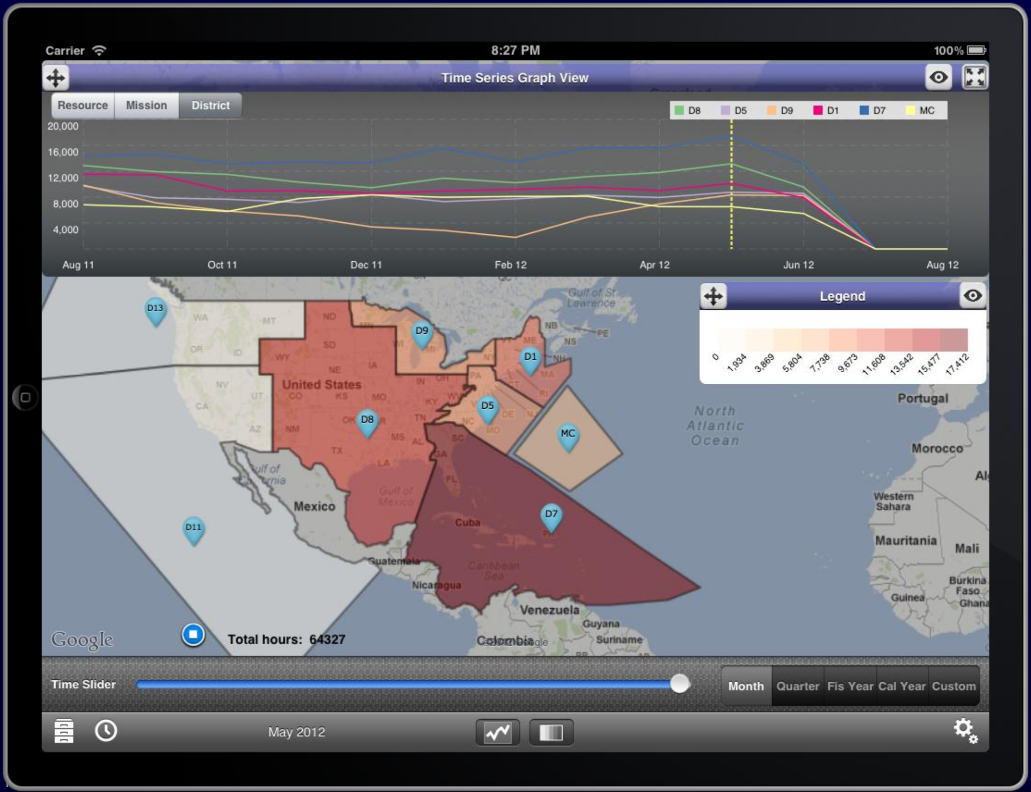
Top 2 Residual SAR districts highlighted (D7 and D9)

Interactive Operational Performance Assessment Report (iOPAR)

- OPAR
  - Standard report within the Coast Guard Business Intelligence (CGBI) system
  - Displays resource use and performance by core CG mission areas
- iOPAR – interactive iPad version to allow interactive visualization and analysis and inform decision making
  - Anaylze/visualize performance, targets, seasonal trends, predictions by boat, aircraft, cutter type, region levels

iOPAR Background

- Provide an interactive, mobile display of OPAR information
- Enable real time or near real time assessment of hourly expenditure and performance
- For senior level decision makers to have access to OPAR information while away from a workstation
- Assist operational decision making



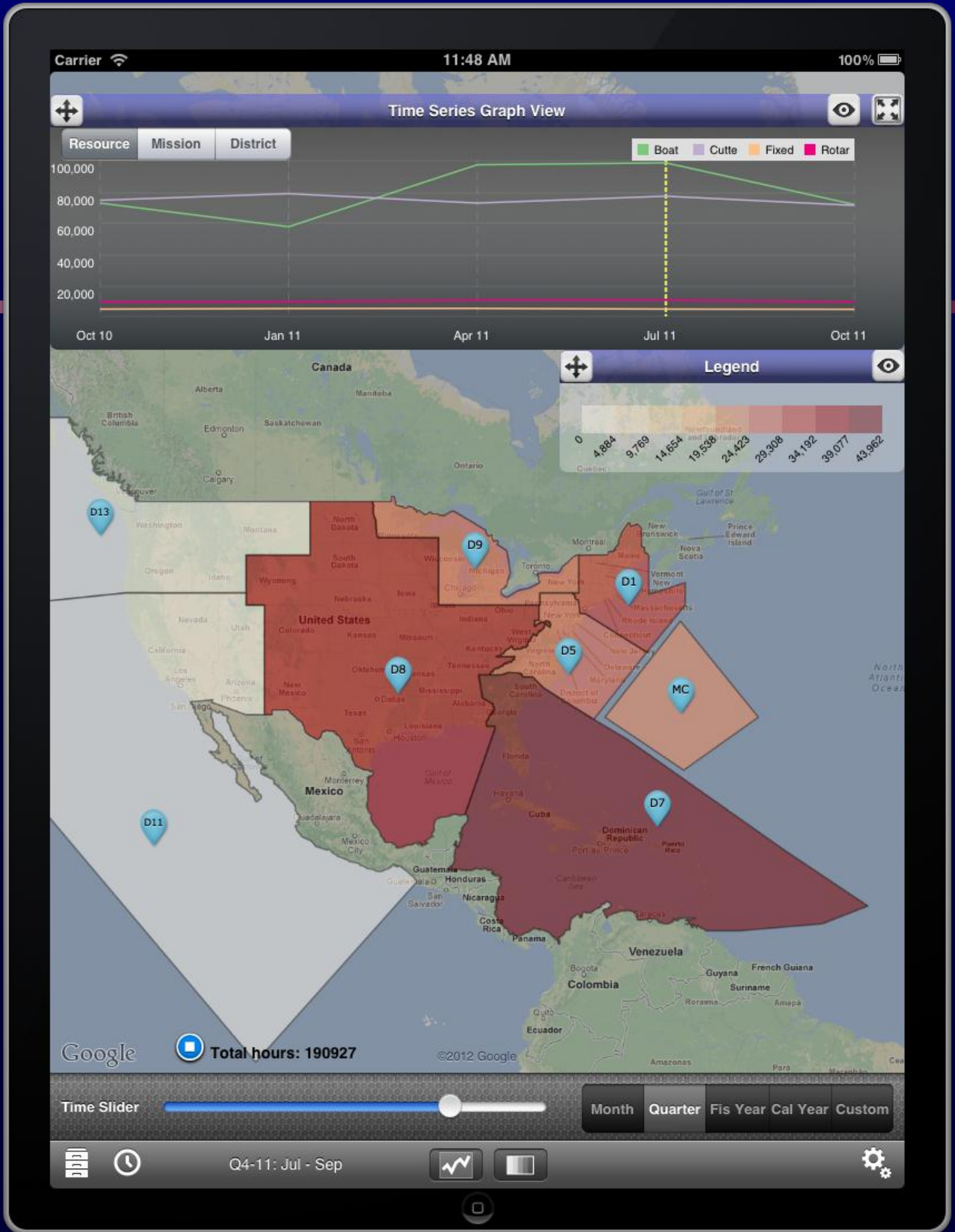
VACCINE

iOPAR goals

- Provide interactive filtering tools to filter data by:
  - Resources and subclasses (e.g., boats, aircrafts, cutters, types of boats)
  - Region (e.g., area, district, sector)
  - Mission area (e.g., SAR, MEP)
- Display hourly expenditure across missions, regions, resources
  - Drill down using interactions
- Measure performance
- Analysis capabilities
  - Seasonal trends
  - Predictive utilization rates

iOPAR

- iOPAR map & time series view
- Showing mission hour distribution over districts for Q4-2011
- One look tells which district is driving the hours for the selected time period
- Time series show hours of resources usage across time
- (All mission, resources and districts selected)



VACCINE

iOPAR Grouped Time Series

- Group time periods (months, quarters) and compare across years.
- Here, looking at how each quarter uses the resources across years

VACCINE

Table of Contents

For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu

ebertd@purdue.edu

VACCINE

Table of Contents

VACCINE

Visual Analytics for Command, Control, and Interoperability Environments

A U.S. Department of Homeland Security Center of Excellence

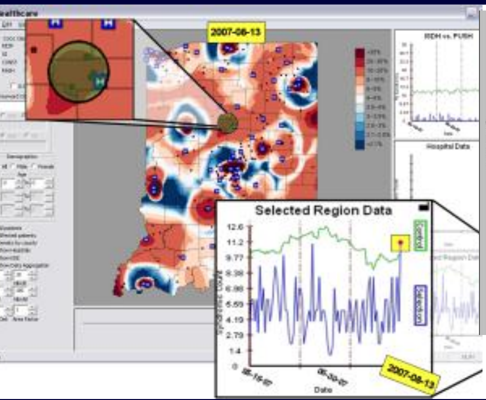
Visual Analytics in Public Health

Dr. David S. Ebert
Purdue University, VACCINE
ebertd@purdue.edu

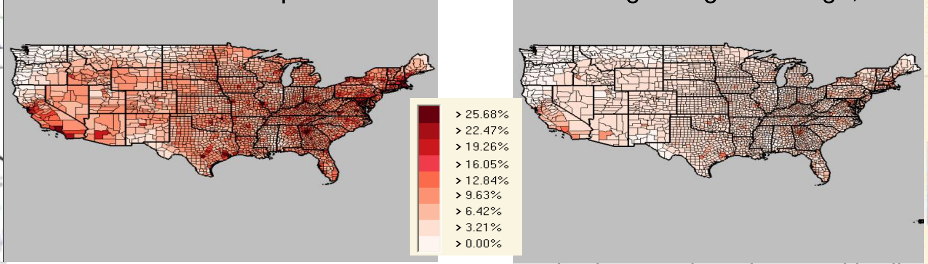
Health Surveillance

Pandemic Simulation and Planning

Zoonoses Planning and Response

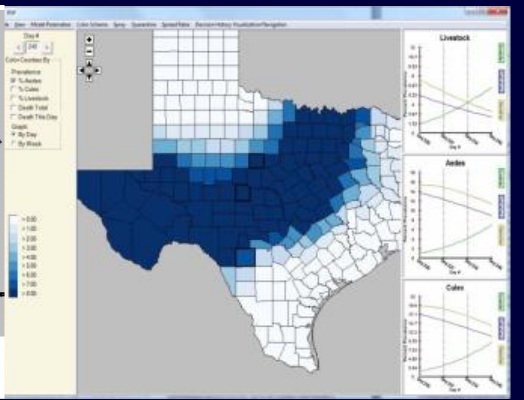


Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.



Pandemic spread on day 37 with no decision measures implemented

Pandemic spread on day 37 with all decision measures implemented



Background on Our Work

- Purdue leads the Visual Analytics for Command, Control, and Interoperability Environments (VACCINE) Center of Excellence in Visualization Sciences for the U.S. Department of Homeland Security
- 26 schools across the globe
- Research, education, and engagement to solve real world problems

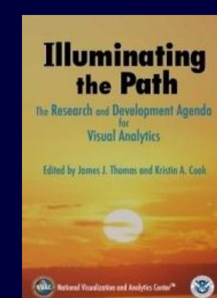
Definition

Visual Analytics[1] is the science of analytical reasoning facilitated by interactive visual interfaces

People use visual analytics tools and techniques to

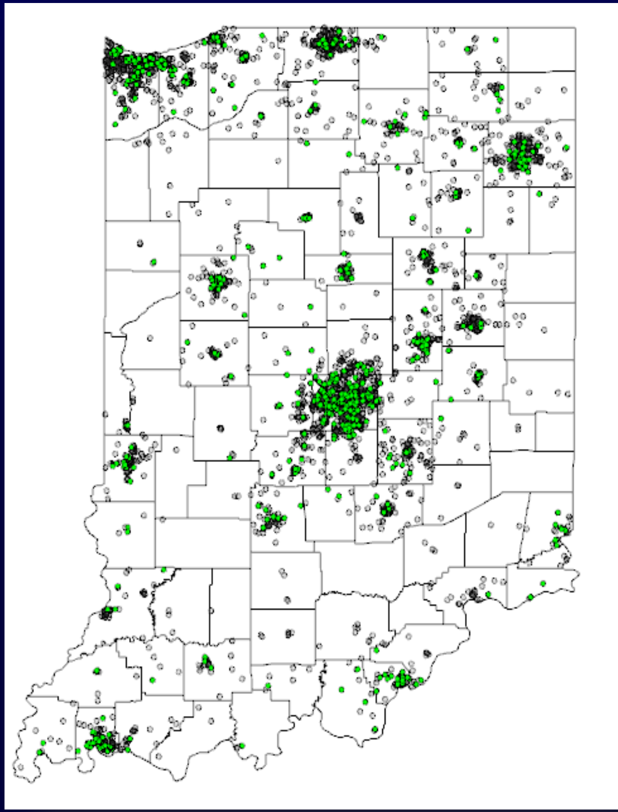
- *Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data*
- *Detect the expected and discover the unexpected*
- *Provide timely, defensible, understandable assessments*
- *Communicate assessment effectively for action*

1. *Illuminating the Path: The R&D Agenda for Visual Analytics, Editors: Thomas and Cook*



Spatiotemporal Proactive and Predictive Visual Analytics

- Most activities involve detecting patterns, clusters, and anomalies in space, time, people, and events
- Existing analytical methods remain poor for identifying real-world place-time features as well as for understanding the space-time dynamics
- The methods are also infeasible computationally for most real-time scenarios



Syndromic Surveillance data: Emergency Department complaints reported to the Indiana Public Health Emergency Surveillance System.

VACCINE

Syndromic Surveillance

- Syndromic surveillance is the detection of adverse health events focusing on pre-diagnosis information to improve response time
- Disease surveillance enables public health officials and their partners to achieve such purposes by monitoring changes in the population's health status.
- Pre-diagnosis information can consist of multiple data sources:
  - *Over the counter medicine sales*
  - *News reports on emerging diseases*
  - *Pro-med news feeds*
  - *Emergency department chief complaints*

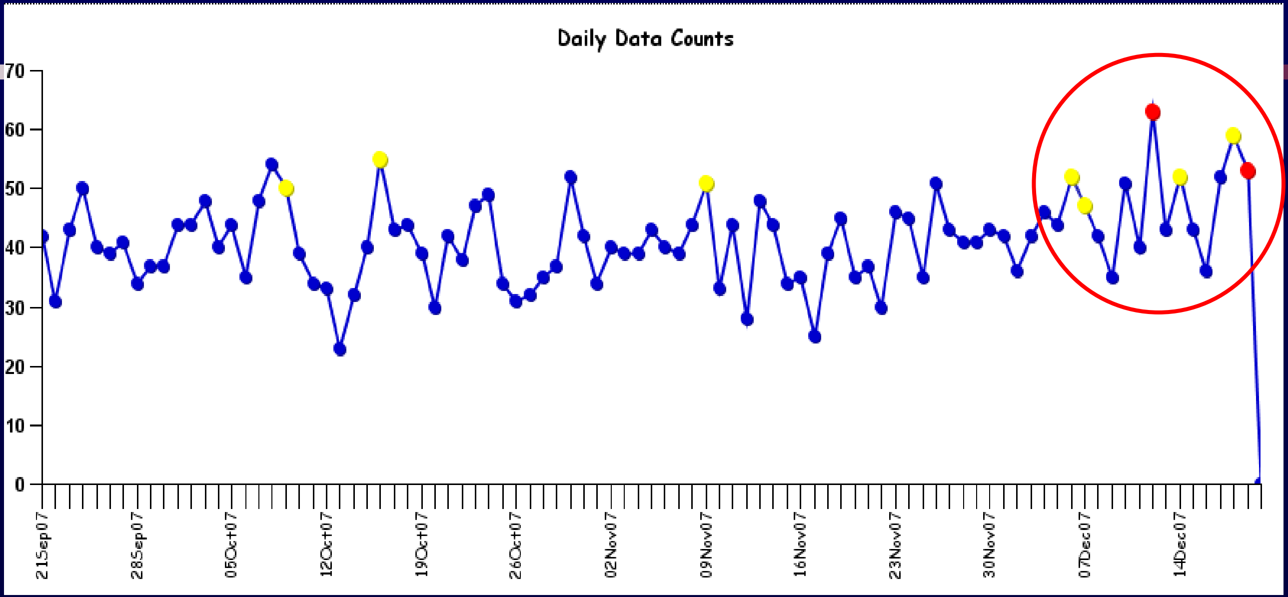
Syndromic Surveillance

- How do we conduct early identification of adverse health events with high potential to cause high morbidity and mortality?
- In response to the need for earlier recognition of significant health events, public health institutes have developed modern surveillance applications based on the world wide web.
- US CDC funds development of ESSENCE: Electronic Surveillance Syndrome for the Early Notification of Community-based Epidemics.

Basic ESSENCE Functions:

- Early event detection (Alert List):
  - Analyzes time-sensitive data for the purpose of detecting and flagging outbreaks as early as possible.
- Situational awareness (Query Portal):
  - Queries data sources to obtain size, location, spread of an imminent health event.
  - Tracks ongoing health events to assess impact in terms of time, geography, and demography.

Alert Generation



- Alerts are detected by ESSENCE using specially developed statistical algorithms.
- Red flags: the difference in count and expected frequency to be statistically significant at p-value ≤ 0.01.
- Yellow flag: difference at statistical significance of a p-value between 0.01 and 0.05.

VACCINE

Data Sources

- Data sources:
  - Emergency department data
  - Over-the-counter drug sales (non-traditional data source)

- Why does this work?
  - In USA, Emergency Departments are the primary care provider for the uninsured population

ESSENCE Syndromes

- **Syndromes (subsyndromes):**
  - Respiratory (cough, pneumonia, influenza)
  - Gastrointestinal (vomiting, diarrhea)
  - Neurological (meningitis, altered mental status, dizzy)
  - Fever
  - Rash (vesicular rash, chicken pox)
  - Botulism-like (weakness, blurred vision, speech)
  - Shock/Coma (syncope)
  - Hemorrhagic Illness (bleeding)

Primary Benefits of ESSENCE for Hospitals

- Offers excellent "slicing and dicing" capabilities for analysis.
- Delivers valuable hospital view of data for Infection Control Practioners.
- Allows collaboration with public health while protecting hospital/patient confidentiality.

Primary Benefits for Public Health Officials

- Early event detection.
- Outbreak case investigation and follow-up management.
- Assessing impacts of natural disasters or severe weather.
- Exposure contact tracing.
- Exposure source investigation and linking of cases and contacts to exposure sources.
- Good to know that no alerts have been generated.

VACCINE

Indiana's Public Health Emergency Surveillance System (PHESS)

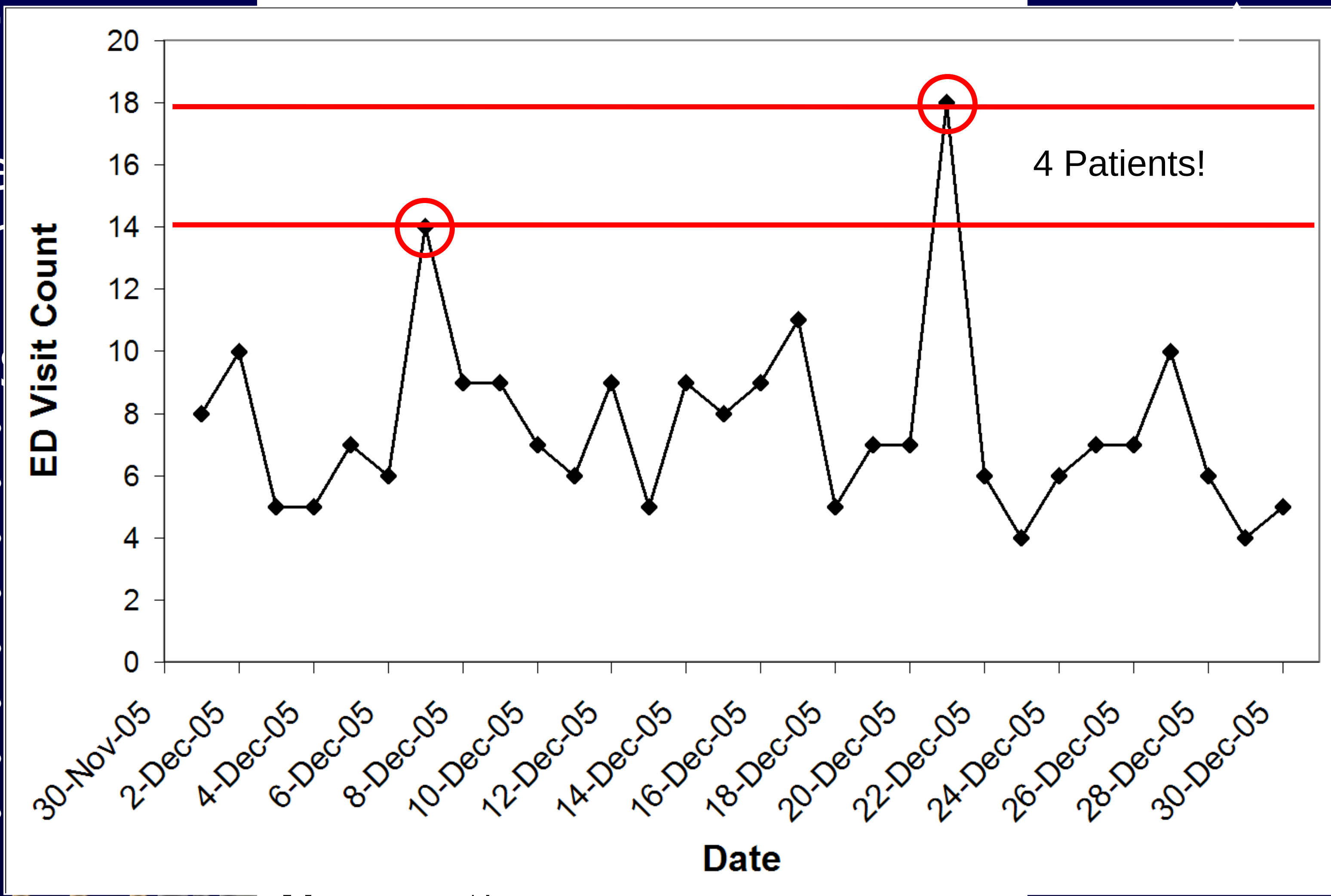
• 77

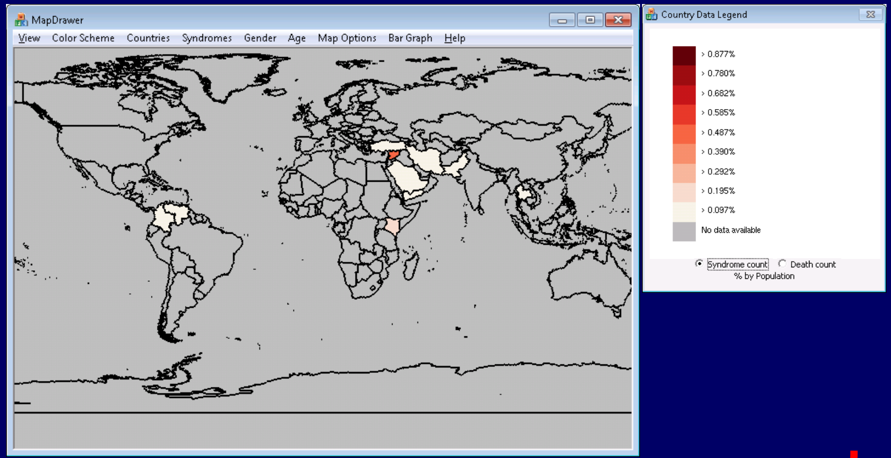
• 6,

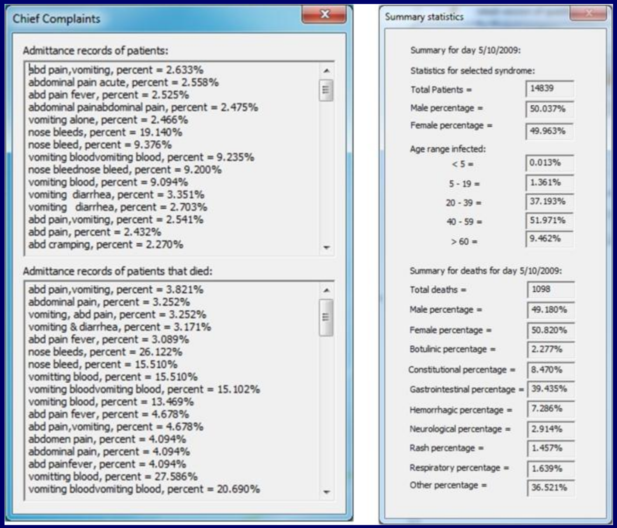
• Re

com

• 8

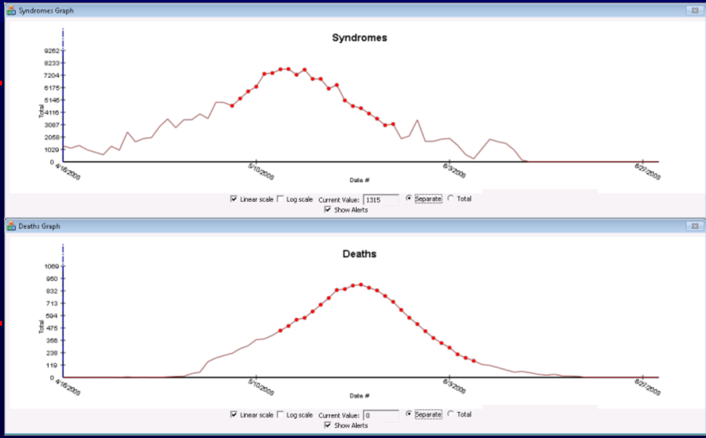


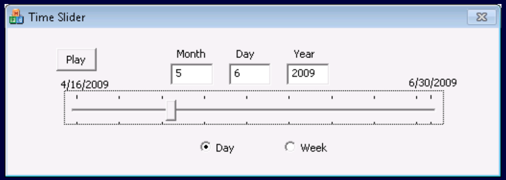




Patient ID	Date	Chief Complaints	Location
9398	4/16/09	Ear pain	Karachi
10816	4/16/09	Stuffy nose	Lebanon
1491	4/16/09	Fever	Allepo
16237	4/16/09	Head bleed	Yemen

CoCo Classifier





VAST 2010 Challenge Award: Support for Future Detection, IEEE Conference on Visual Analytics Science and Technology, October 2010 for "VACCINATED," (Malik, A., Afzal, S., Hodgess, E., Ebert, D., Maciejewski, R.).

Improving Syndromic Surveillance

Interactive visual analytic environment for effective syndromic surveillance and response

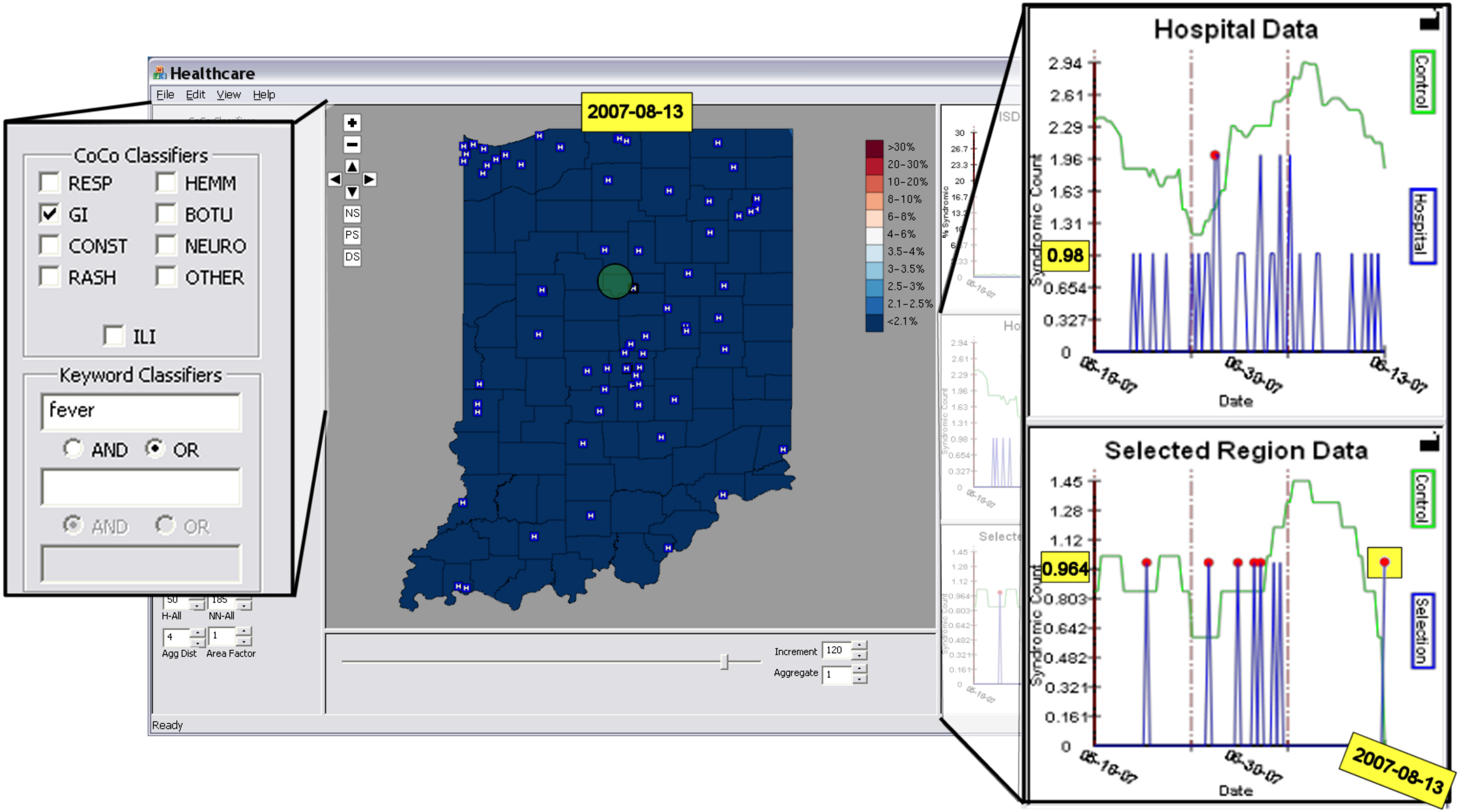
- System designed based on collaboration and feedback with state epidemiologists
- Integrated temporal, geospatial, multi-source, multi-scale analytic capability
- Density estimation for data exploration
- Syndromic control charts for temporal alerts
- Demographic filter controls for advanced analysis

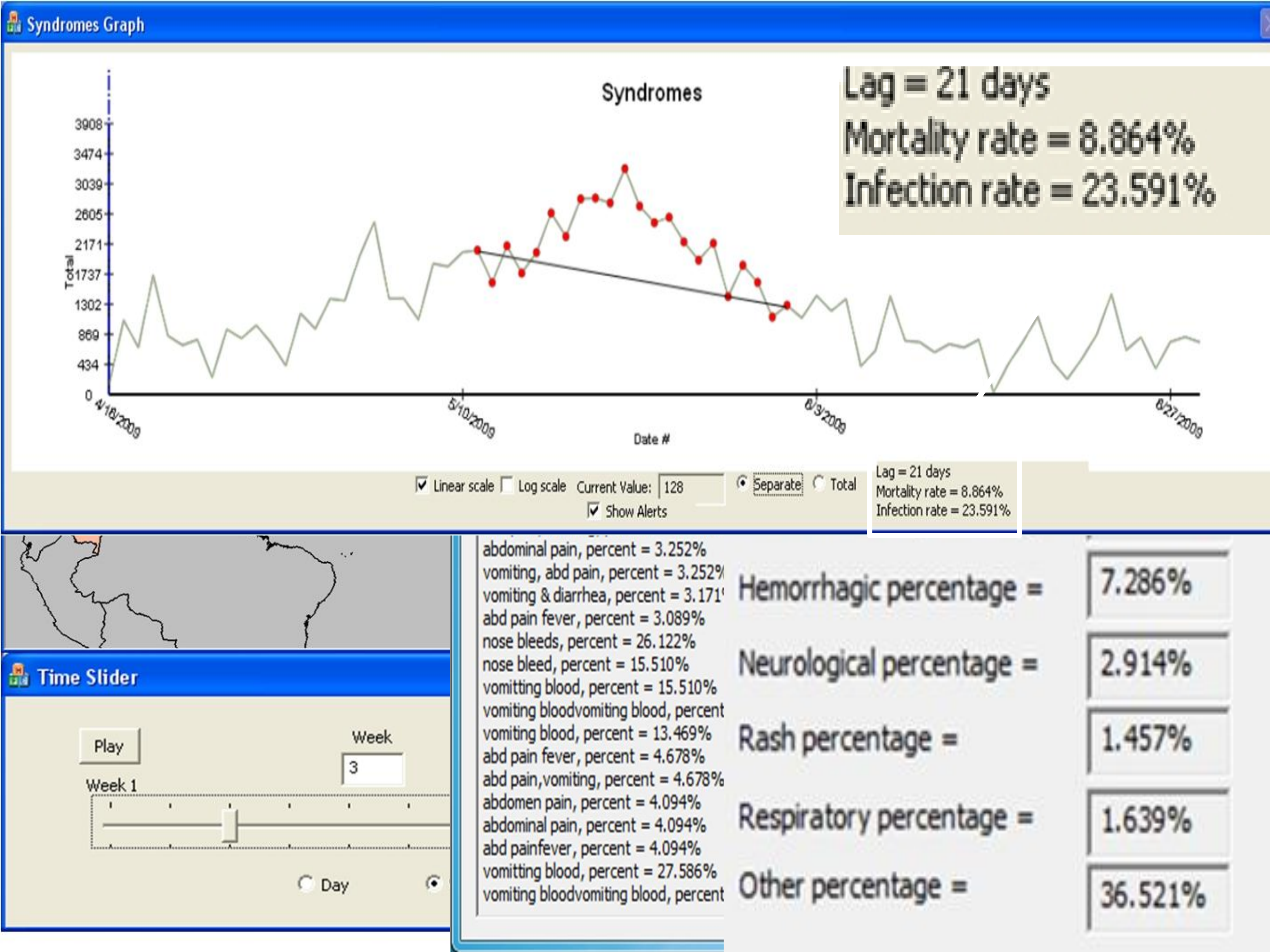
Improving Syndromic Surveillance

Benefits/ impact

- Enhanced hypothesis testing capabilities
- Linked views allow quicker cross validation of hypothesis
- Less time investigating false positives
- Systemic biological pandemic, syndromic, chem/bio/nuclear surveillance, management, and response

Spatiotemporal Hypothesis Generation and Visual Analysis





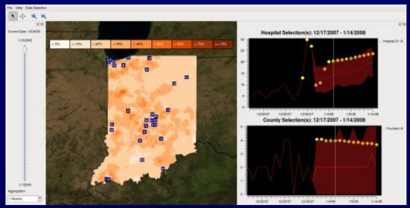
Syndromes Graph
Syndromes
Lag = 21 days
Mortality rate = 8.864%
Infection rate = 23.591%
3908
3474
3039
2605
2171
1737
1302
869
434
0
Total
4/18/2009
5/10/2009
6/3/2009
6/27/2009
Date #
Linear scale
Log scale
Current Value: 128
Show Alerts
Separate
Total
Lag = 21 days
Mortality rate = 8.864%
Infection rate = 23.591%
Time Slider
Play
Week
3
Week 1
Day
abdominal pain, percent = 3.252%
vomiting, abd pain, percent = 3.252%
vomiting & diarrhea, percent = 3.171
abd pain fever, percent = 3.089%
nose bleeds, percent = 26.122%
nose bleed, percent = 15.510%
vomitting blood, percent = 15.510%
vomiting bloodvomiting blood, percent
vomiting blood, percent = 13.469%
abd pain fever, percent = 4.678%
abd pain,vomiting, percent = 4.678%
abdomen pain, percent = 4.094%
abdominal pain, percent = 4.094%
abd painfever, percent = 4.094%
vomitting blood, percent = 27.586%
vomiting bloodvomiting blood, percent
Hemorrhagic percentage = 7.286%
Neurological percentage = 2.914%
Rash percentage = 1.457%
Respiratory percentage = 1.639%
Other percentage = 36.521%

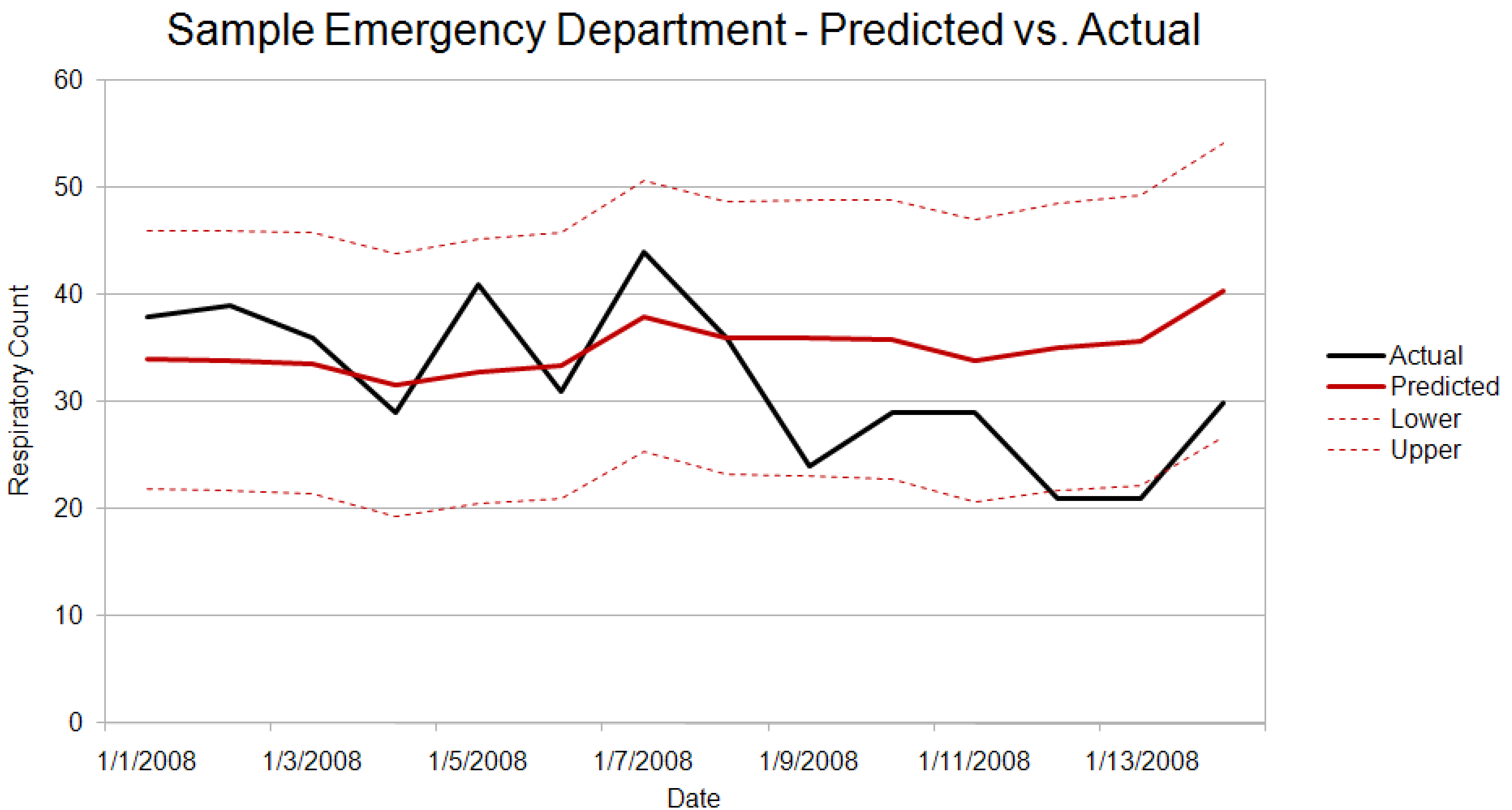
Situational Surveillance and Predictive Visual Analytics

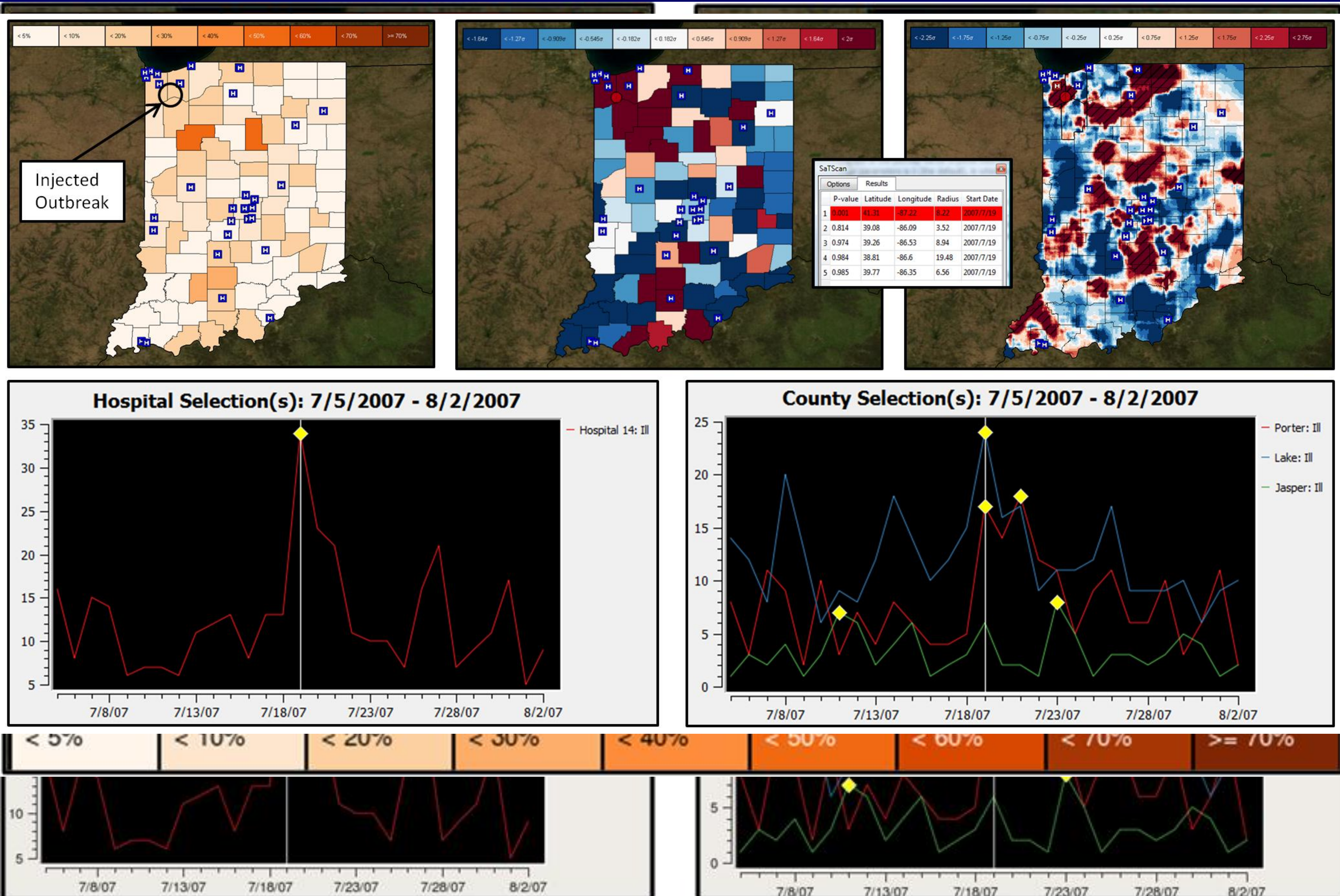
- Focus is on categorical spatiotemporal event data
- Utilizing time series and density estimations we want to create an interactive environment for predicting future event magnitudes and locations
- We utilize seasonal trend decomposition with Loess smoothing
- 3D Kernel density estimation for spatiotemporal probability distributions

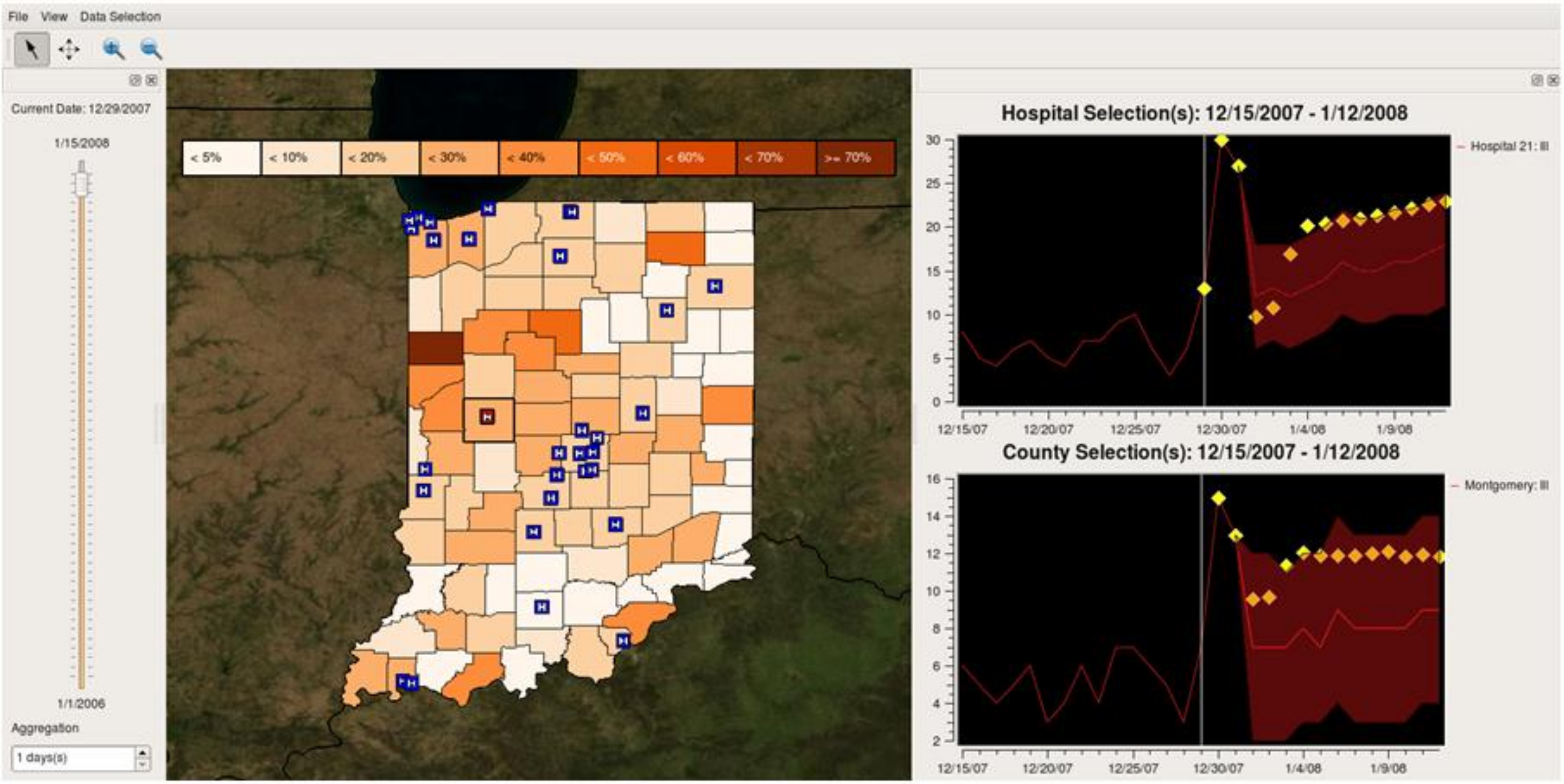
VACCINE

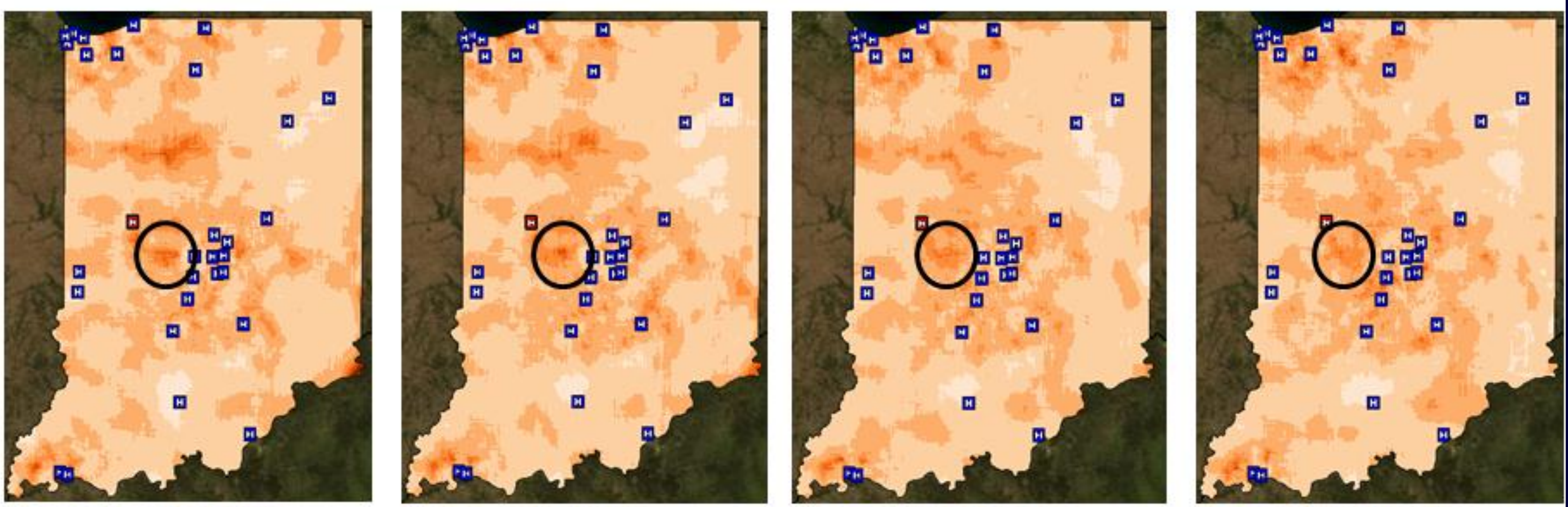
Predictive Visual Analytics





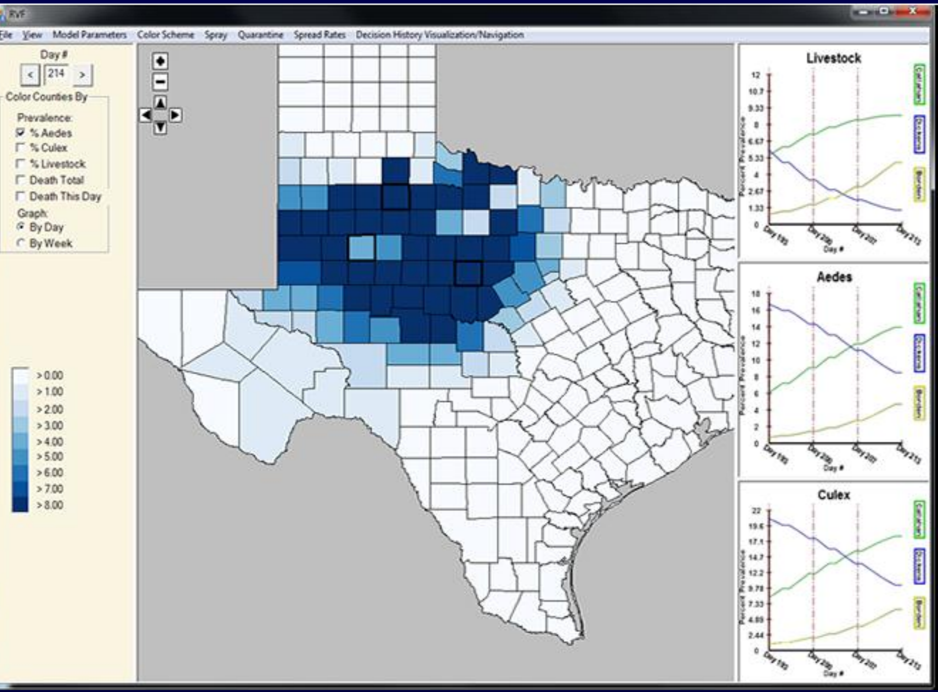

Injected
Outbreak
SaTScan
Options
Results
P-value Latitude Longitude Radius Start Date
1 0.001 41.31 -87.22 8.22 2007/7/19
2 0.814 39.08 -86.09 3.52 2007/7/19
3 0.974 39.26 -86.53 8.94 2007/7/19
4 0.984 38.81 -86.6 19.48 2007/7/19
5 0.985 39.77 -86.35 6.56 2007/7/19
Hospital Selection(s): 7/5/2007 - 8/2/2007
Hospital 14: Ill
7/8/07
7/13/07
7/18/07
7/23/07
7/28/07
8/2/07
County Selection(s): 7/5/2007 - 8/2/2007
Porter: Ill
Lake: Ill
Jasper: Ill
< 5%
< 10%
< 20%
< 30%
< 40%
< 50%
< 60%
< 70%
>= 70%

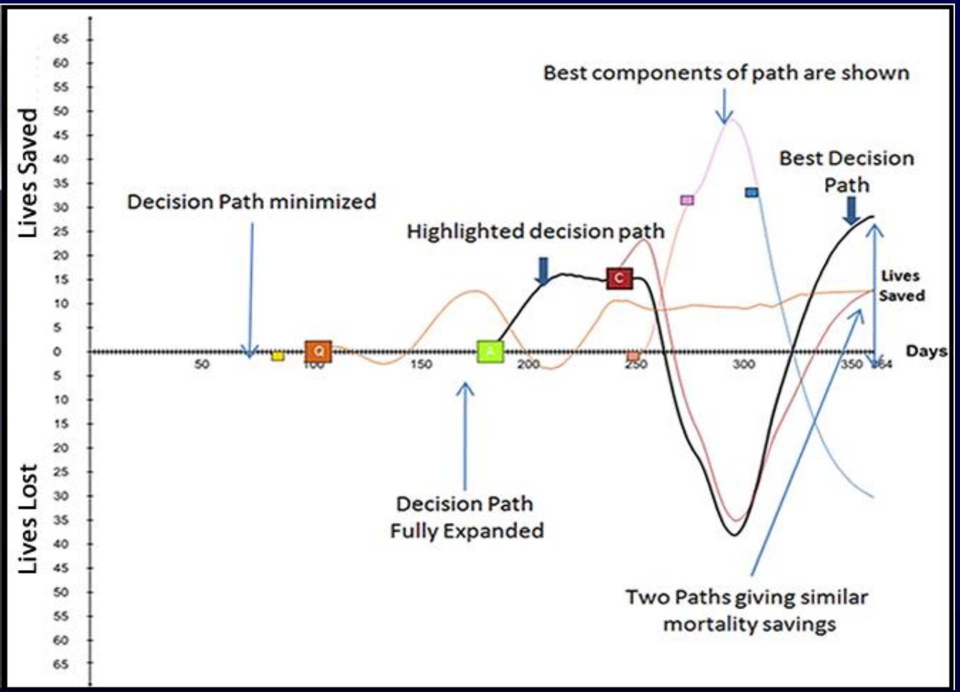




12-31-2007 1-1-2008 1-2-2008 1-3-2008

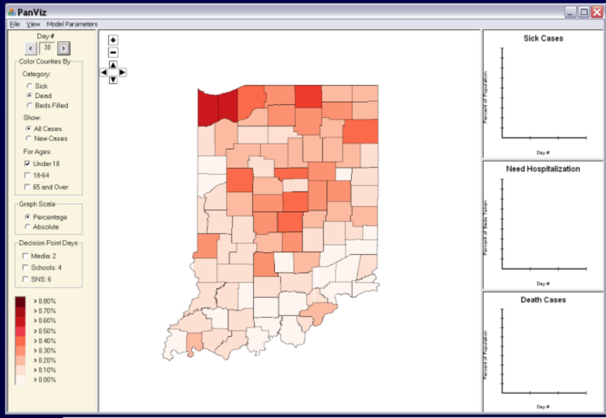
Advanced Decision Support Tools: Rift Valley Fever

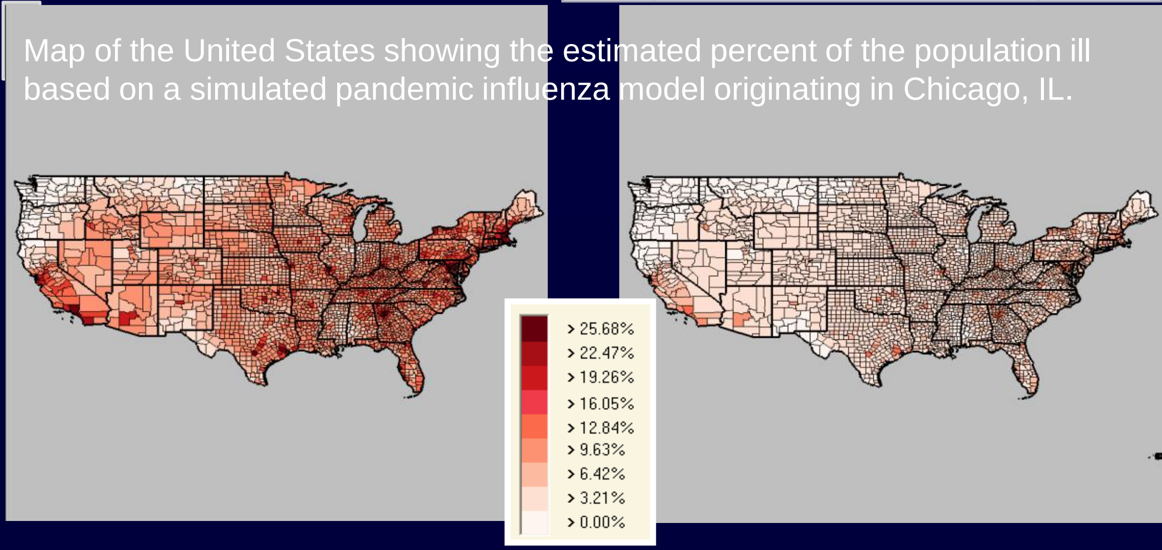




Modeling a Pandemic

- Pandemic Influenza Planning Tool
- Models user specified:
  - Pandemic influenza characteristics
  - County population, demographics, hospital beds
- Decision measures
  - Strategic National Stockpile deployment
  - School Closures
  - Media Alerts



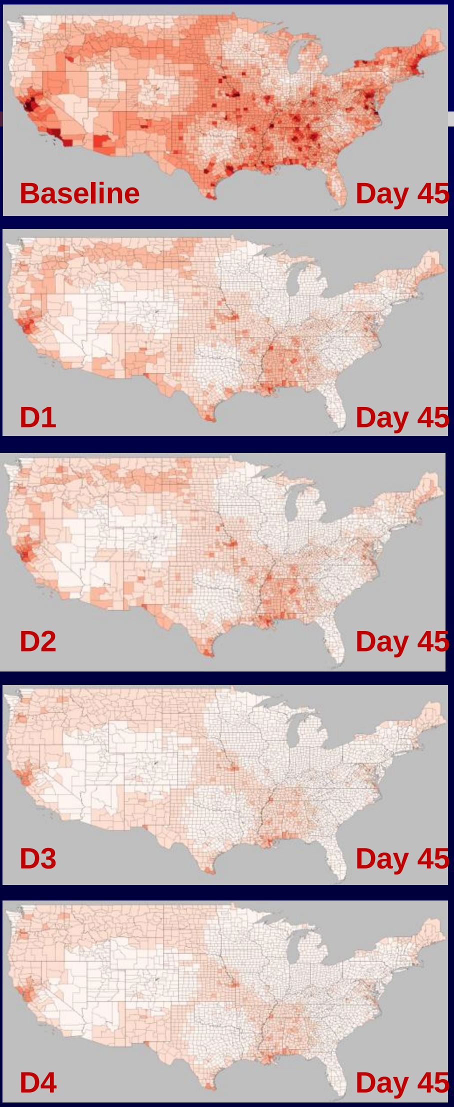


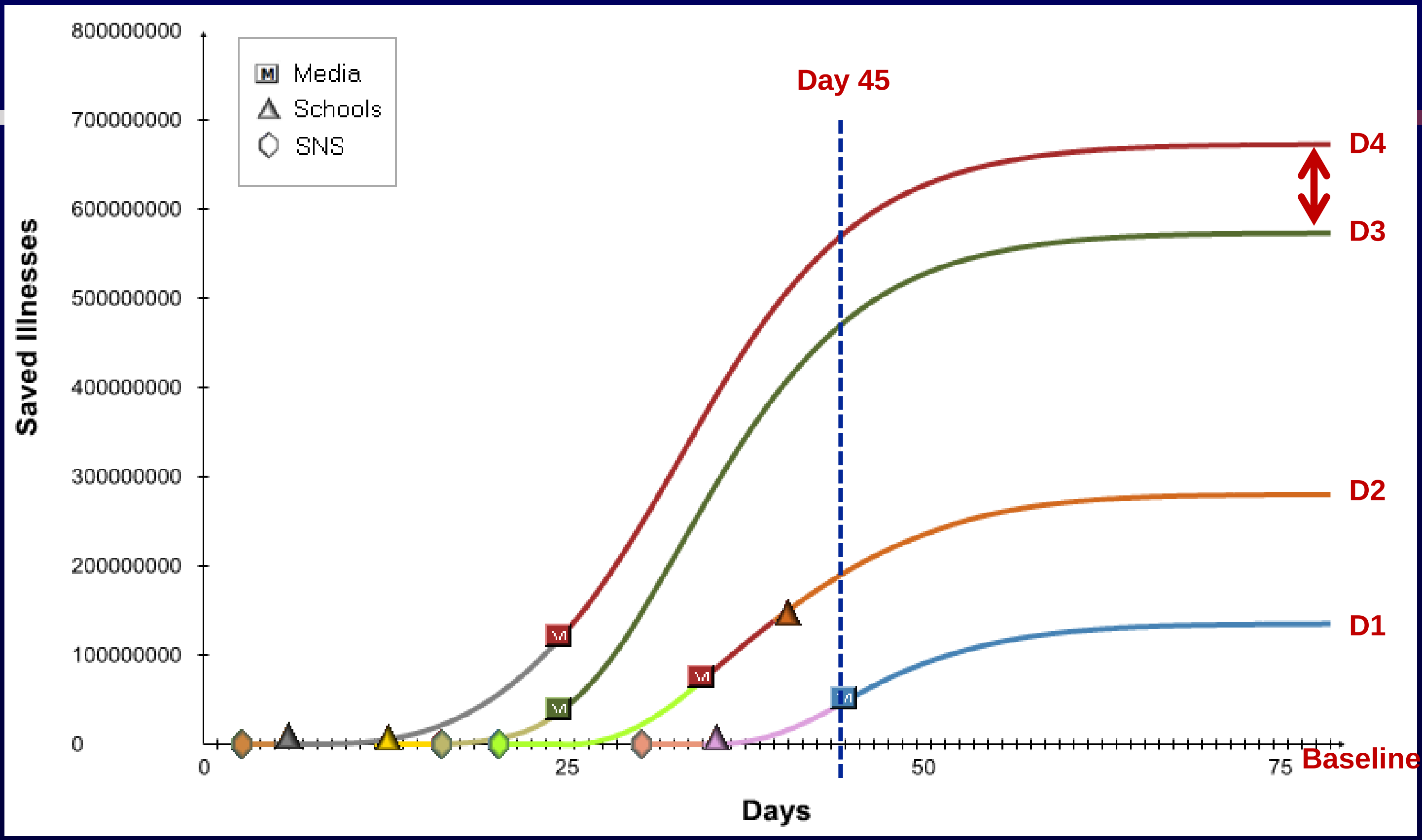
Map of the United States showing the estimated percent of the population ill based on a simulated pandemic influenza model originating in Chicago, IL.

Pandemic spread on day 37 with no decision measures implemented

Pandemic spread on day 37 with all decision measures implemented

Case Study – Pandemic Influenza





Decision Path D4

SNS Day 2 **School Closure Day 6** **Media Day 25**

VACCINE

For Further Information

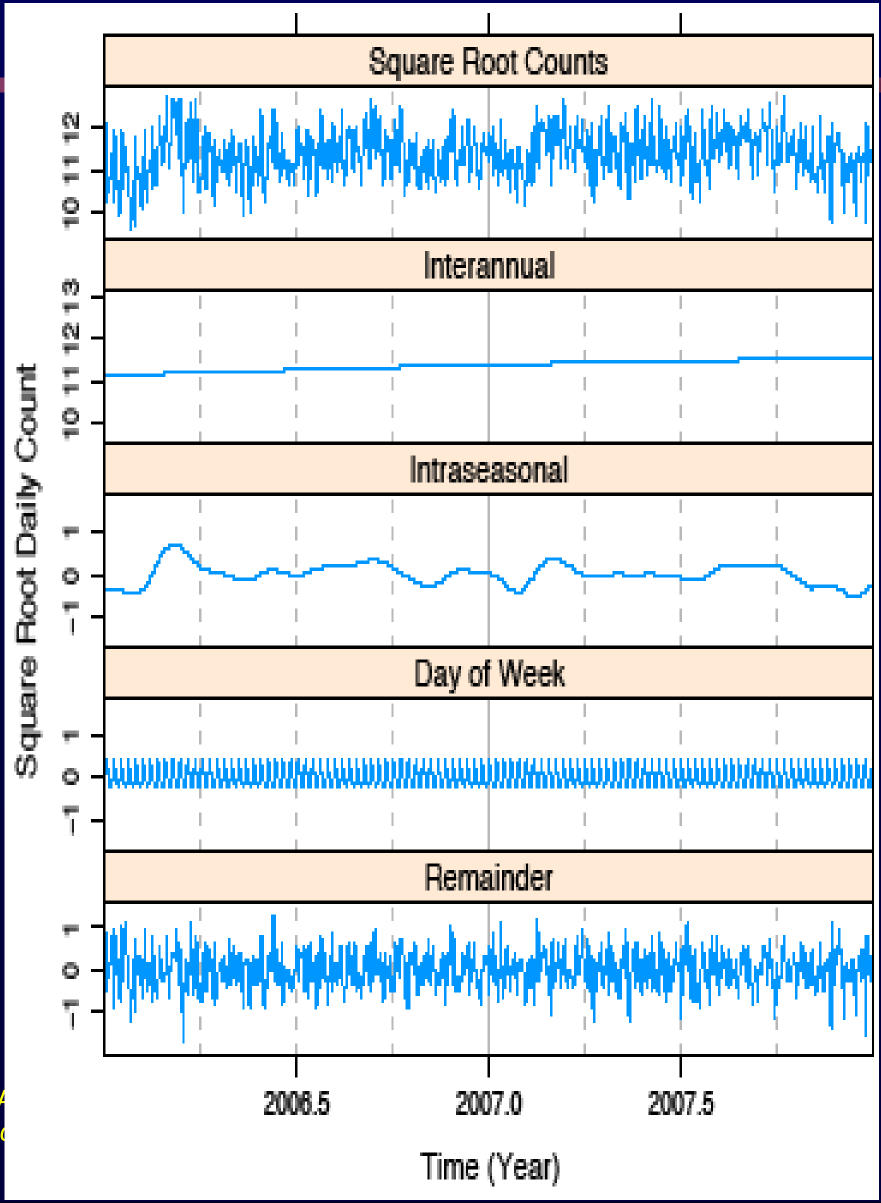
www.VisualAnalytics-CCI.org

vaccine@purdue.edu
ebertd@purdue.edu

VACCINE

Time Series Modeling

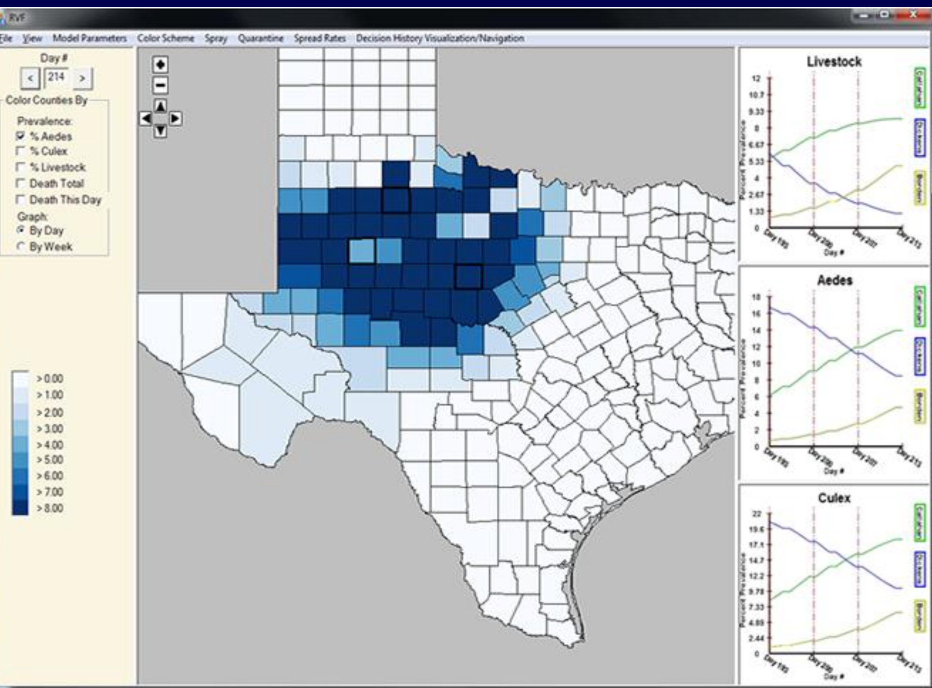
- Seasonal-Trend Decomposition Based on Loess
  - Time series can be viewed as the sum of multiple trend components
  - For each data signal, components are extracted
  - Can then analyze correlation between components



Hafen, R., Anderson, D., Cleveland, W., Maciejewski, R., Ebert, D., Abusalah, A
Visualizing, and Monitoring Disease Counts," *BMC Medical Informatics and Dec*

Advanced Decision Support Tools: Rift Valley Fever

Table of Contents



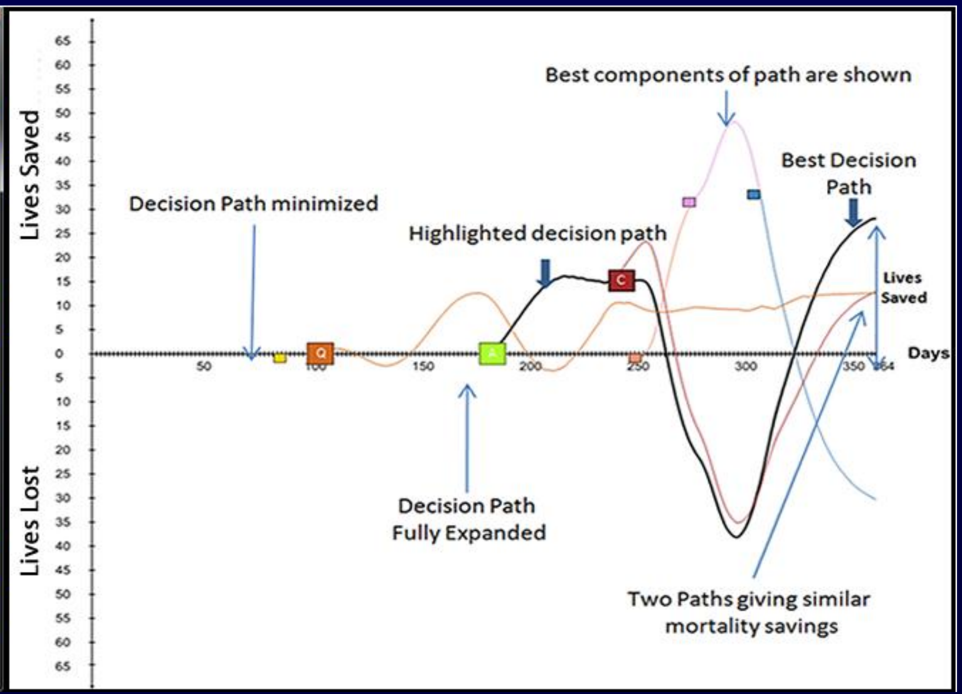


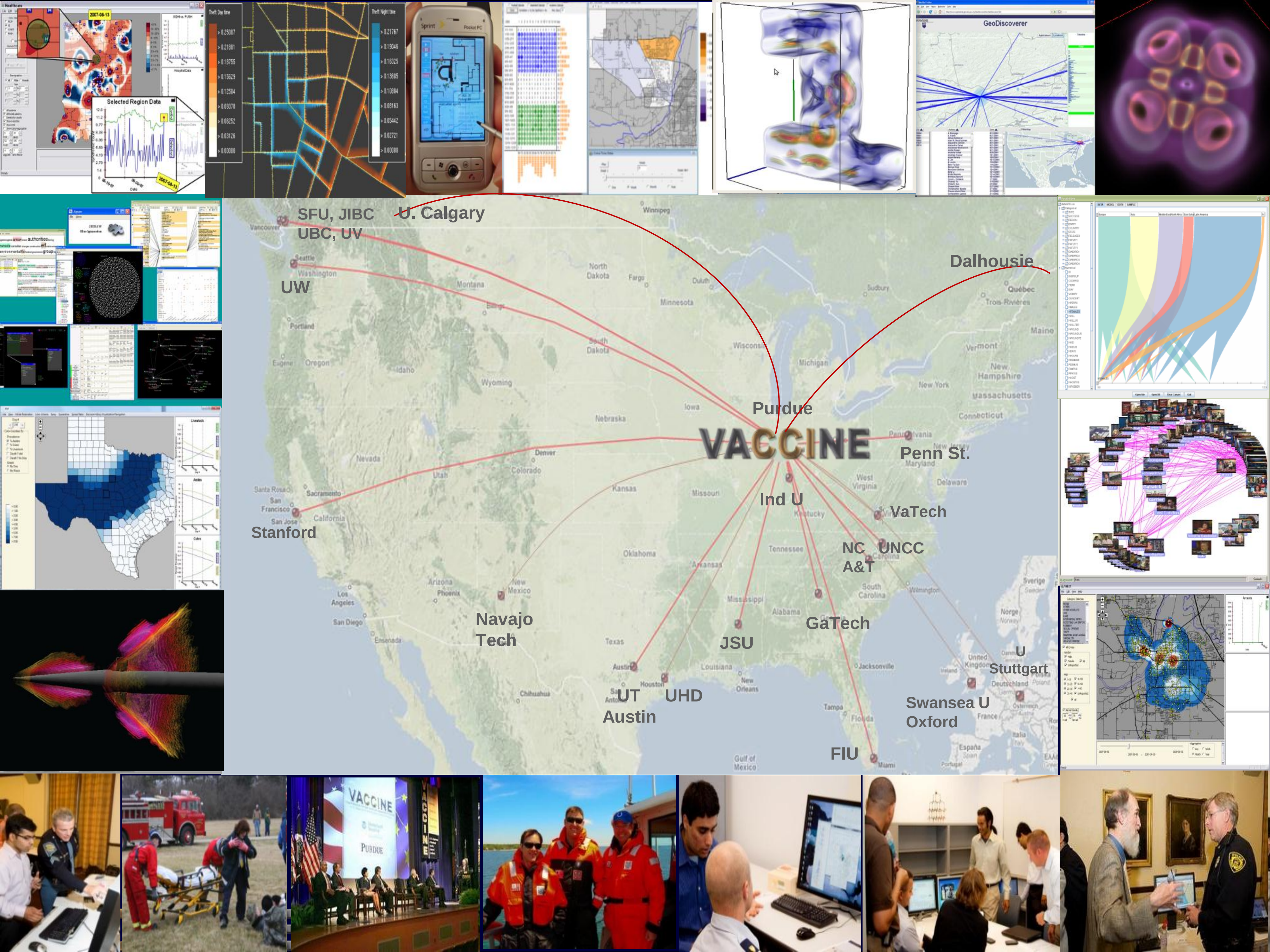
Table of Contents

CVADA
Center for Visual Analytics and Data Analytics
A U.S. Department of Homeland Security Center of Excellence

Visual Analytics for Decision Making

David S. Ebert

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


VACCINE
SFU, JIBC
UBC, UV
U. Calgary
UW
Dalhousie
Purdue
Penn St.
Stanford
Ind U
VaTech
NC A&T
UNCC
Navajo Tech
GaTech
JSU
UT Austin
UHD
U Stuttgart
Swansea U
Oxford
FIU
GeoDiscoverer
Selected Region Data

Who We Are:
International Team of Experts
- 75+ Faculty, 25 institutions

- **Purdue University**
- **Georgia Institute of Technology**
- **Pennsylvania State University**
- **Stanford University**
- **University of North Carolina at Charlotte**
- **University of Washington**
- **Simon Fraser University**
- **University of British Columbia**
- **Justice Institute of British Columbia**
- **Ontario Institute of Technology**
- **Dalhousie University**
- **University of Houston, Downtown**
- **Virginia Tech**
- **Indiana University**
- **Florida International University**
- **University of Texas at Austin**
- **Morgan State University**
- **Navajo Technical College**
- **University of Stuttgart**
- **University of Swansea, UK**
- **Oxford**
- **University of Calgary**
- **University of Manitoba**
- **Carleton University**
- **University of Victoria**

VACCINE

VACCINE Benefit

Enable users to be more effective through innovative interactive visualization, analysis, and decision making tools

- Provide the right information, in the right format within the right time to solve the problem
- Turn data deluge into a pool of relevant, actionable knowledge
- Enable user to be more effective from planning to detection to response to recovery
- Enable effective communication of information

Approach: Partner-driven solutions and research

Engaged End-Users

- Federal Operating Compoents:
  - US Coast Guard
  - US Transportation Security Agency
  - US Federal Emergency Management Agency
  - US Customs and Border Patrol
  - US CERT
- Law Enforcement
  - Over 40 local and state agencies
- Fusion Centers
  - Ohio and Indiana

VACCINE

What We Do in a Nutshell

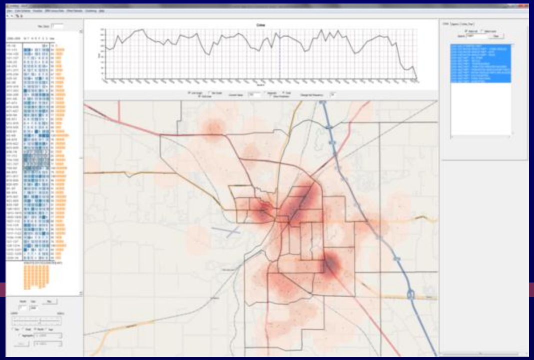
- Help people find relevant, useful information from massive, messy, conflicting, complicated data
- Enable effective communication of information
- Provide quantitative, reliable, reproducible evidence
- Enable user to be more effective from planning to detection to response to recovery

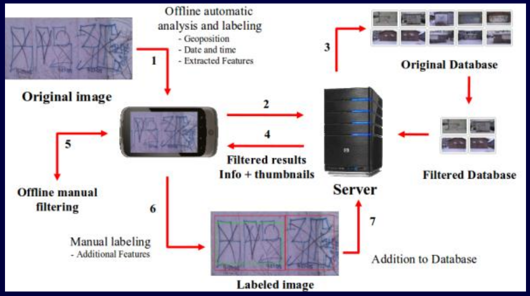
Example Projects

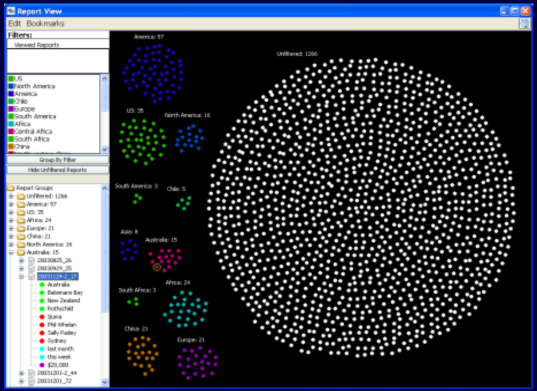
- TSA
- Law Enforcement
- USCG

Law Enforcement

- Law enforcement visual analytics
  - VALET, iVALET (iPhone/iPad), CrimeViz
    - Visual Analytics Law Enforcement Toolkit
    - Analyzing crime patterns and time of day problems
- Gang activity analytics
  - GARI
    - Gang Graffiti Recognition and Interpretation using a mobile telephone
    - Allowing police to catalog and analyze gang graffiti images, better track and determine gang activities
- Document visual analytics
  - JIGSAW
    - Visualization for investigative analysis
    - Discovery of hidden relationship and threats across documents





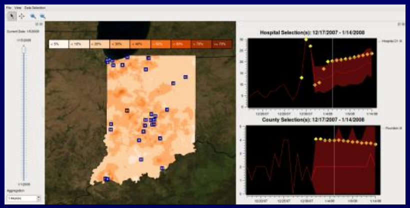


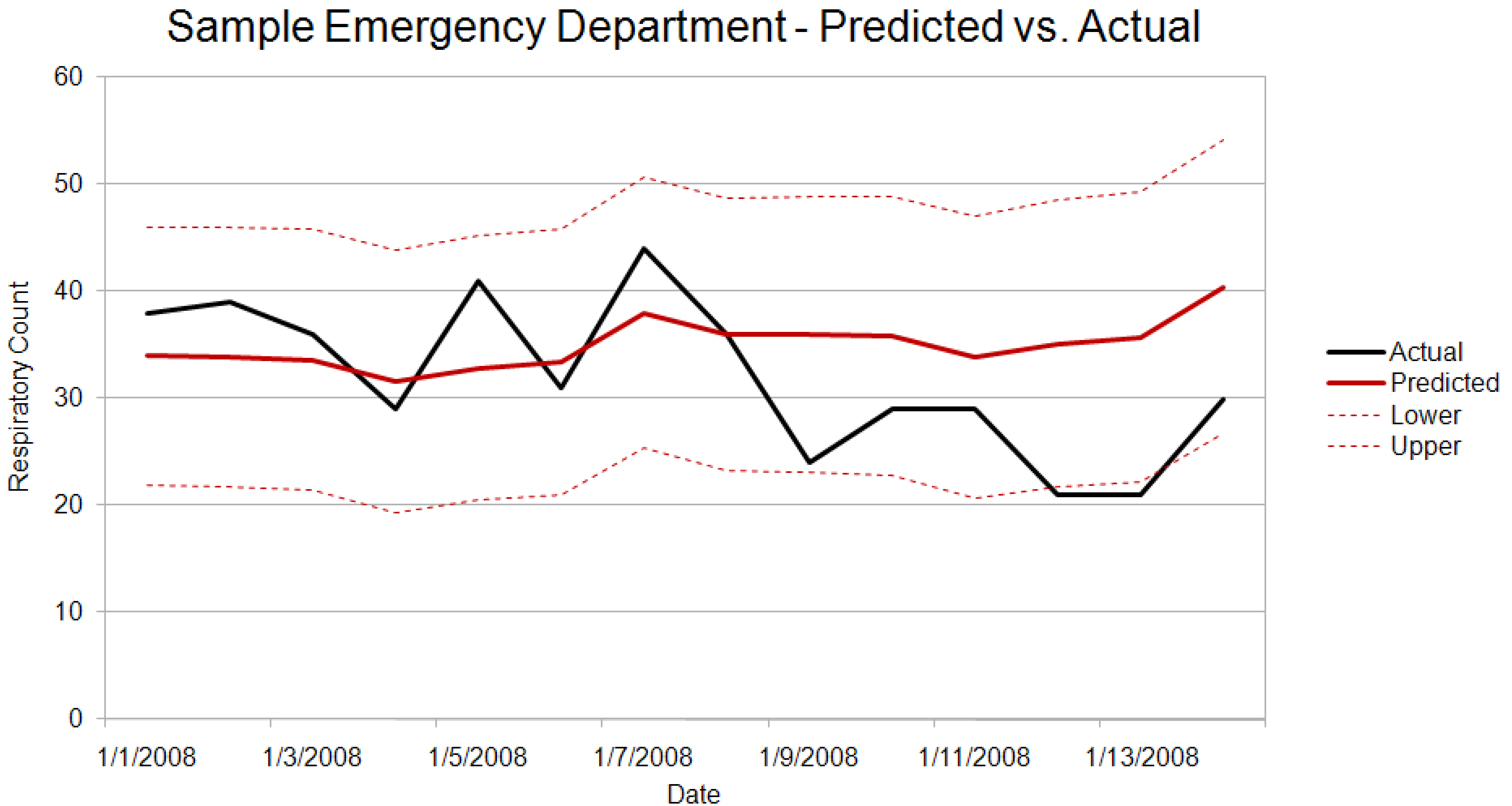
VACCINE

Temporal Predictive Analysis

- Temporal trend analysis and prediction:
  - Seasonal Trend decomposition based on *loess* (locally weighted regression) (STL)
  - Time series data viewed as the sum of multiple components
  - Filter out the noise component
  - Predict in the future

Predictive Visual Analytics

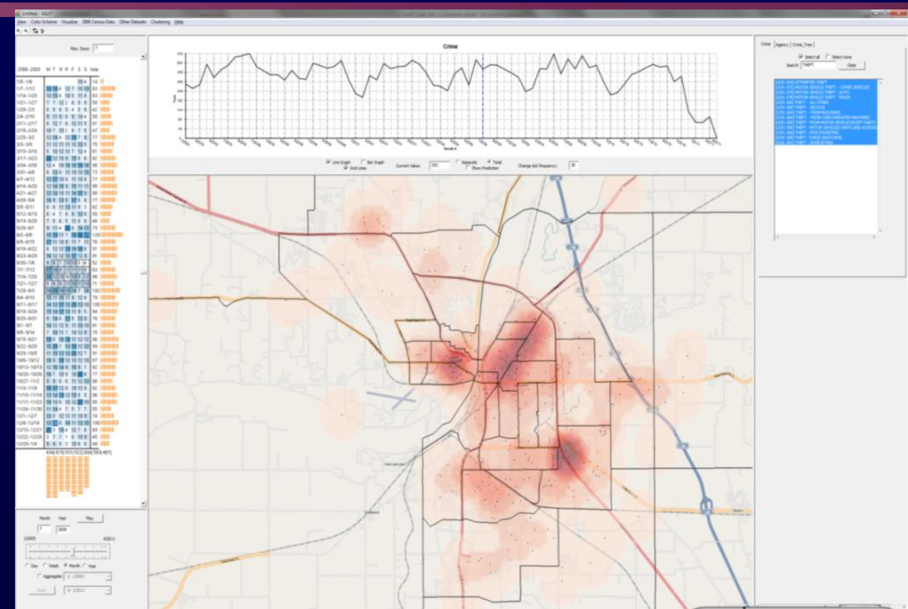




Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

Impacts:

- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (September 2012) for community-based policing
- Investigating correlation of bus routes and crime, street lights and crime
- Analyzing time of day problems and improving accuracy of police record management system
- Novel statistical predictive model incorporated for planning



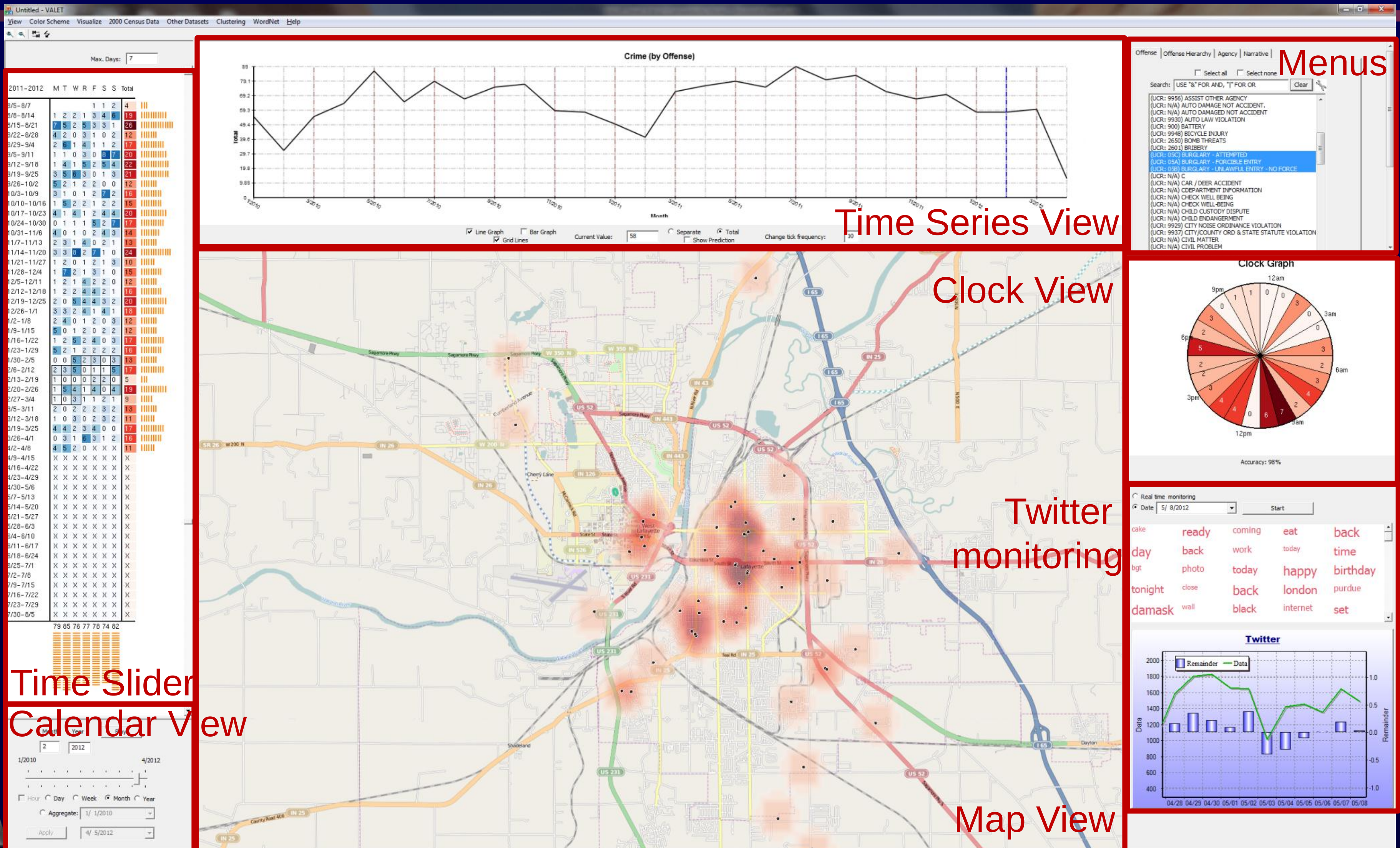
VALET delivered:

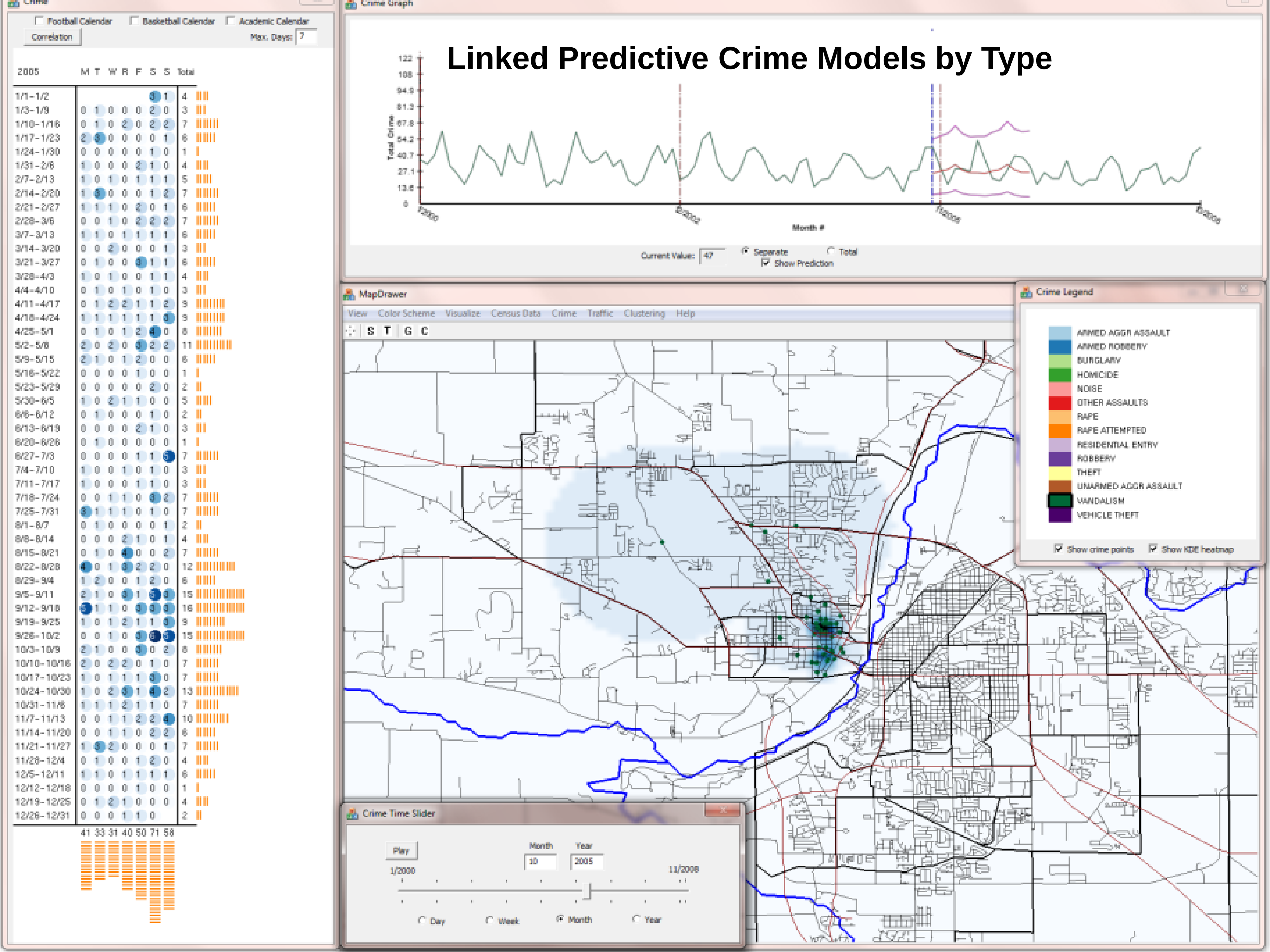
- Spring 2011: WL, Lafayette Police

iVALET delivered:

- October 2011: Purdue, WL Police

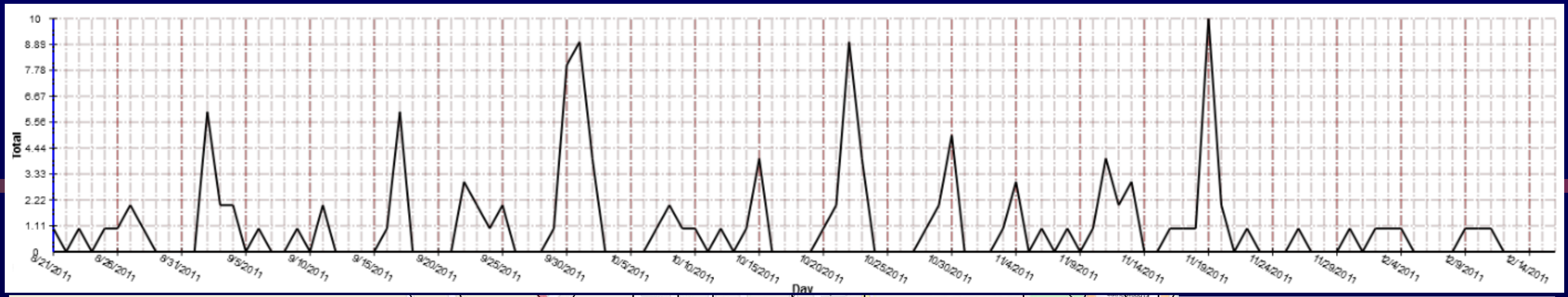
VALET Overview

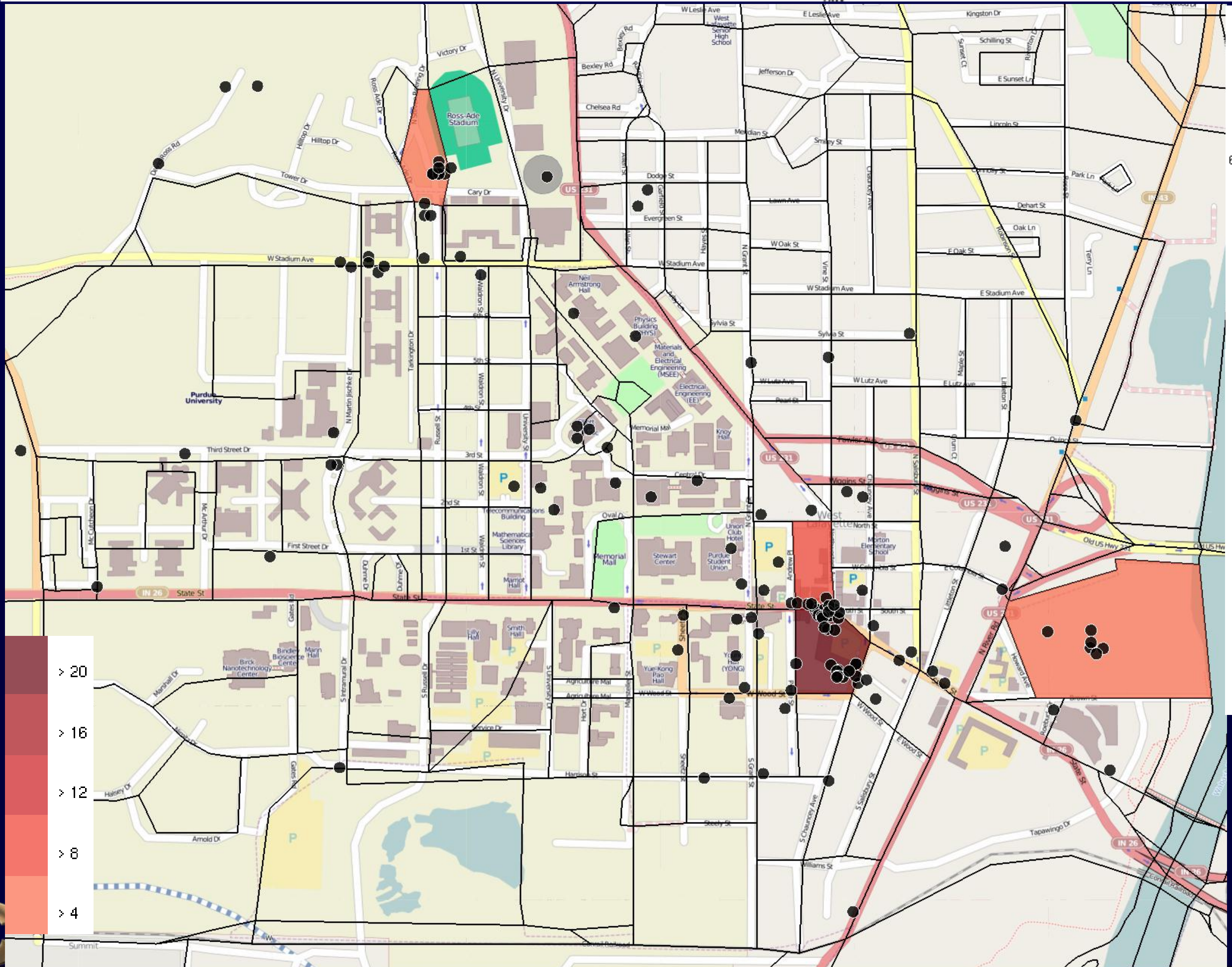


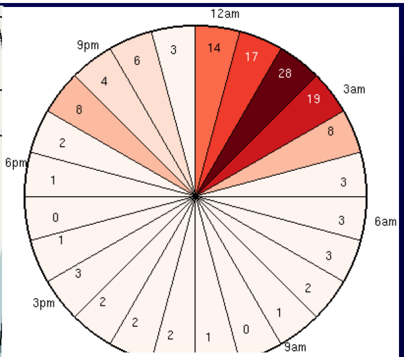


Linked Predictive Crime Models by Type
Crime
Football Calendar
Basketball Calendar
Academic Calendar
Correlation
Max. Days:
2005
M T W R F S S Total
Crime Graph
Total Crime
Month #
Current Value:
Separate
Total
Show Prediction
MapDrawer
View
Color Scheme
Visualize
Census Data
Crime
Traffic
Clustering
Help
Crime Legend
ARMED AGGR ASSAULT
ARMED ROBBERY
BURGLARY
HOMICIDE
NOISE
OTHER ASSAULTS
RAPE
RAPE ATTEMPTED
RESIDENTIAL ENTRY
ROBBERY
THEFT
UNARMED AGGR ASSAULT
VANDALISM
VEHICLE THEFT
Show crime points
Show KDE heatmap
Crime Time Slider
Play
Month
Year
1/2000
11/2008
Day
Week
Month
Year

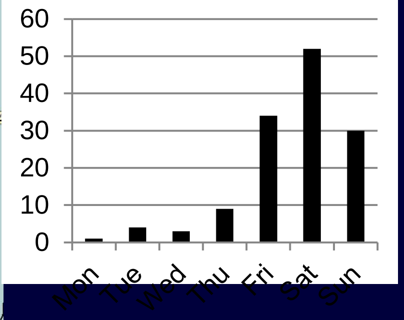
Example: Drunkenness / Public Intoxication





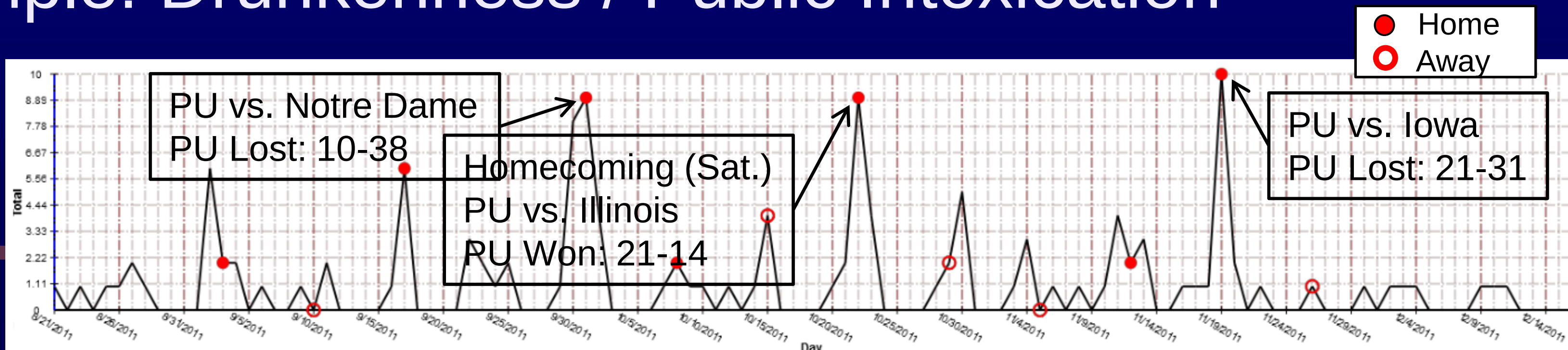


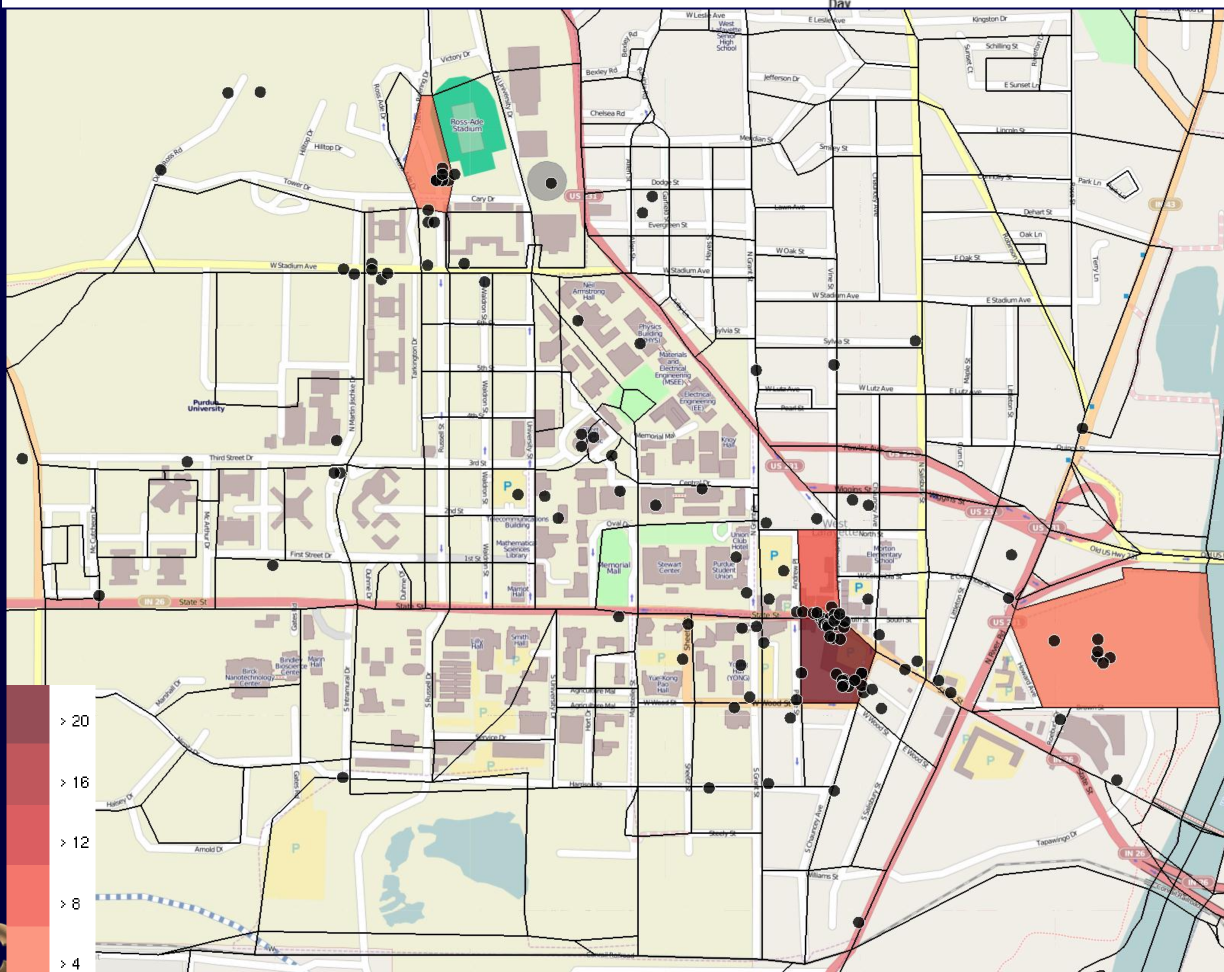
Day-of-the-Week

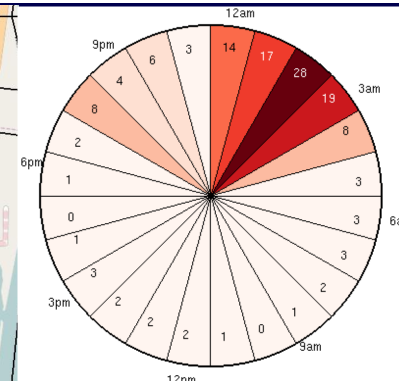


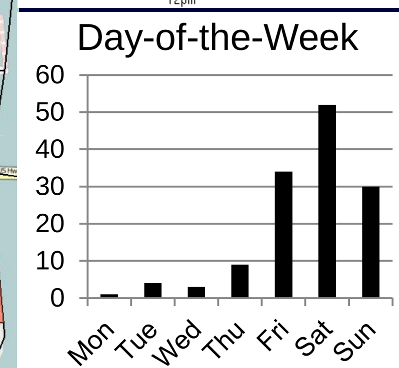
Example: Drunkenness / Public Intoxication

Football season



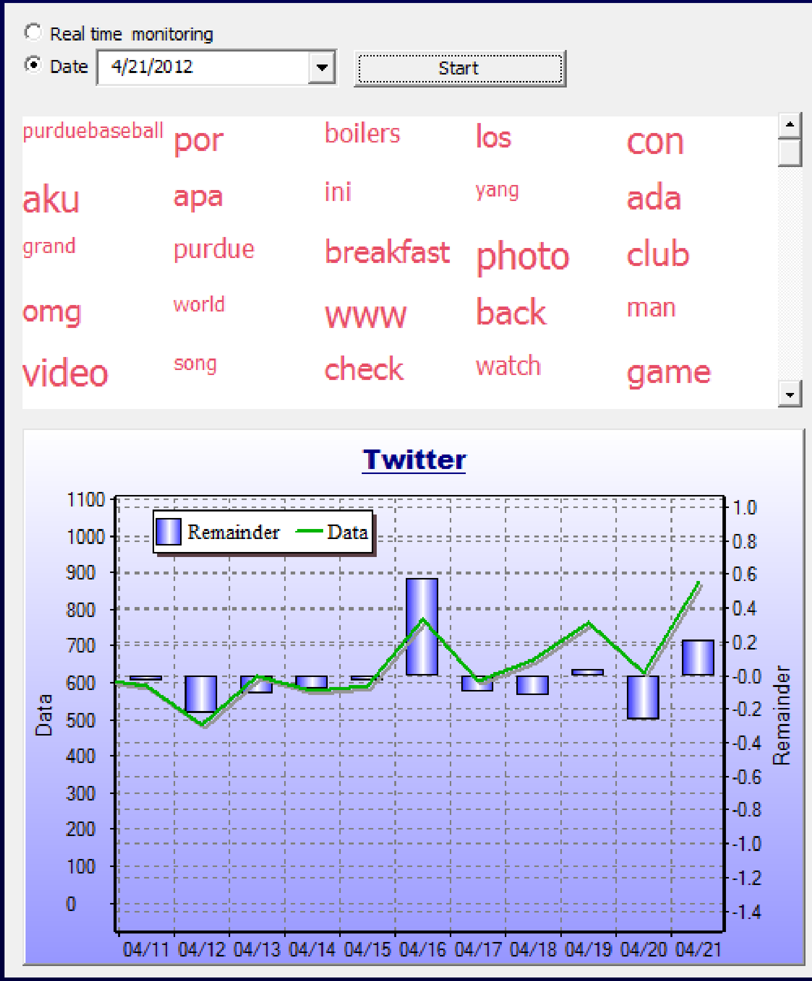






VALET Feature:
Real-time Twitter Monitoring and Integration

- Topic extraction using novel STL based remainder estimation technique
- Dynamically linked views providing options to monitor emerging / emergent twitter feeds
- Topics extracted shown as a dynamic word cloud

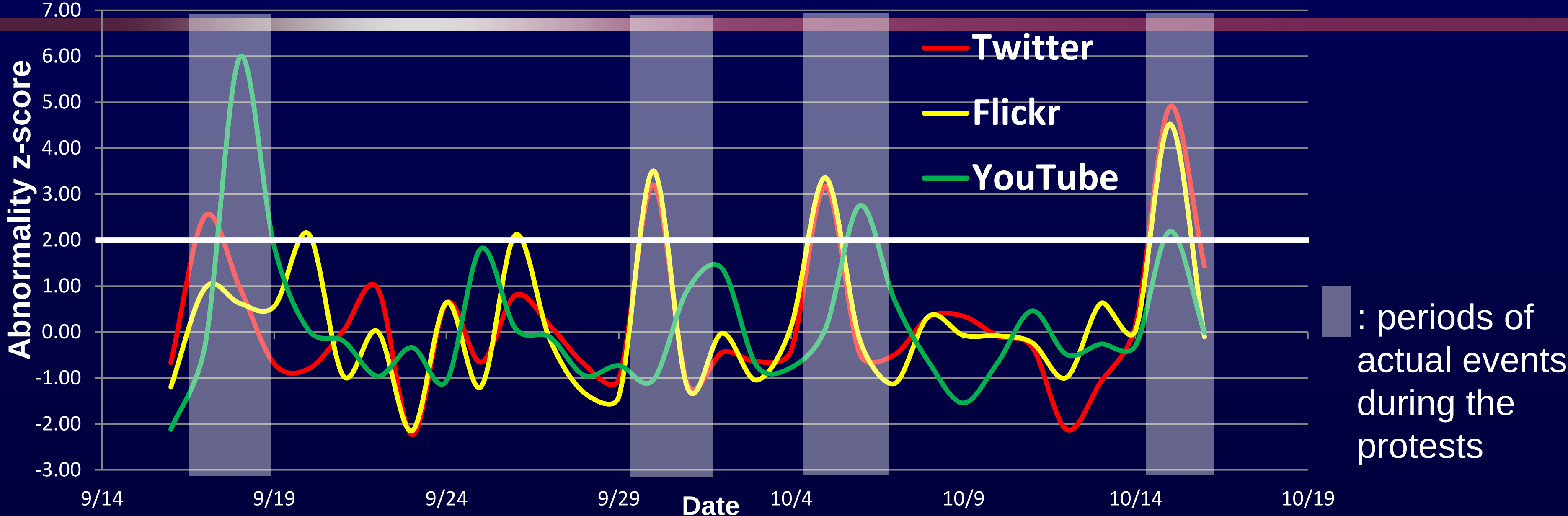


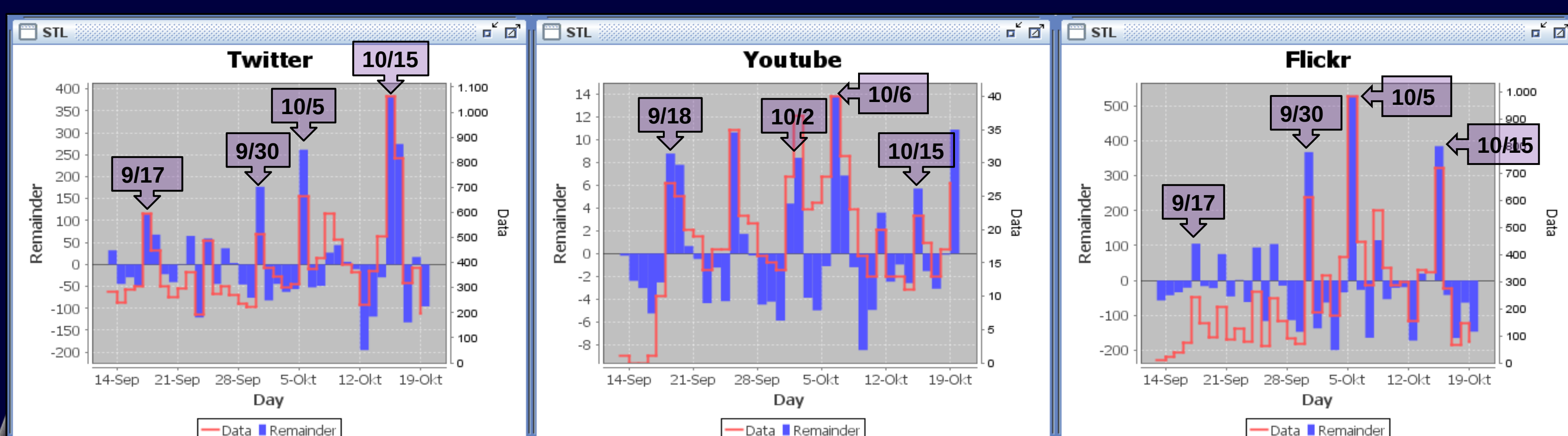
Grand Prix Weekend, Purdue University

Correlation b/w Multiple Social Media (1/2)

- Occupy Wall Street
  - Protest movement starting on September 17, 2011, in New York City
- Utilization of social media
  - Communication and reports about the movement in forms of text, images and videos

Correlation b/w Multiple Social Media (2/2)





Discussion

- Different temporal trends according to event types

	Abrupt short event (e.g., Earthquake)	**Social and planned event (e.g., Occupy wall street)**
Text type (Twitter)	**Strong change**	**Gradual increase and decrease**
Multimedia type (YouTube, Flickr)	**No significant signal**	**Strong change, but delay**

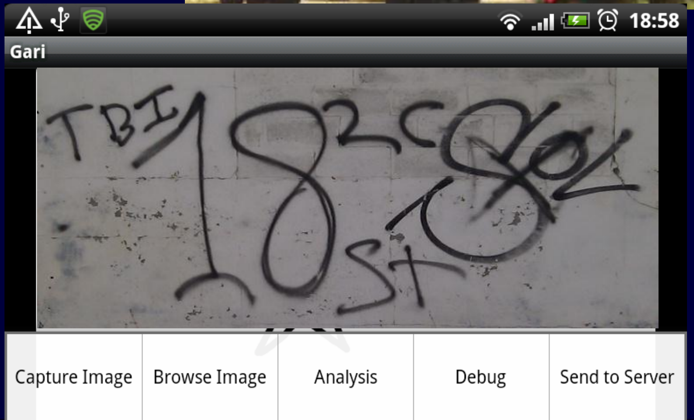
- Data limitation
  - Only geo-located messages by GPS enabled devices, but fast increase
  - The most relevant group for situational awareness scenarios

Carrier
10:11 AM
100%
Time Series Graph View
Time Series
Count
Predicted value
600
480
360
240
120
11/14/2010
01/01/2011
02/13/2011
03/27/2011
Calendar View
M T W Th F Sa S
1/1/10 36 46 29
1/4/10 38 48 44 43 42 32 30
1/11/10 62 53 53 56 59 45 34
1/18/10 62 49 53 72 61 52 49
1/25/10 98 66 57 66 64 55 49
2/1/10 58 63 50 51 73 51 42
2/8/10 50 46 48 63 52 39 35
2/15/10 48 65 48 54 55 67 51
2/22/10 51 66 48 56 57 46 47
3/1/10 50 59 42 54 82 45 41
3/8/10 50 57 64 78 63 36 30
Clock View
12Hrs
10:00
Accuracy: 64%
Choro
Plume
Moran
Region
Route
Hanna
Hedgewood
Centennial
Perrin
Lafayette
Wabash
Highland Park
©2012 Google
Time Slider
Day
Week
Month
Year
Custom
2010-08-14 to 2010-08-20

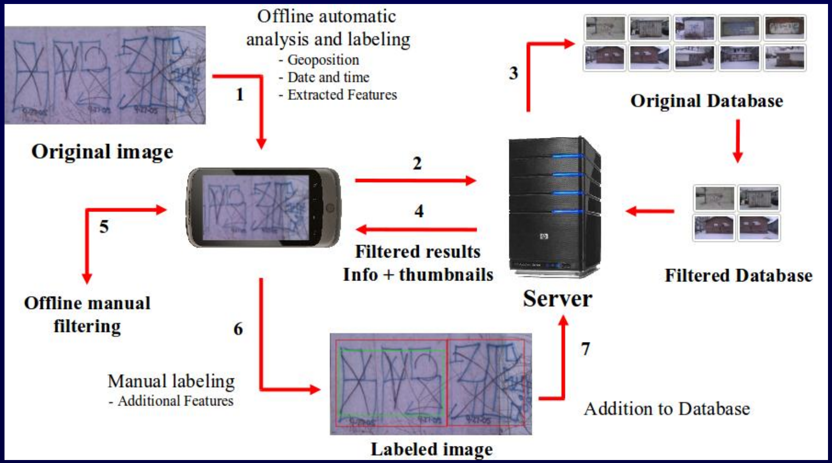
We Have an App For That!

Our Mobile Tools

- VALET
- Evacuation Planning
- Rosetta Phone
- Hazmat app
- Gang graffiti app
- Tatoo app



Gang Graffiti Recognition and Analysis Using a Mobile Telephone (GARI)



IMPACT:

- Allows police to catalog and analyze gang graffiti images into a database system to better track and determine gang activity throughout a region
- Will allow the graffiti images to be “interpreted”
- More than 60 users and 500 graffiti images acquired

GARI delivered:

- Summer 2011:
  - IPD gang detectives
- August 2011:
  - IPD at large
  - Ind Fusion Gang Task Force
- September 2011:
  - Gang detectives across Indiana

VACCINE

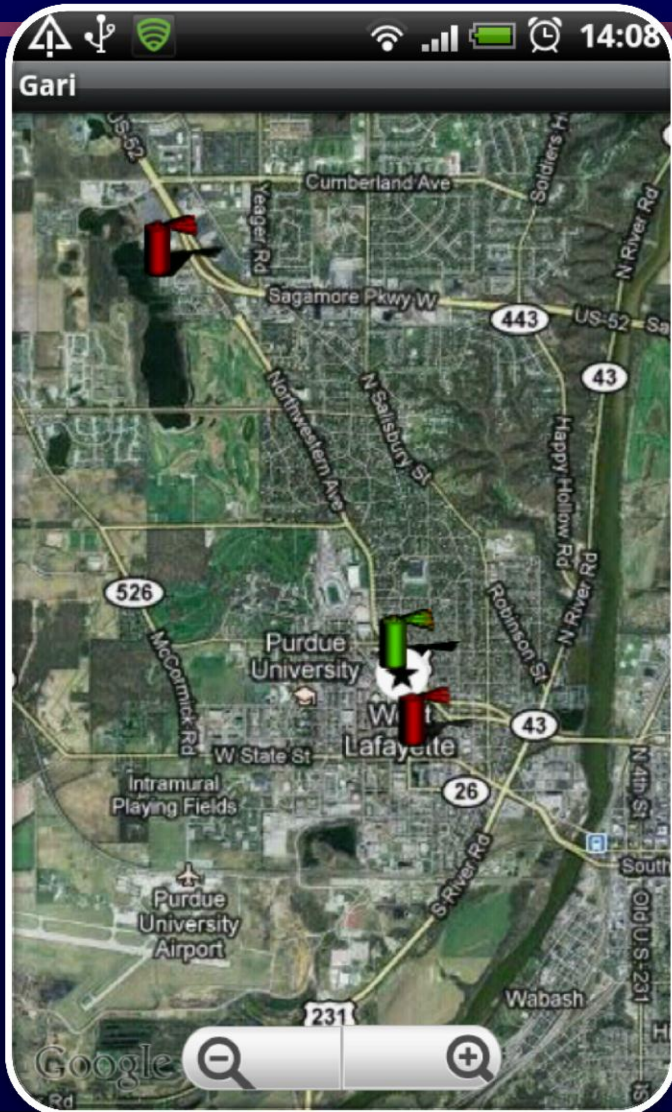
Gang Graffiti Recognition and Analysis (GARI)
PI: Ed Delp

There's An App For That

- What is GARI?
  - A mobile device application that analyzes gang graffiti
- How does GARI work?
  - User takes an image of the graffiti
  - User receives an analysis
    - Geographic locations
    - Graffiti colors, shapes, meaning
  - Database search of similar tags

Why It Works

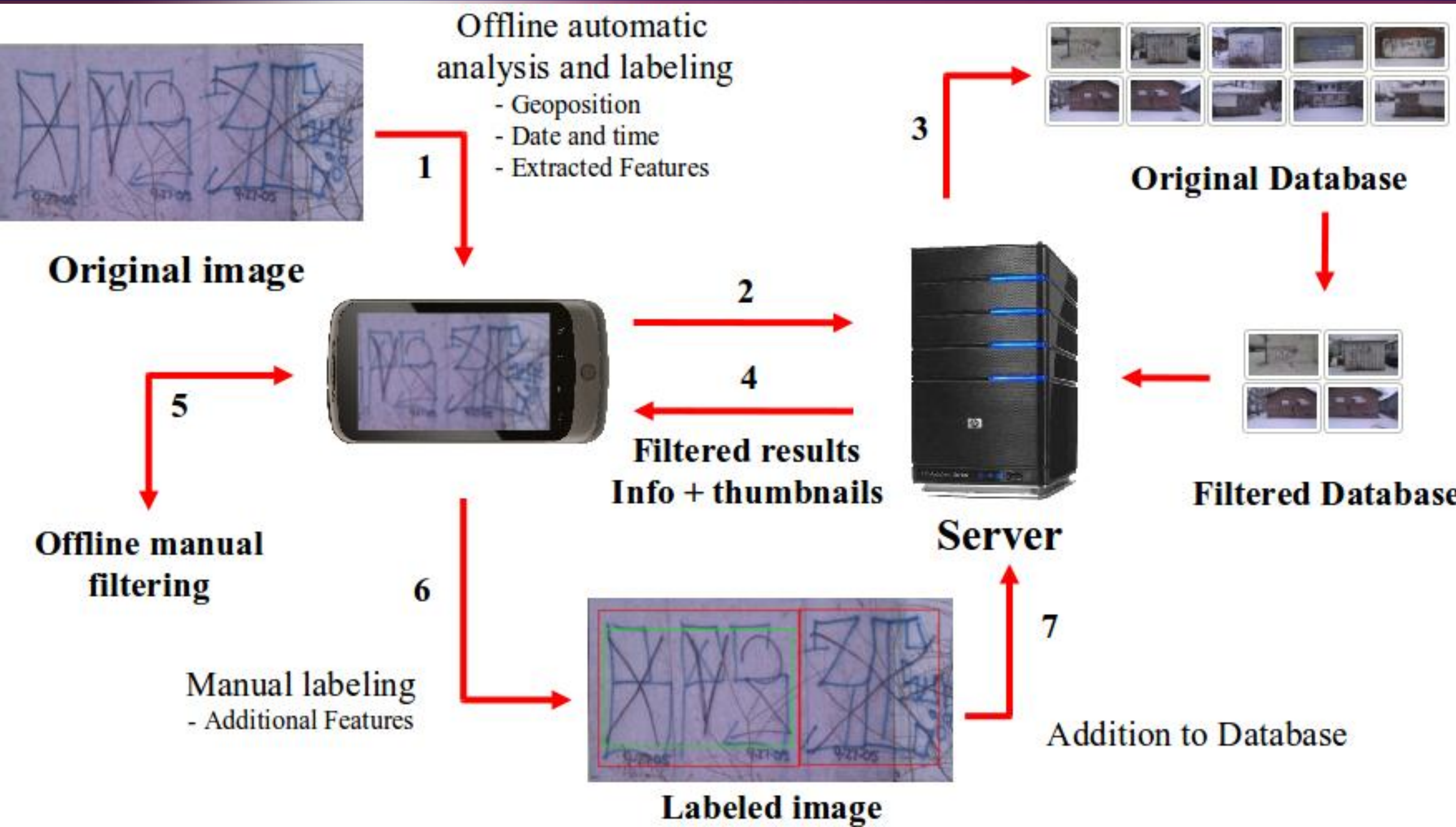
- Catalogs and analyzes graffiti images
  - More than 500 graffiti images acquired (plus 506 for research testing)
- Search of related image
- Search of local area
- Partially interprets graffiti images
  - Colors, symbols
  - Recorded gang analyst notes
- Future analysis
  - Image meaning
  - Which gang
- Helps law enforcement track and identify gang activity throughout a region



GARI On The Street

- Over 50 law enforcement officials are using or field testing GARI
- Delivered to:
  - IMPD gang detectives (Summer 2011)
  - IMPD at large, Indiana Fusion Gang Task Force (August 2011)
  - Gang detectives across Indiana (October 2011)
- Homeland Security applications

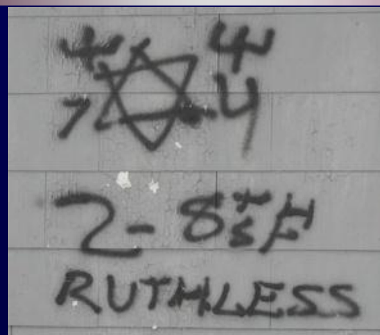
GARI System Overview



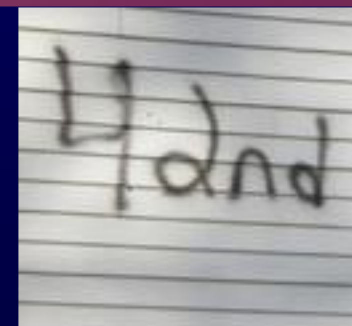
Project Status

- Deployed on Android Phone
- Main phone application is image database system and browsing (acquisition tool) with searching and some limited analysis capabilities
- Desktop backend
  - Browse by radius
  - Upload images
  - View and edit image details
  - Interact with map
  - Searching for similar images

Gang Graffiti Interpretation

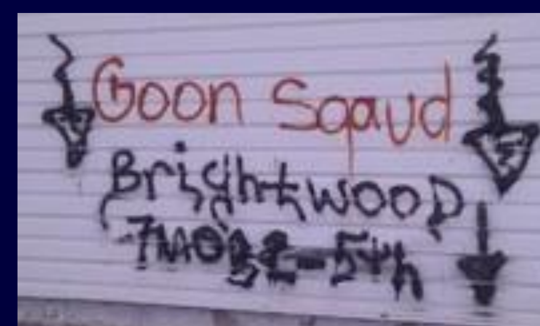


SHAPE
Simple, Straightforward



NUMBERS
42nd street gang

SYMBOLS
6-point star, pitchforks



COLOR
Goon Squad: Red/Black

Gang Graffiti Interpretation

LETTERS

East Side Gang

TIME

Black: 18 ST (18th Street Gang)

Red: 13 SUR (Sureños 13)

POSITION/ALIGNMENT

Letters at star points

Numbers in the middle

Letters at the bottom

Pitchforks upright

Image Analysis – Color Recognition

→ MEXICANOS MALDITOS
SUREÑOS 13

→ 18 STREET GANG

→ SUREÑOS 13

Image Analysis – Scene Analysis

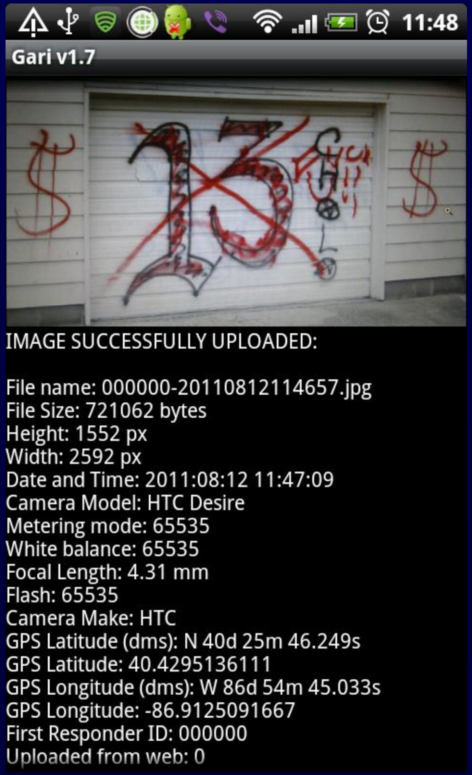
GANG RIVALRY TRACKING

- Date: 08/19/2010
- Time: 3.25 PM
- Geo: 41.387917, 2.169919

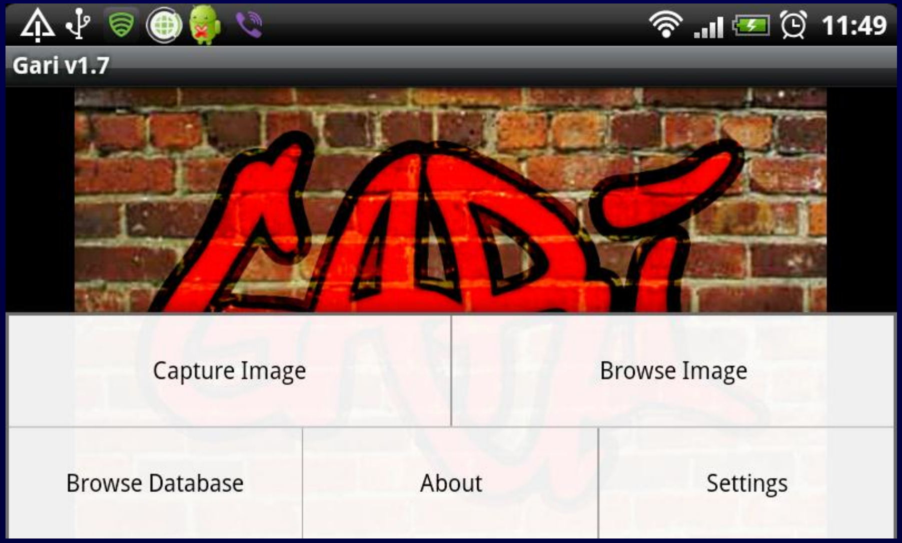
- Date: 01/03/2011
- Time: 5.11 PM
- Geo: 41.387917, 2.169919

GARI App Features

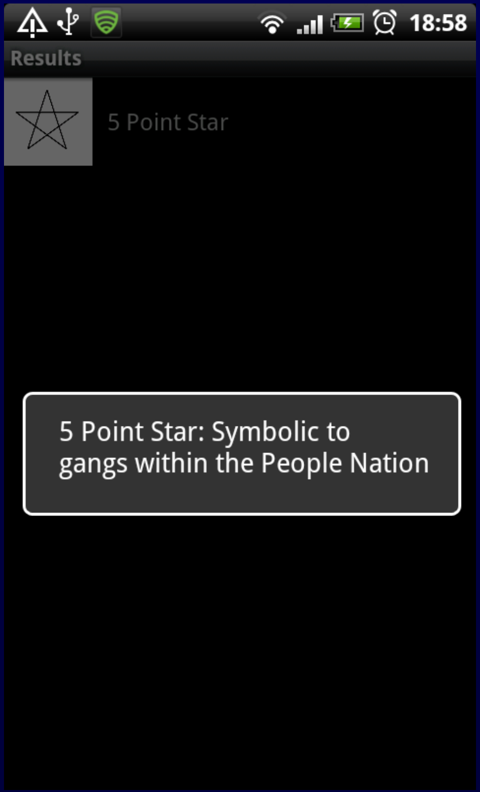
Gang Graffiti –User Interface



Send to server



Main menu



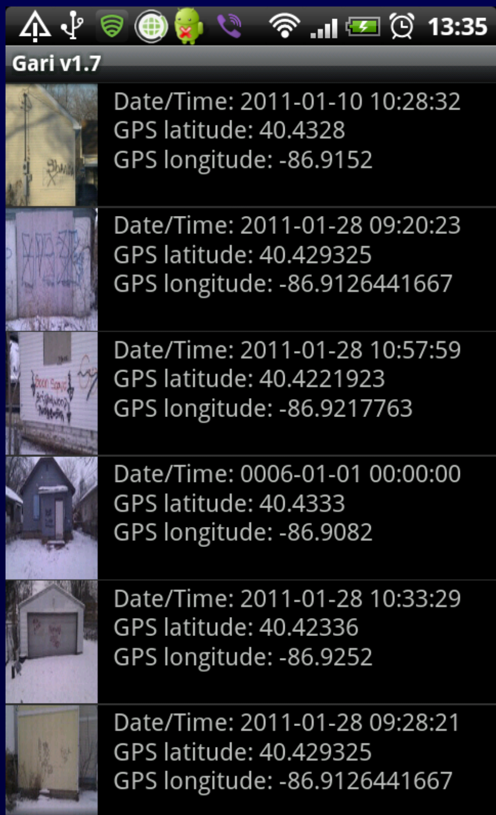
Analysis

Gang Graffiti – Browse Database

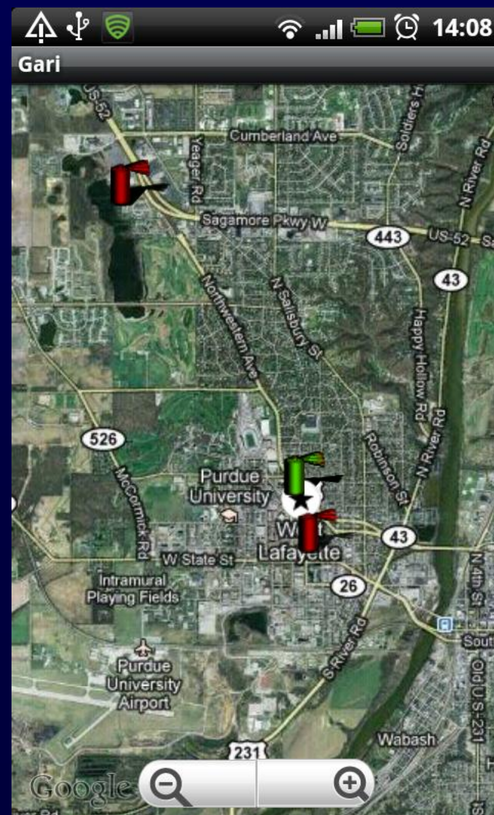
- Browse by radius from current position
- Download images and information from server (image EXIF tags, gang related information)
  - Compare results
  - Track graffiti
- Browse map (Google Maps API)

- Network connection required

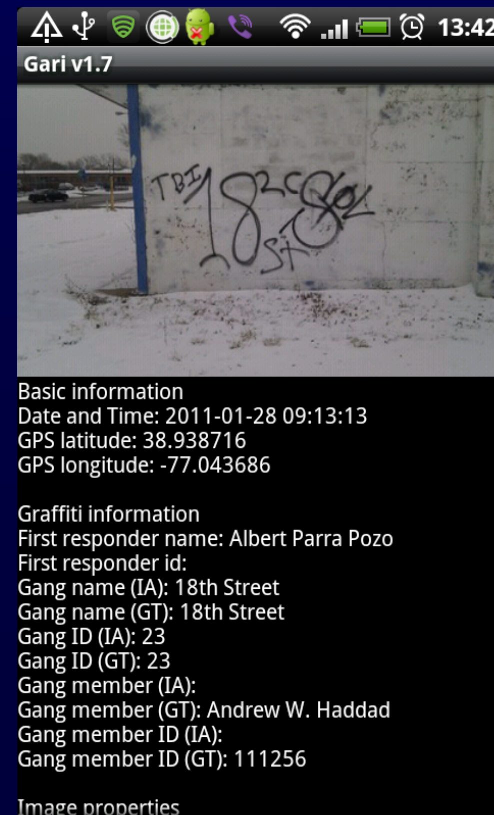
Gang Graffiti – Browse Database



Browse general results



Show graffiti on map



Inspect specific graffiti

VACCINE

GARI Progress and Directions

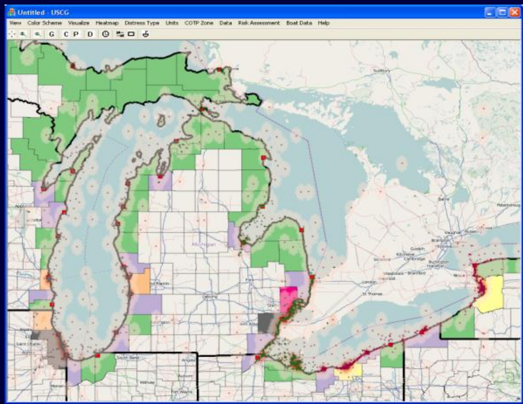
- Detective Steve Shafer, IPD adding annotations to database
- Shape and tag analysis always improving
- Transitioning to Indiana State Police / IIFC Gang Task Force
- Creating linked, satellite database and tool for Navajo Nation (using Navajo Technical College)
- Investigating use by ICE, CBP, FBI
- Two-way linkage with VALET / iVALET

How to Get The App?

- Contact gari@ecn.purdue.edu
  - We assign user IDs and initial passwords
  - Database server

Visual Analytics Uses for Risk-Based Decision Making

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Interactive, multi-day, month, type investigation
- Multisource, multimedia data integration & analysis



USCG: Effective Risk-based Decision Making and Resource Allocation Visual Analytics

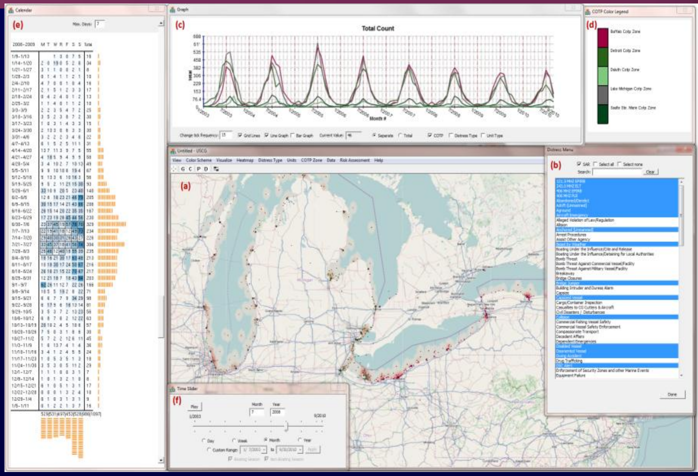
- Evaluate current and historical mission area:
  - Demands
  - Risks (total, mitigated, residual)
  - Resource allocation
  - Return on investment
- Evaluate courses of action
- Evaluate above at both Strategic and Tactical/Operational level

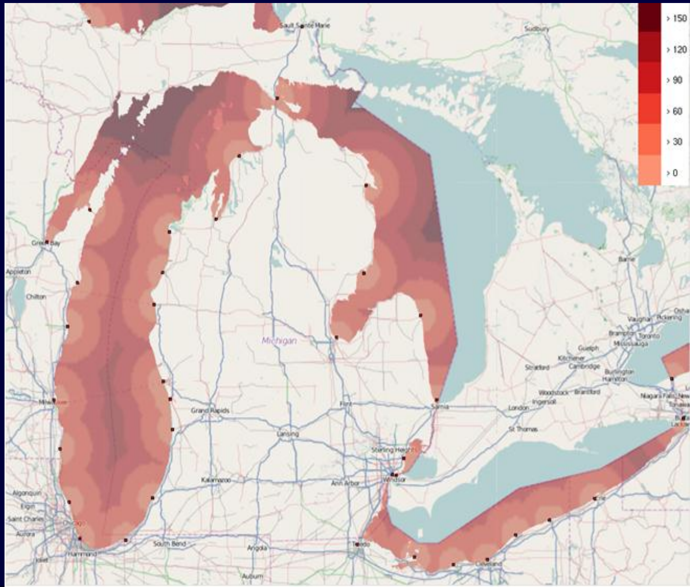
Effective Risk-based Decision Making and Resource Allocation Visual Analytics (including finished projects)

Project Components	USCG Mission Areas	Level	Resources	End-users	Interesting questions
cgSARVA / COAST	SAR, LE, others in future	Tactical	Boats, Aircraft	USCG HQ 771, USCG LANT-7 D9, D5	• Effect of closing a boat station on SAR mission •Outcome of SAR case based on initiation-type •Cost of SAR mission by case type & station •Effect of next-gen boats on station utility
ORAM VA, iOPAR	All	Strategic	Boats, aircraft, personnel hours	USCG LANT-7	Operational performance across mission areas by district and sector. Performance target evaluation Strategic resource reallocation evaluation
PROTECT (Completed)	PWCS	Tactical	Boat patrols	USCG RDC, LANT-7, D1	• What predictable patterns in PWCS patrols existed in Boston and does the new randomized PROTECT system remove patterns and optimize PWCS schedules?
Port Closure VA (Completed)	Port Security	Strategic	Ports, industry by sector, local and national	USCG RDC, LANT-7	•Evaluate economic impact of port closures by economic sector at local and national level •Evaluate effect of different mitigation strategies

U.S. Coast Guard Search and Rescue VA (cgSARVA)

Partners: USCG LANT 7, USCG D9, USCG D5, USCG HQ 771

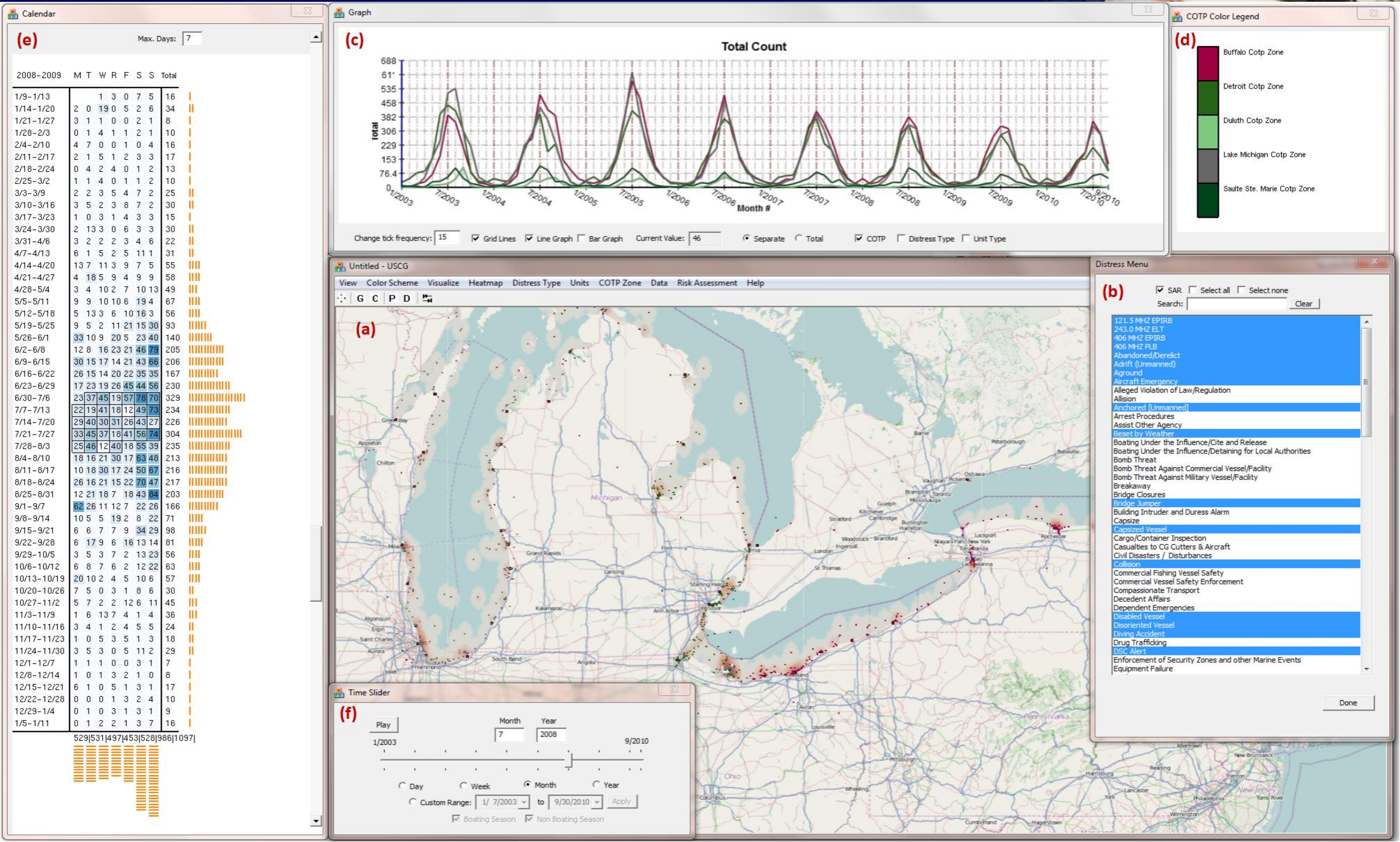




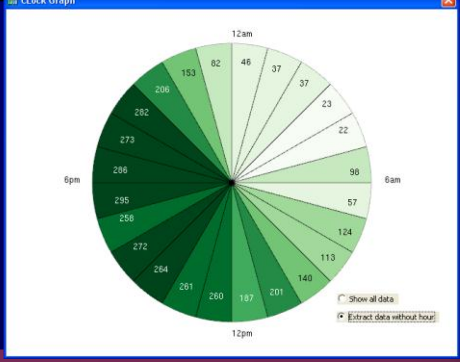
IMPACTS:

- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Hurricane Irene resource allocation decision based on cgSARVA analysis and visualization
- Informed Commandant's budget testimony and recommendations to Congress
- Key component of USCG D9 reallocation plan for 2011-2012 based on decreased budget
- Key component of Coastal Operations Allocation Suite of Tools (COAST) – USCG HQ

Example Screenshot: cgSARVA

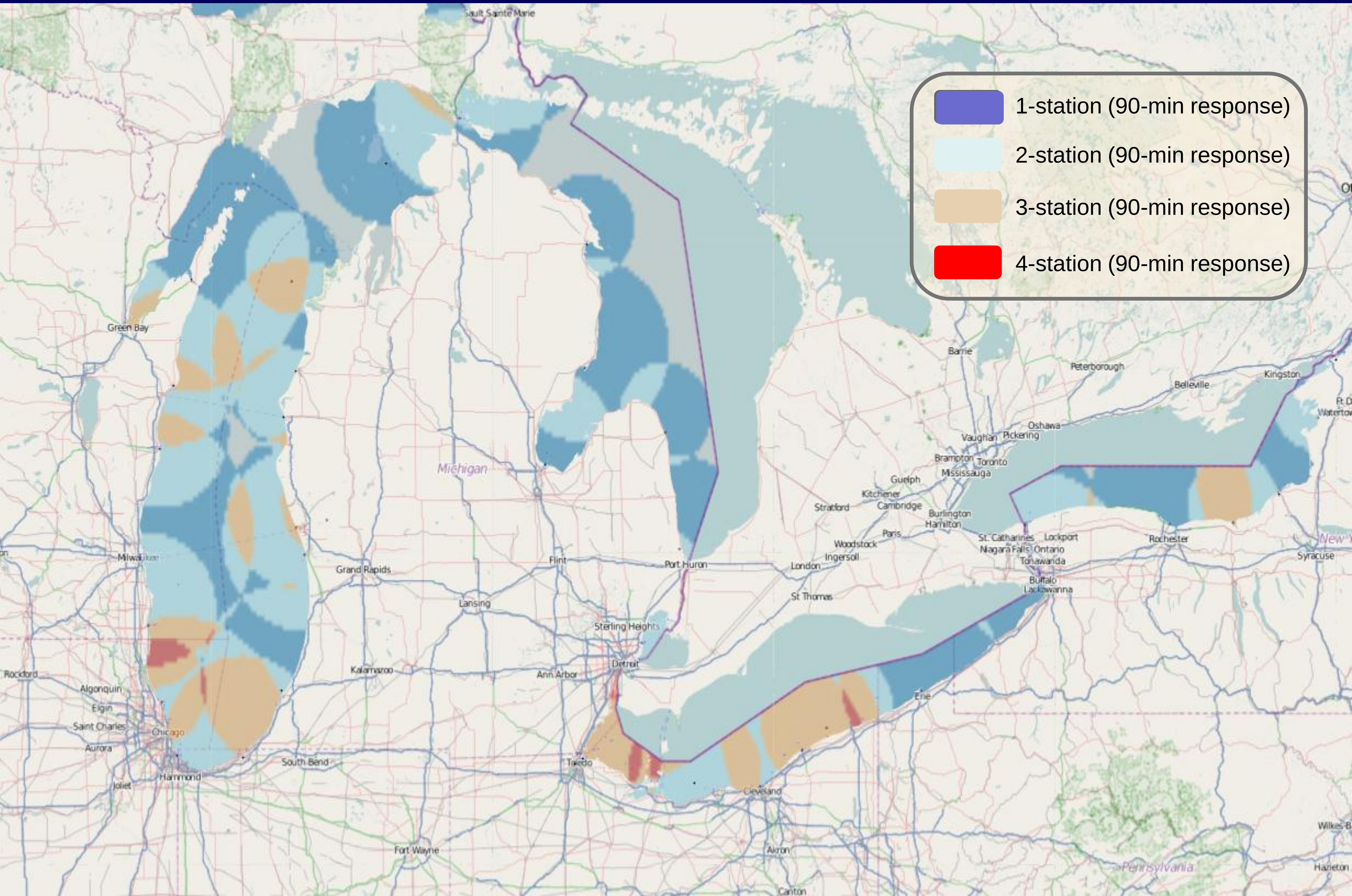


Risk-Based Allocations

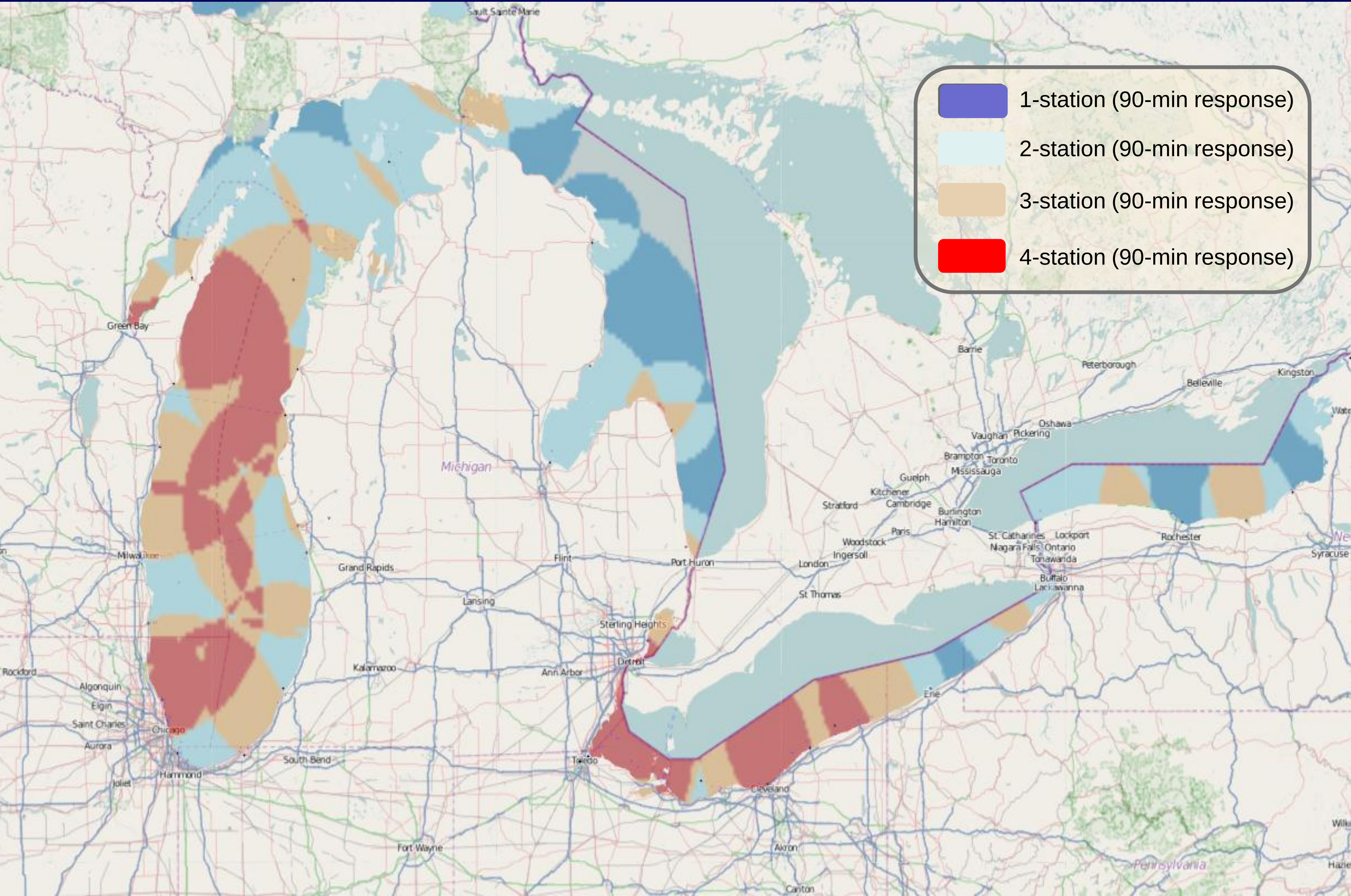


- Comparative visual analysis of mission cases/hours vs. staffing hours
- Comparative visualization of resources vs. risk
- Trend visual analytics
  - Increase/decrease in resource allocation
  - Increase/decrease in risk (total, mitigated, residual)
  - Increase/decrease in incidents
- Exploration of alternatives and effect on risk
- Predictive analytics based on historical data (STL and EWMA)

Response Efficiency – Current Assets



Response Efficiency – Potential Future Assets



For Further Information

www.VisualAnalytics-CCI.org

vaccine@purdue.edu

ebertd@purdue.edu

VACCINE

USCG PWCS PROTECT Project

Partners: USC CREATE, USCG RDC, USCG D1, USCG LANT, USCG PAC

Impact:

- Provided insight and analysis into historical security patrols in Boston
- Analyzed developing PROTECT model and provided pattern analysis to improve model during operational deployment in Boston and for improvements for deployment in NYC
- Developed end-user tool for evaluation and validation of patrol routes

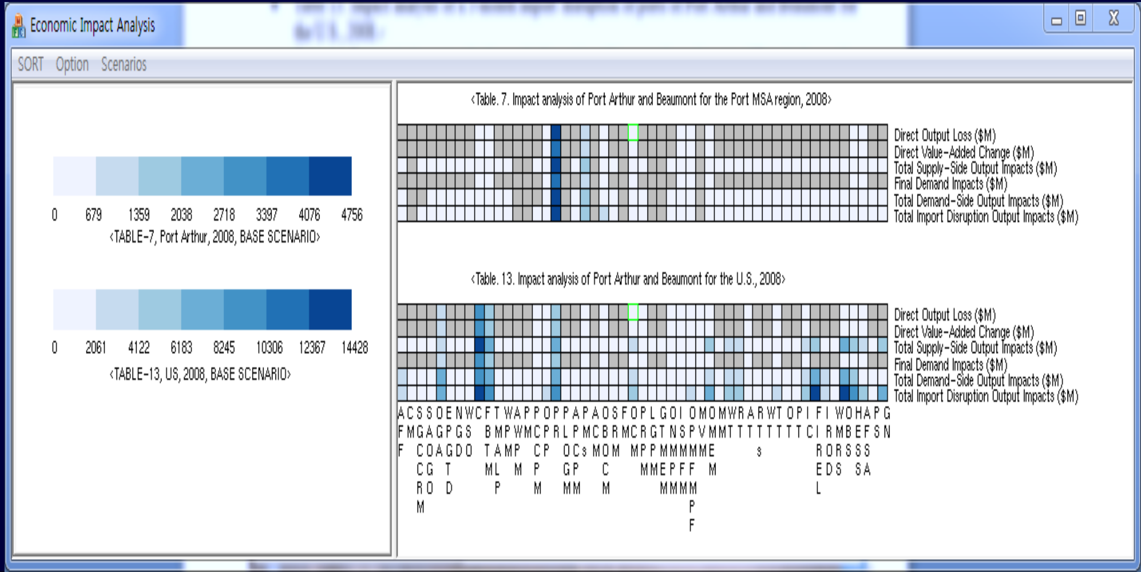
VACCINE

USCG Port Closure Economic Impact VA

Partners: USC CREATE, USCG RDC, USCG D7, USCG LANT

IMPACT:

- Provided tool for use analysis and planning for impact of port closure in Port Arthur, Tx
- Economic sector impact, local and national impact
- Impact and effectiveness of alternative mitigation strategies



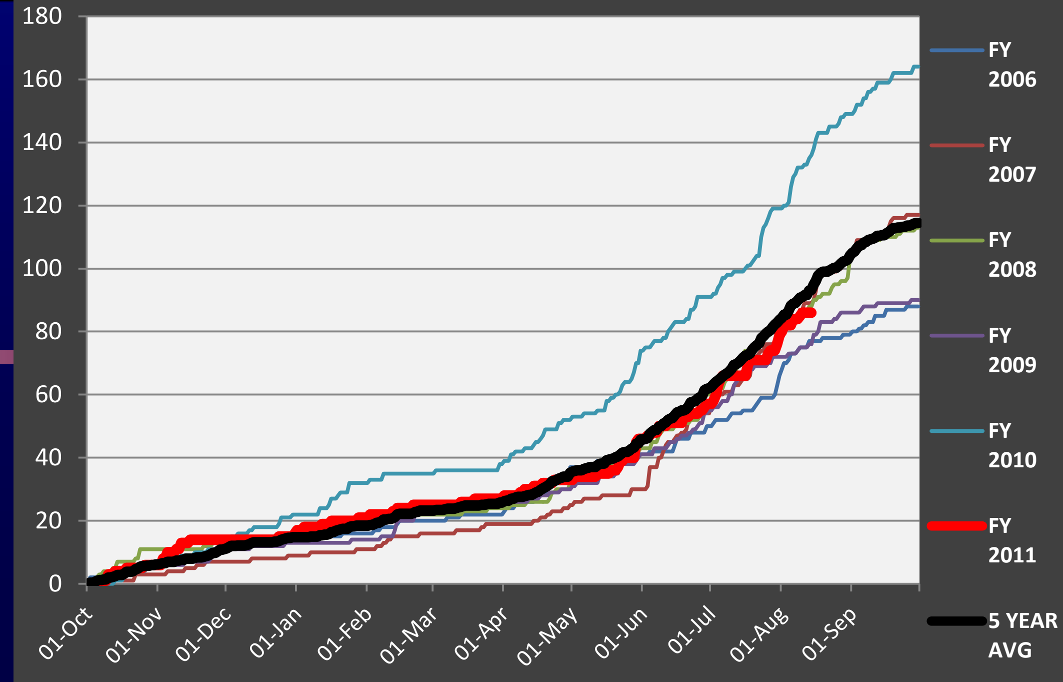
VACCINE

U.S. Coast Guard Swimmer Death Analysis

PI: Ebert, Maciejewski

End-User(s): USCG District 9

Delivered: May 2010



Impact:

- Analyzed spatial and temporal patterns of shore-based and boat-based swimmer deaths to understand death dramatic increase in D9 in Summer 2010
- Provided information and visualizations used for public information campaign 2011 and for patrols 2011
- Significant decrease in deaths in 2011

Findings:

- Swimmer deaths
  - August highest frequency
  - Late afternoon highest frequency
  - Lake Michigan (south and west shore) have high concentration
- Boating deaths
  - Fri, Sat, Sun account for almost all deaths
  - Mid July to Mid August have highest frequency (only 1 week significantly high)
- 2009-2010 from MISLE Data
  - Large increase on Mon, Thu, Fri, Sun
  - Early and late season increase

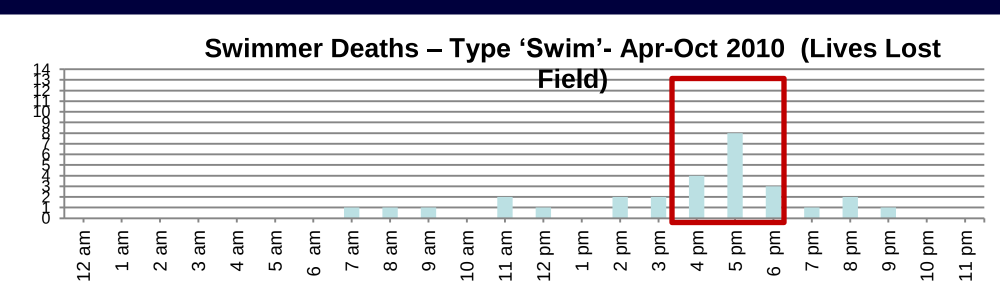


Table of Contents

VACCINE

Table of Contents

Fisher, B (2012) Opportunities in Data Visualization and Visual Analytics for Behavioral and Social Science Research. National Institutes of Health (NIH) Office of Behavioral and Social Science Research (OBSSR), National Cancer Institute, and National Institute of Drug Abuse. Bethesda, MD.

video online at http://videocast.nih.gov/summary.asp?Live=10949