

Empowering Discovery and Decision Making Through Visual Analytics

David S. Ebert

Purdue University

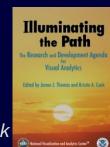
ebertd@purdue.edu



What We Do: Visual Analytics

Visual Analytics¹ is the science of analytical reasoning facilitated by interactive visual interfaces

Interactive visualization, data analysis, exploration, and decision making with human in the loop!







Big Data: Solution to Global Challenges?



- 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



Why VA For Big Data?

(Or 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 incorporates all relevant information
 Humans have extensive information and experience that never make it into the collected data
- 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



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 this process
- Adapt to available hardware for the problem



What Our Visual Analytic Solutions Offer

- **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



Our Teams' Benefits

- 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
- Innovative Fielded Solutions: We provide innovative visual analytic and scalable solutions to the extended homeland security community
- People and Partnerships

Interdisciplinary world-leading team of researchers and students

"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

VADM Robert Parker with VACCINE student researchers (cgSARVA, COAST, iOPAR)



Our Visual Analytics Process

(extended and adapted from Bertin, Normal, Tversky)

Process:

- Define problem / question
- Determine:
 - Workflow and interaction
 - Relevant data
 - Data characteristics (e.g., types, quality)
 - Abstraction/representation level
 - Natural scales
- Map to appropriate visual representation for task
- Utilize iterative refinement with user feedback from the start

Guiding Principles:

- Appropriateness principle
 - Neither more nor less
- Naturalness principle
 - Match with cognitive models
- Matching principle
 - Match the task, offer action affordances
- Principle of apprehension
 - Content should be accurately and easily perceived



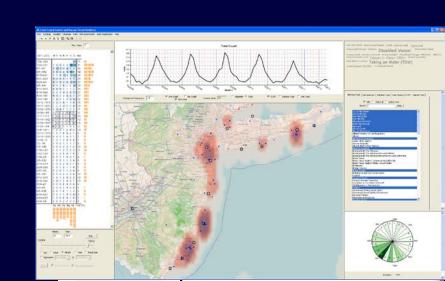
Example of Users With a Problem: Law Enforcement - Tasks and Uses

- Police Chief/Sheriff strategic goals
 - Resource allocation, return on investment
 - Yearly effectiveness (Are we safer?)
 - Emerging trends
- Police Captain short-term effectiveness
 - Force allocation for the day/week
 - Emerging trends and anomalies
- Officer improved situational awareness
 - What happened near me around the same time in the past?
 - What will make me better prepared when I respond to this call?
- Crime analyst/investigator connections
 - Correlations, dependencies, trends



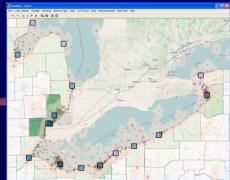
Research Area Overview and Example Projects

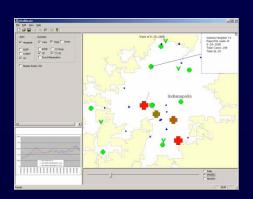
- Research applications areas
- Example projects
 - Public safety and law enforcement
 - Social media visual analytics
 - US Coast Guard
 - Additional projects

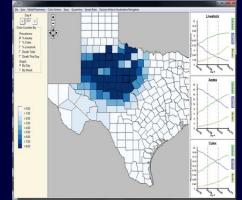


Public Safety & Health

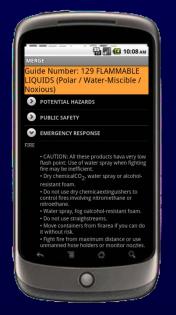
- Public safety visual analytics
 - cgSARVA
 - Coast Guard Search And Rescue Visual Analytics
 - MERGE
 - Mobile Emergency Response Guide
- Public health visual analytics
 - LAHVA
 - Linked Animal-Human Health Visual Analytics
 - RVF
 - Rift Valley Fever
 - Decision support environment for epidemic modeling and responses
 - PanVis
 - Pandemic influenza modeling and visualization tool
 - Cancer Care Engineering









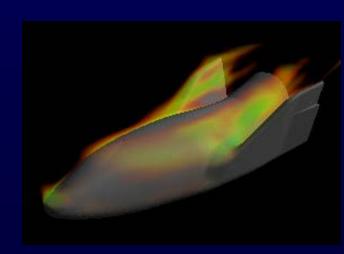


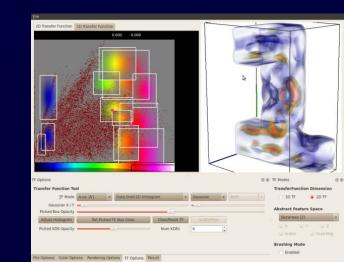




- Flow dynamics visualization (Purdue, TACC)
 - Providing insights on large flow data
 - Visualization linked with simulations
 - Innovative feature visualization
- Nanohub
 - Information-assisted data analysis and visualization of nanoelectronic models



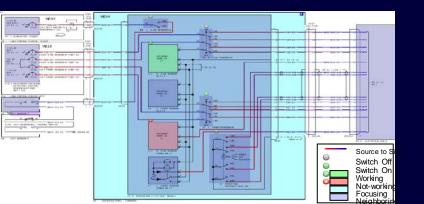






Scientific Discovery

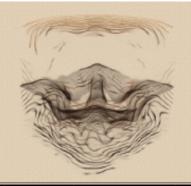
- Illustrative Visualization
 - Mechanical assemblies
 - Schematics and wiring diagrams
 - Illustrative flow visualization
 - Illustrative medical visualization
- Interactive multivariate atmospheric science data visual analytics



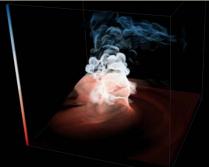






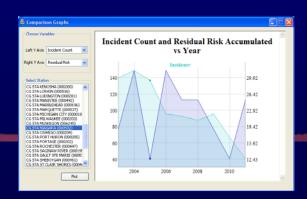


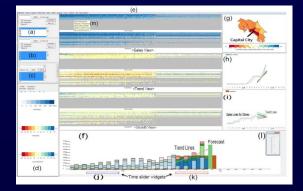




Business Visual Analytics

- 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



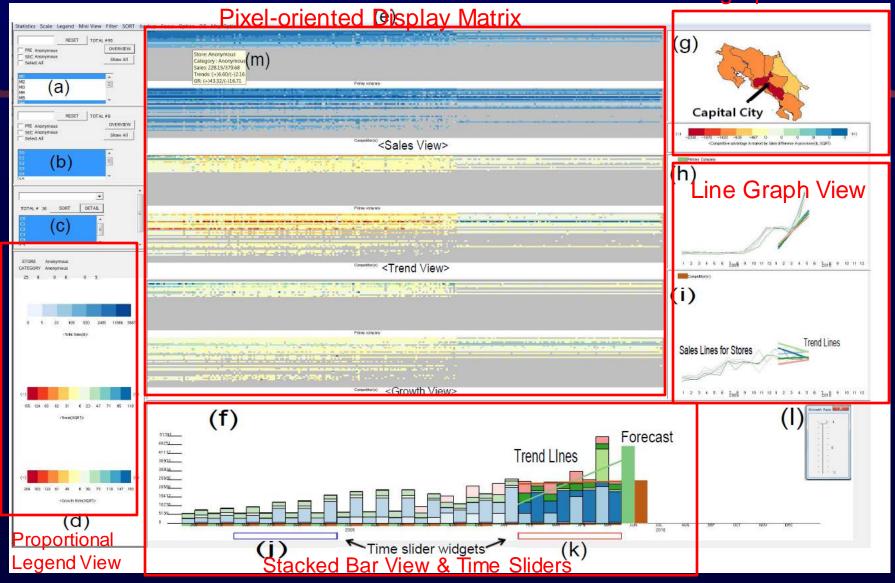






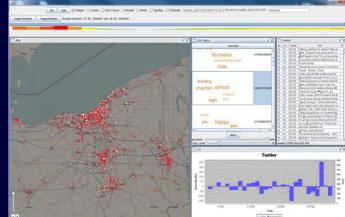
Market Analyzer

Geographical View

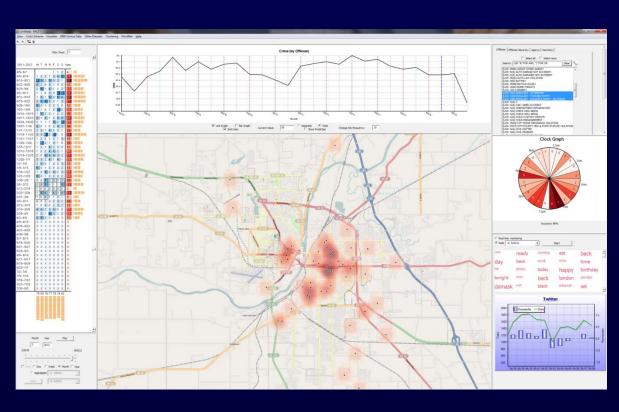


Visual Analytics Uses for Public Safety

- Risk visualization and analysis
- Predictive analytics
- Uncertain decision making
- Alternative evaluation and consequence investigation
- Trend analysis, clustering, anomaly detection
- Multisource, multimedia massive data
 - integration & analysis
- Purpose: Planning for resiliency, training, detection, investigation, response, recovery, remediation



Visual Analytics Law Enforcement Toolkit (i) VALET

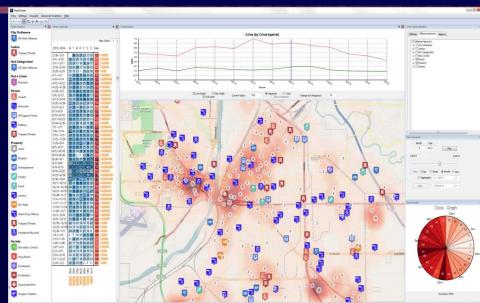




Visual Analytics Law Enforcement Toolkit (VALET, iVALET)

Impacts:

- In use to analyze crime patterns in Lafayette, Indiana and to connect strings of activities
- Mobile version being released to public (June 2014) for communitybased 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



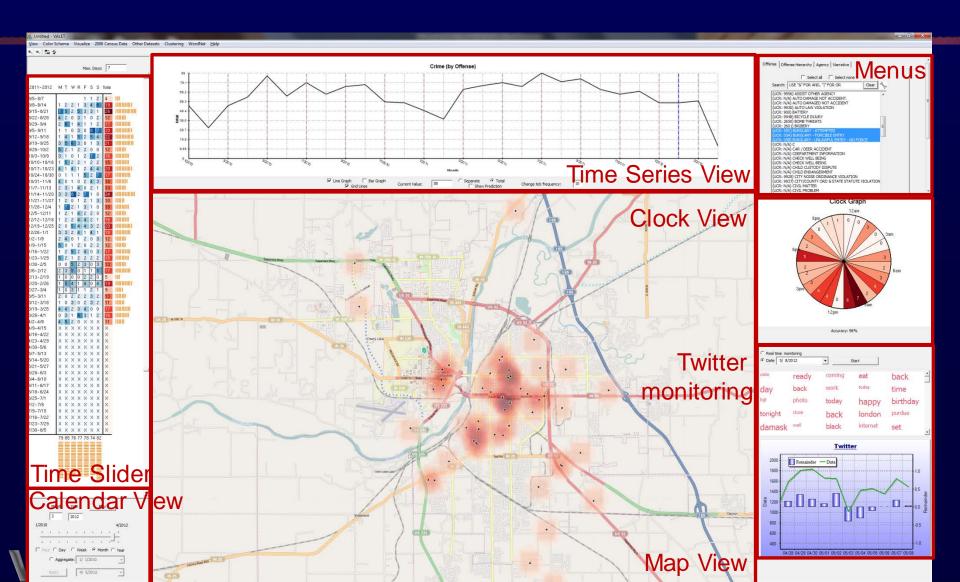
VALET delivered:

- Spring 2011: WL, Lafayette Police
- Fall 2013: NYPD <u>iVALET delivered</u>:
- October 2011: Purdue, WL Police

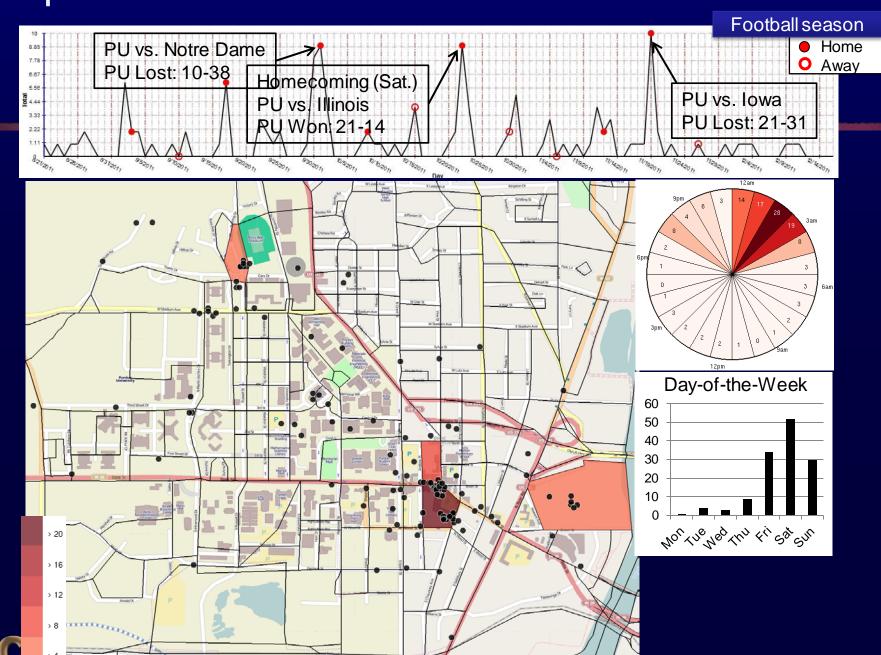




VALET Overview

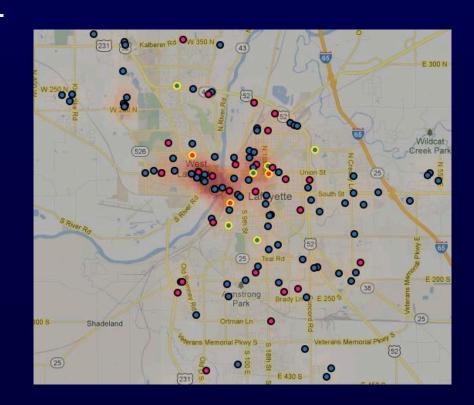


Example: Drunkenness / Public Intoxication



Top 10 Hot Incidents

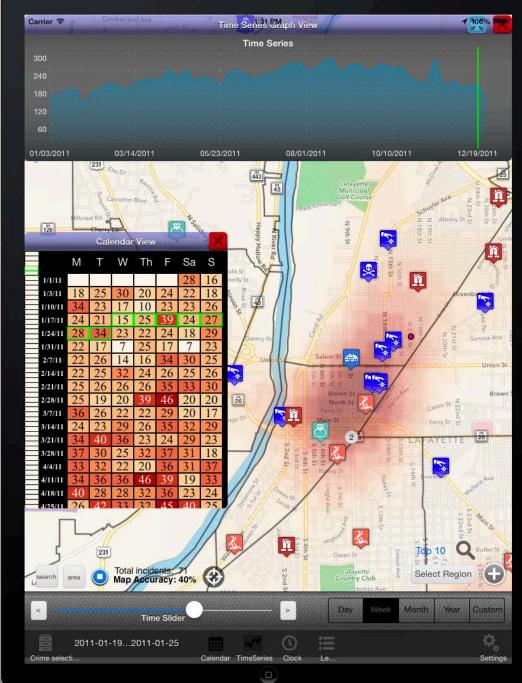
- Identify unusual localized highfrequency patterns of crimes (near repeats)
- Each data entry is checked for other crimes with similar properties within a 1 block radius of the incident location and a 14-day time period;
- Top 10 incidents with the most number of related incidents in this space-time window are highlighted





iVALET









lore criminal, fic and civil a on-the-go k assessment current tial + poral context analysis



Educational Use of VALET: Greensburg, KS City Rebuilding (2007 EF5 Tornado)

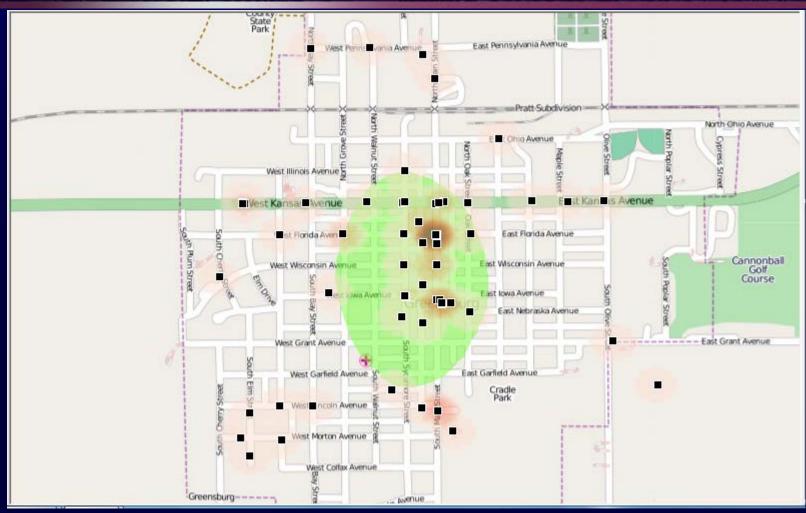
95% of town levelled

Groundbreaking data

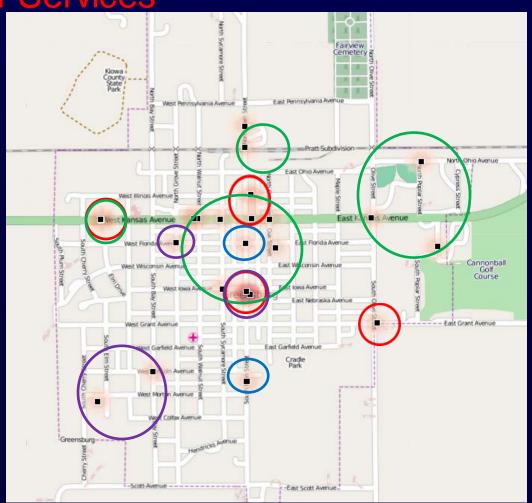




Statistical Analysis

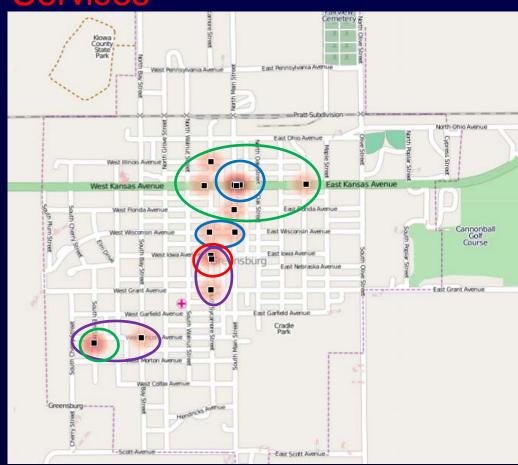


- Basic and Social Services
 - 7
- Business
 - 16
- Home
 - 7
- Third Space
 - 2



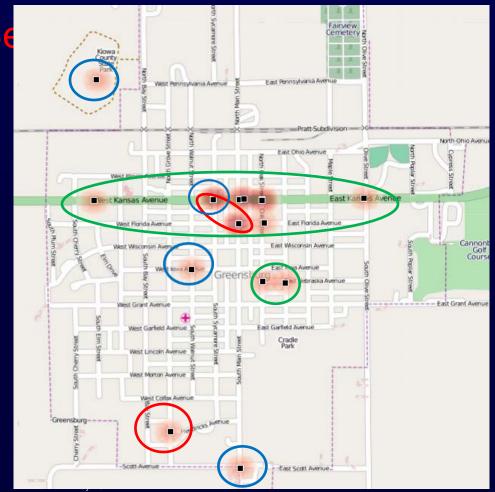


- Basic and Social Services
 - 1
- Business
 - 8
- Home
 - 3
- Third Space
 - 3

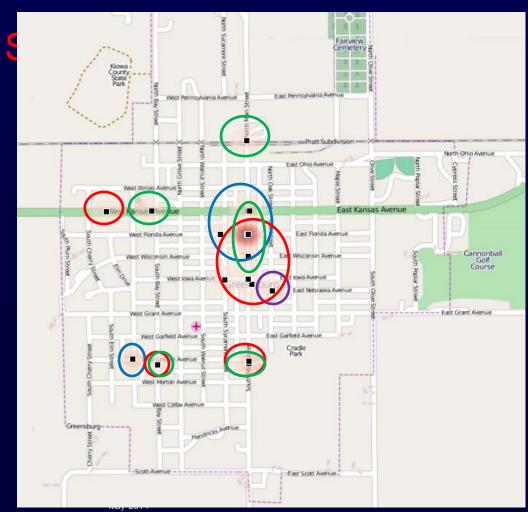




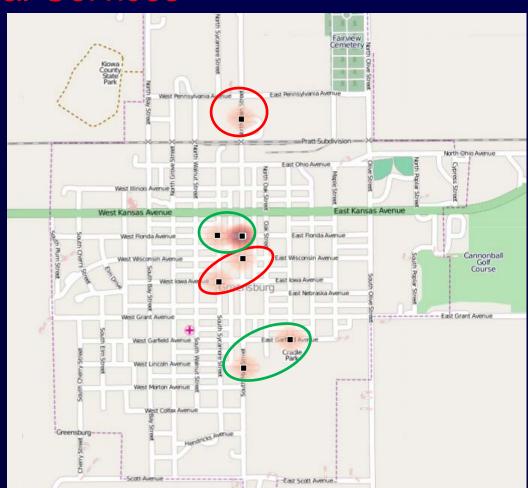
- Basic and Social Se
 - 3
- Business
 - 10
- Home
 - O
- Third Space
 - 4



- Basic and Social S
 - 14
- Business
 - **2**6
- Home
 - 1
- Third Space
 - 5

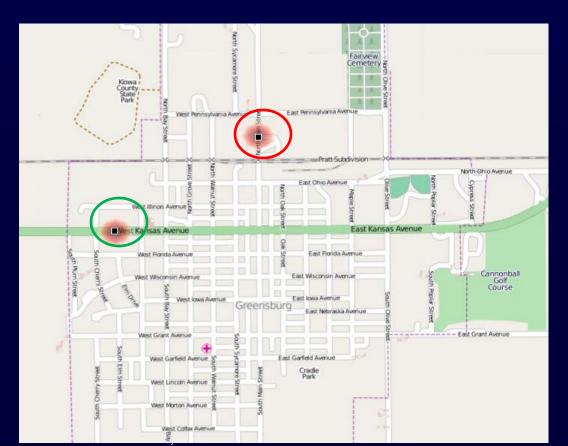


- Basic and Social Services
 - 4
- Business
 - 6
- Home
 - O
- Third Space
 - 0





- Basic and Social Services
 - 1
- Business
 - 1
- Home
 - 0
- Third Space
 - 0

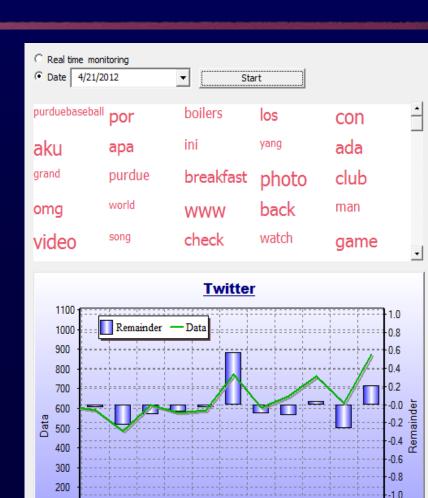




Social Media: Real-time Twitter Visual Analytics

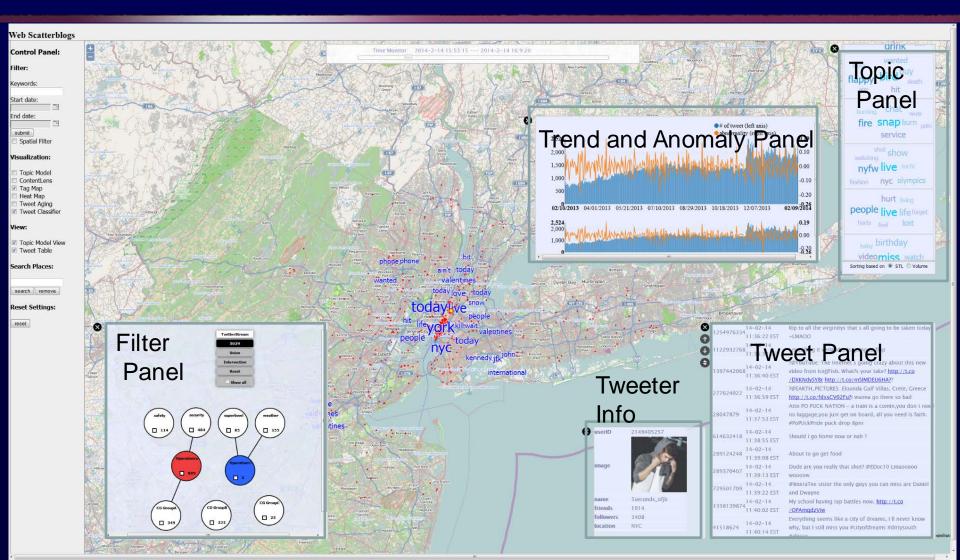
(Purdue, Stuttgart, Penn St.)

- Anomalous topic extraction using LDA and 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

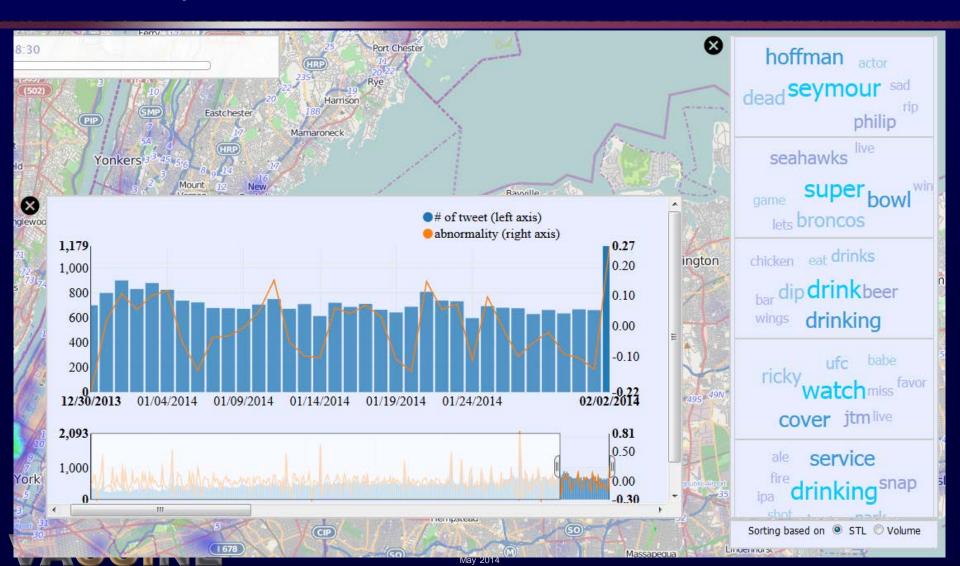




Social Media Visual Analytics: Scatterblogs



Anomaly Detection



Explosion Area in Boston Charlestown Cambridge. was 22.6% running happen explosion 12.7% finish bostonmarathon what and they 11.5% report everywhere okay Keywords that have awful Text **Tweets** Oh my god what just happened 10.2% Something happened at the end of the Boston Marathon. Something bad and there is a lot of chatter on Twitter. What's going on? now Multiple people are injured near the Boston Marathon finish line after two explosions. The #BostonMarathon has been stopped. two bombs just went off on boylston BREAKING NEWS: Two powerful explosions detonated in quick succession right next to the Boston Marathon finsh line this afternoon. What the FUCK was that



two sound

6.5%

Literally what the fuck get me out of here

@FRANCESCalciO I figured 3 people would get the joke

can someone tell me what that explosion was!? #boston #bostonmarathon

This is crazy i seen that blow up #bostonmarathon

@DTenenbaum my office right next to it

Back in Sept, @croon 1 solicited me for \$2000. He now has a music video with William Shatner. If you watch it (god forbid) keep that on mind,

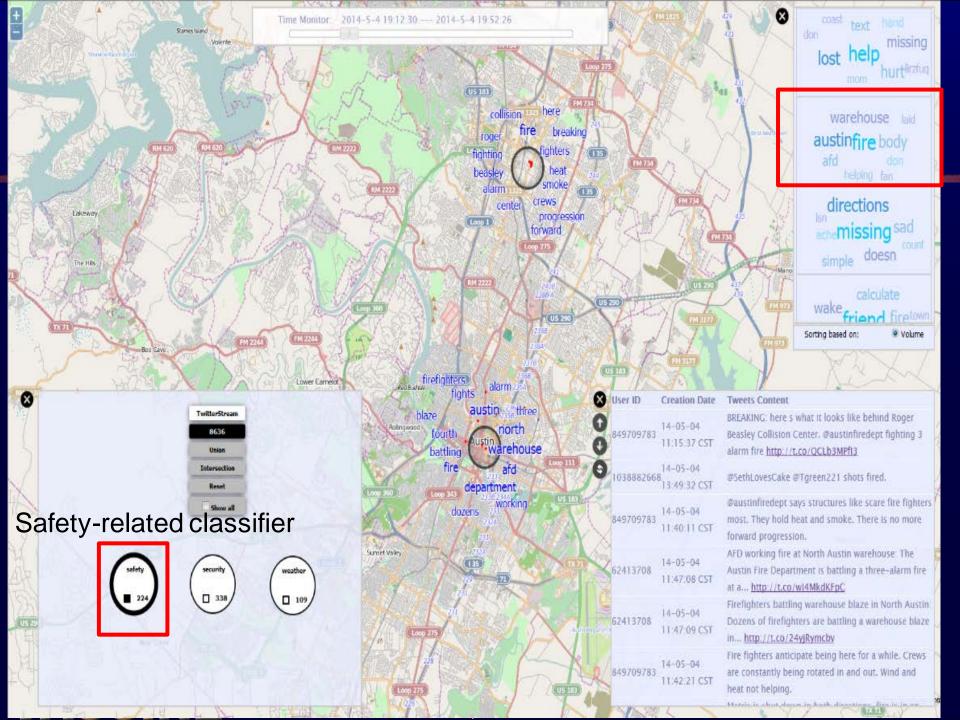
Two explosions just rocked the finish line of the Boston Marathon, Sirens galore, People running in fear, Wonder what happened,

First Response (Tweet & Picture)

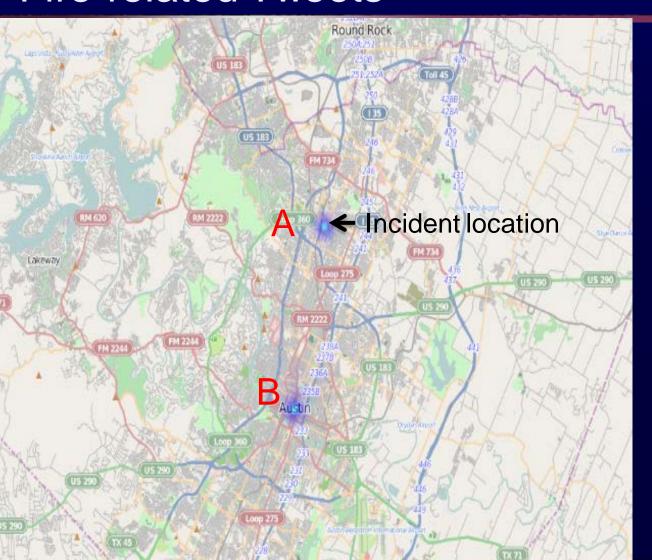


Austin, Tx Sunday Morning May 4, 2014

Safety Classifier: Fire Event Detected Location: Roger Beasley Collision Center

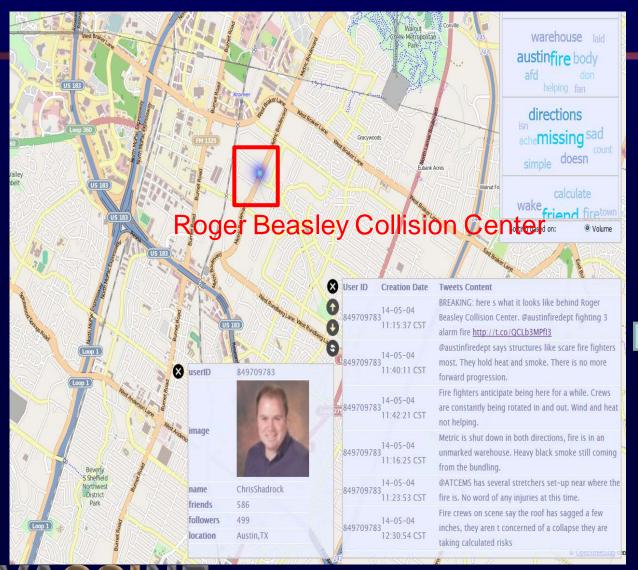


Spatial Distribution (Heat Map) and Keywords (Content Lens) of Fire-related Tweets



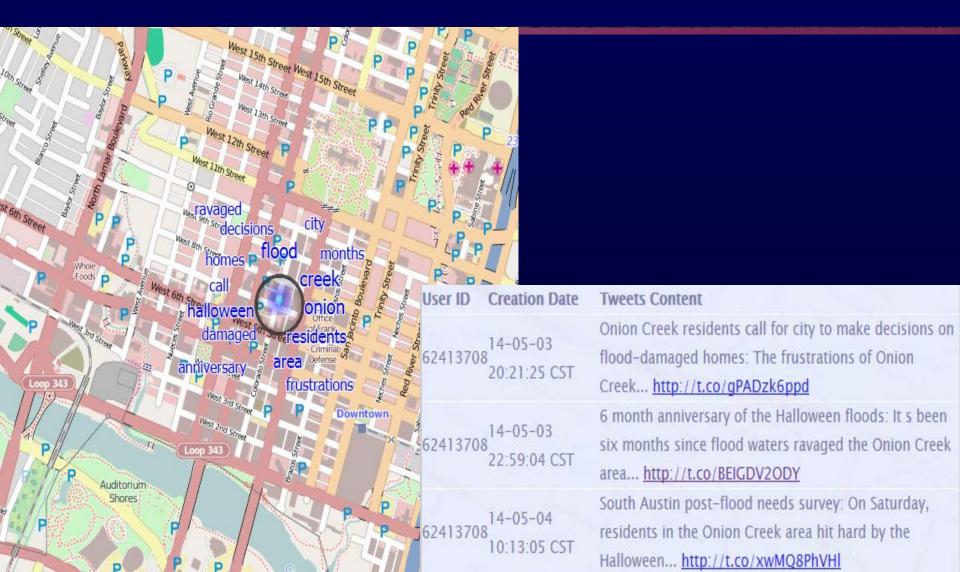


A Local Journalist Kept Updating The Event On The Spot





Also Detected This Weekend: Halloween Flood Related Tweets

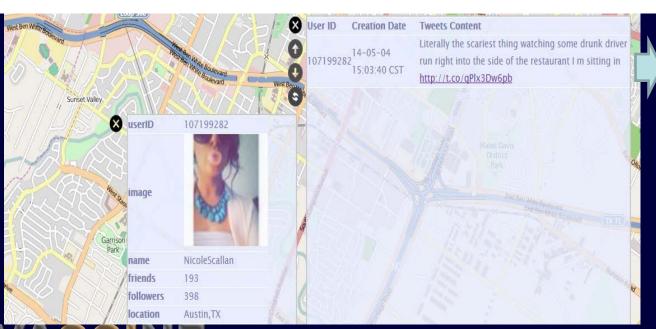


Also Detected This Weekend: Drunk Driving Accident





Tweet: Literally the scariest thing watching some drunk driver run right into the side of the restaurant I'm sitting in.





Cascading Critical Infrastructure Resiliency Modeling and Analytics (VASA)



- 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. Texas (TACC), 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, food supply

Purdue

UNCC

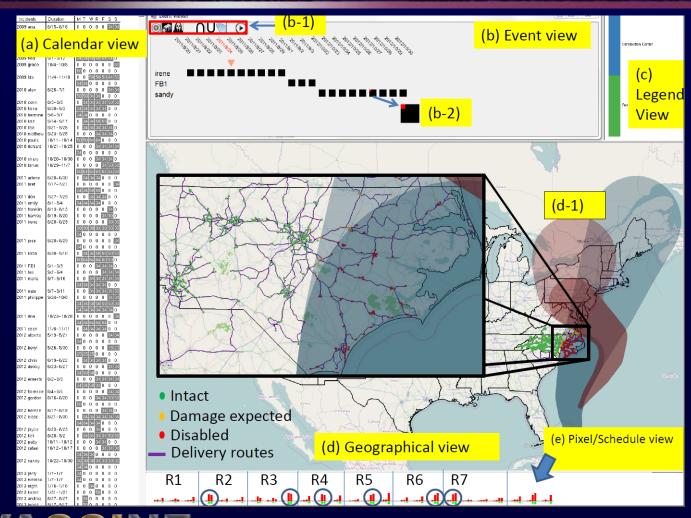


U Minn





System Interface



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



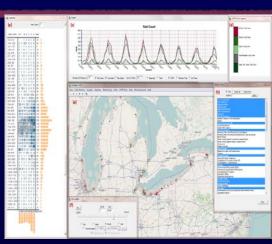
U.S. Coast Guard Search and Rescue VA

(cgSARVA)

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

IMPACTS:

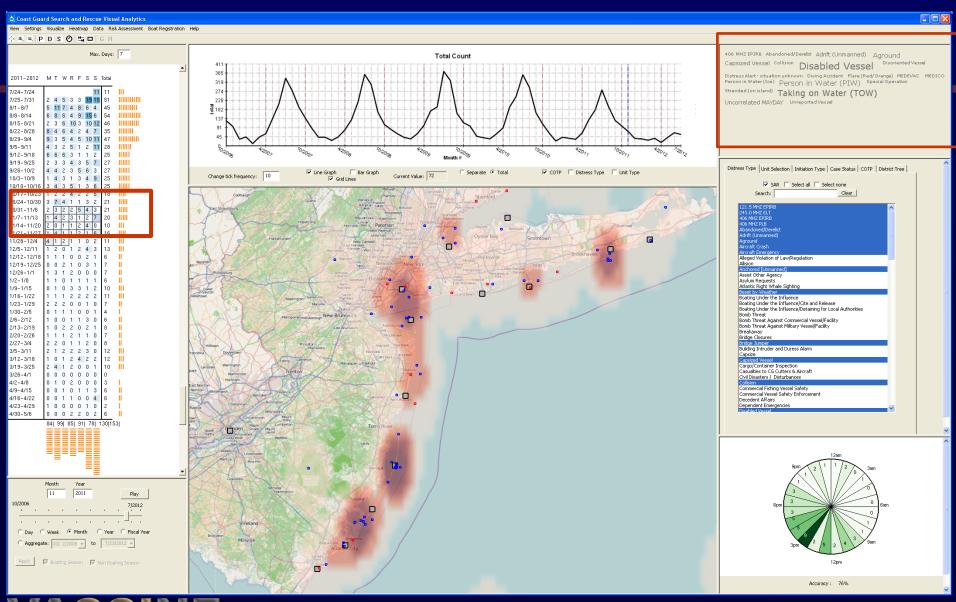
- Analyzed impact of CG auxiliary stations on search and rescue mission in Great Lakes
- Used for resource allocation for SAR
- Provided new insights to SAR mission
- Hurricanes Sandy and Irene resource allocation decisions based on cgSARVA analysis and visualization
- Informed Commandant's budget testimony to Congress
- Key component of USCG D9 reallocation plan for 2011-12
- Key component of Coastal Operations Allocation Suite of Tools (COAST) – USCG HQ



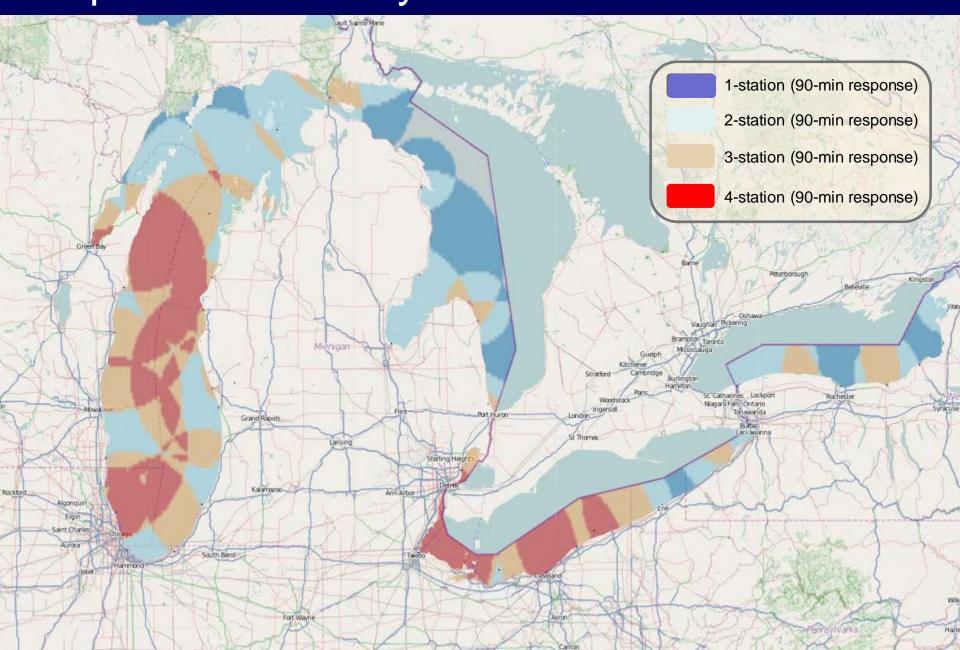


Example: Risks and Consequences From Sandy:

SAR Cases November 2011 NJ/NYC Area



Response Efficiency – Potential Future Assets

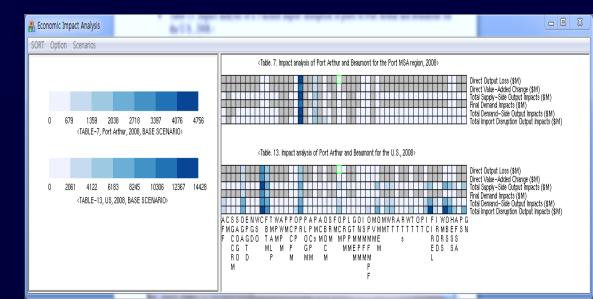


USCG Port Closure Economic Impact VA

Partners: USC CREATE, USCG RDC, USCG D7, USCG LANT

IMPACT:

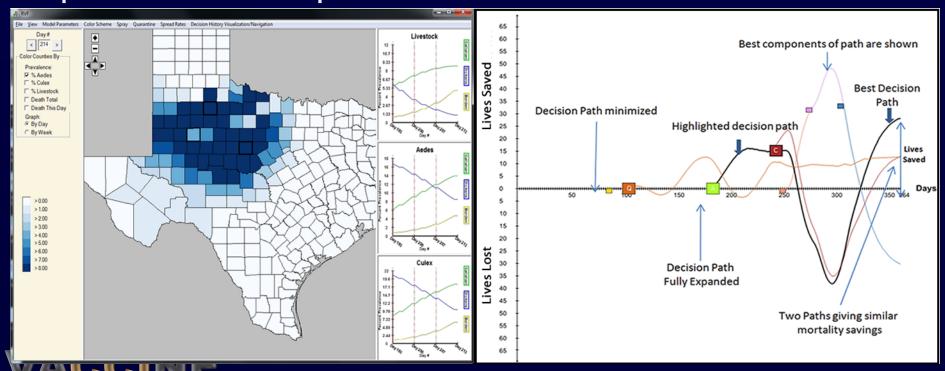
- Provided tool for use analysis and planning for impact of port closure in Port Arthur, Tx
- Economic sector impact, local and national impact
- Impact and effectiveness of alternative mitigation strategies





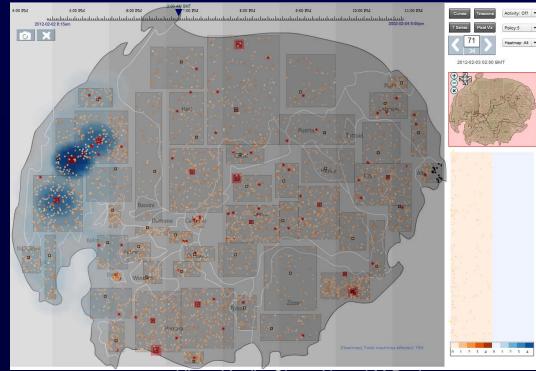
Complex Decision Making: Advanced Decision Support Tools (e.g., Rift Valley Fever)

- Integrated simulation models
- Explore alternative courses of action in decision space and real space

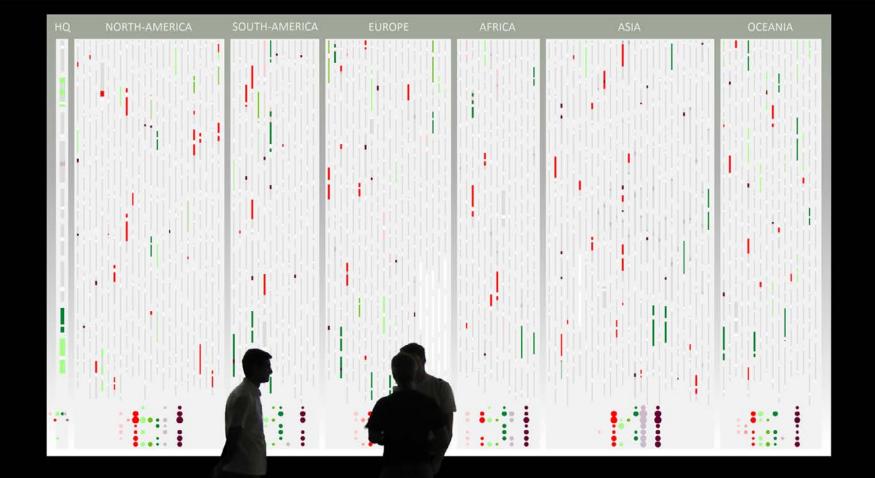


SemanticPrism: A Multi-aspect View of Large High-dimensional Data (Purdue University)

- VAST 2012 Mini Challenge 1 Award:
 Outstanding Integrated Analysis and Visualization
- Geo-Temporal
- Time-serial
- Pixel-based
- Semantic Zoom







2013 VAST Challenge MC2 Award Spring Rain Outstanding Creative Design

www.interactiondesign.us/vast2013/SpringRain

A Visual Analytics System with an Ambient Information Display +



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



