



RVAC
Regional Visualization
and Analytics Centers

North-East Visualization
and Analytics Center

www.geovista.psu.edu/NEVAC



Geo-temporal information discovery

Alan M. MacEachren – PI
GeoVISTA Center/Geography, Penn State
maceachren@psu.edu



NEVAC: Core goal and approach

- **Where** and **When** matter ... in relation to *threat and vulnerability analysis, situation assessment and mitigation, & disaster response and recovery*



NEVAC: Core goal and approach

- **Where** and **When** matter ... in relation to *threat and vulnerability analysis, situation assessment and mitigation, & disaster response and recovery*
- geo-temporal references in heterogeneous info are both the **object of analysis** (determine where & when) and the **glue to connect** otherwise independent data fragments



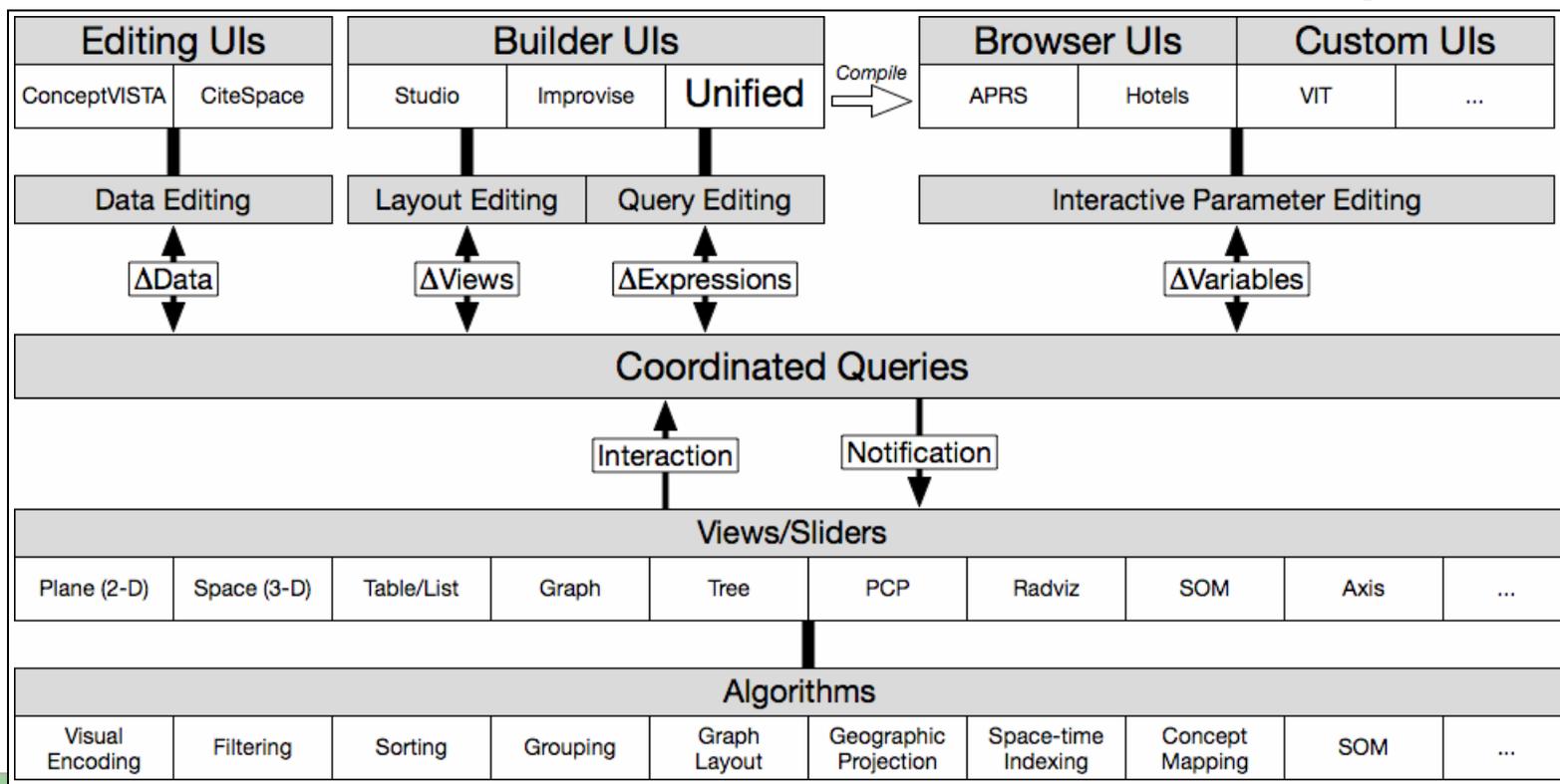
NEVAC: Core goal and approach

- **Where** and **When** matter ... in relation to *threat and vulnerability analysis, situation assessment and mitigation, & disaster response and recovery*
- geo-temporal references in heterogeneous info are both the **object of analysis** (determine where & when) and the **glue to connect** otherwise independent data fragments
- Goal: exploit complex, heterogeneous information that contains both *implicit* and *explicit* geospatial and temporal references ... by supporting an iterative process of *geo-temporally grounded information foraging, analysis, sensemaking, and decision-making*



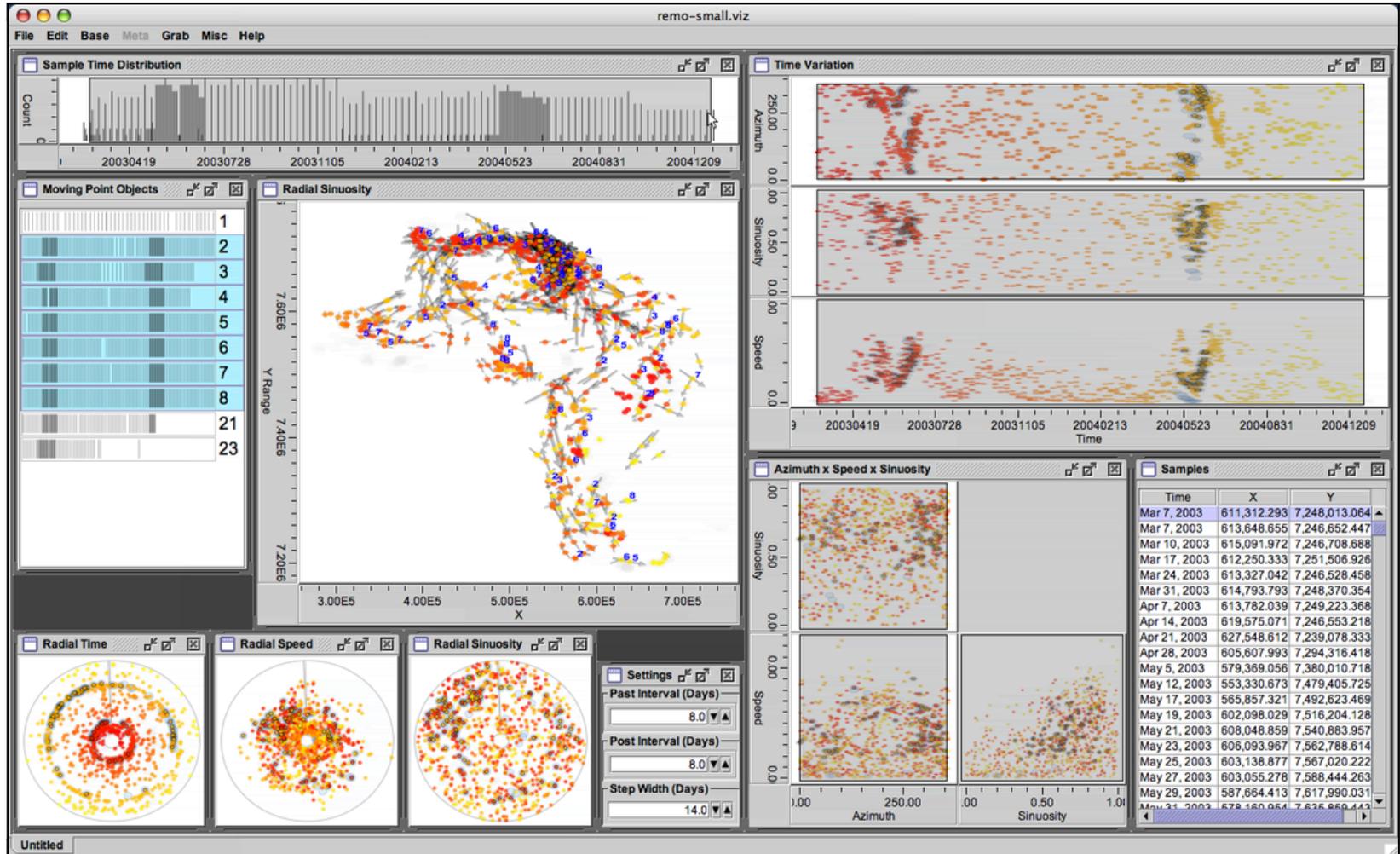
Toward An Integrated Geovisual Analytics System

- GeoVISTA *Studio*, ConceptVISTA, CiteSpace, FactXtractor, and others have sophisticated views and/or data processing algorithms
- Improvise has a flexible coordination architecture and query language, but a limited library of components
- Build a unified infrastructure that will have the advantages of all



Analysis of group behavior

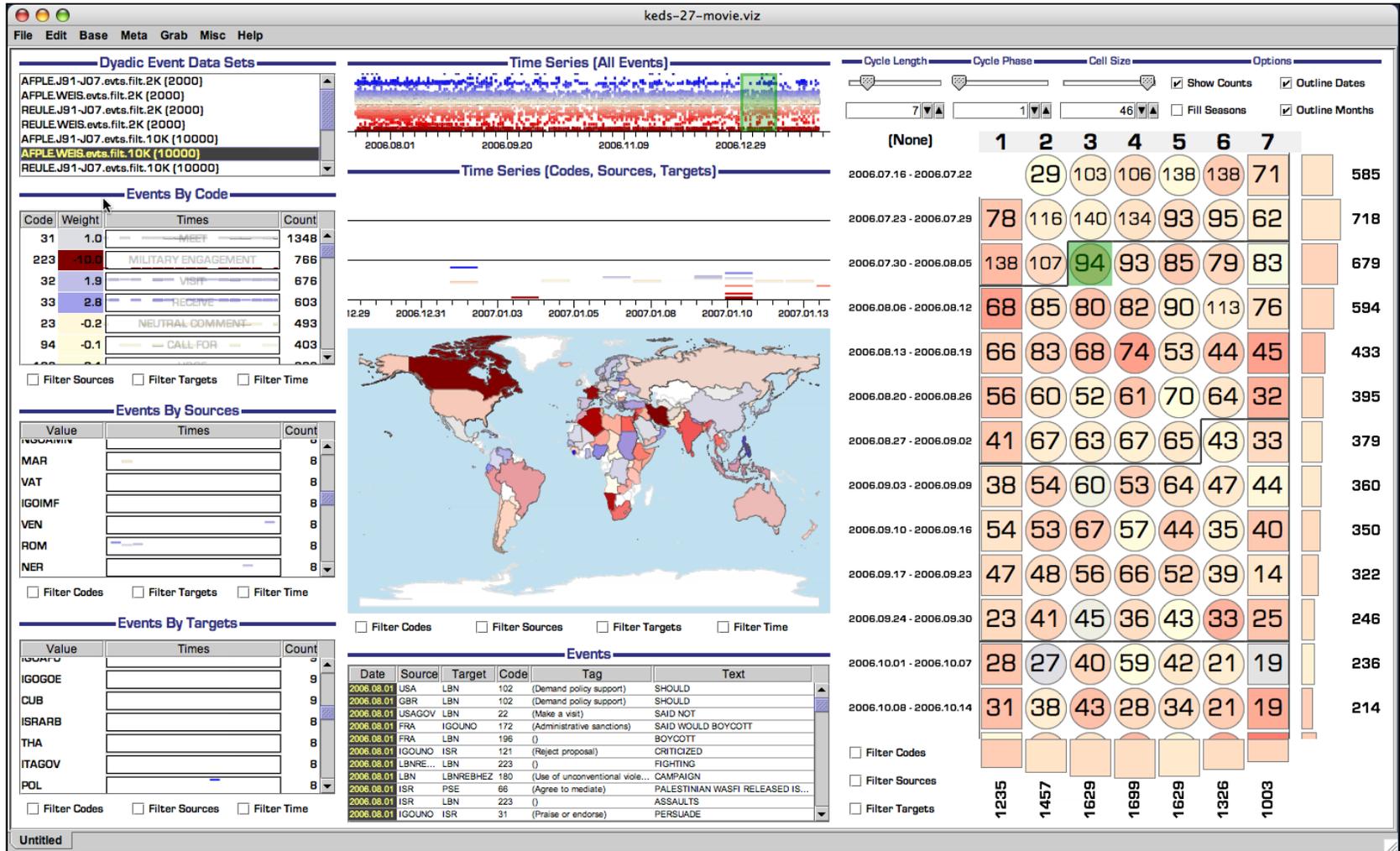
Problem: Given the dramatic increases of data depicting moving point objects, how can “relevant” group behaviors be recognized?



Chris Weaver & Patrick Laube © Alan M. MacEachren, 2007

Newswire-extracted international events between state-level actors

Problem: What geographic and temporal patterns of political activity can be discovered in international events reported by major news services?



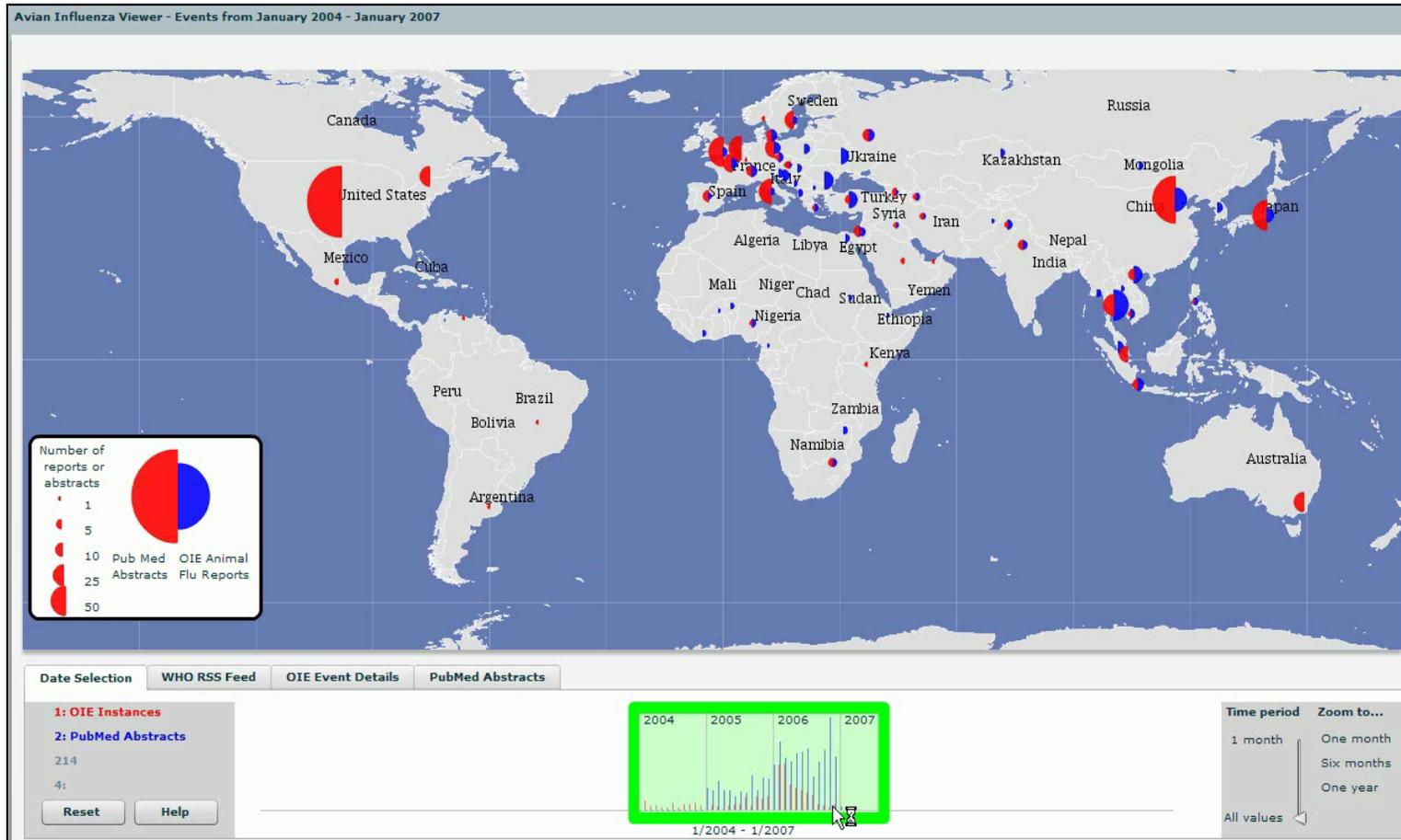
Using the Kansas Event Data System (KEDS) Chris Weaver & Phil Schrodt (U of Kansas)

© Alan M. MacEachren, 2007

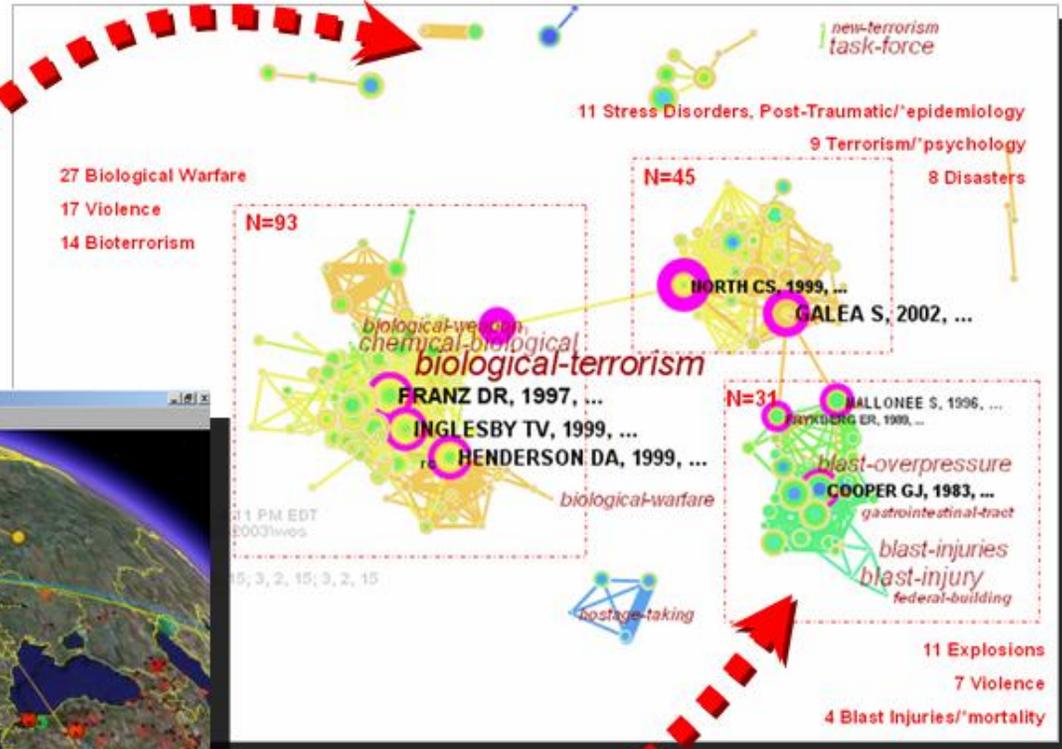
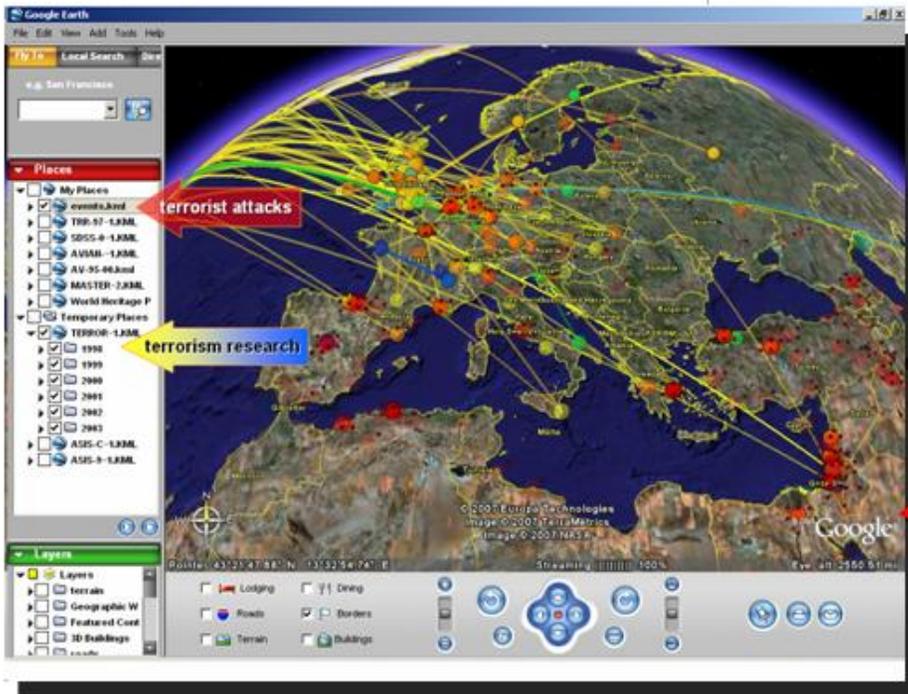
Approach: visual analysis using group-select-filter method

Linking events with science through web services: AvianFlu Viewer

Problem: How can we contextualize events and risks, geo-temporally, in evolving scientific knowledge?



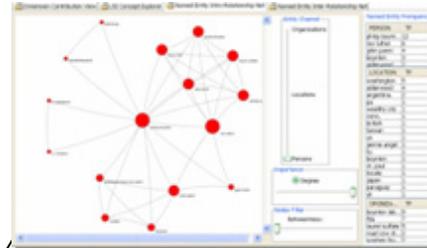
Linking events with science: Tracing Conceptual and Geospatial Diffusion of Knowledge



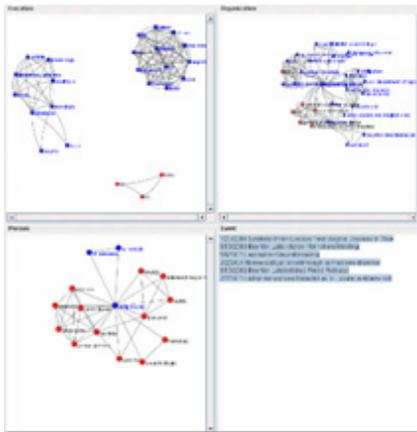
Chen, C., Zhu, W., Tomaszewski, B., & MacEachren, A. in press, Tracing Conceptual and Geospatial Diffusion of Knowledge. *HCI International 2007, Lecture Notes in Computer Science*, Beijing, P.R. China, 22-27 July, 2007

Storylines: named entity extraction + topic detection

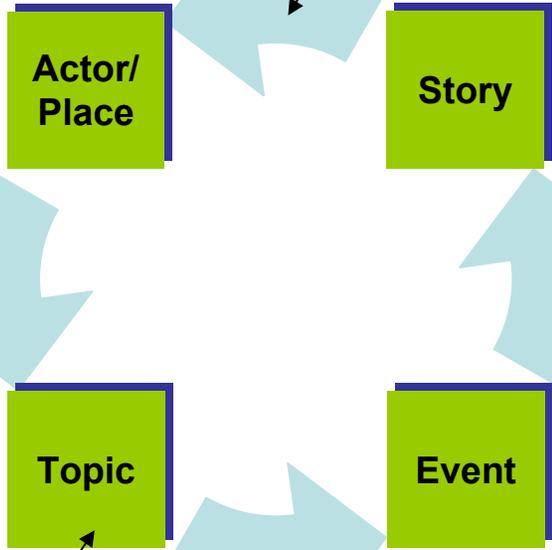
Problem: How can activities of interest be identified across a large set of independent intelligence reports?



Social Network Analysis



Relation Explorer
people, locations and events

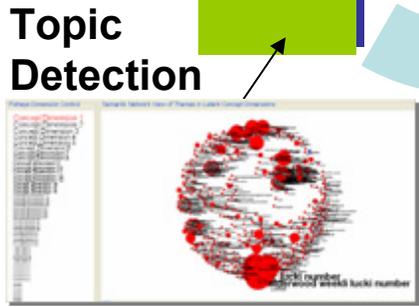


Story Formation



Unstructured Text

Natural Language Processing



Topic Detection



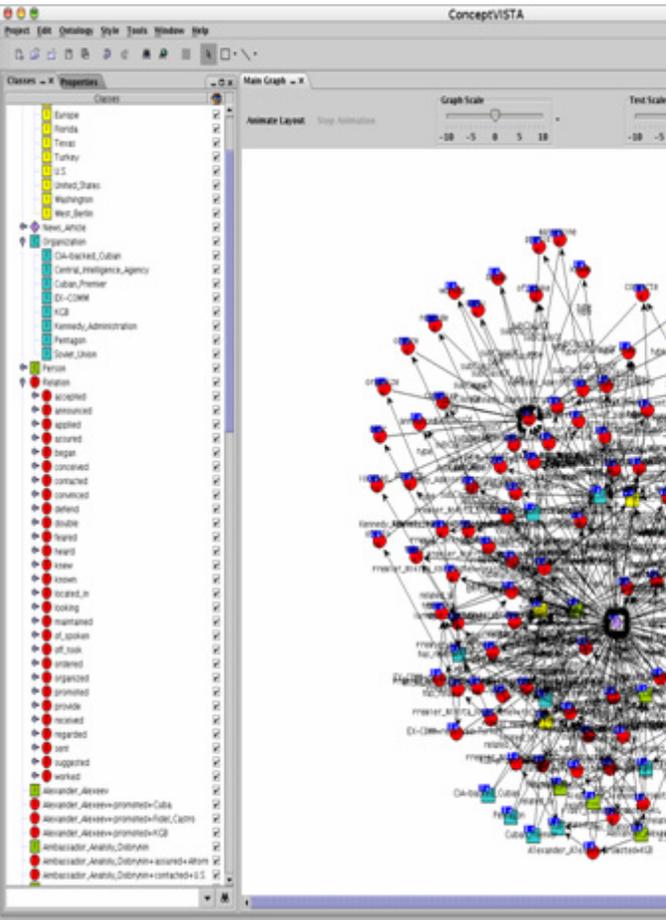
Time Line Information Filter

see poster

Weizhong Zhu, Chaomei Chen

FactXtractor: Named entity recognition (NER) + entity-relation extraction (RE)

Problem: Extract relevant geo-temporal (and other) info from text documents:
Named entities, locations, time; Common concepts



- Approach:
 - FactXtractor processes text documents using GATE and identifies entity relations using Stripped Dependency Tree kernels.
 - Output is a concept network formatted in OWL; which can be visualized with ConceptVista.
- Uniqueness: Multi-strategy information extraction
 - Machine-learning: high manual cost (training), high precision, low speed
 - Unsupervised rule-based: low cost, lower precision, higher speed
- Impact:
 - Can process large corpus of text documents
 - Visualize and convey info about named entities and extracted concepts and their relationships
 - Allow navigation to explore the relationships obtained from multiple heterogeneous documents

Presenjit Mitra & ChiChun Pan

FEMARepViz: geo-temporal situation monitoring

- FEMARepViz is an information extractor Web service for FEMA Situation Reports.
- FEMARepViz processes situation reports using FactXtractor and GeoTagger.
- Output is a Google Earth network link providing dynamic update / interactive visualization.



see poster

Presentij Mitra & ChiChun Pan



FEMARepViz: geo-temporal situation monitoring

The screenshot shows a Mozilla Firefox browser window with the URL `http://julian.mine.nu:8080/femaviz2.html#`. The page content includes:

FEMAViz is an information extractor Web service for the [FEMA Situation Reports](#). FEMAViz processes situation reports using [Texttractor](#) and [GeoTagger](#). The output is a [Google Earth](#) network link that provides dynamic updates and interactive visualization.

Please follow the steps below to use the online demo of FEMAViz.

Step 1:
pick a start date and an end date

Start Date:
End Date:

Step 2:
Click the following link and open google earth
[network link](#)

Step 3:
Change the start date and end date, the content in the google earth will be updated accordingly.

Below the instructions is a calendar for February 2007:

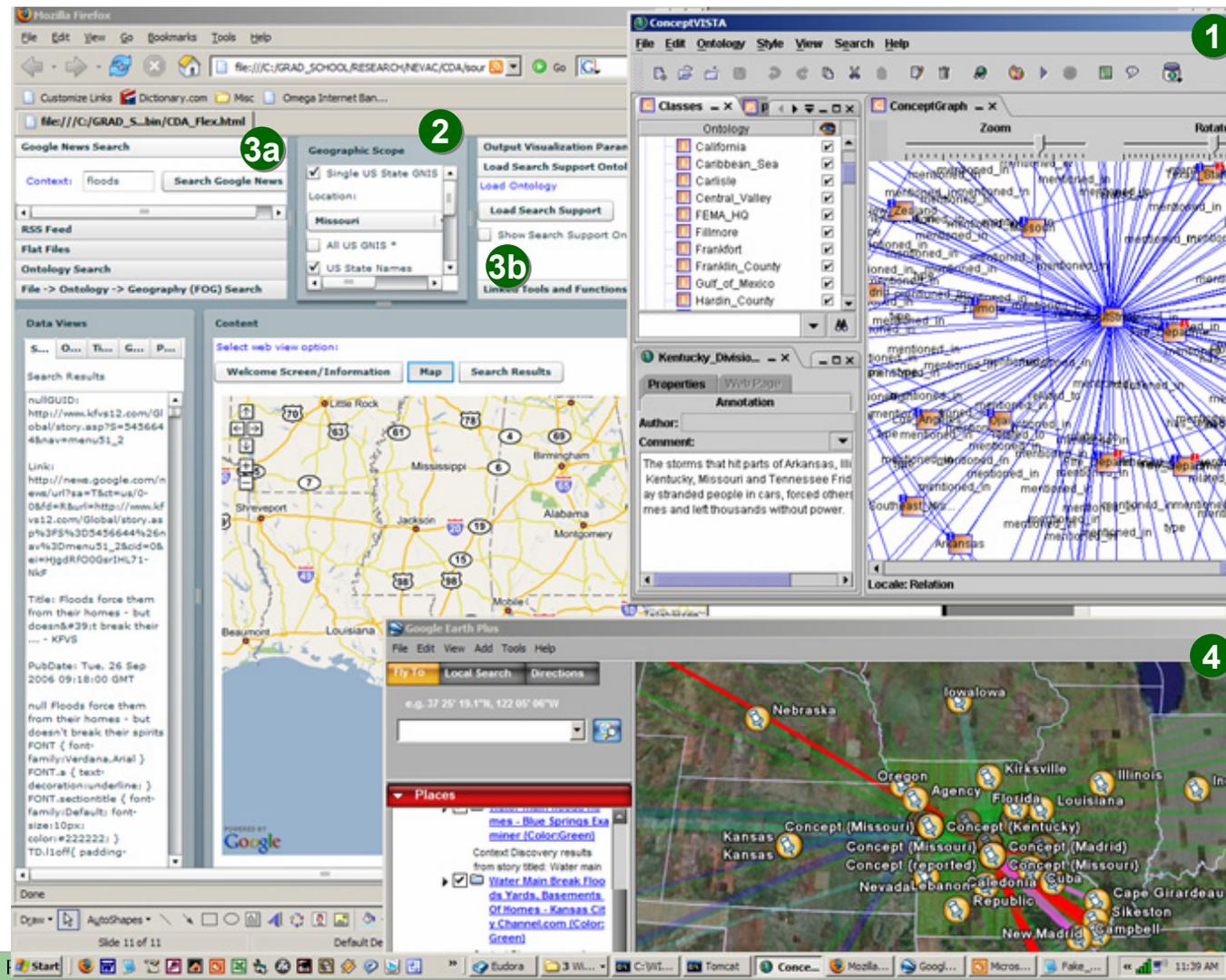
| February, 2007 | | | | | | | |
|----------------|-----|-----|-----|-----|-----|-----|-----|
| Today | | | | | | | |
| wk | Sun | Mon | Tue | Wed | Thu | Fri | Sat |
| 4 | | | | | 1 | 2 | 3 |
| 5 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 6 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 7 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| 8 | 25 | 26 | 27 | 28 | | | |
| Tue, Feb 13 | | | | | | | |

The Google Earth interface shows a map of the United States with a yellow outline. The status bar at the bottom displays coordinates: `Pointer 36°46'11.17" N 99°38'14.26" W`, `Streaming`, `100%`, and `Eye alt 3958.94 mi`. Copyright information for Europa Technologies, TerraMetrics, and NASA is also visible.



Problem: Help analysts retrieve relevant info and contextualize complex events

Approach: Concept Discovery Application (CDA) – knowledge enabled information foraging and geographical contextualization



CDA process

1. apply entity-relation extractor (ERE) to FEMA National Situation Report → generate: people, places, organizations
2. use ERE results to ID places of interest
3. supplement Google News query on floods, in located places, with ERE derived Owl (web ontology language) file
4. map results & relations

see poster

Problem: Help analysts retrieve relevant info and contextualize complex events

Approach: Concept Discovery Application (CDA) – knowledge enabled information foraging and geographical contextualization

Context Discovery Web Version 2.0 - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.geovista.psu.edu/context_discovery_web/CDA_Flex.html

Getting Started Latest Headlines

Google News Search

RSS Feed

Flat Files

Ontology Search

File -> Ontology -> Geography (FOG) Search

Select File and run FOG run Download OWL file after processing

Geographic Scope

- US State Names
- US County Names
- World Cities
- World Country Names

Output Visualization Parameters

- Spider Diagram
- Aggregated

Data Views

Sear... Onto... Time Map ... Previ...

- > <owl:Class rdf:ID="assisted" xmlns="http://www.geovista.psu.edu/ontology#"><owl:Class rdf:ID="stranded" xmlns="http://www.geovista.psu.edu/ontology#"><rdfs:subClassOf rdf:resource="#Reported" /></owl:Class>
- > <owl:Class rdf:ID="work_for" xmlns="http://www.geovista.psu.edu/ontology#"><rdfs:subClassOf rdf:resource="#Reported" /></owl:Class>
- > <owl:Class rdf:ID="reported" xmlns="http://www.geovista.psu.edu/ontology#"><rdfs:subClassOf rdf:resource="#Reported" /></owl:Class>
- > <owl:Class rdf:ID="swept" xmlns="http://www.geovista.psu.edu/ontology#"><rdfs:subClassOf rdf:resource="#Reported" /></owl:Class>

Content

Select web view option:

Welcome Screen/Information Map Search Results

Performing Search

Please Wait...

Transferring data from www.geovista.psu.edu...

Data-to-knowledge and back: Integrating knowledge management with visual analytics

Problem: How can an analyst assemble place-time-concept information needed to identify an unexpected disease event?

The screenshot shows a complex data visualization interface with several panels:

- Map of Pennsylvania:** A geographical map showing county boundaries and a color-coded overlay.
- Scatterplot of Demographics:** A plot showing the relationship between population density and other demographic variables.
- Disease Timeline:** A horizontal bar chart showing the progression of disease cases over time.
- Demography Table:** A table with columns for Counties, FIPS, County, pctpoor99, crowded2k, MDration95, and hosp95. It lists data for various counties including York, Philadelphia, and Delaware.
- Resources Panel:** A text window displaying information about West Nile Virus, including its first isolation in Uganda in 1937 and its recognition as a cause of severe human meningoencephalitis in Israel in 1957. It also includes a photograph of a mosquito.
- Concept Map of Emerging Disease Understanding:** A network diagram showing relationships between various diseases and concepts like 'Vector_Borne_Infectious_Disease', 'West_Nile_Virus', 'Yellow_fever', 'AIDS', etc.
- Case Details Table:** A table with columns for County, Age, Gender, and Date, listing specific cases such as 'York 83 Male 0' and 'Philadel... 54 Female 1'.
- Ontological Resources:** A list of classes and properties from an ontology, such as 'AIDS', 'Yellow_fever', and 'Viral_pneumonia'.

Callouts on the left side of the interface identify the following features:

- Counties in Pennsylvania
- Map of Pennsylvania
- Variable Selection Tool
- Scatterplot of Demographics
- Disease Timeline
- Demography Table

Callouts on the right side of the interface identify the following features:

- Ontological Resources
- Concept map of Emerging Disease Understanding
- Case Details
- Disease Case Table

Approach: visual group, select, and filter of data linked to disease-focused concept maps, ontologies, and web resources



Summary

- place & time are critical components of knowledge relevant to homeland security – and provide the glue to connect diverse information fragments
- geovisual analytics combines visual interfaces, computational methods, and knowledge management strategies to support exploitation of both explicit and implicit place & time information ... and to enable analytical reasoning with that information

contact:

Alan M. MacEachren
maceachren@psu.edu

