



NEVAC R&D – GeoVisual Analytics

June, 2008

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









North-East Visualization and Analytics Center

www.geovista.psu.edu/NEVAC



GeoDiscoverer — Searching social networks + geo info in web and text



ConceptVista — Visualizing, manipulating, and managing knowledge (perspectives about chemical and biological weapons)



Health GeoJunction — Web portal analysis of placetime-concept info extracted from documents

GeoVisual Analytics Tools for:

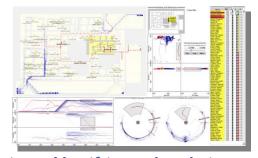
Public Health
Facility Security
Study of Terrorism
Intelligence Analysis
Emergency Management



GeoViz Toolkit — Rapid construction of custom analysis tools (crime rates in PA)



Improvise — Querying multidimensional data (world events in newswire reports)



Improvise — Identifying and analyzing movement behavior (synthetic health facility evacuation)

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







NEVAC Goal: geo-visual analytics for public health, intel analysis, study of terrorism, & emergency mgt

- Place & Time are Fundamental ... to all components of Homeland Security
- NEVAC foci: R&D to assemble time-relevant info about places at risk and to use the information effectively to decrease threats and increase resilience ... three topics highlighted:
 - finding info quickly: browse, query, & filter
 - foraging & contextualizing info: retrieve info & build place-time context → achieve situation awareness
 - analyzing place-time info: get answers about what, where, when, who, how, and why



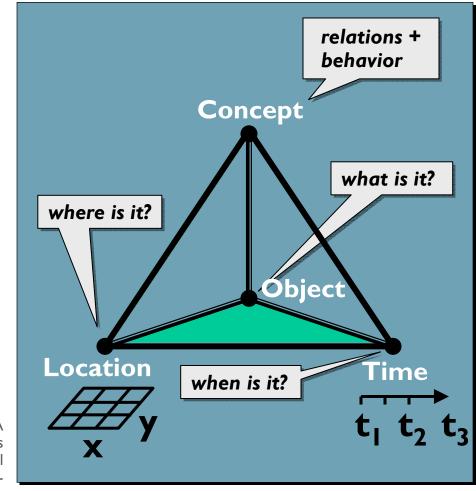


A multi-representation info-model





Mennis, J. L., **D. J. Peuquet**, & L. J. Qian. 2000. 'A conceptual framework for incorporating cognitive principles into geographical database representation', International Journal of Geographical Information Science, 14(6), pp. 501-520.







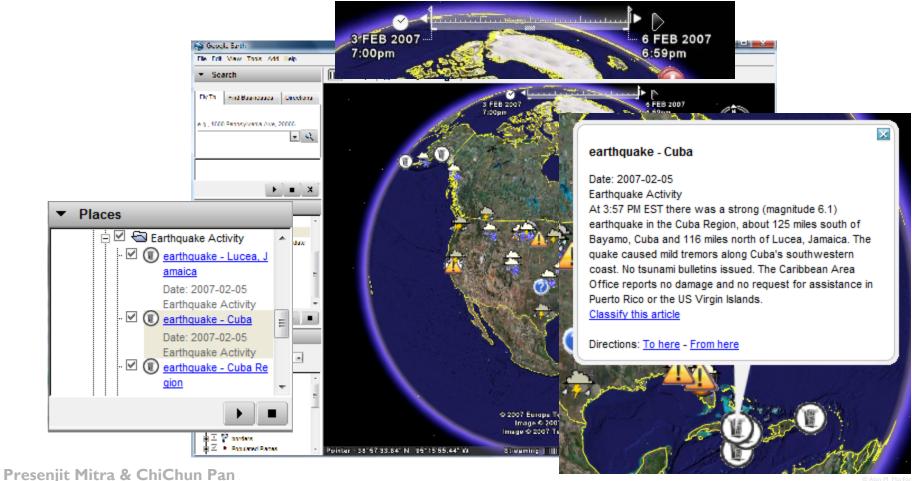
Foraging and contextualizing info → Situation Awareness

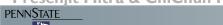
- knowledge enabled information foraging and geo-historical contextualization
- FEMARepViz:



FEMARepViz

example info review session

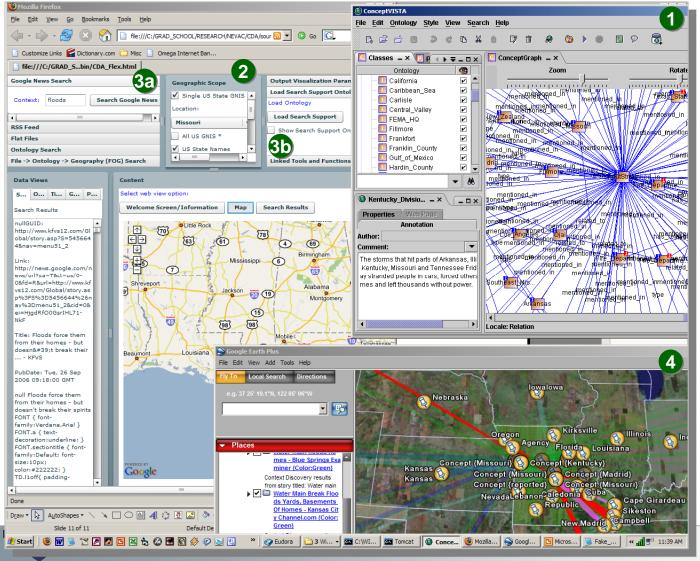






Problem: Help analysts retrieve relevant info and contextualize complex events

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



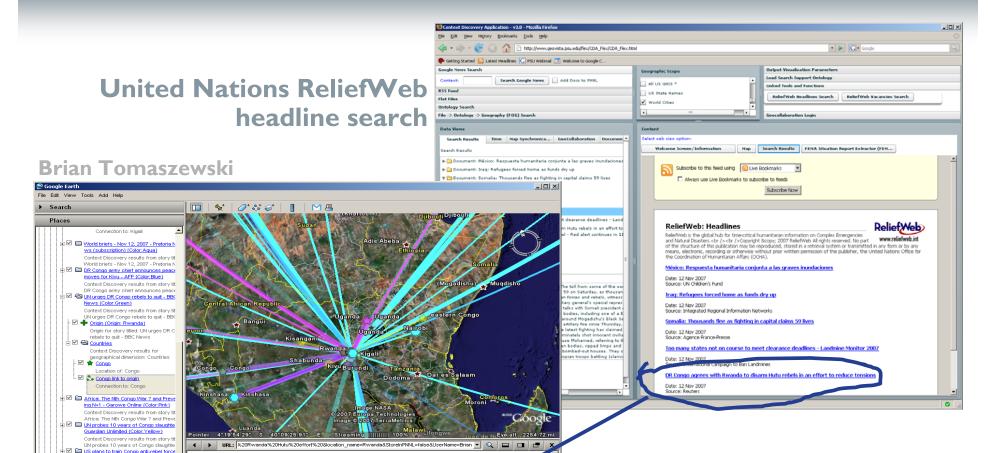
CDA process

- I. 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

Brian Tomaszewski, Alan MacEachren

Concept Discovery Application (CDA) – knowledge enabled information foraging and geo-histrical contextualization



ReliefWeb headline followup query in Google News using nouns from RW headline as search terms



diplomat - Reuters South Africa (Color.)

US plans to train Congo anti-rebel forc

Context Discovery results from story tit

Context Discovery results from story tit Africa: Making Sense of the DRC -

Congo army readies to take on Mai Mai

Context Discovery Application (CDA) Output Results

Content: Rwanda: Congo Agrees to FDLR Disarmament Plan This Page This Page on this article (Kigali) 12 November 2007 James Munyaneza Nairobi Rwanda and the Democratic Republic of Congo (DRC) have reached an agreement that will see the latter disarm and

espel genocidal forces, grouped in what is known as the Democratic Forces for the Liberation of Rwanda (FDLR). The deal was reached on Friday between Foreign Affairs ministers from the two neighbours in Nairobi, Kenya, at a ceremony that was facilitated by the UN and witnessed by the US and the European Union. The Government of the Democratic Republic of Congo commits to launch military

operations, as a matter of urgency, to dismantle the ex-FAR Internhamme as a genocidal military organisation in the DRC | a joint communique signed by among others, Rwanda's Foreign Affairs minister Dr Charles Munigande and Congolese State Minister of Foreign Affairs, Antipas Mbusa Nyamwisi, reads in part. DRC agreed to prepare a detailed plan to disamm the militia, while the UN Mission in

Rwanda: Congo Agrees to FDLR Disarmament Plan - AllAfrica.com

Publish Date: Mon, 12 Nov 2007 04:14:53 GMT



computational methods with visual interface support applied to semi/unstructured text

FEMARepViz

GeoJunction

Jerusalem, Israel



Tags related to Jerusalem, Israel

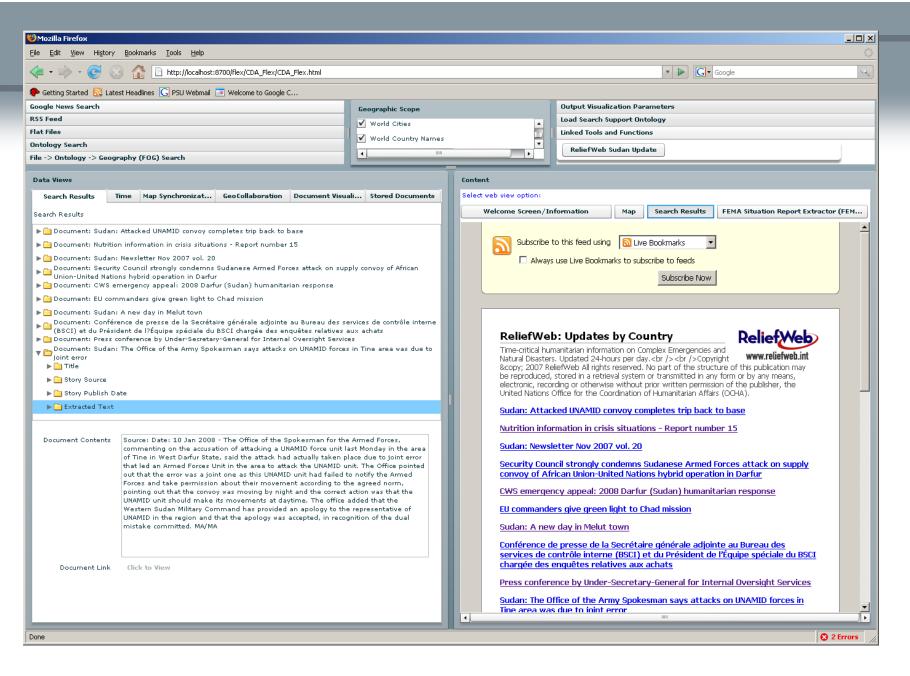
Fly To Find Businesses Direction e.g., Tokyo, Japan Tornadoes Rip Through AL ornadoes ripped through Alabama and killed several people Thursday, including some at a high school where students became pinned under debris when a roof collapsed, state officials said. Crews dug through piles of rubble beneath portable lights at Enterprise High School well into the night, looking for other victims. The burst of ornadoes was part of a larger line of thunderstorms and snowstorms that stretched from Minnesota to the Gulf Coast. More than 50 people were hospitalized as the violent storm front crossed the state. Officials opened shelters for those whose homes were damaged. The state sent in National Guardsmen, along with emergency personnel, lights and generators. The high school, about 75 miles south of Montgomery, appears to have been **▼** Places right in the path said a meteorologist with National Weather Service in Tallahassee, Fla., which monitors President Bush was briefed on the tornadoes by senior staff and called Riley and Missouri Gov. Matt Blunt Original Report White House spokeswoman Dana Perrino said. The Federal Emergency Management Agency was working with ☑ Tornadoes Rip Throu officials in both states, she said. As the system pushed eastward Thursday night, tornado watches remained in effect in eastern Alabama and were posted in Florida. Georgia and South Carolina. The tornadoes were the Alabama Tornadoes second to devastate a portion of the South this year. In early February, tornadoes ripped through a 30-mile patr Georgia Tornado in central Florida, killing 21 and destroying hundreds of homes and businesses Inwa Severe Winter S Severe Storms in Mis: Directions: To here - From here Primary Database Terrain Geographic Web Preatured Conten 3D Buildings roads Google borders 🕩 Eve alt 1658.04 r

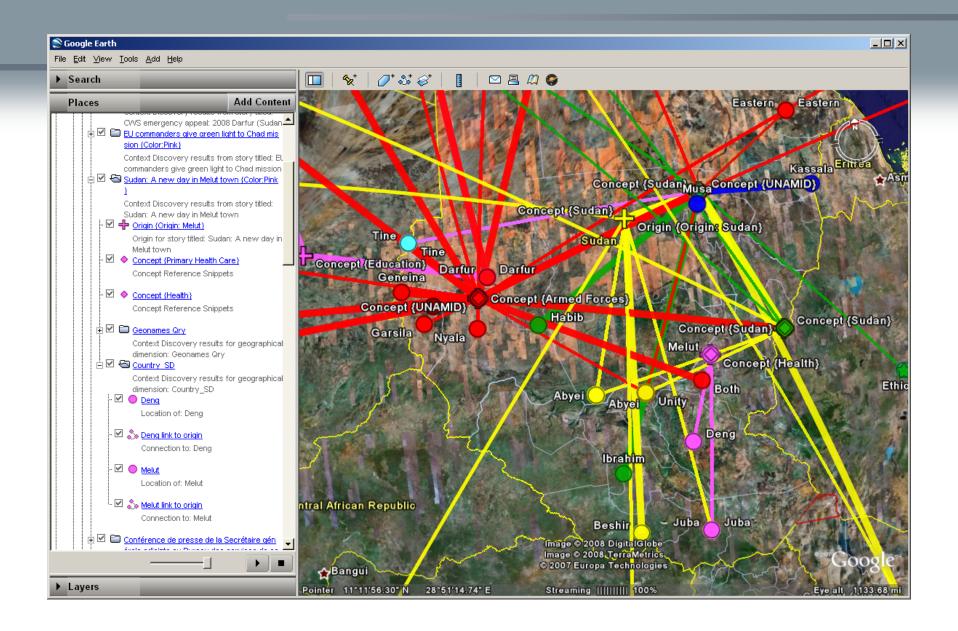
Tags: measles, household, vaccine coverage, index case, contact, households, mumps, mmr vaccine, rubella, national immunization, populations, 5 months, immunization coverage, outbreak, offspring, controlling, subsequently, varicella, genotypes, d4, communicable disease, d8, five months, pregnancy, kindergarten, infectious hepatitis, orthodox communities, jerusalem, trisomy 18, jerusalem israel, west jerusalem, menstrual periods, born babies, autosomal,

Papers about Jerusalem, Israel

FactXtractor used in collaboration with PNNL on SRS & in new NGA GeoCAM project ...

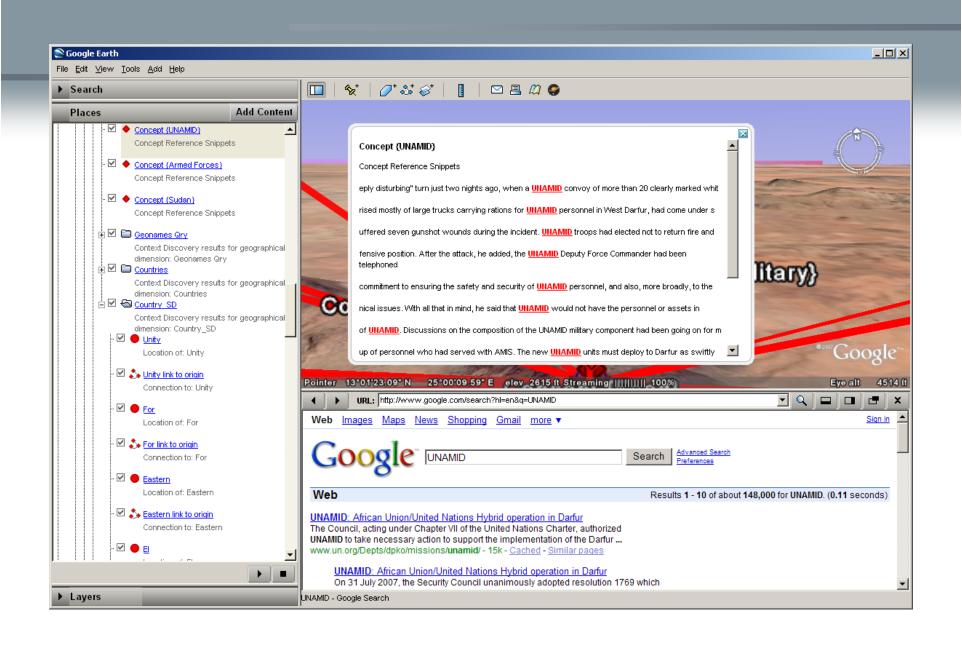






PENNSTATE





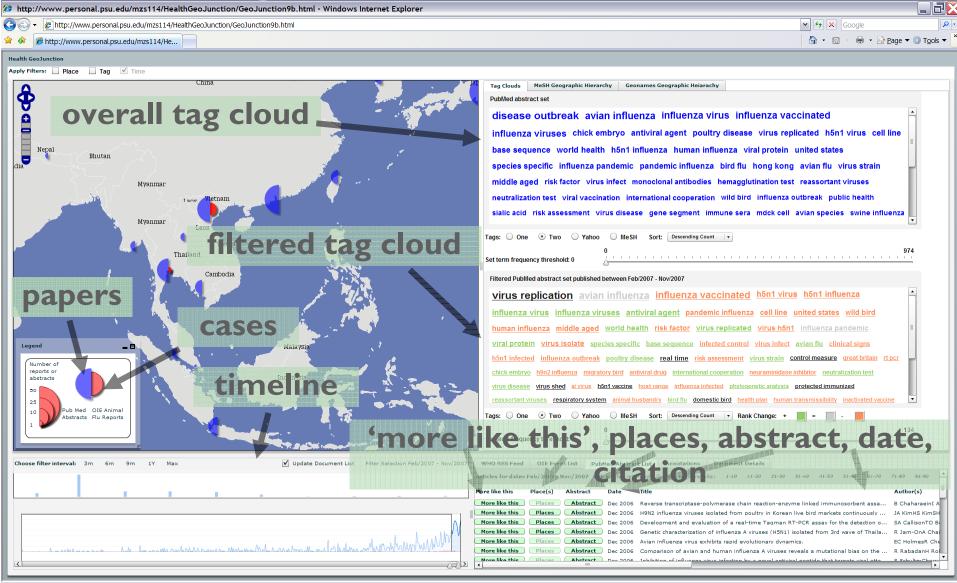
finding info quickly \rightarrow browse, query, & filter

- space-time questions: given X → determine Y
 - when + where → what: Describe the objects or set of objects that are present at a given location or set of locations at a given time or set of times.
 - when + what → where: Describe the location or set of locations occupied by a given object or set of objects at a given time or set of times.
 - where + what → when: Describe the times or set of times that a given object or set of objects occupied a given location or set of locations



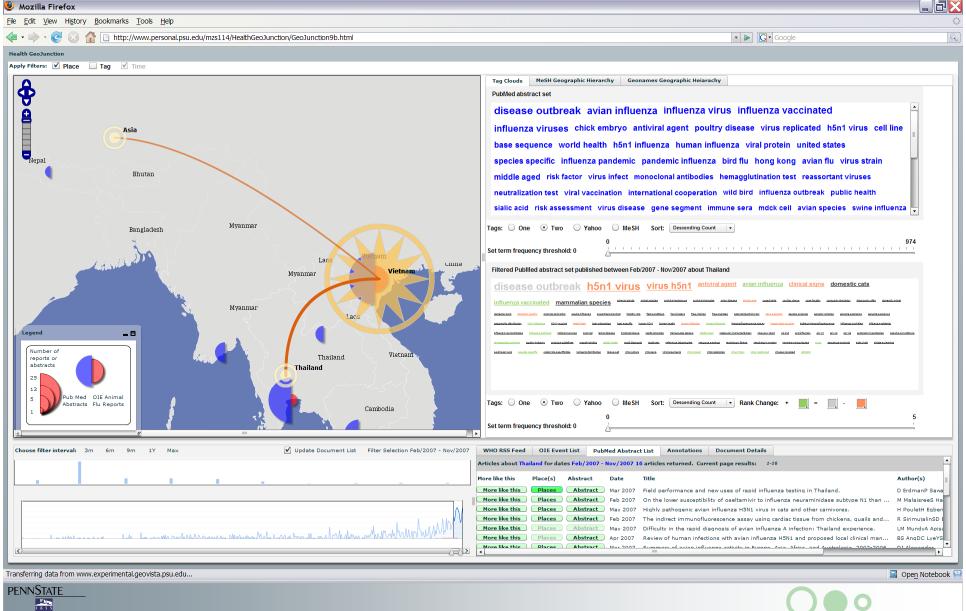


Health GeoJunction: place, time, concept filtering of health science in PubMed



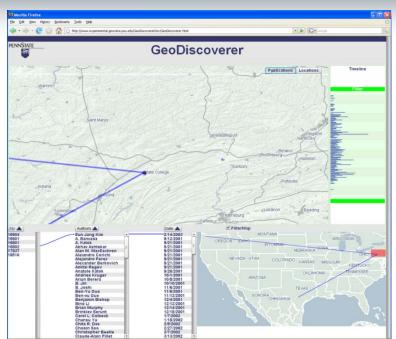
PENNSTATE

Health GeoJunction: place, time, concept filtering of health science in PubMed



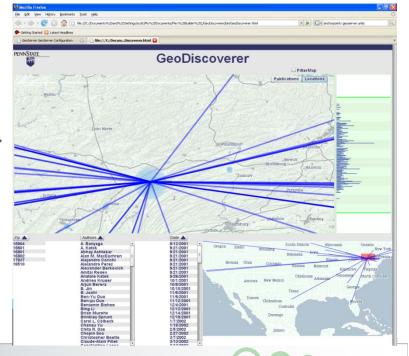
GeoDiscoverer: search engine to integrate social networks with geospatial information

in collaboration with Lee Giles



An analyst is interested in determining when a particular individual has 'published' one or more documents, thus when the content had the potential to make an impact. To pose this question, he simply scrolls to the author name, clicks on it, and views a display of publication dates associated with this author. Clicking on a particular publication shows (on the map) where the co-authors are in relation to the main author.

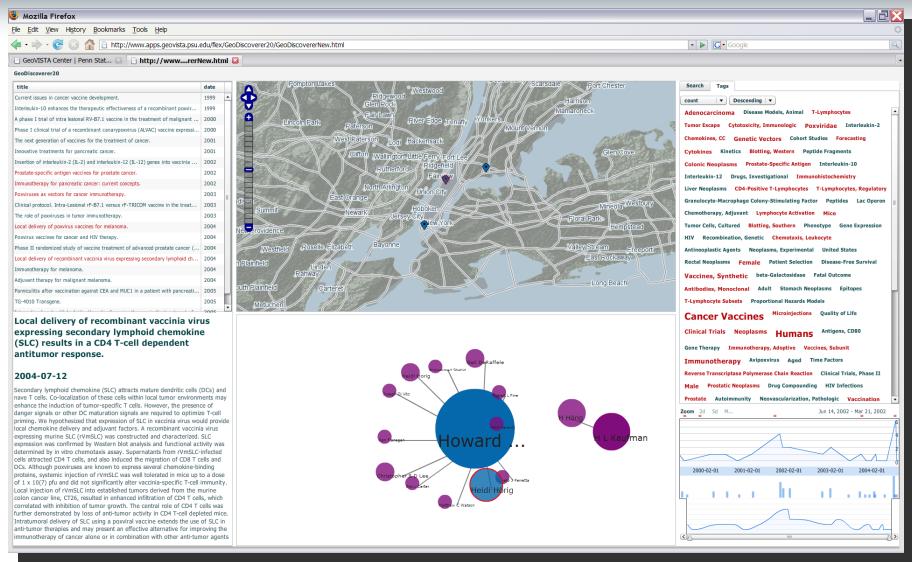
A second analyst notices a series of publications of interest generated from one institution/ location. She is curious about who the authors are collaborating with. To investigate, he uses the geographic view into the data and selects a place by clicking its location on the map. The map shows the number of publications (represented by circle size) and the locations of all collaborators from other locations (shown with radiating lines).





GeoDiscoverer: search engine to integrate social networks with geospatial information

in collaboration with Lee Giles



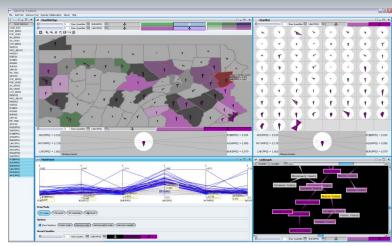


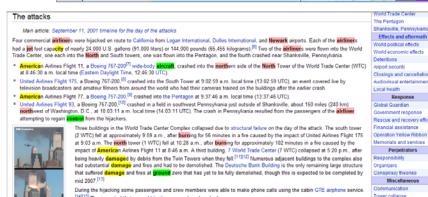


Analyzing Place-Time

homeland security goal: address the need for an orders of magnitude increase in ability to extract situation-relevant info from diverse data sources

- integrated visual + computational methods applied to multivariate data
- finding the where, when, what, and who of events and relations in unstructured text





(5) They reported that several hijackers were aboard each plane. The terrorists reportedly took control of the aircraft by using knives and box-cutter knives to kill flight attendants and at least

one pilot or passenger, including the captain of Flight 11, John Ogonowski.[16] Some form of noxious chemical spray, such as tear gas or pepper spray, was reported to have been used on American 11 and United 175 to keep passengers out of the first-class cabin. [17] Bomb threats were made on three of the aircraft, but not on American 77. According to the 9/11 Commission Report, the bombs were probably fake. The 9/11 Com two of the hijackers had recently purchased Leatherman multi-function hand tools.

On United Airlines Flight 93, black box recordings revealed that crew and passengers attempted to seize cor plane from the hijackers after learning through phone calls that similarly hijacked planes had been crashed into buildings that morning. According to the transcript of Flight 93's recorder, one of the hijackers gave the order to roll the plane once it became evident that they would lose **control** of the plane to the passengers. Soon afterward, the **aircra** crashed into a field near Shanksville in Stonycreek Township, Somerset County, Pennsylvania, at 10:03:11 a.m. loca time (14:03:11 UTC). Al-Qaeda leader Khalid Shaikh Mohammed mentioned in a 2002 interview with an Al Jazeera ournalist that Flight 93's target was the United States Capitol, [18] which was given the code name "the Faculty of Law." [19]

The attacks created widespread confusion across the United States. All international civilian air traffic was banned from landing on US soil for aft already in flight were either turned back or redirected to airports in eastern Canada. Unconfirmed and often contradictory reports were aired and published throughout the day. One of the most prevalent of these reported that a car bomb had been detonated at the U.S. State Department's head quarters, the Truman Building in Foggy Bottom, Washington, D.C. This erroneous report, picked up by the wire services was reported on CNN and in a number of newspapers published that day. Soon after reporting for the first time on the Pentagon crash, CNN and



ogans and term

9/11 Commission Report

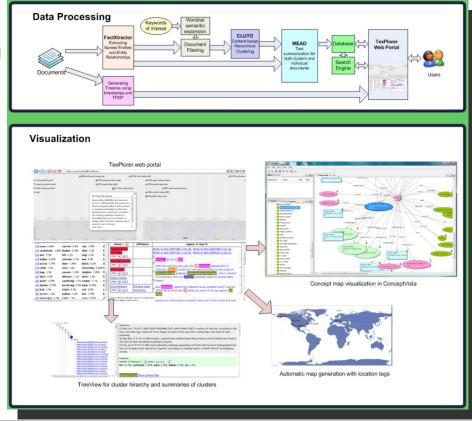


TexPlorer

 An integrated system for exploring and analyzing large amounts of text documents.

TexPlorer includes:

- Named Entity Recognition
- Entity Relation Extraction
- Text Clustering
- Text Summarization
- Visualization

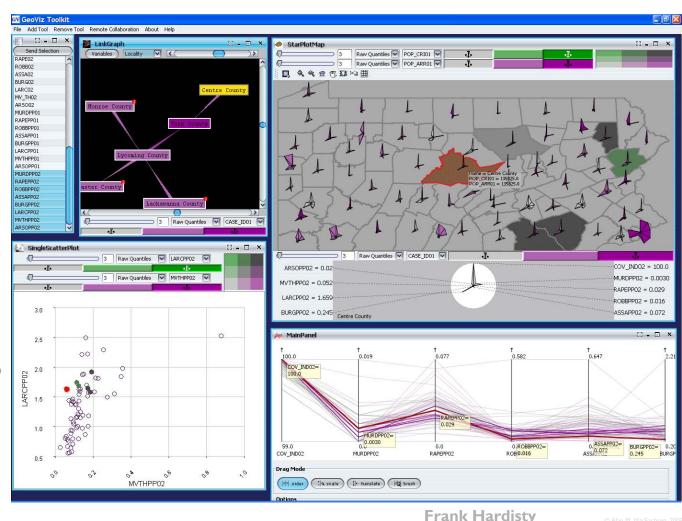






GeoViz Toolkit: finding patterns and relations in multivariate, georeferenced data

- Enables analysts to easily understand multivariate patterns
- Applied here to crime pattern analysis
- Funding by CDC for application to cancer & covariate analysis





Research: Visual Relationship Analysis

Goals

Support a wide variety of general and specific visual analysis tasks.

Allow users to follow complex lines of inquiry using sequences of simple

interactions.

Relevance

 Provide visual means to find and compare the 'needles' that connect attribute values in 'haystacks' of high-dimensional data.

Accomplishments

Cross-Filtered Views: Pose complex drill-down set queries.

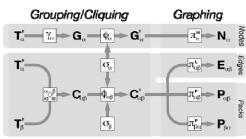
Attribute Relationship Graph: Filter a multigraph of attribute value

co-occurrences across data columns.

Generalizable to different information sources

Extendable to multiple interrelated data sets

Combinable into functionally-rich visual tools



Cross-filterina

Chris Weaver

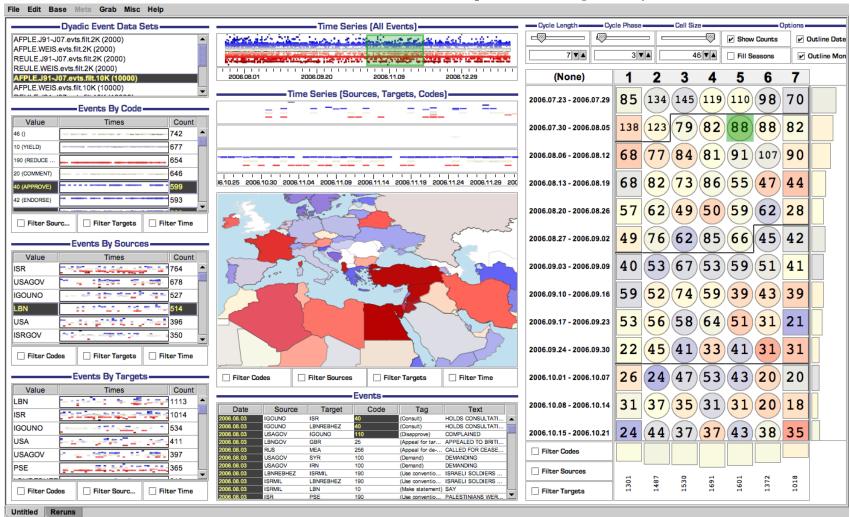






Newswire-extracted international events between state-level actors

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

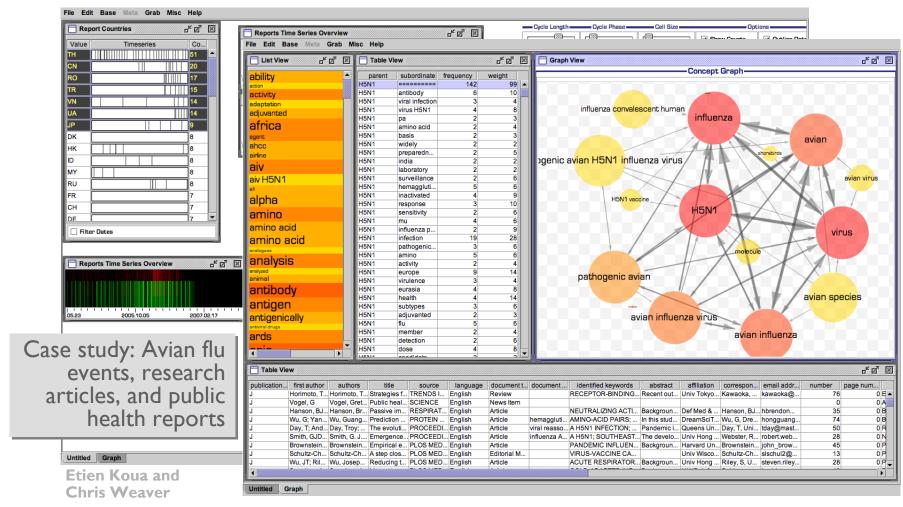


Using the Kansas Event Data System (KEDS) Chris Weaver & Phil Schrodt (U of Kansas) M. MacEachren, 2001



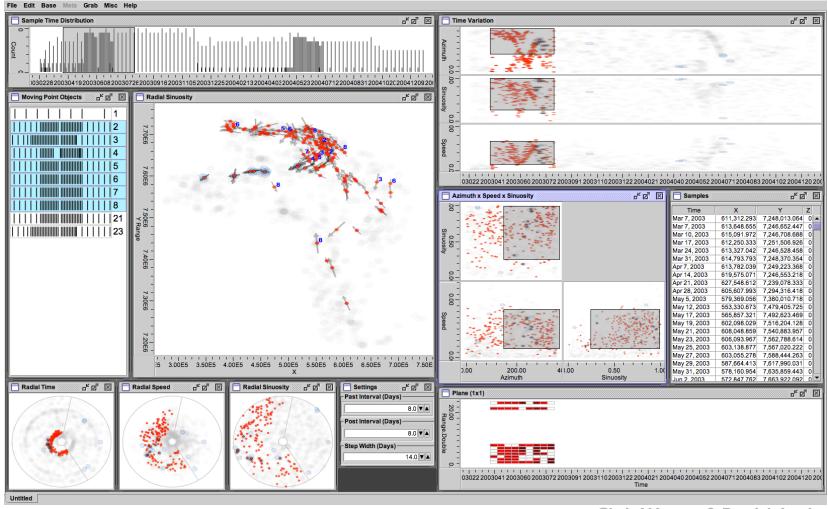
Linking events with science: Geo-temporal relationships among multiple data channels

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



Space-Time Analytics: Analysis of group behavior

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



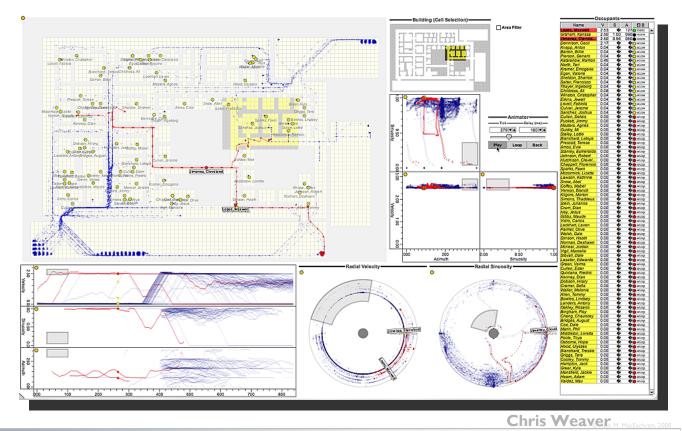
Chris Weaver & Patrick Laube





Space-Time Analytics: building evacuation

 evacuation analysis: looking for unusual behaviors, bottlenecks, impediments







Conclusions

- Place & Time are Fundamental ... to all components of Homeland Security
- NEVAC foci: R&D to assemble time-relevant info about places at risk and to use the information effectively to decrease threats and increase resilience



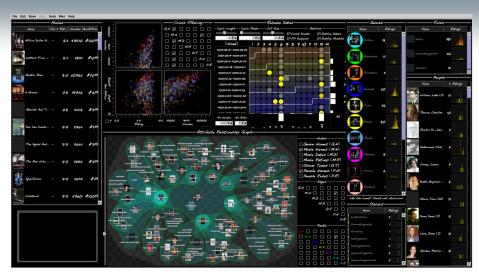
maceachren@psu.edu



Impacts: Visual Relationship Analysis

Prototype tool: Cinegraph

- InfoVis 2007 Contest Winner
- Multiple data sets from IMDB
- Built in 2 days by one person
- Drill-down into movies, genres, release dates, people, awards, ratings, box office numbers



Movies for each actor who has worked with Kevin Bacon since 2001.

Future Impact

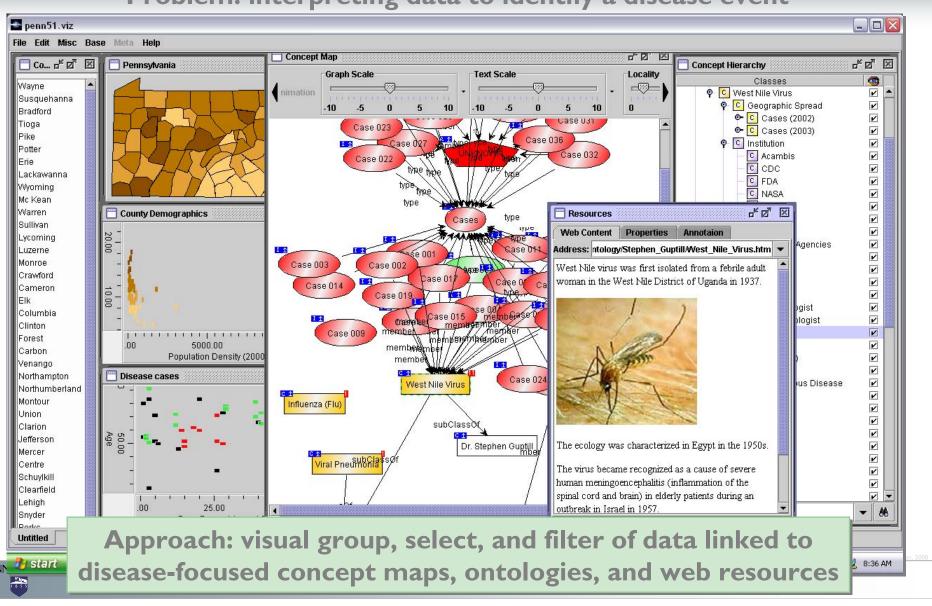
- Analogous application to problems of interest to the community
- Drill-down into events and activities in terms of occurrences, categories, dates, actors, responses, measurable effects, etc.
- Plans
 - Demonstrate on data sets extracted from synthetic intelligence reports in Frank Hughes' exercises, using NEVAC text extraction research
 - One possible element of our VAST 2008 Contest solution strategy [?]



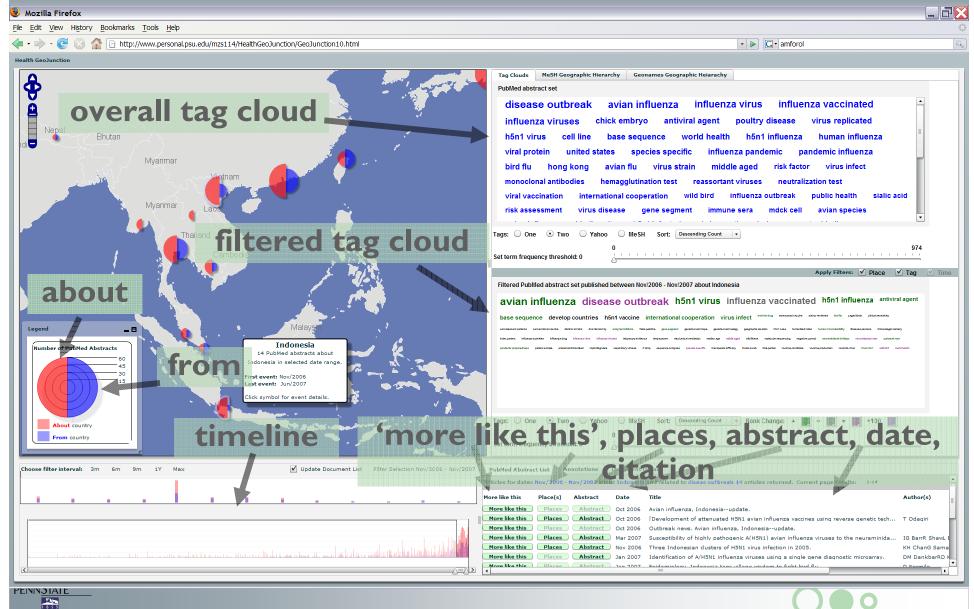


Data-to-knowledge and back: Integrating knowledge management with visual analytics

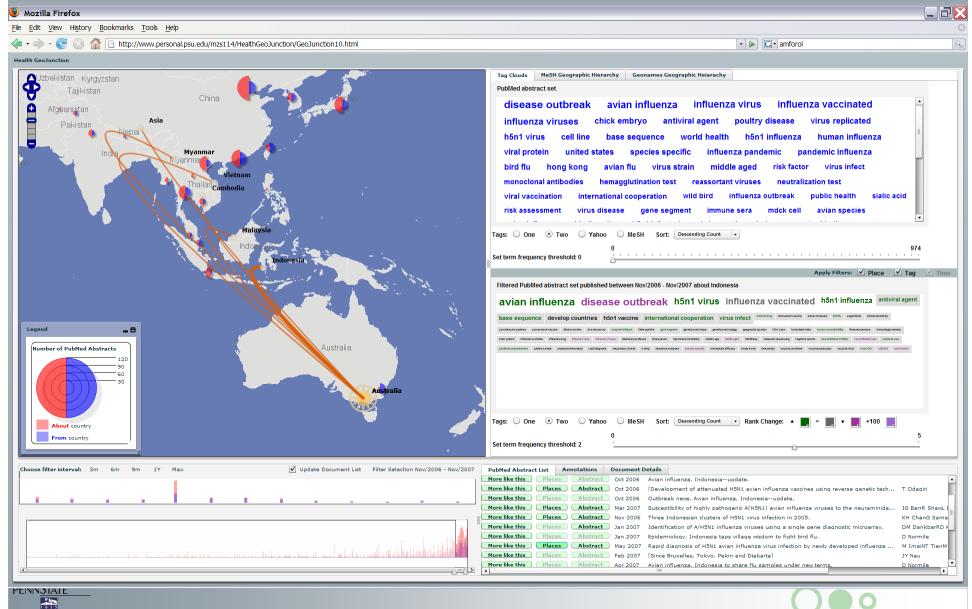
Problem: interpreting data to identify a disease event



Health GeoJunction: place, time, concept filtering of health science in PubMed



Health GeoJunction: place, time, concept filtering of health science in PubMed



GeoDiscoverer

- A Flexible map-based interface that supports overview + detail analysis of cross connections of social networks
- Web-based UI using Flex/Flash on the client
 and PHP and Java web services on the server
- Multi-view interface including Tag Clouds for displaying search results, map, network graph, timeline and detailed publication view of data





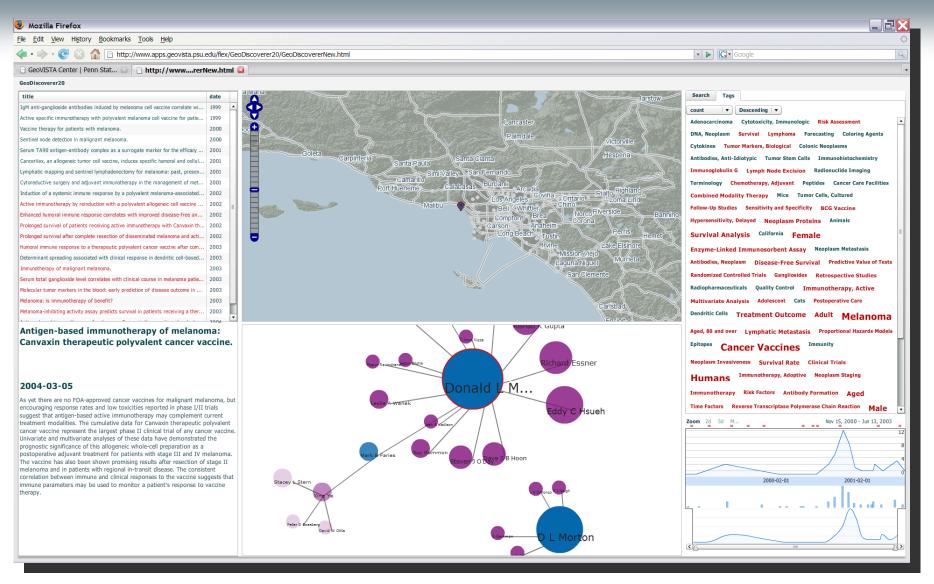
GeoDiscoverer - Analysis Scenario

- Search for an Individual through a flexible search tool
- Show names/aliases which match the search
- Select individual names/aliases or multiple to visualize network graphs, maps, timelines of individual locations and attributes
- Resolve names/aliases which are believed to be the same person





GeoDiscoverer







applying the typology

structure: given $X \rightarrow determine Y$ example — elementary <u>when</u> $\rightarrow what + where$

- given when → determine what + where
 - *ID* task: given 1999 →
 - determine which county has the highest rate
 - Compare task: given 1997 and 2001
 - determine which county ... increased the most





general tasks (with respect to time)

when → what + where

- identify behaviors over time interval(s)
 - elementary for what, where, or both
 - general for what, where, or both
- Compare behaviors over same interval
 - two instances of one "thing"
 - two different, potentially related "things"
- Compare behaviors between/among time intervals: e.g., one year to another





applying the typology

structure: given X → determine Y

example – general <u>what</u> + <u>where</u> → <u>when</u>

- given what + where > determine when
 - ID task: given ? →
 - determine?
 - Compare task: given ?
 - determine ?





general tasks (with respect to time)

what + where → when

- determine time interval(s) during which a specific behavior took place
 - elementary what + where: during what time interval did stork X fly to location Y
 - general what + where: during what time interval did most of the storks fly north





Integration of geospatial and abstract structuraltemporal visualization



overlapping layers depicting terrorist attacks (event location) and terrorism research (expert location)

Avian Flu
Knowledge Diffusion

movement of co-citation behavior related to avian flu research







What is the CDA?

- Context Discovery Application (CDA)
- Geovisual Analytic environment
- Production of geographical and historical context from implicit sources
- Context analysis and situation assessment activities in humanitarian crisis management



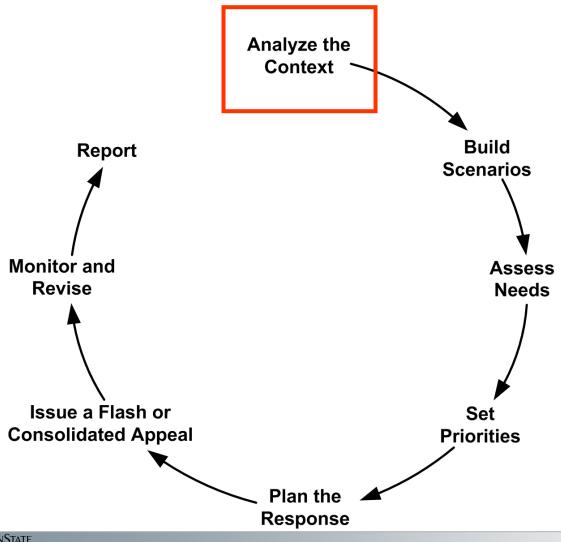


Tools

- The Context Discovery Application (CDA)
- Visual representation of and reasoning with GHC
- Geocollaboration
- Geographic Information Retrieval
 - RSS feeds from Google News
- Crisis Management application domain



Consolidated Appeals Process (CAP) Cycle OCHA 2007

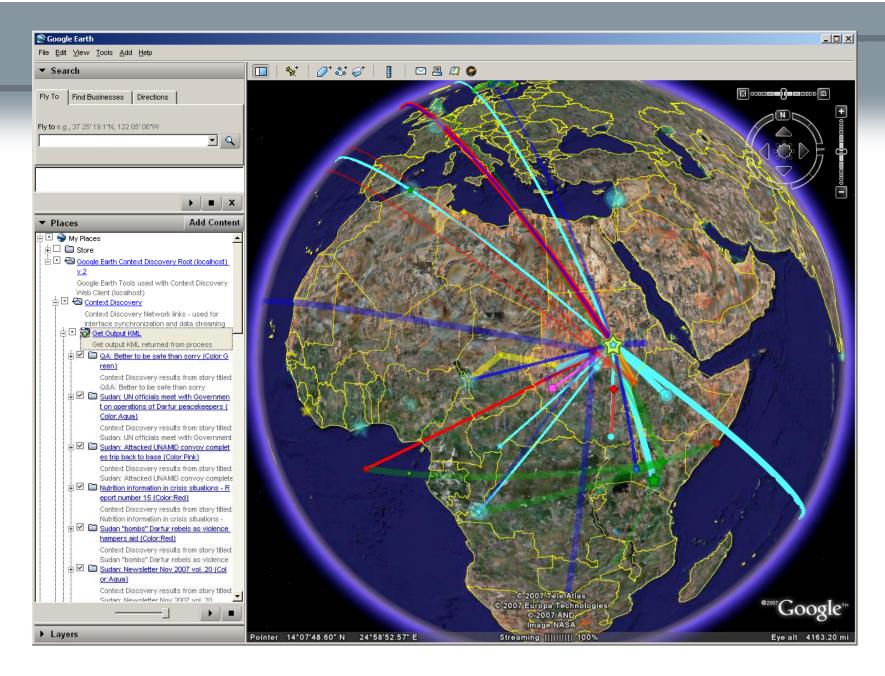


Context analysis is meant to "outline the root cause of the crisis and its evolution, including any regional or historical dimensions"

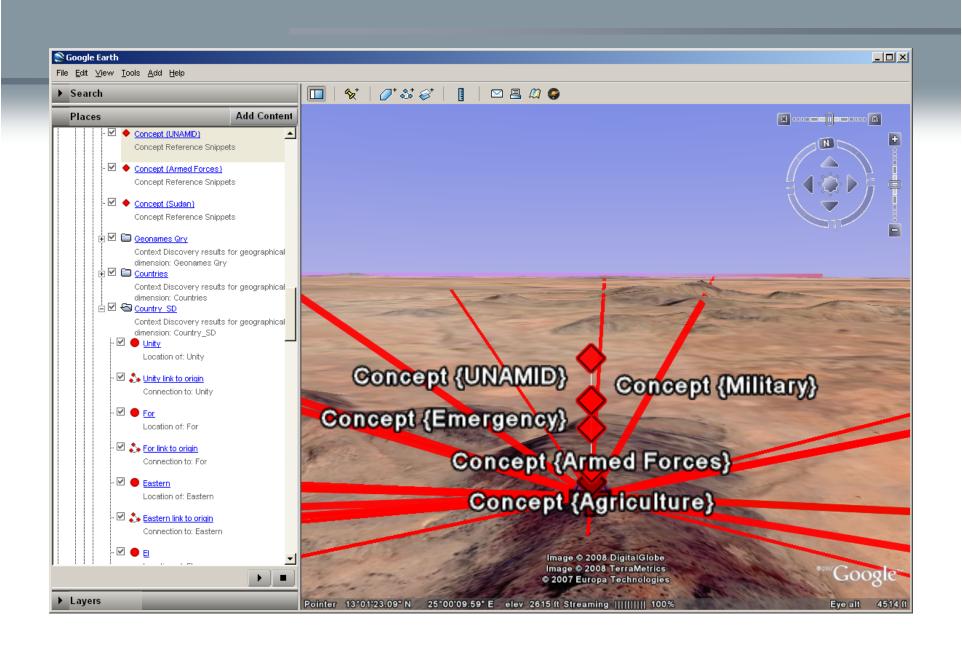




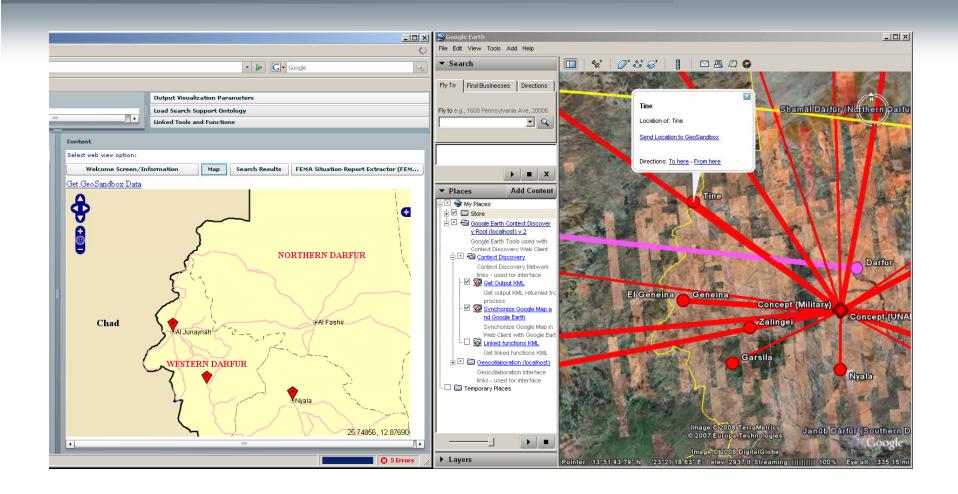




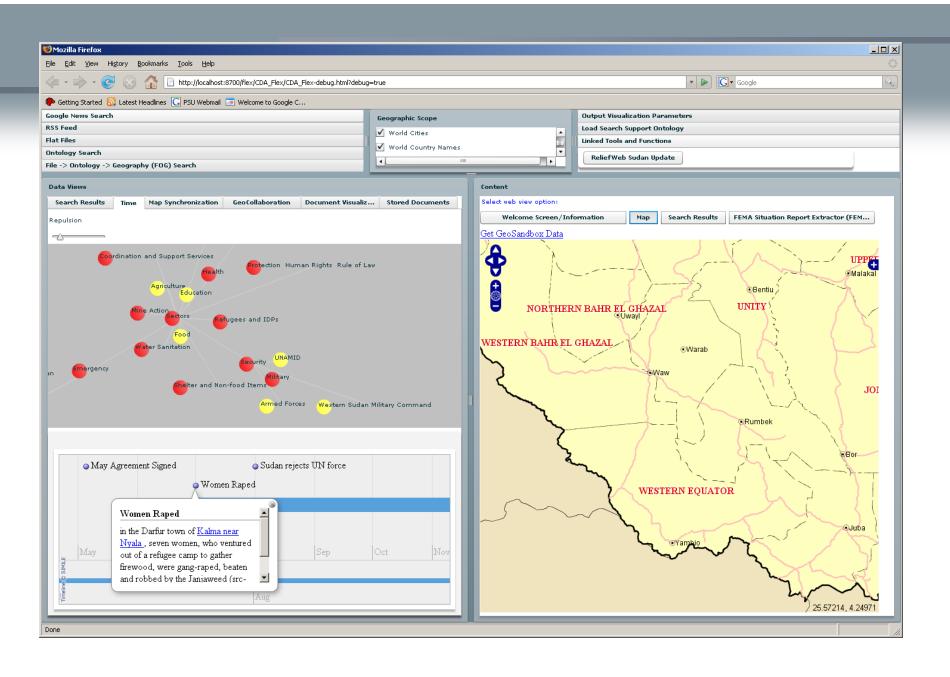














Geo-contextualizing to enable situation assessment → Situation Awareness

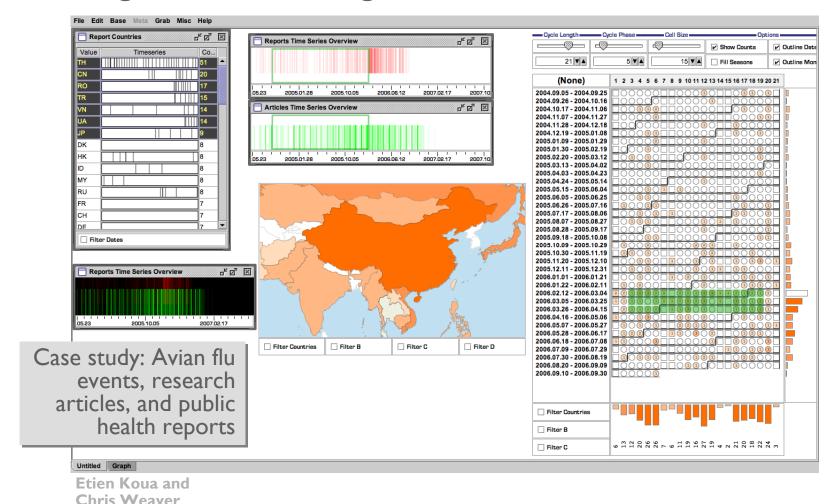
- cross-filtering & attribute-relation graphs
- geocontextualization of social networks
- supporting exploration of space-time-attribute connections in document collections:
 TexPlorer
- knowledge enabled information foraging and geographical contextualization





Geo-temporal relationships among multiple data channels

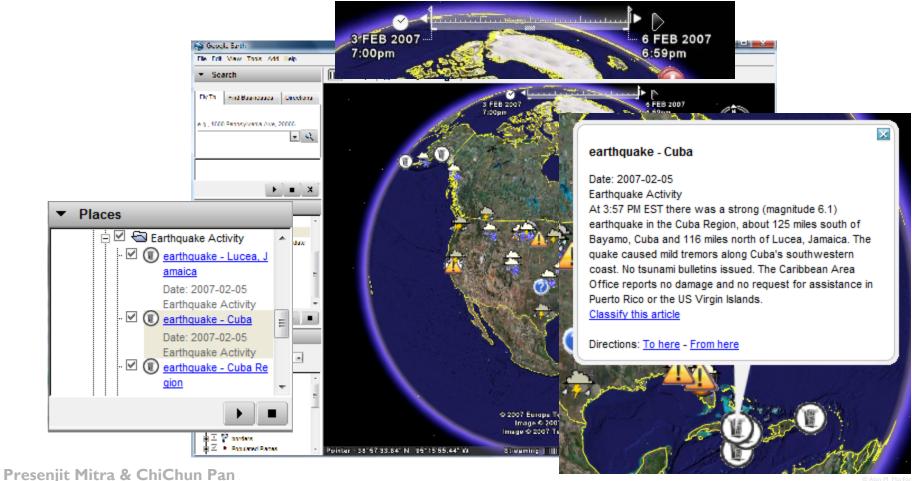
Problem: How can we contextualize events and risks, geo-temporally, in evolving scientific knowledge?

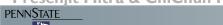


PENNSTATE

FEMARepViz

example info review session

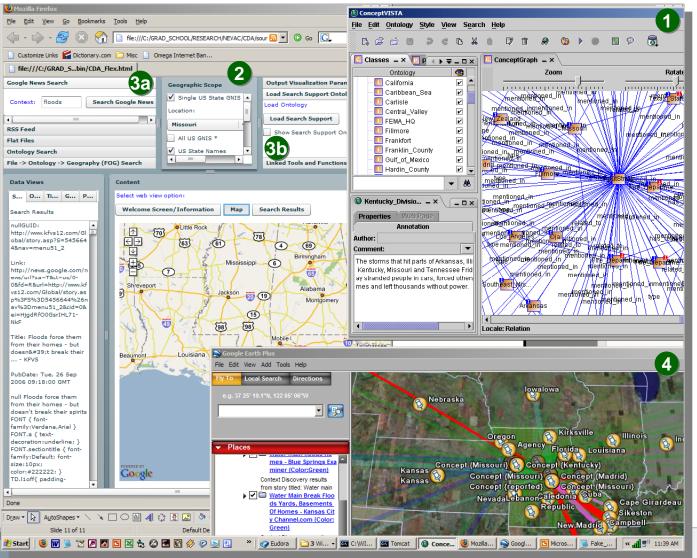






Problem: Help analysts retrieve relevant info and contextualize complex events

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



CDA process

- I. 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

Brian Tomaszewski, Alan MacEachren