RVAC
Regional Visualization and Analytics Centers

North-East Visualization and Analytics Center
www.geovista.psu.edu/NEVAC

NEVAC R&D – GeoVisual Analytics

June, 2008

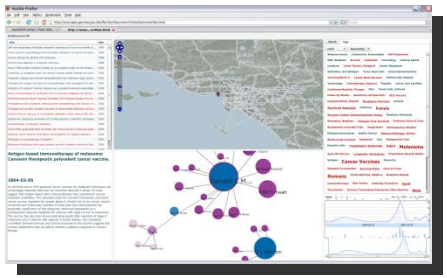
Alan M. MacEachren – PI
GeoVISTA Center/Geography, Penn State
maceachren@psu.edu

RVAC
Regional Visualization and Analytics Centers

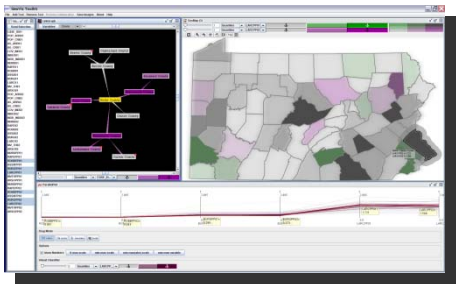
North-East Visualization and Analytics Center
www.geovista.psu.edu/NEVAC

GeoVisual Analytics Tools for:

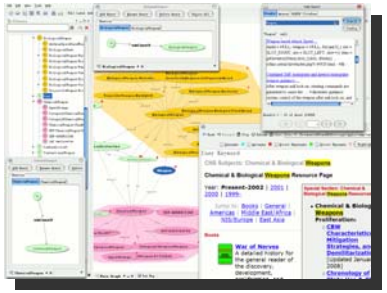
Public Health
Facility Security
Study of Terrorism
Intelligence Analysis
Emergency Management



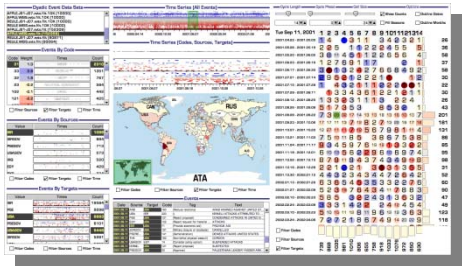
GeoDiscoverer — Searching social networks + geo info in web and text



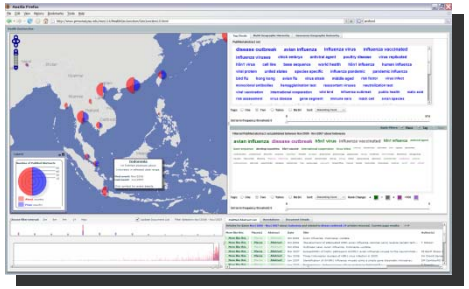
GeoViz Toolkit — Rapid construction of custom analysis tools (crime rates in PA)



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

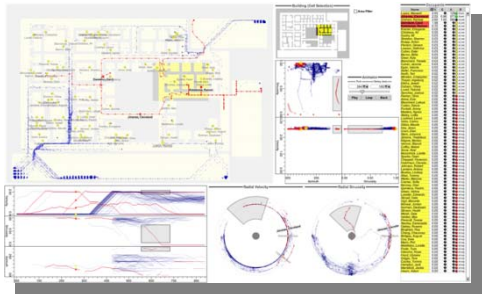


Improvise — Querying multidimensional data (world events in newswire reports)



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

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



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

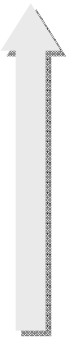
NEVAC Goal: geo-visual analytics for public health, intel analysis, study of terrorism, & emergency mgt

- 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

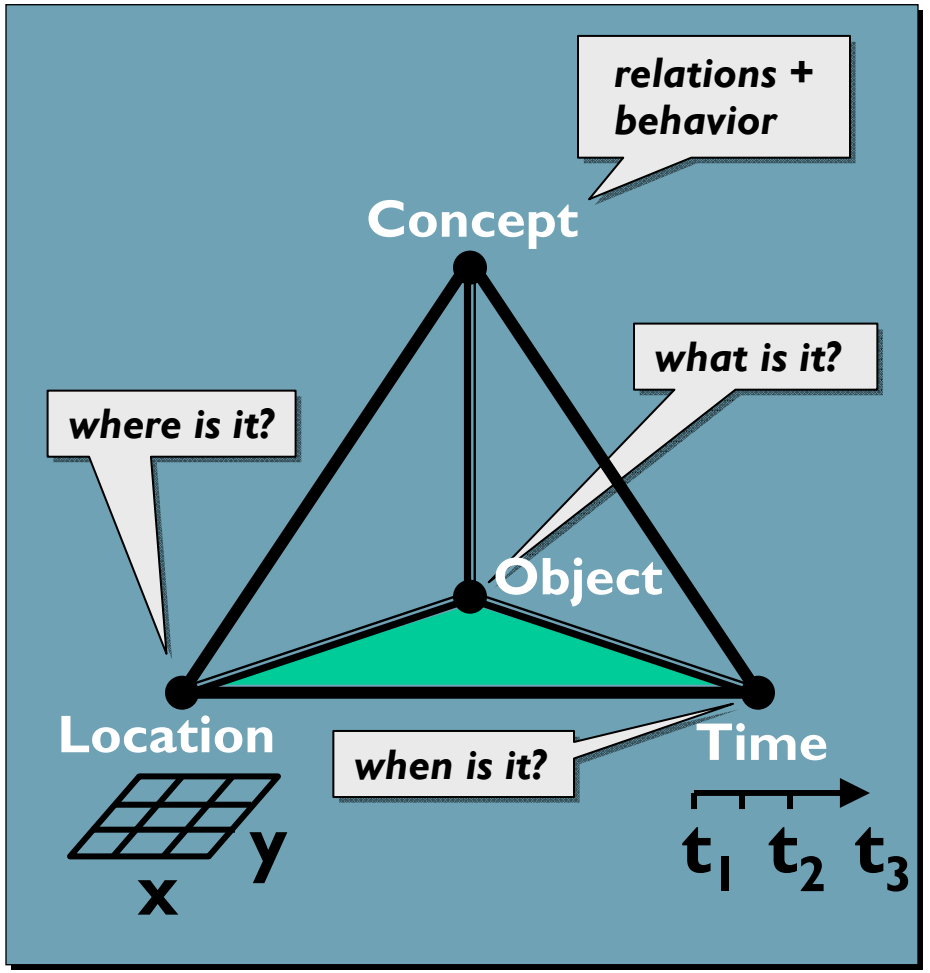
© Alan M. MacEachren, 2008

A multi-representation info-model

Knowledge



Observational data



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.

© Alan M. MacEachren, 2008

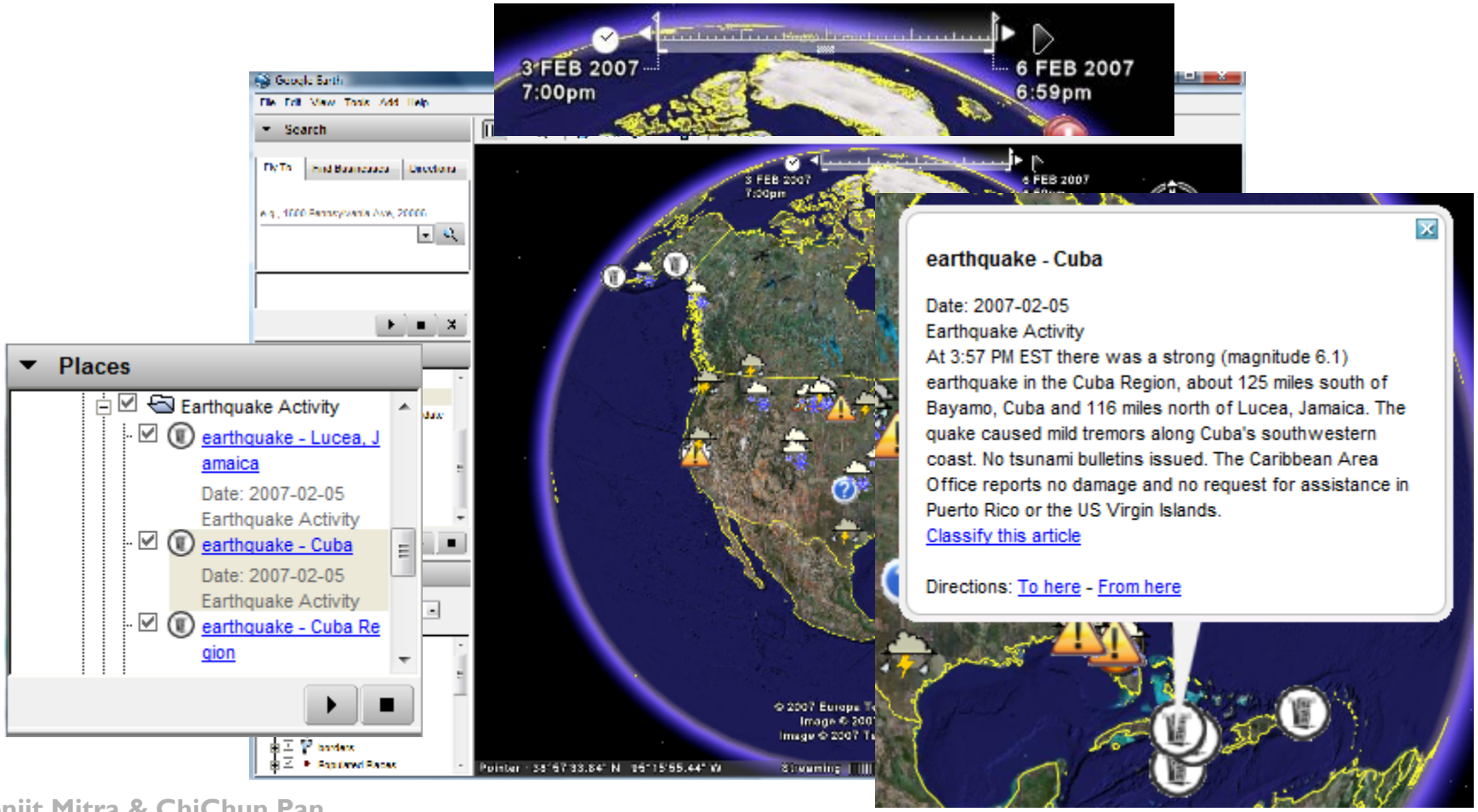
Foraging and contextualizing info → Situation Awareness

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

© Alan M. MacEachren, 2008

FEMARepViz

- example info review session

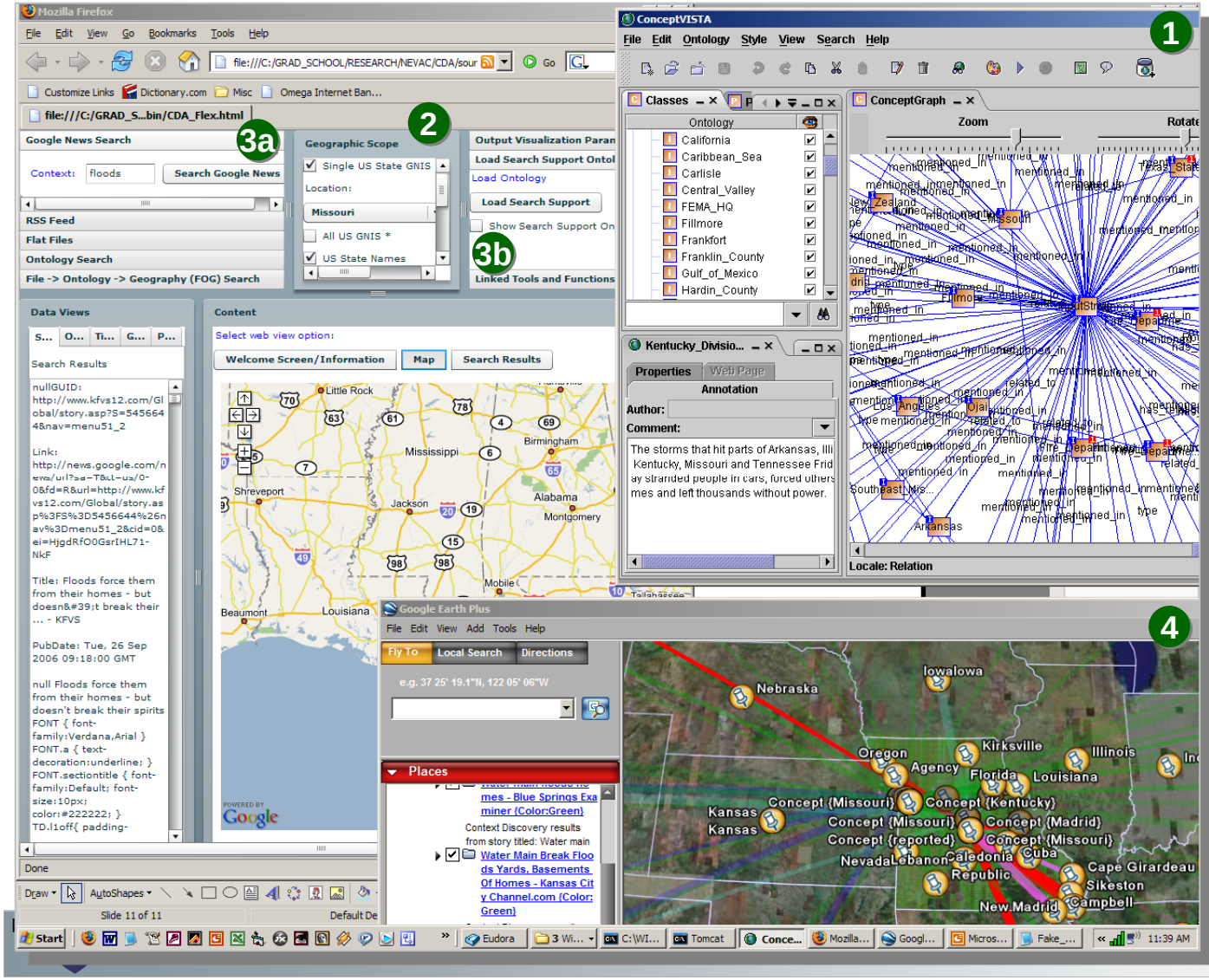


Presenjit Mitra & ChiChun Pan

© Alan M. MacEachren, 2008

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

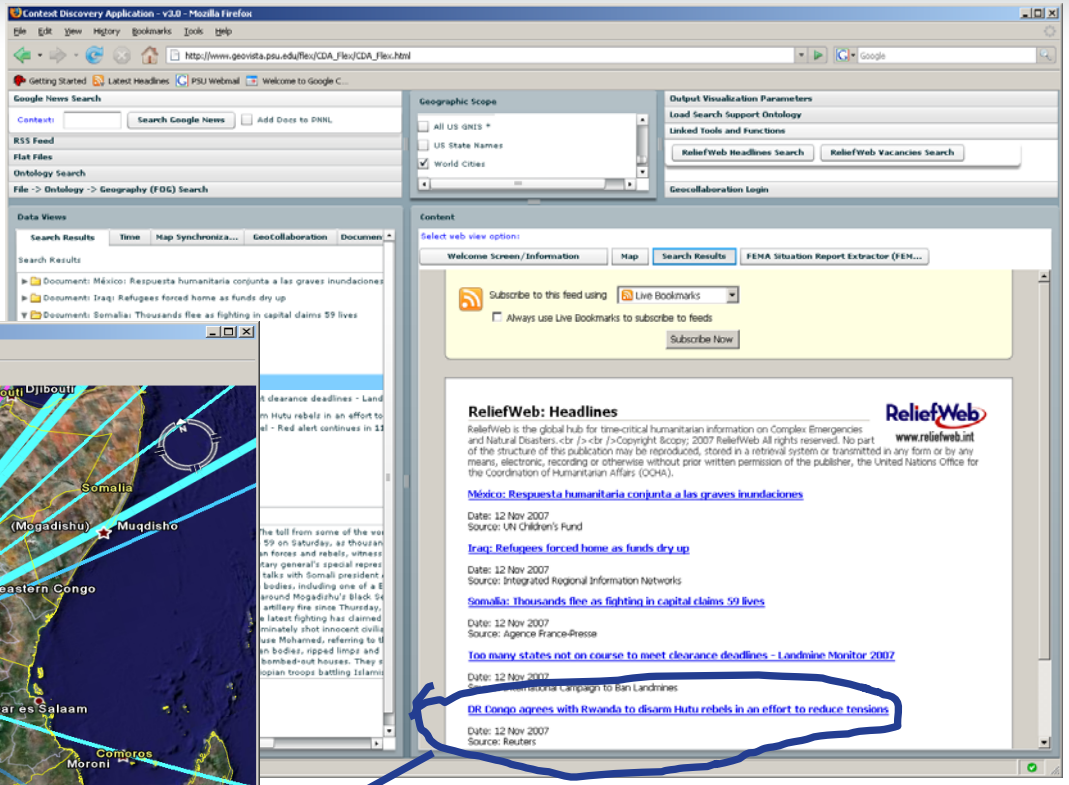
see poster

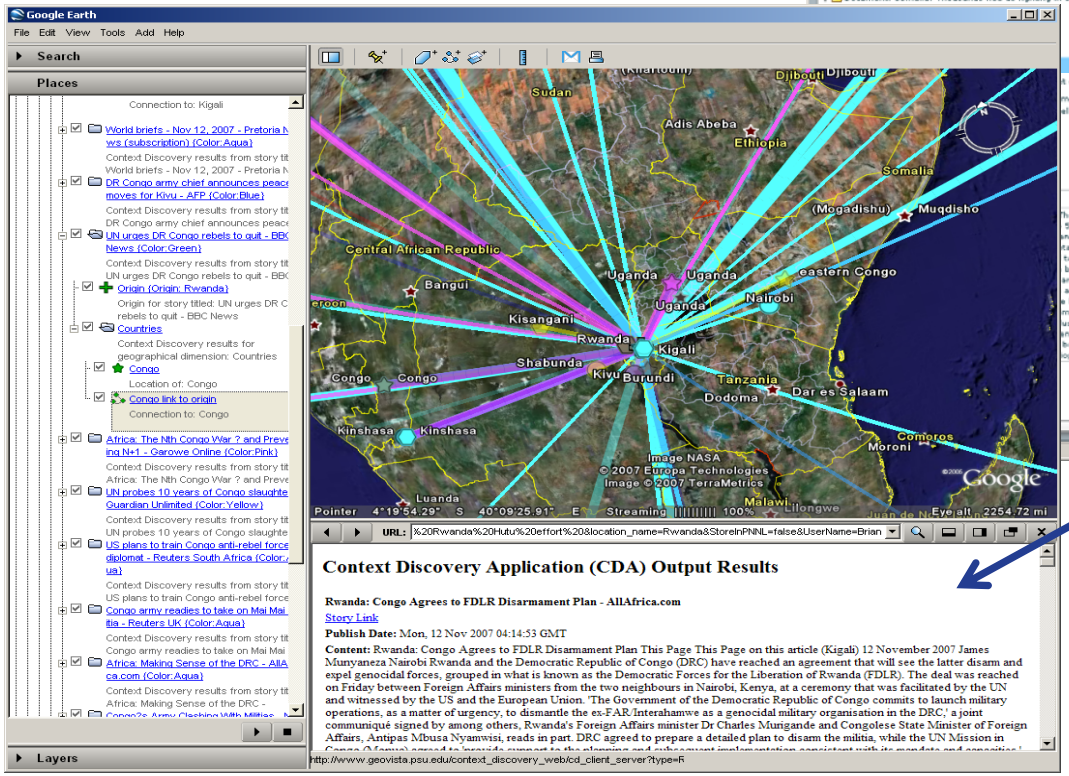
Brian Tomaszewski, Alan MacEachren

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

United Nations ReliefWeb headline search

Brian Tomaszewski





ReliefWeb headline follow-up query in Google News using nouns from RW headline as search terms

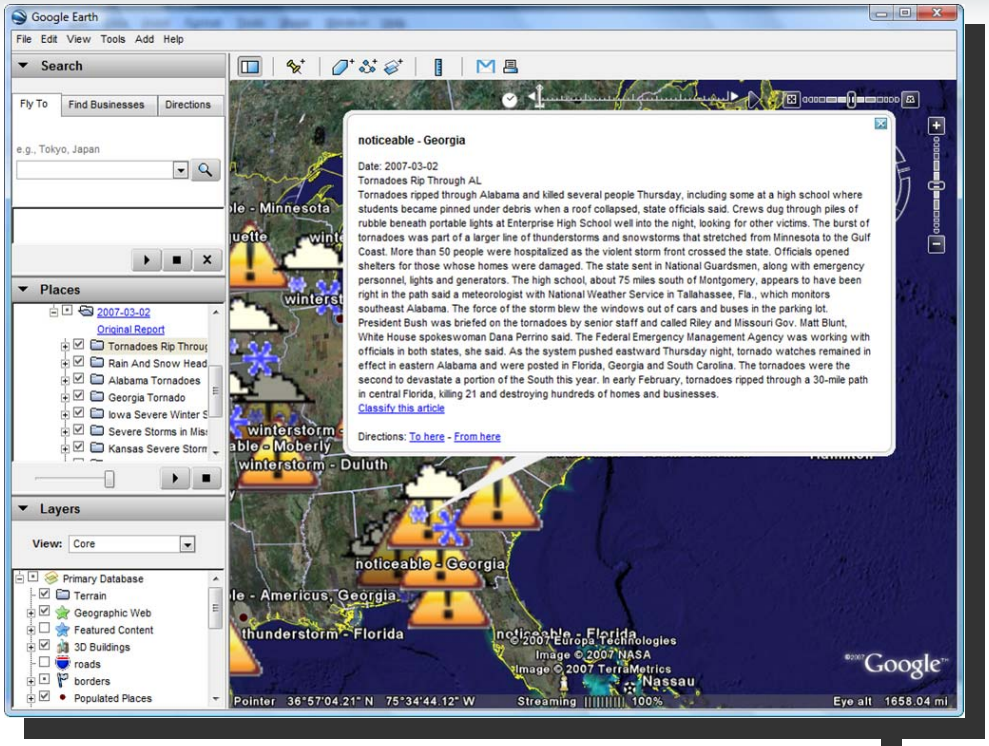
© Alan M. MacEachren, 2008

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

FactXtractor

FEMARepViz

GeoJunction



Jerusalem, Israel

Parents: Israel,

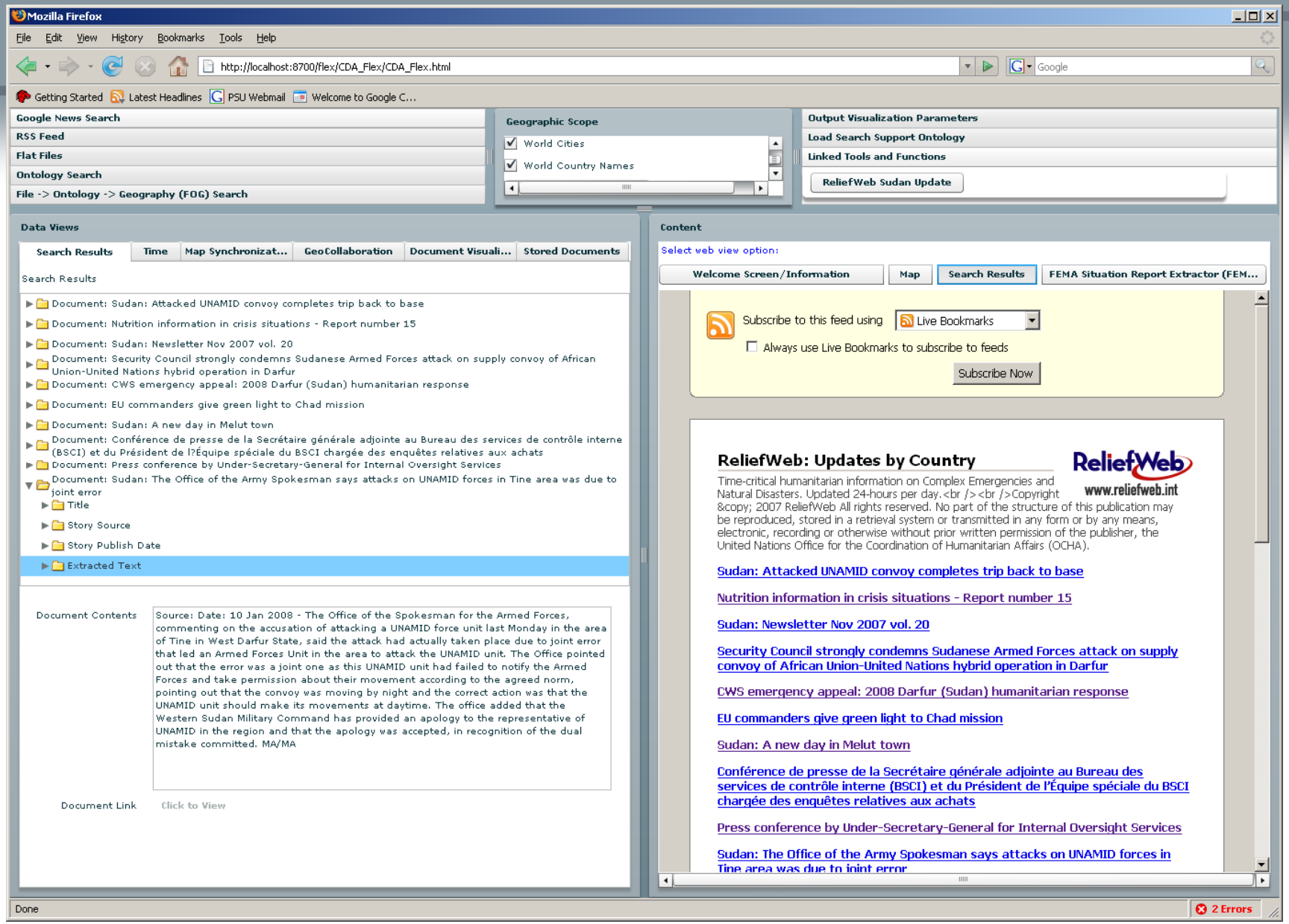
Tags related to Jerusalem, Israel

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,

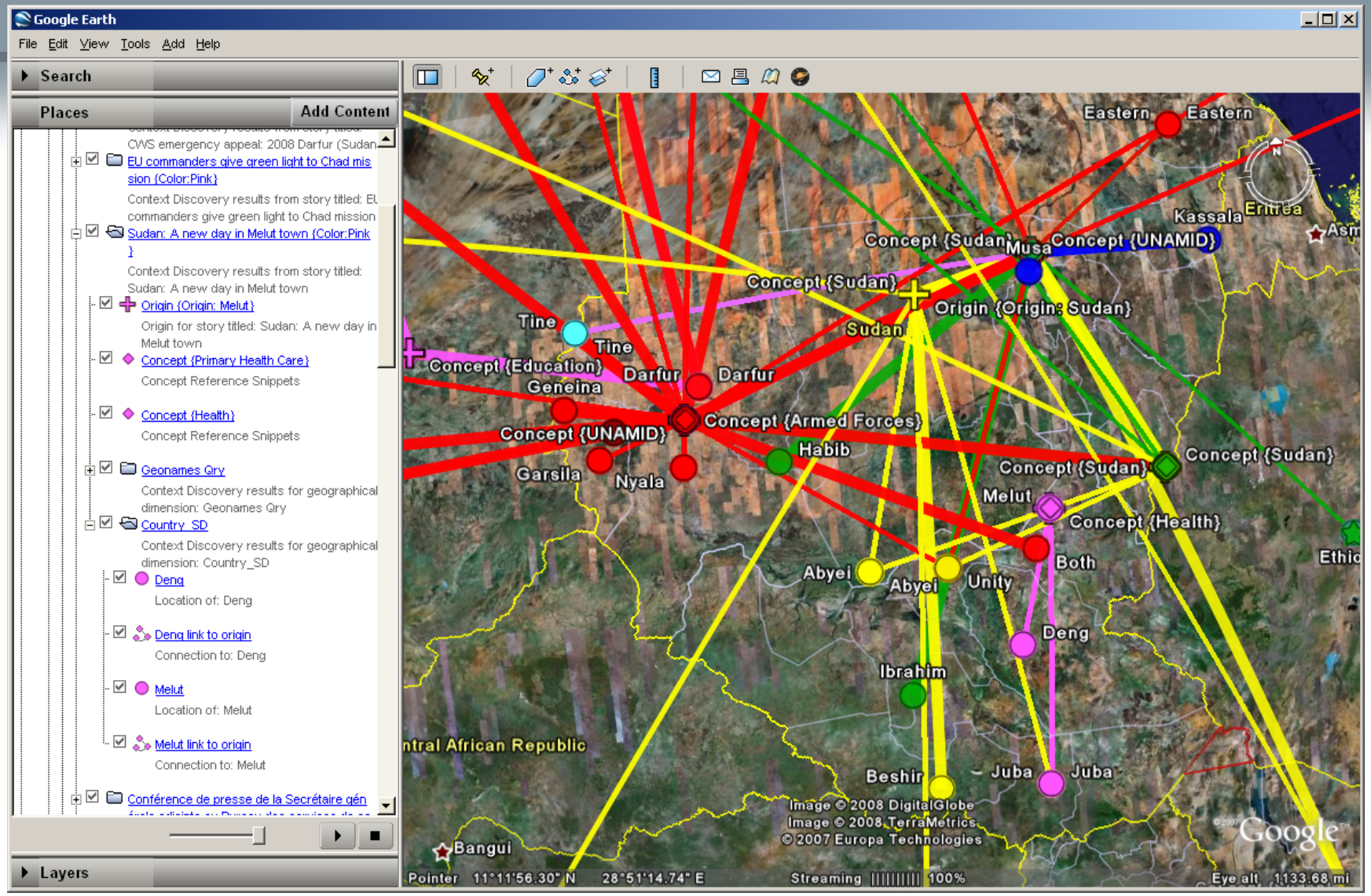
Papers about Jerusalem, Israel

Stein-Zamir,C; Zentner,G; Abram... Shulm	Measles outbreaks affecting children in Jewish ultra-orthodox communities	5 months, controlling, d4, d8, genotypes, households, immunization coverage, index

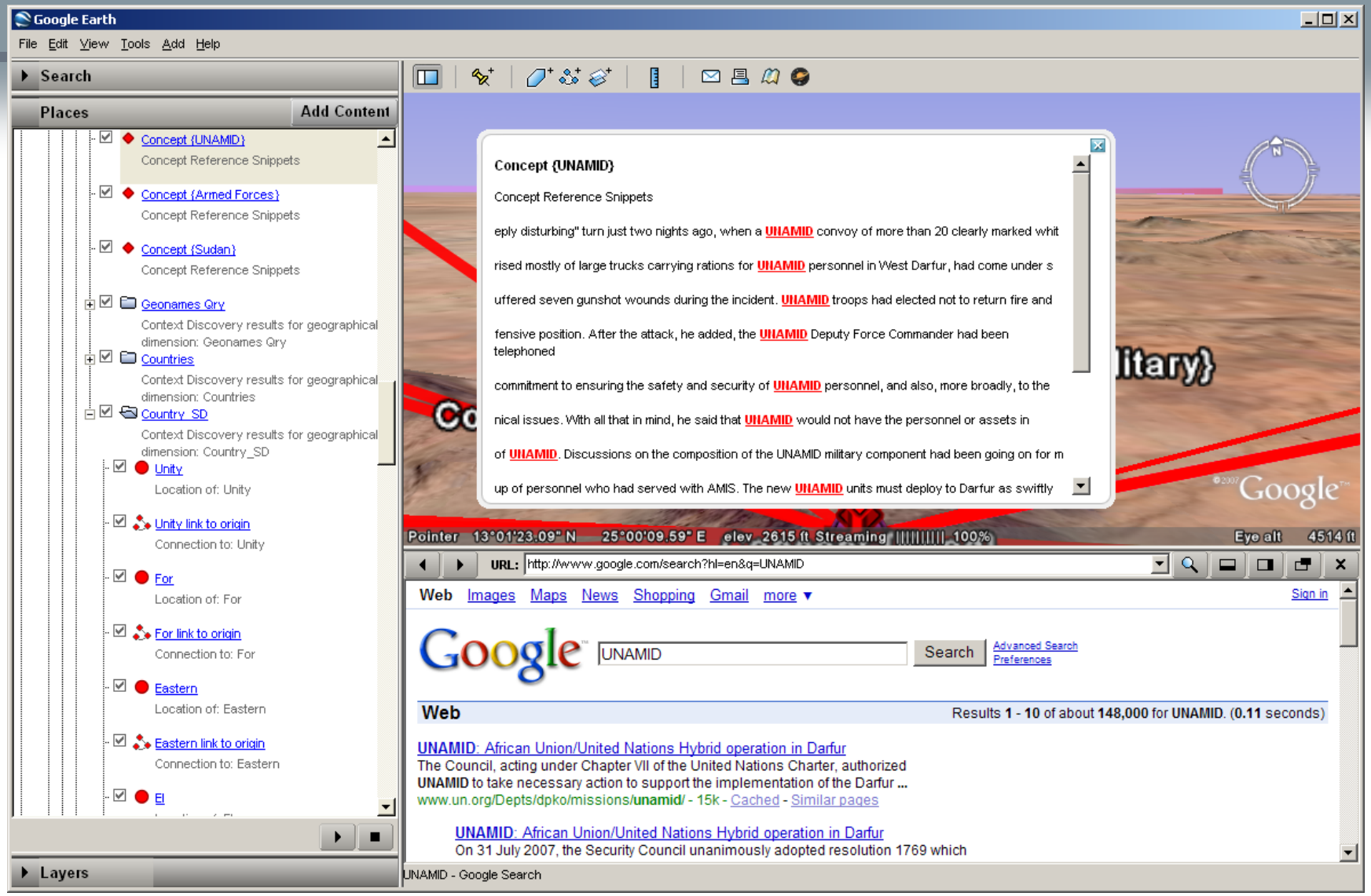
FactXtractor used in collaboration with PNNL on SRS & in new NGA GeoCAM project …



© Alan M. MacEachren, 2008



© Alan M. MacEachren, 2008



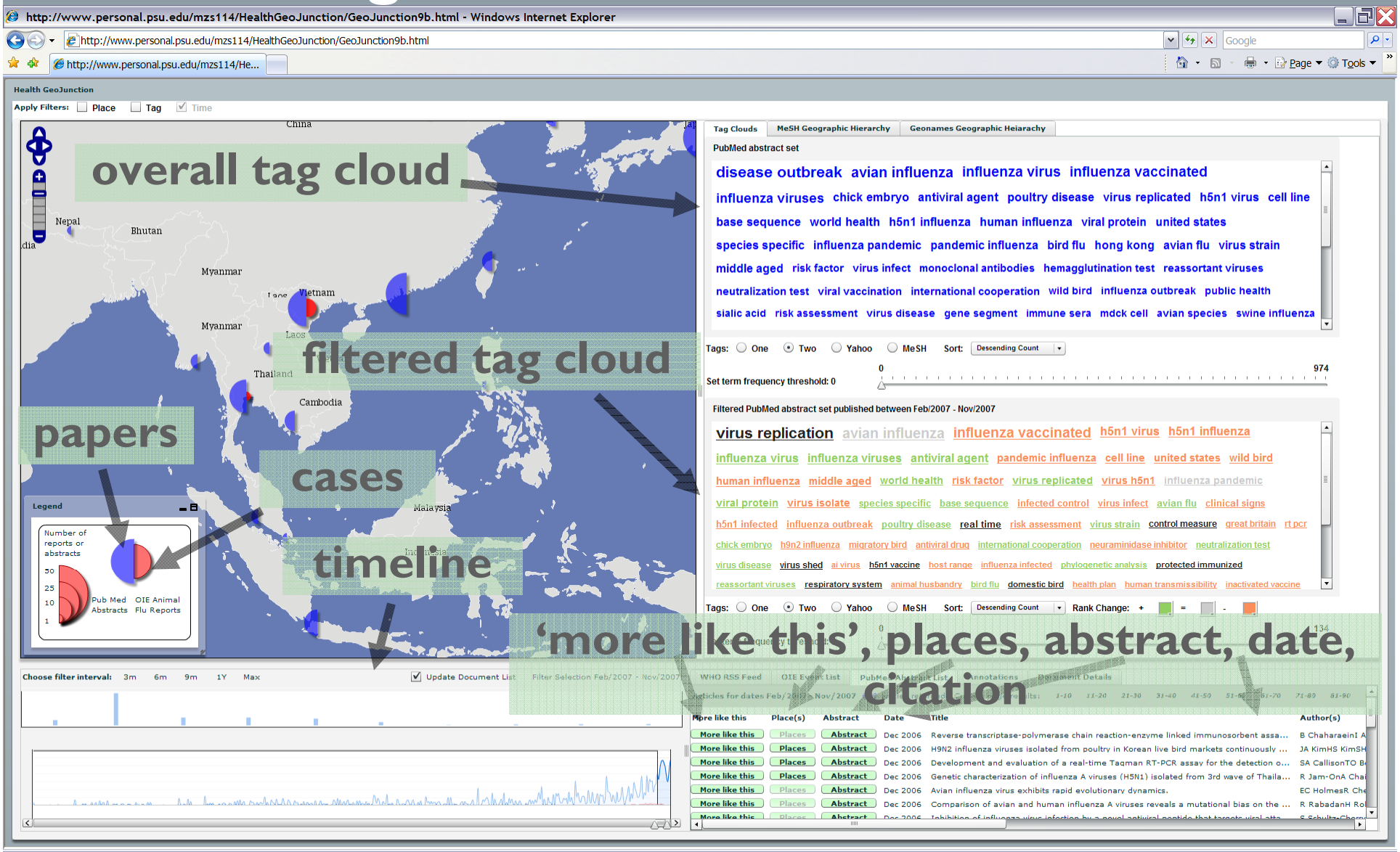
© Alan M. MacEachren, 2008

finding info quickly → browse, query, & filter

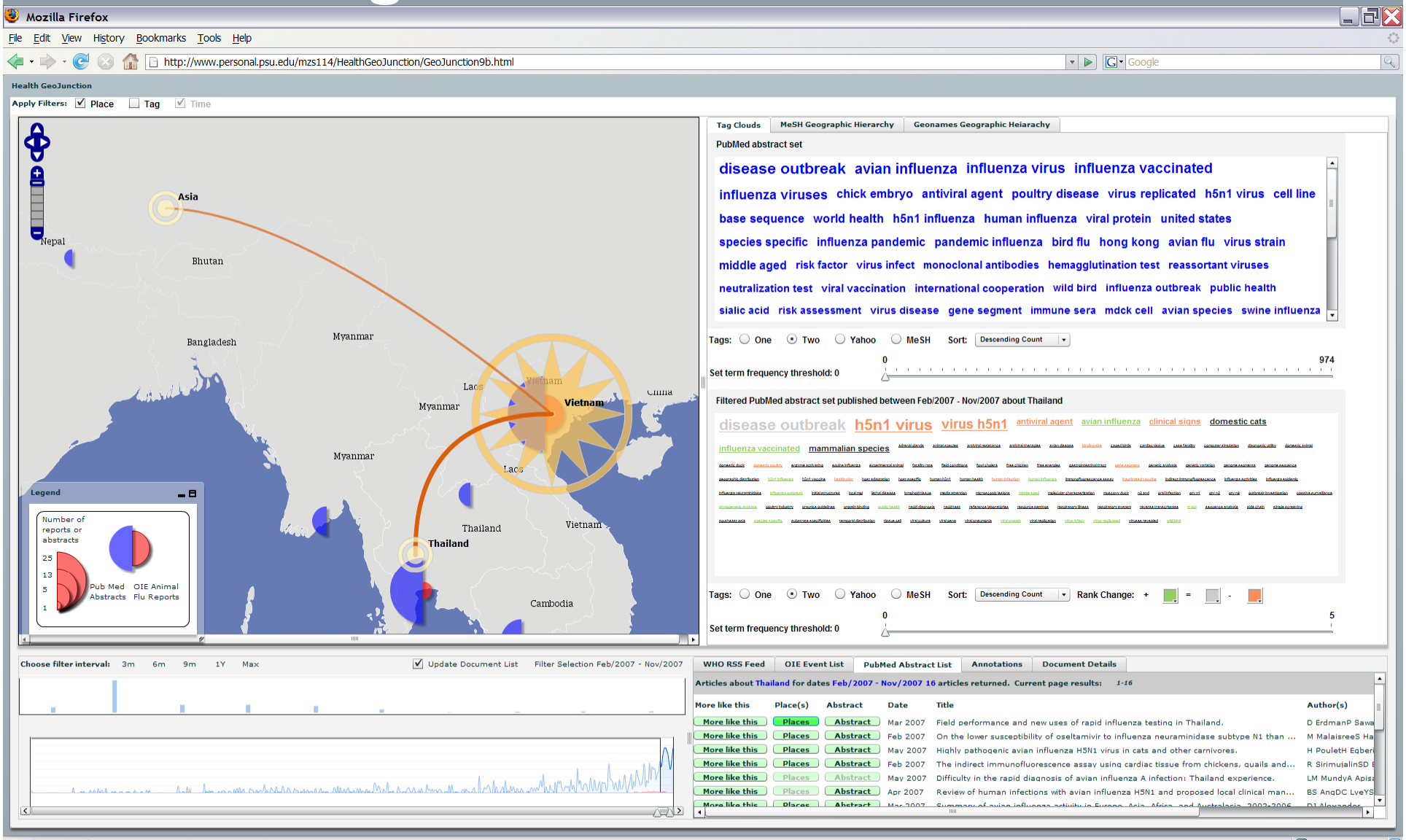
- space-time questions: given X → determine Y
  - when + where → what: Describe the objects or set of objects that are present at a given location or set of locations at a given time or set of times.
  - when + what → where: Describe the location or set of locations occupied by a given object or set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

© Alan M. MacEachren, 2008

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



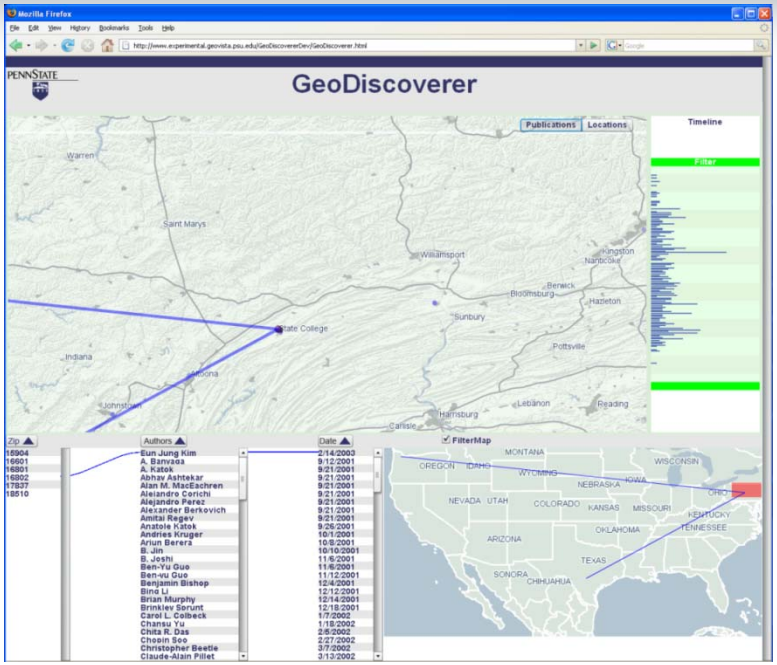
Health GeoJunction: place, time, concept filtering of health science in PubMed



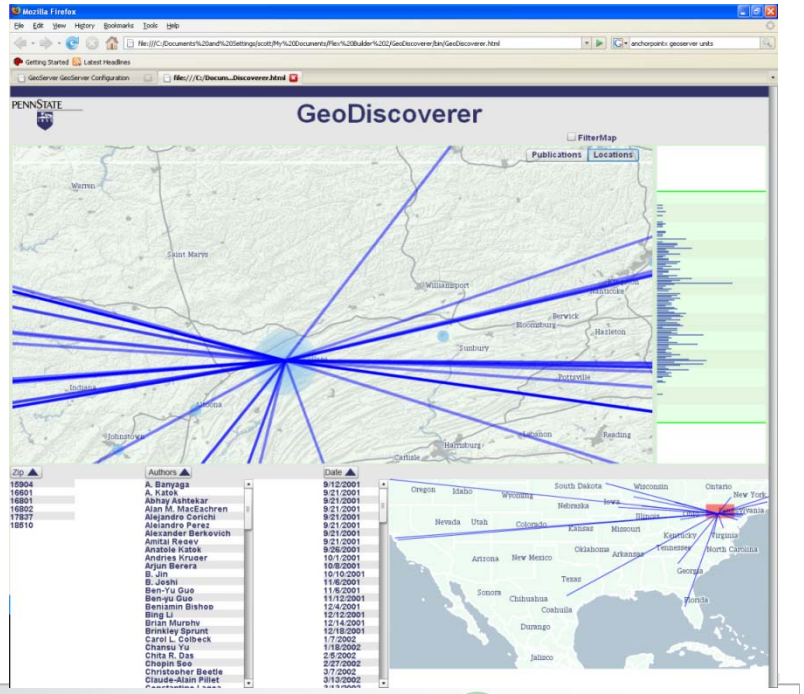
PENNSTATE

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.

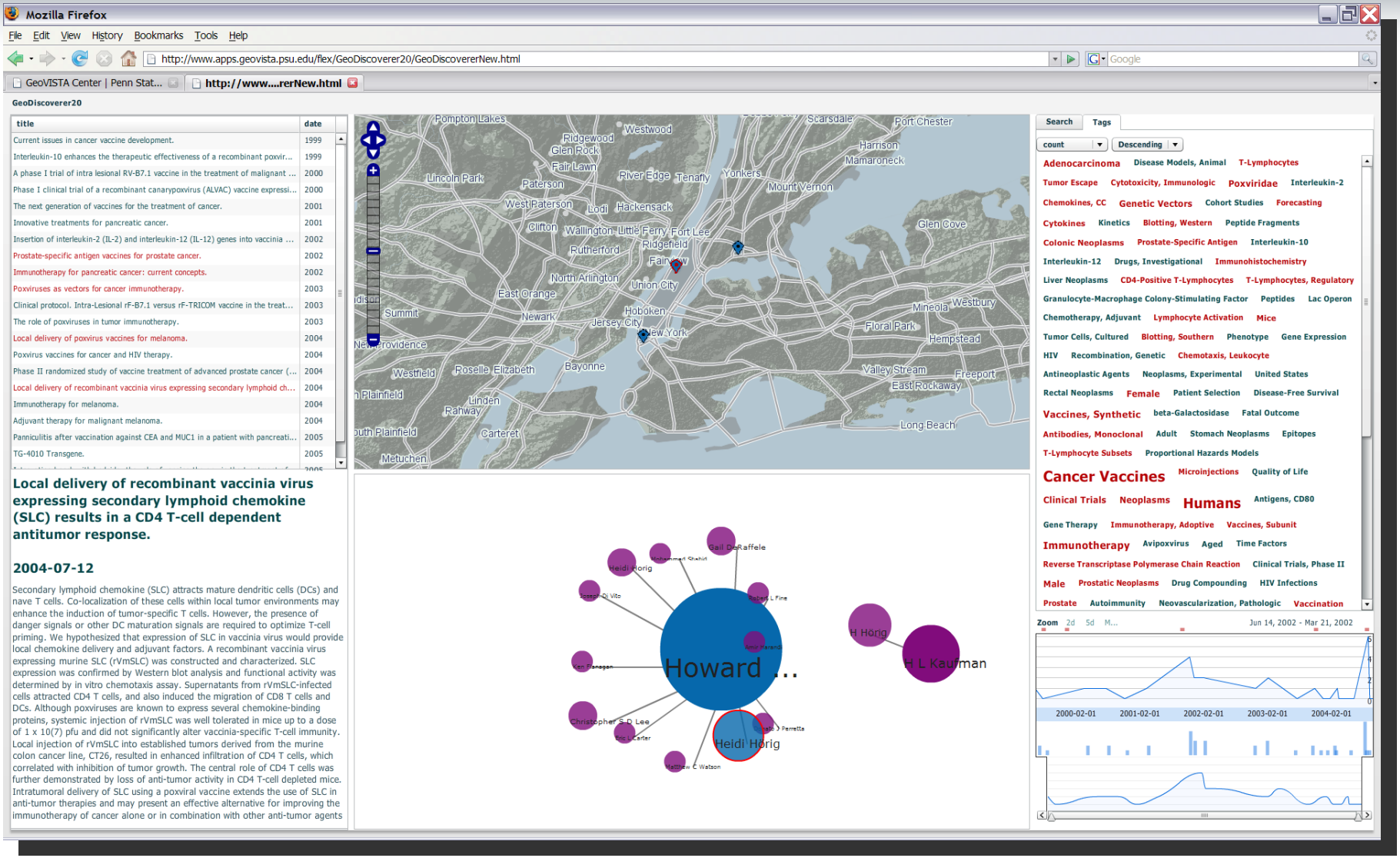


A second analyst notices a series of publications of interest generated from one institution/ location. She is curious about who the authors are collaborating with. To investigate, he uses the geographic view into the data and selects a place by clicking its location on the map. The map shows the number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

PENNSTATE

GeoDiscoverer: search engine to integrate social networks with geospatial information

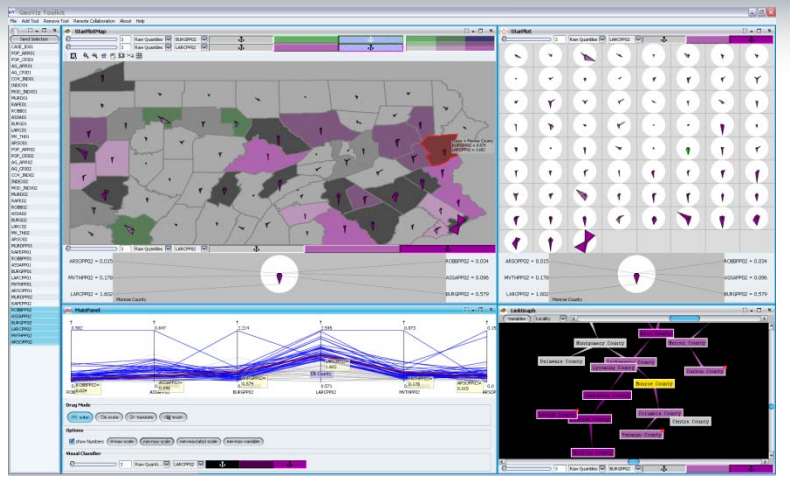
in collaboration with Lee Giles



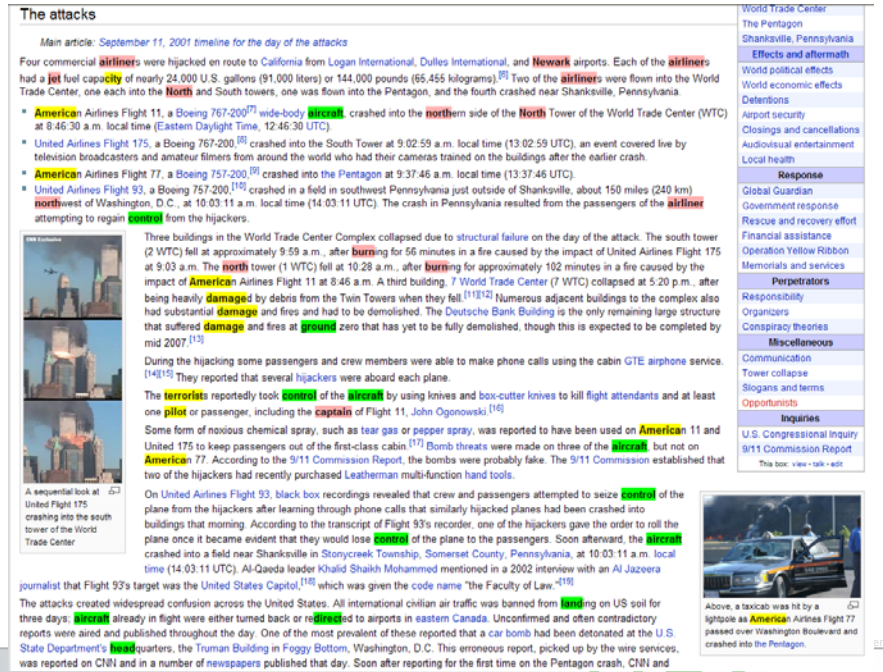
© Alan M. MacEachren, 2008

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

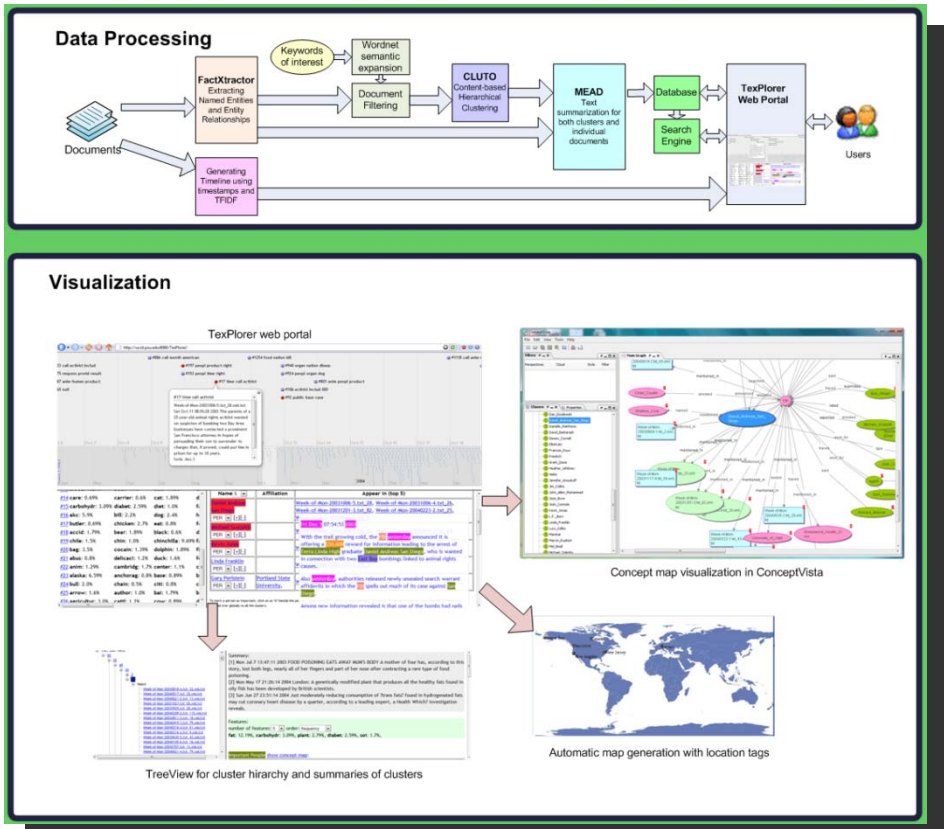


- integrated visual + computational methods applied to multivariate data
- finding the where, when, what, and who of events and relations in unstructured text



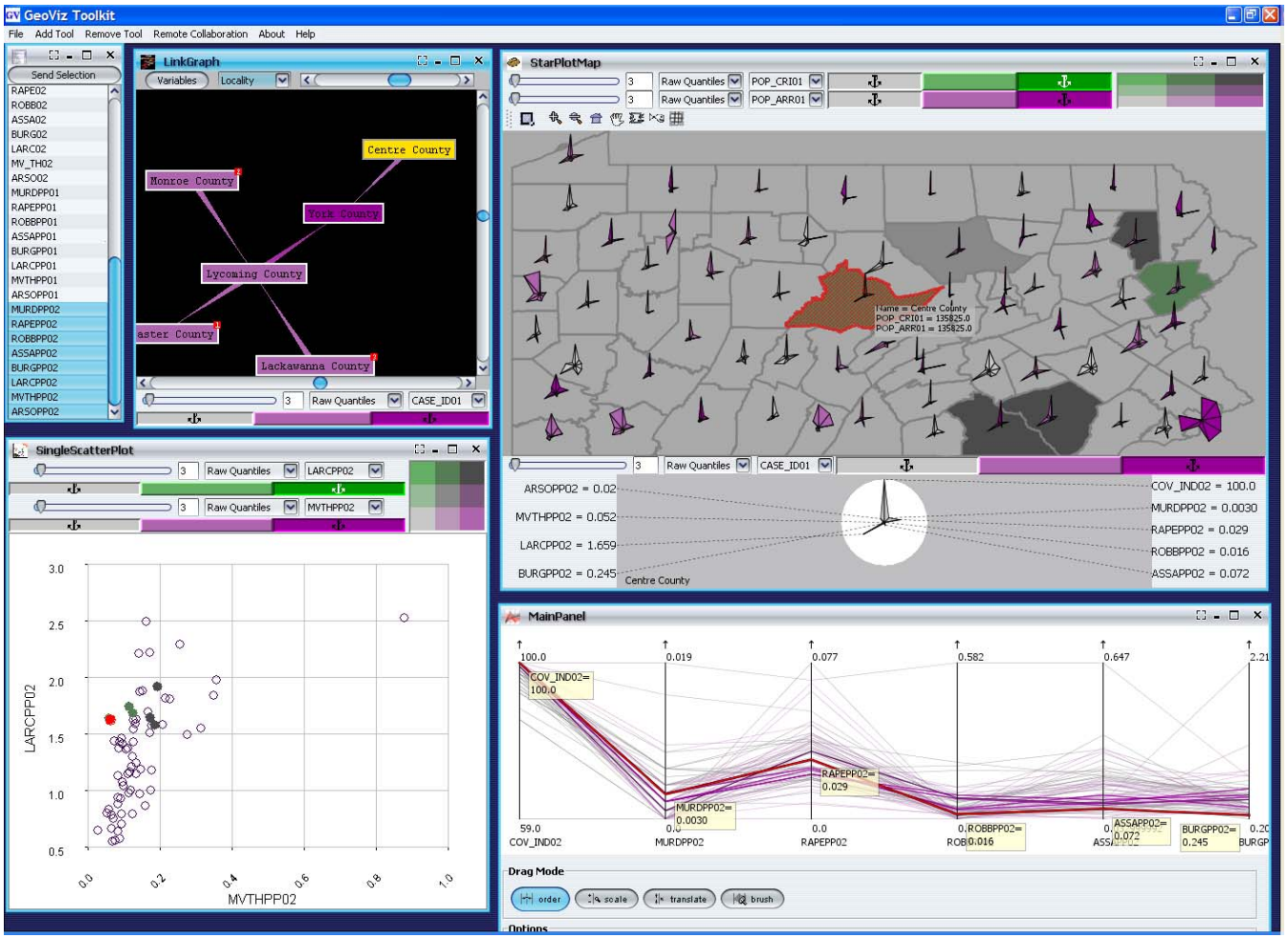
TexPlorer

- 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 & co-variate analysis

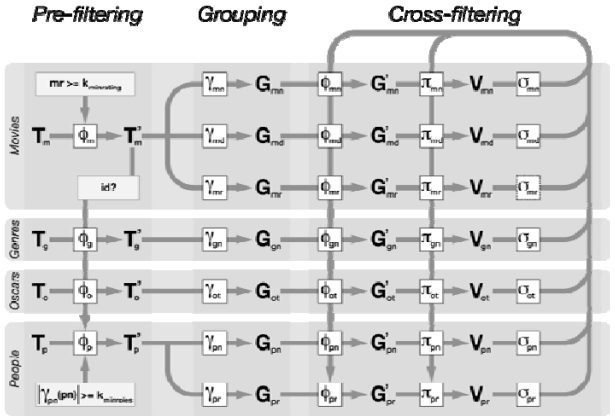


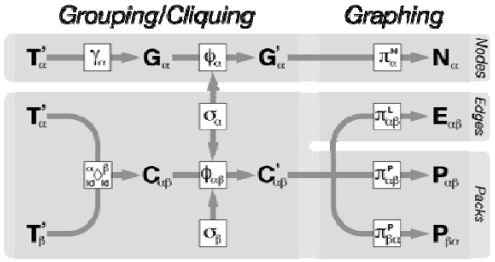
Frank Hardisty

© Alan M. MacEachren, 2008

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



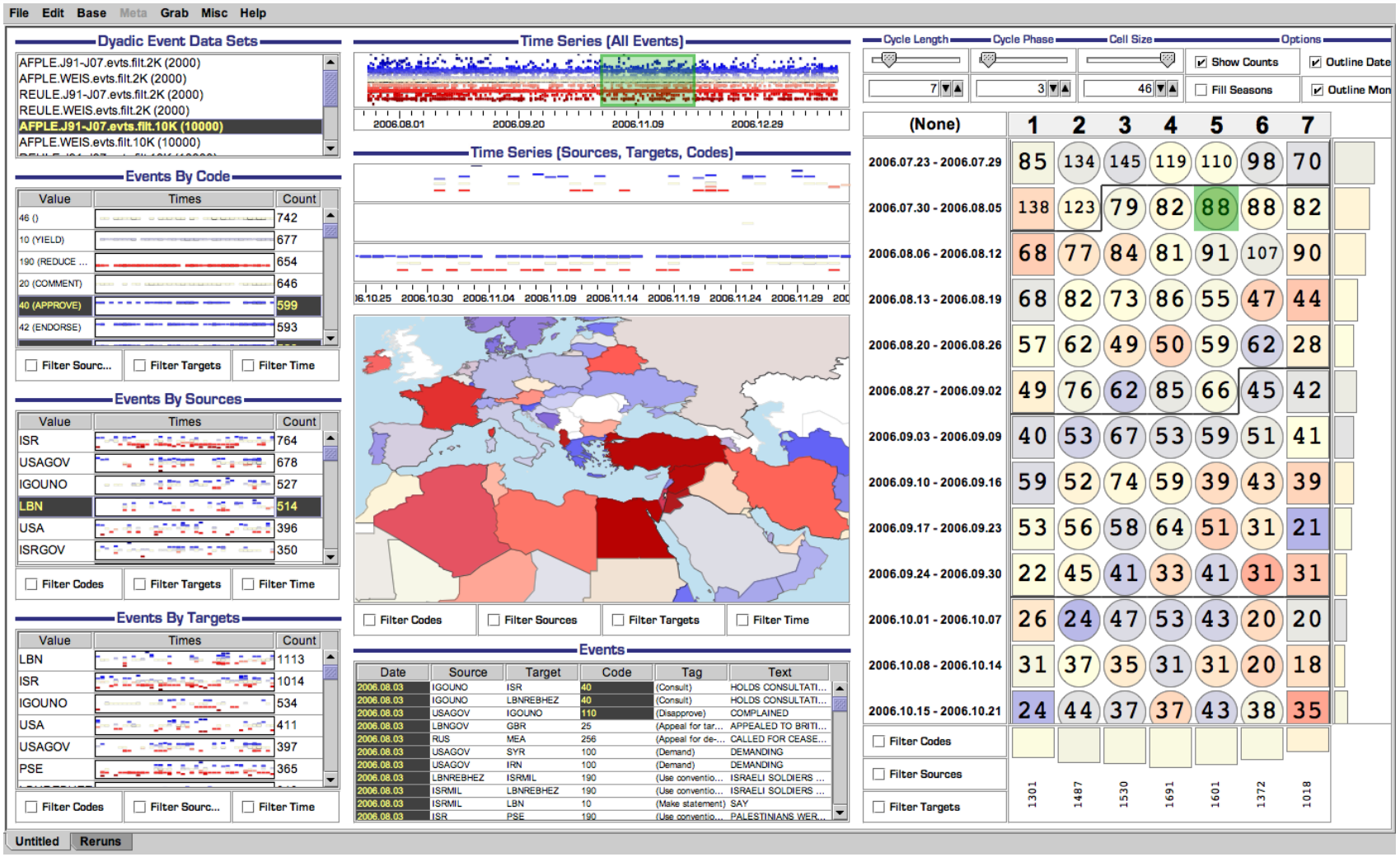


Chris Weaver

© Alan M. MacEachren, 2008

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)

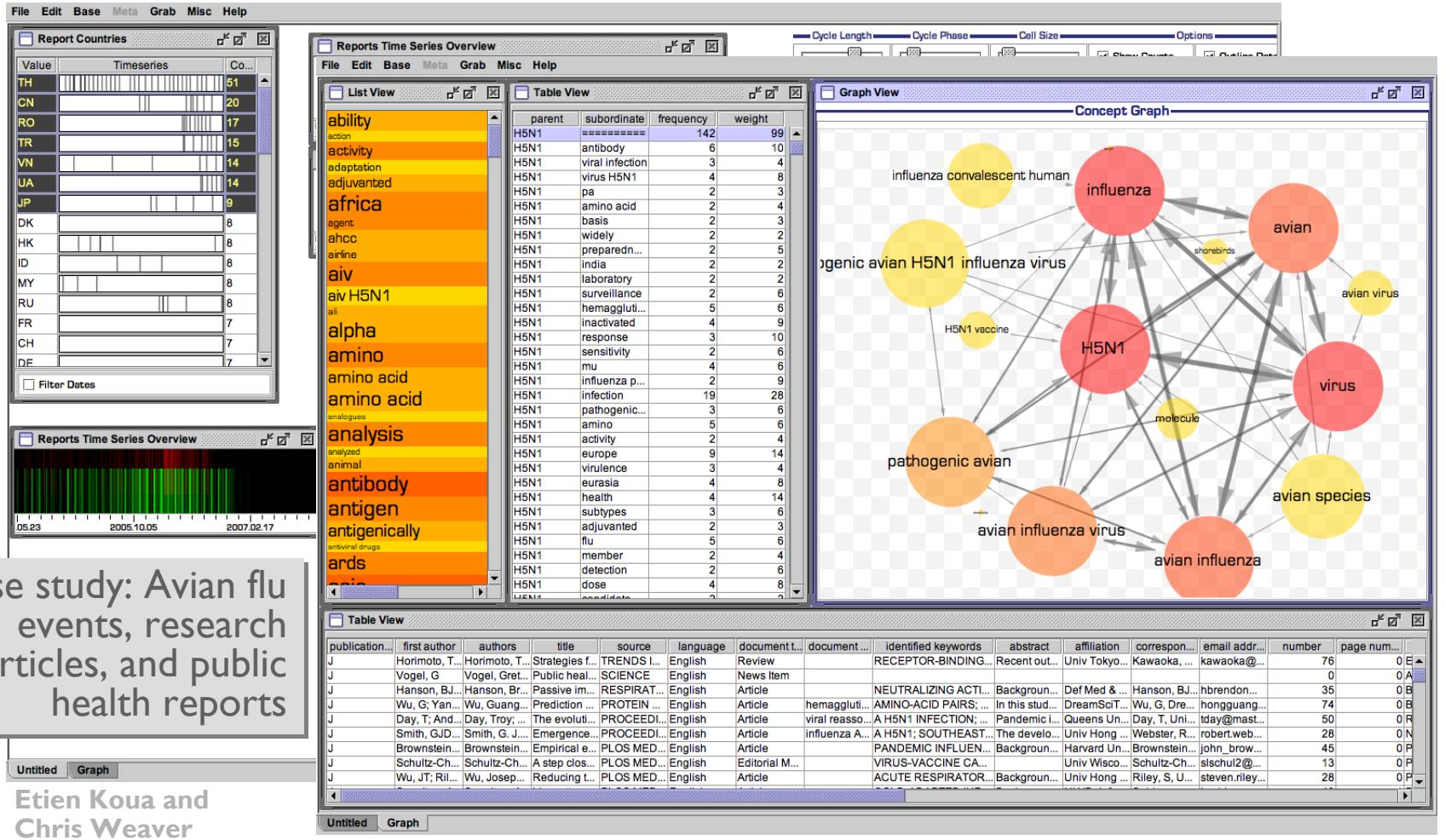
© M. MacEachren, 2008

PENNSTATE

Approach: visual analysis using group-select-filter method

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?



Case study: Avian flu events, research articles, and public health reports

Etien Koua and Chris Weaver

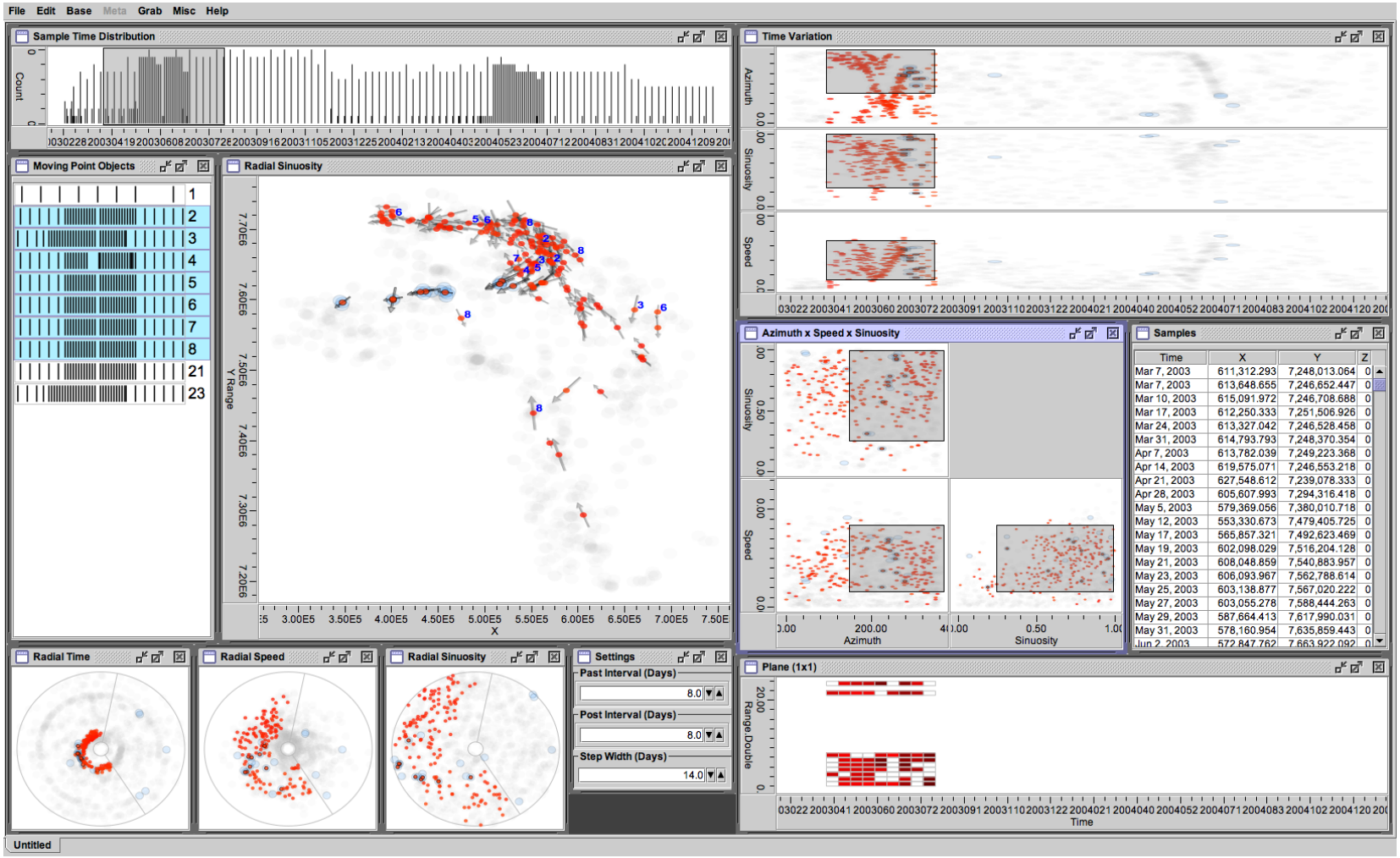
© Alan M. MacEachren, 2008

PENNSTATE

Approach: visual (group, select, filter) + computational analysis

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

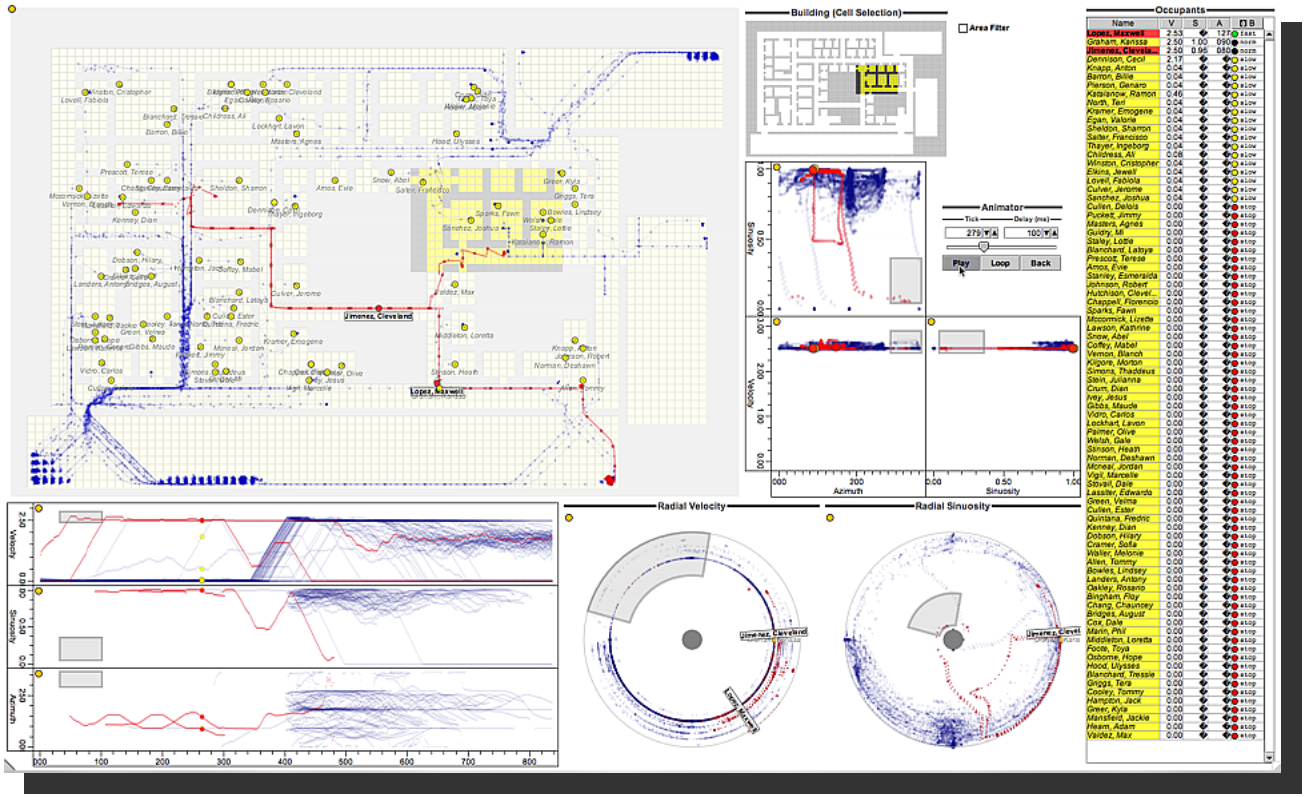
© M. MacEachren, 2008

PENNSTATE

Approach: multi-scale, visual-computational analysis

Space-Time Analytics: building evacuation

- evacuation analysis: looking for unusual behaviors, bottlenecks, impediments



Chris Weaver

© Tim M. MacEachren, 2008

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

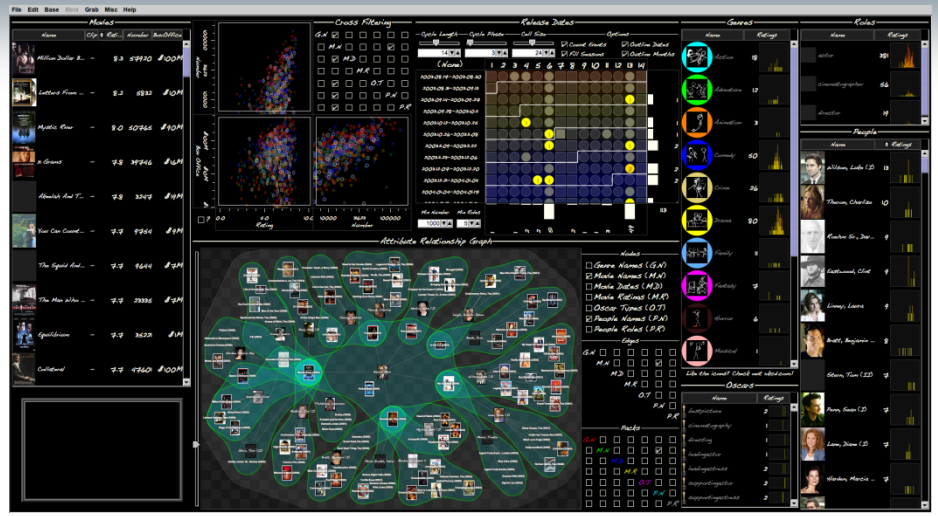
Questions?

maceachren@psu.edu

© Alan M. MacEachren, 2008

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



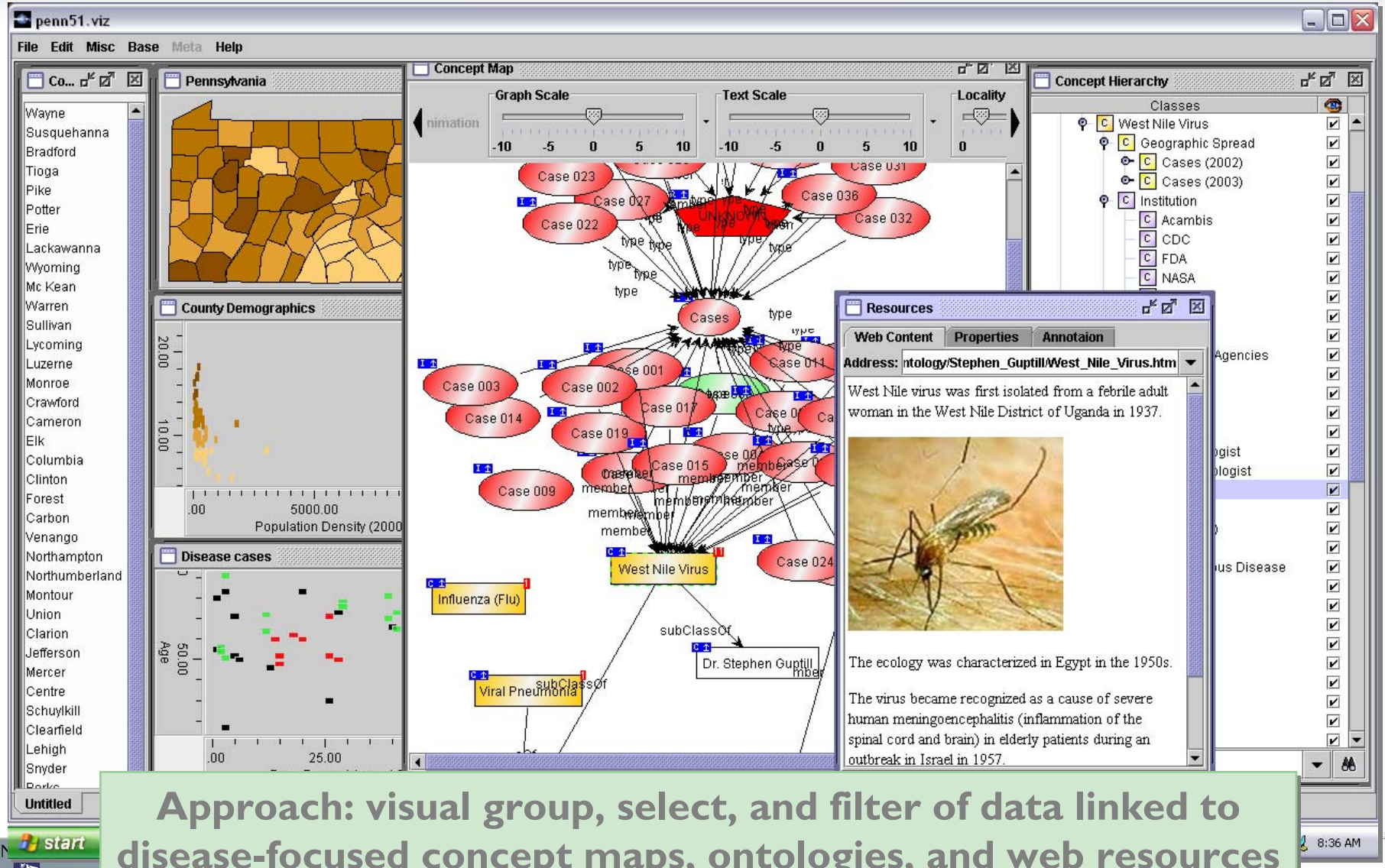
Movies for each actor who has worked with Kevin Bacon since 2001.

- 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 [?]

© Alan M. MacEachren, 2008

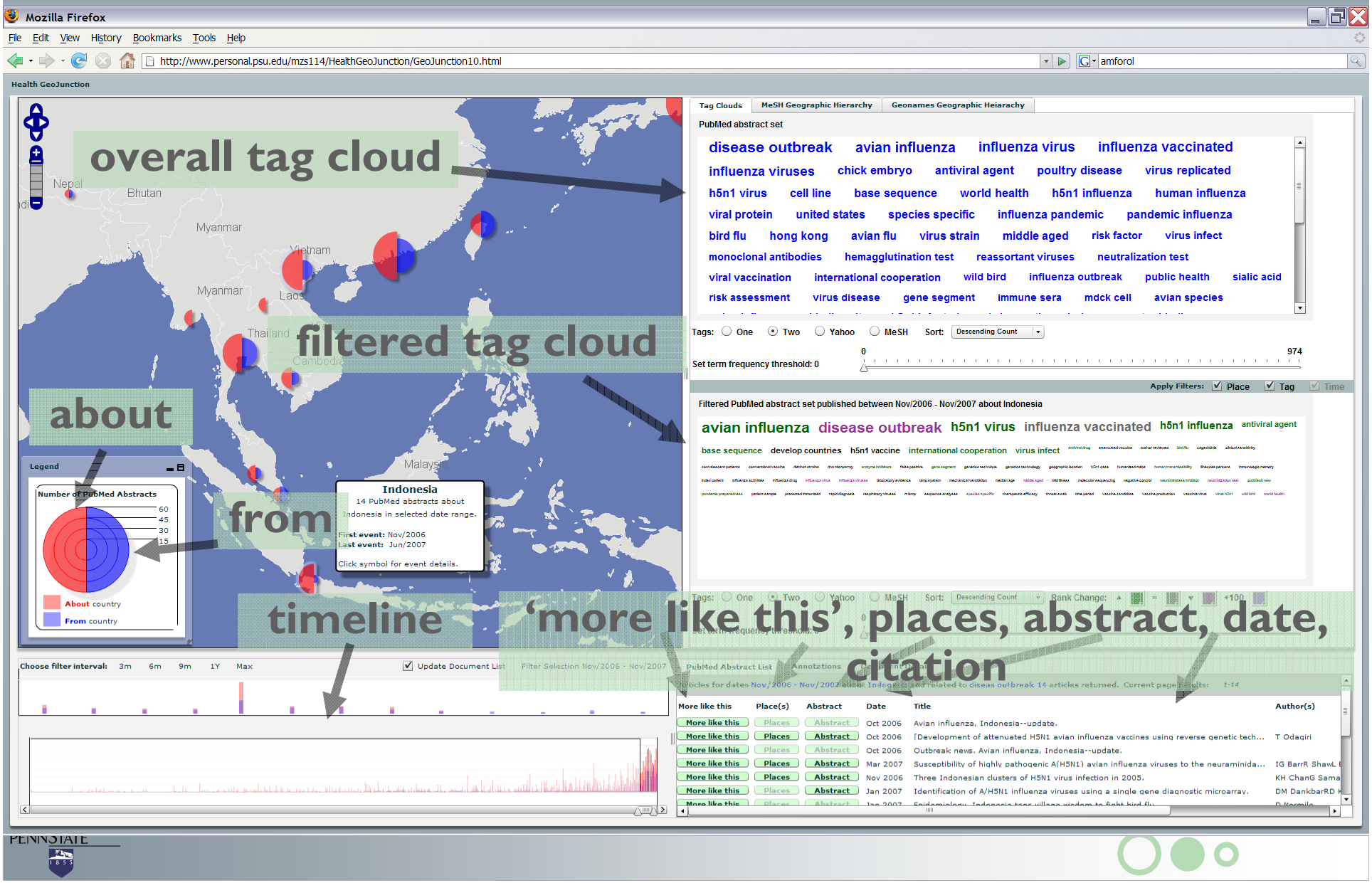
Data-to-knowledge and back: Integrating knowledge management with visual analytics

Problem: interpreting data to identify a disease event

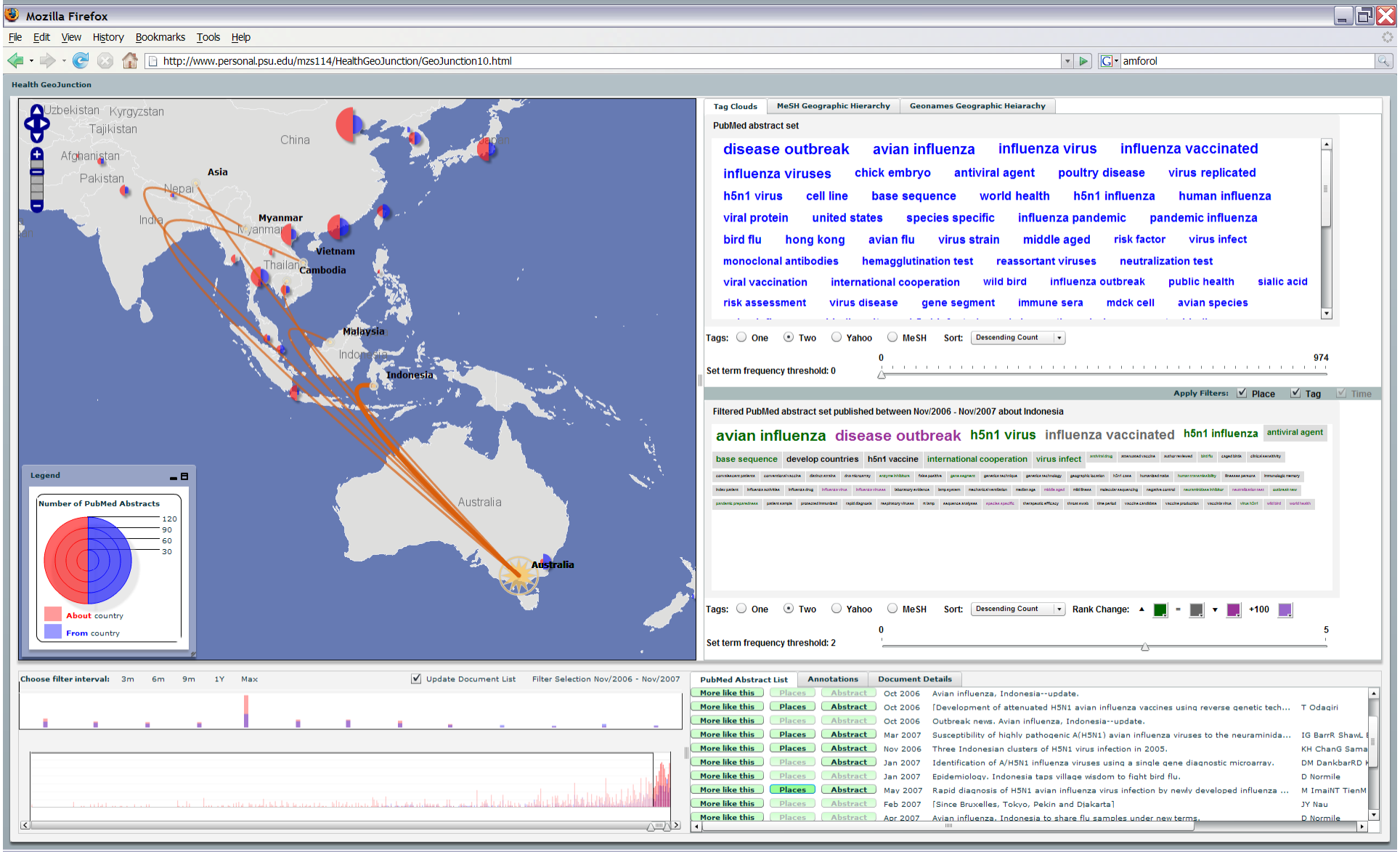


Approach: visual group, select, and filter of data linked to disease-focused concept maps, ontologies, and web resources

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

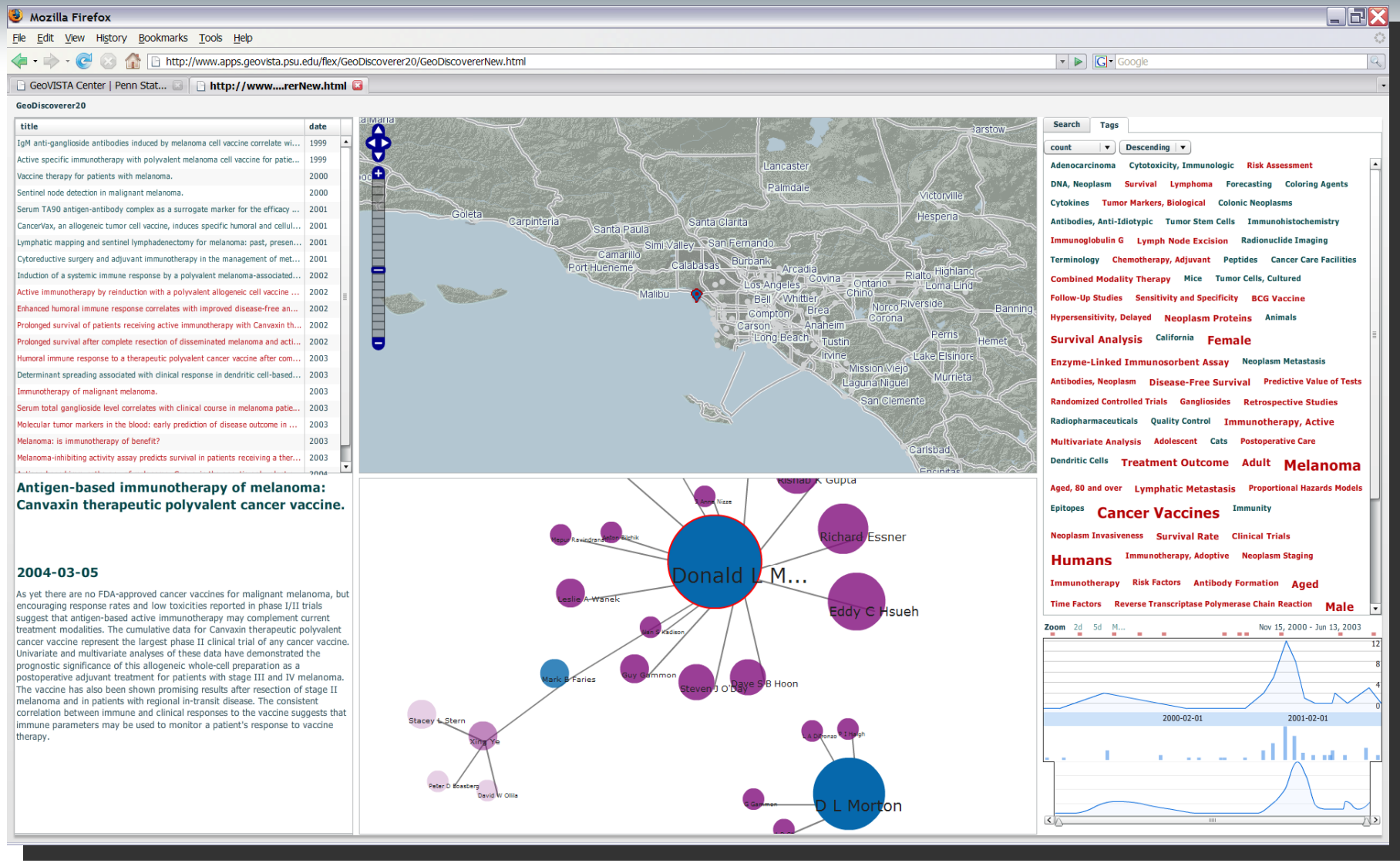
© Alan M. MacEachren, 2008

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

© Alan M. MacEachren, 2008

GeoDiscoverer



© Alan M. MacEachren, 2008

applying the typology

structure: given X → determine Y

example – elementary *when* → *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

© Alan M. MacEachren, 2008

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

© Alan M. MacEachren, 2008

applying the typology

structure: given X → determine Y

example – general *what* + *where* → *when*

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

© Alan M. MacEachren, 2008

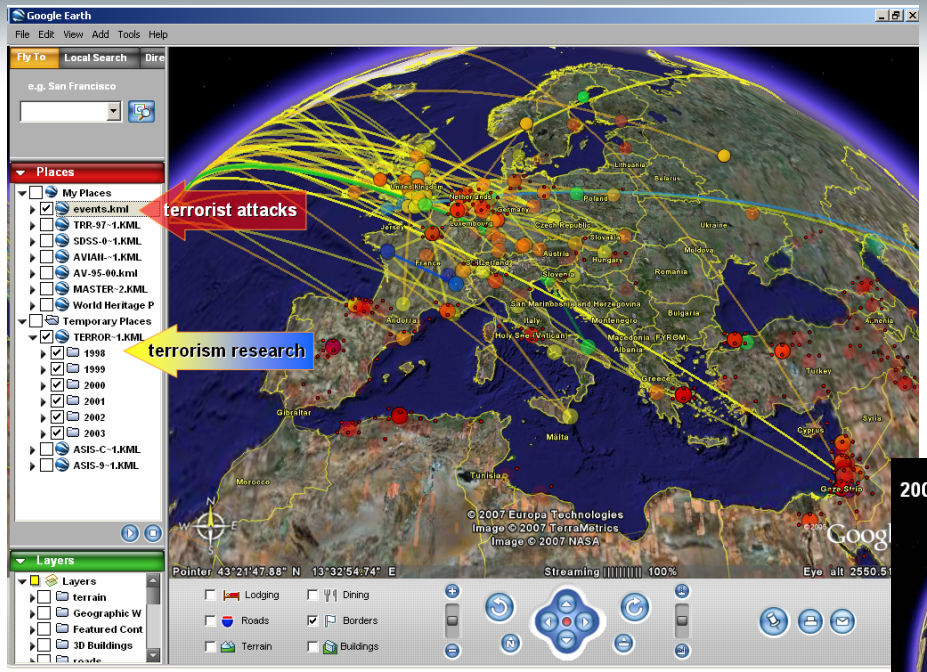
general tasks (with respect to time)

what + where → when

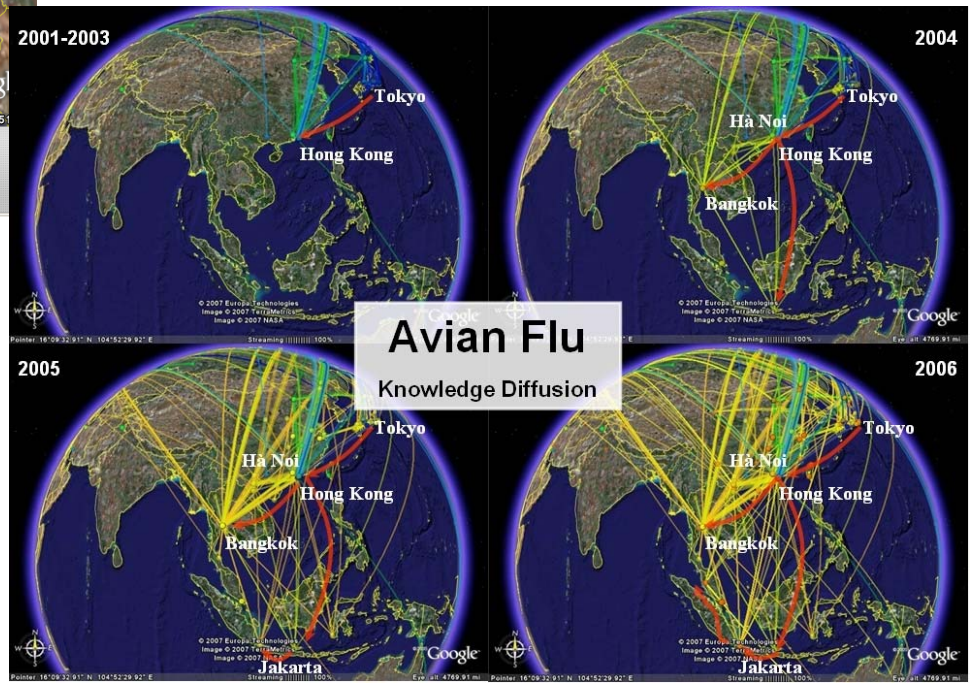
- determine time interval(s) during which a specific behavior took place
  - elementary what + where: during what time interval did stork X fly to location Y
  - general what + where: during what time interval did most of the storks fly north

© Alan M. MacEachren, 2008

Integration of geospatial and abstract structural-temporal visualization



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



movement of co-citation behavior related to avian flu research

© Alan M. MacEachren, 2008

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

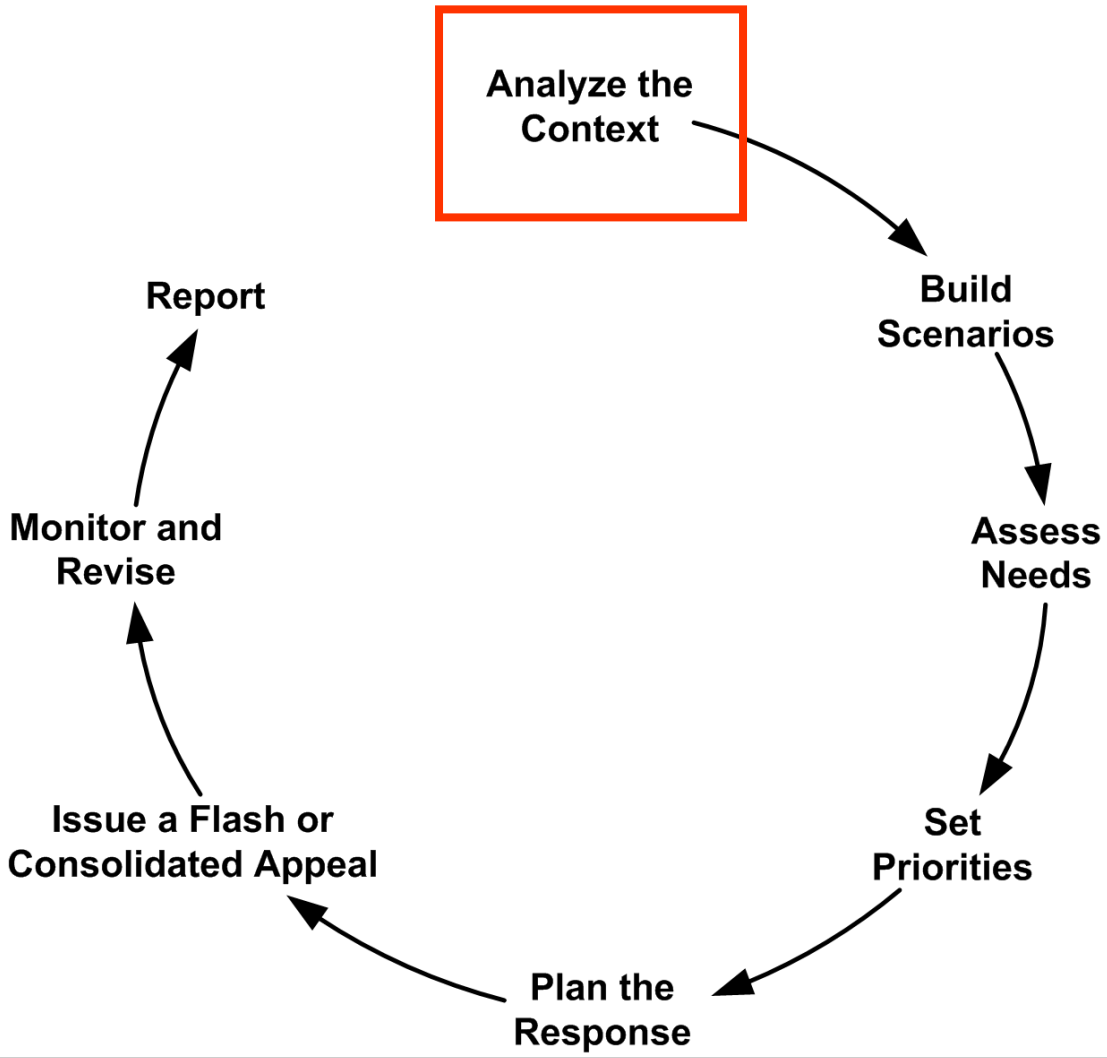
© Alan M. MacEachren, 2008

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

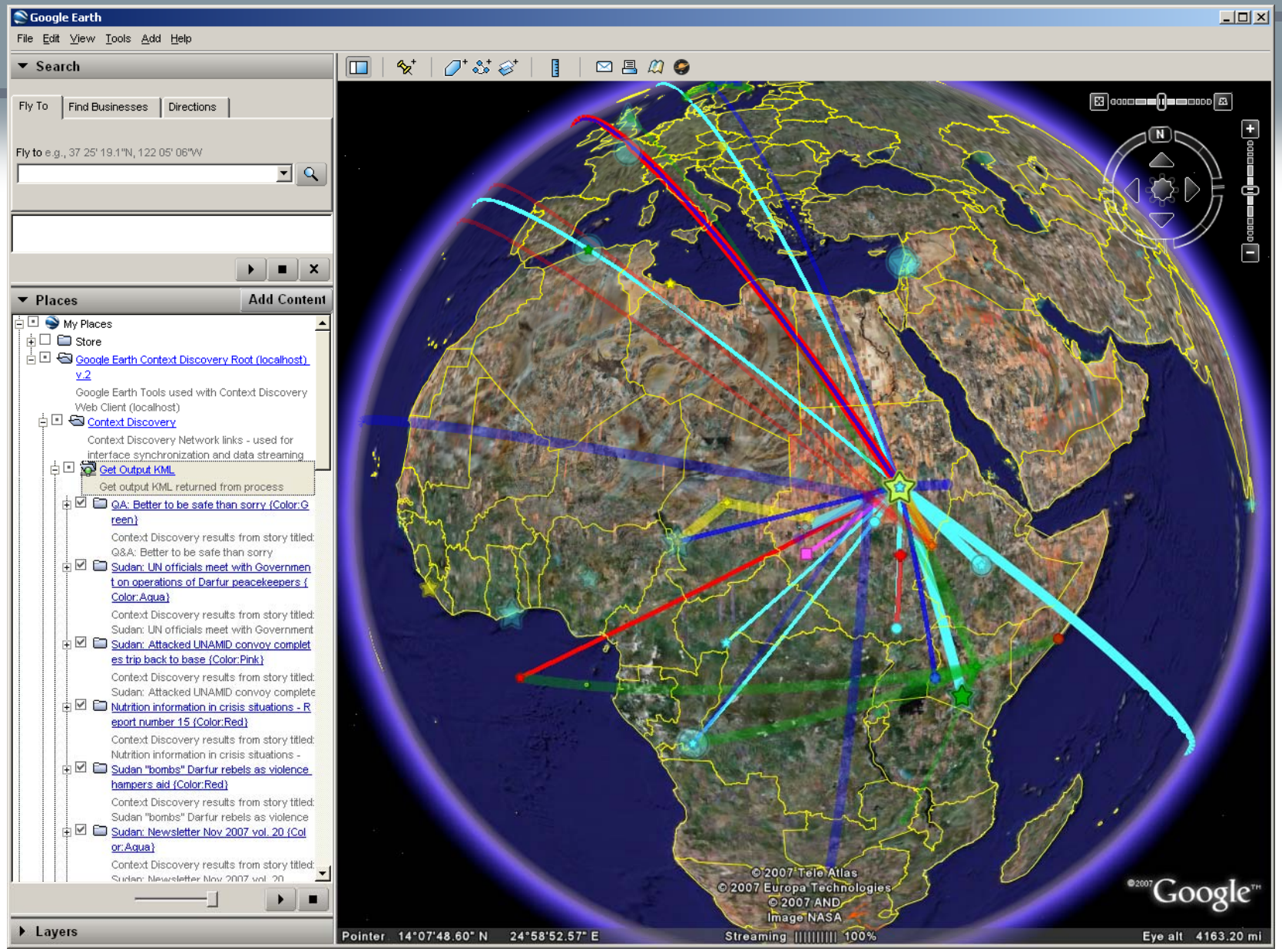
© Alan M. MacEachren, 2008

Consolidated Appeals Process (CAP) Cycle OCHA 2007

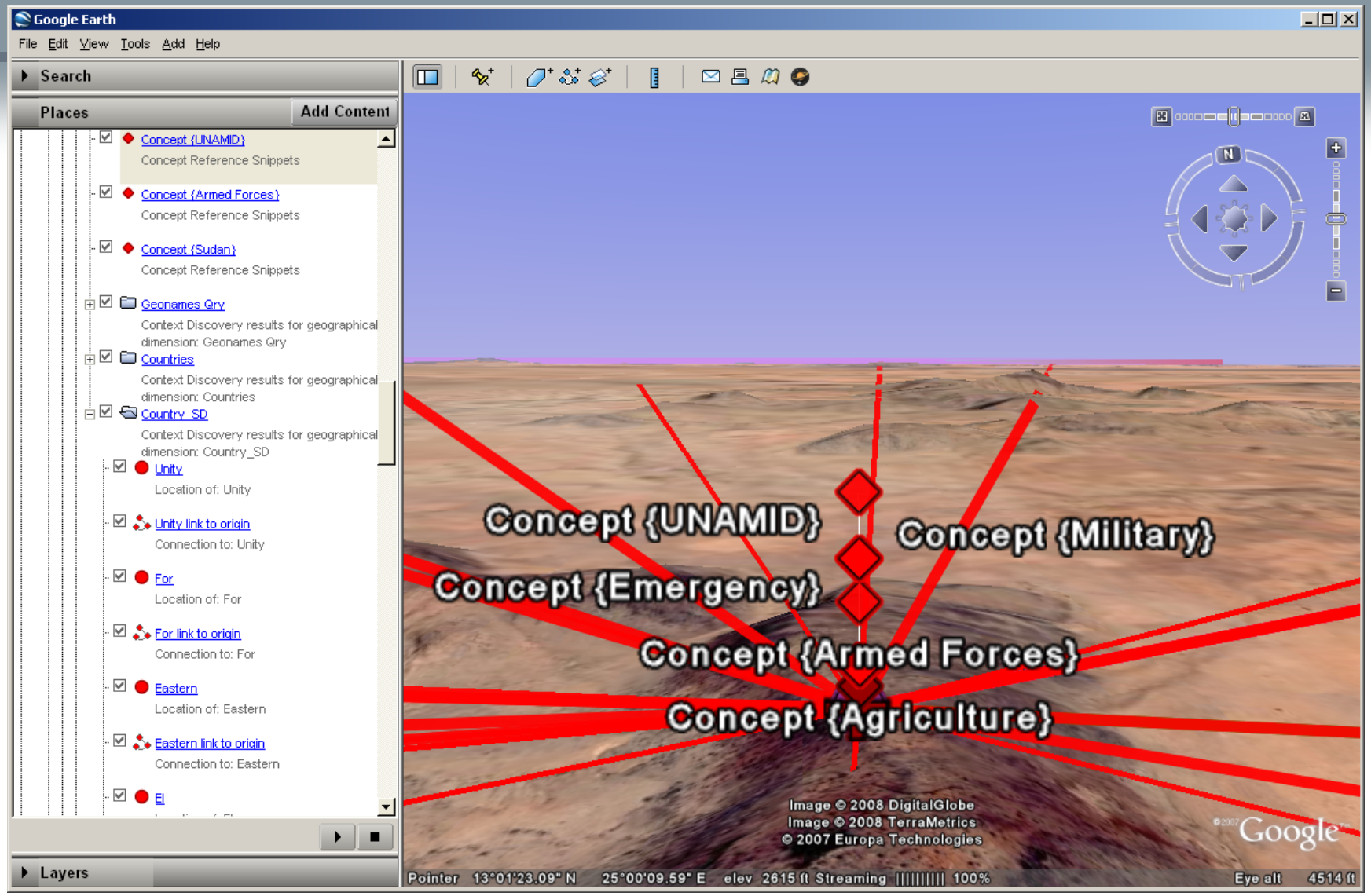


Context analysis is meant to "outline the root cause of the crisis and its evolution, including any regional or historical dimensions"

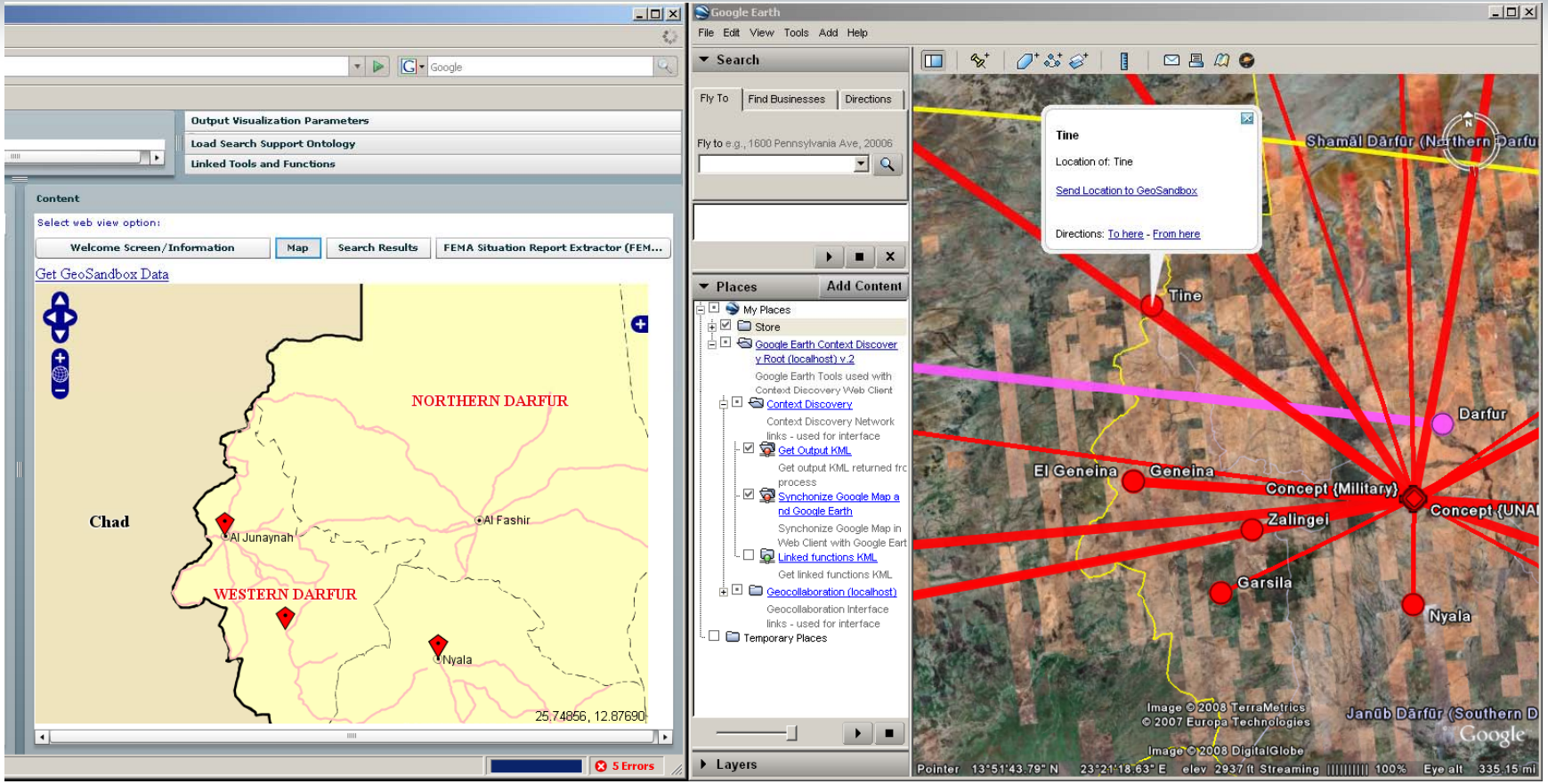
© Alan M. MacEachren, 2008



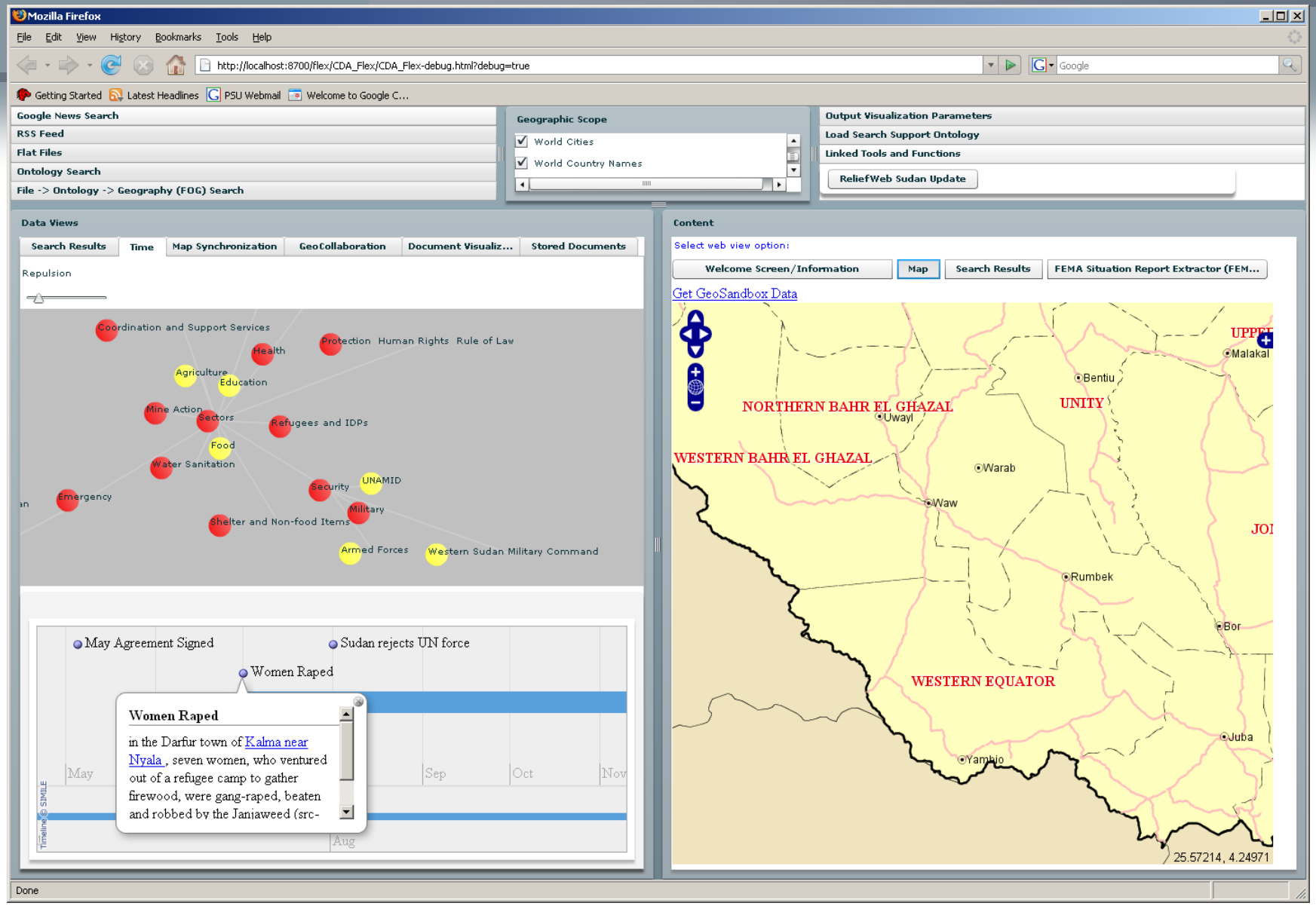
© Alan M. MacEachren, 2008



© Alan M. MacEachren, 2008



© Alan M. MacEachren, 2008



© Alan M. MacEachren, 2008

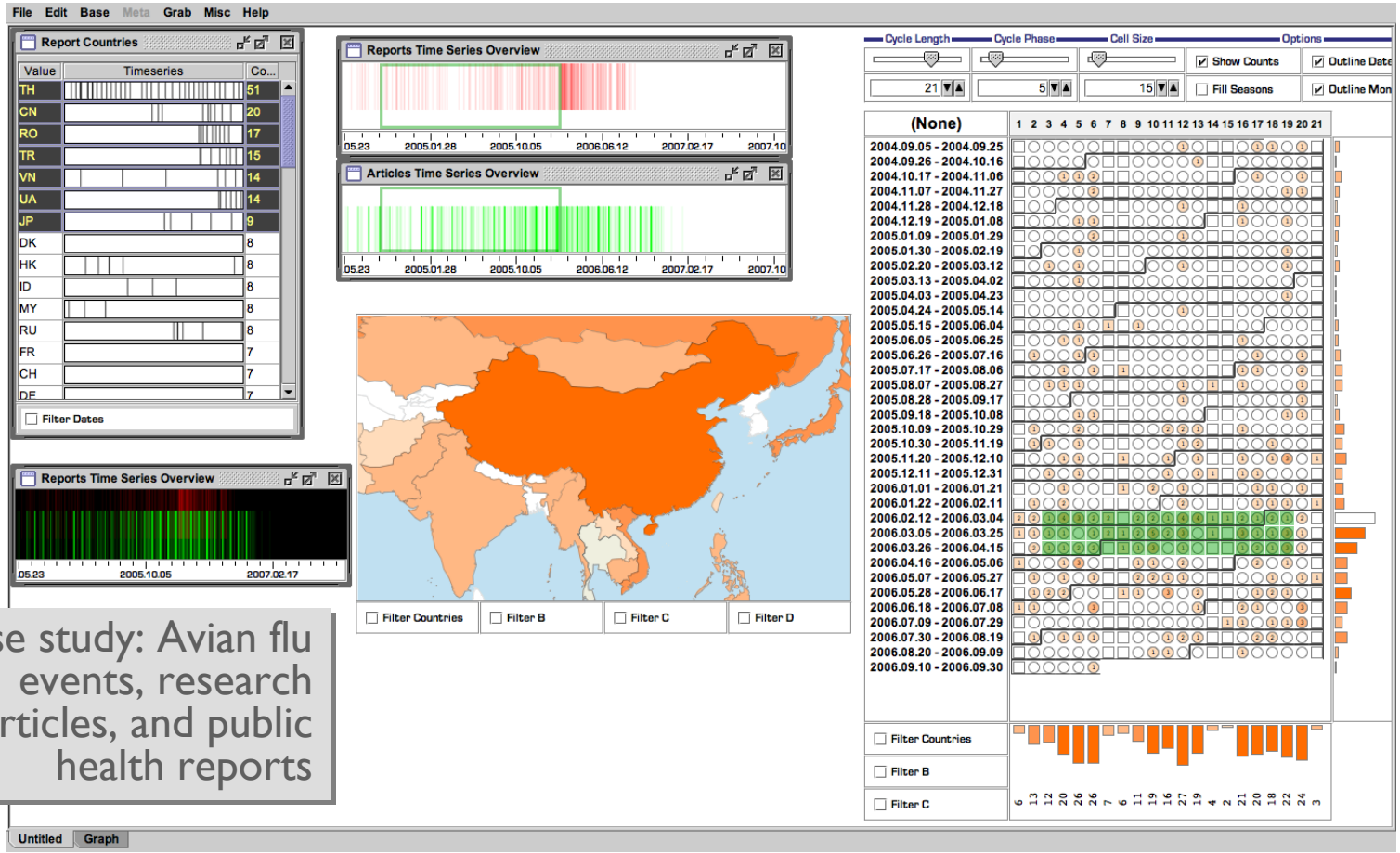
Geo-contextualizing to enable situation assessment → Situation Awareness

- 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

© Alan M. MacEachren, 2008

Geo-temporal relationships among multiple data channels

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



Case study: Avian flu events, research articles, and public health reports

Etien Koua and Chris Weaver

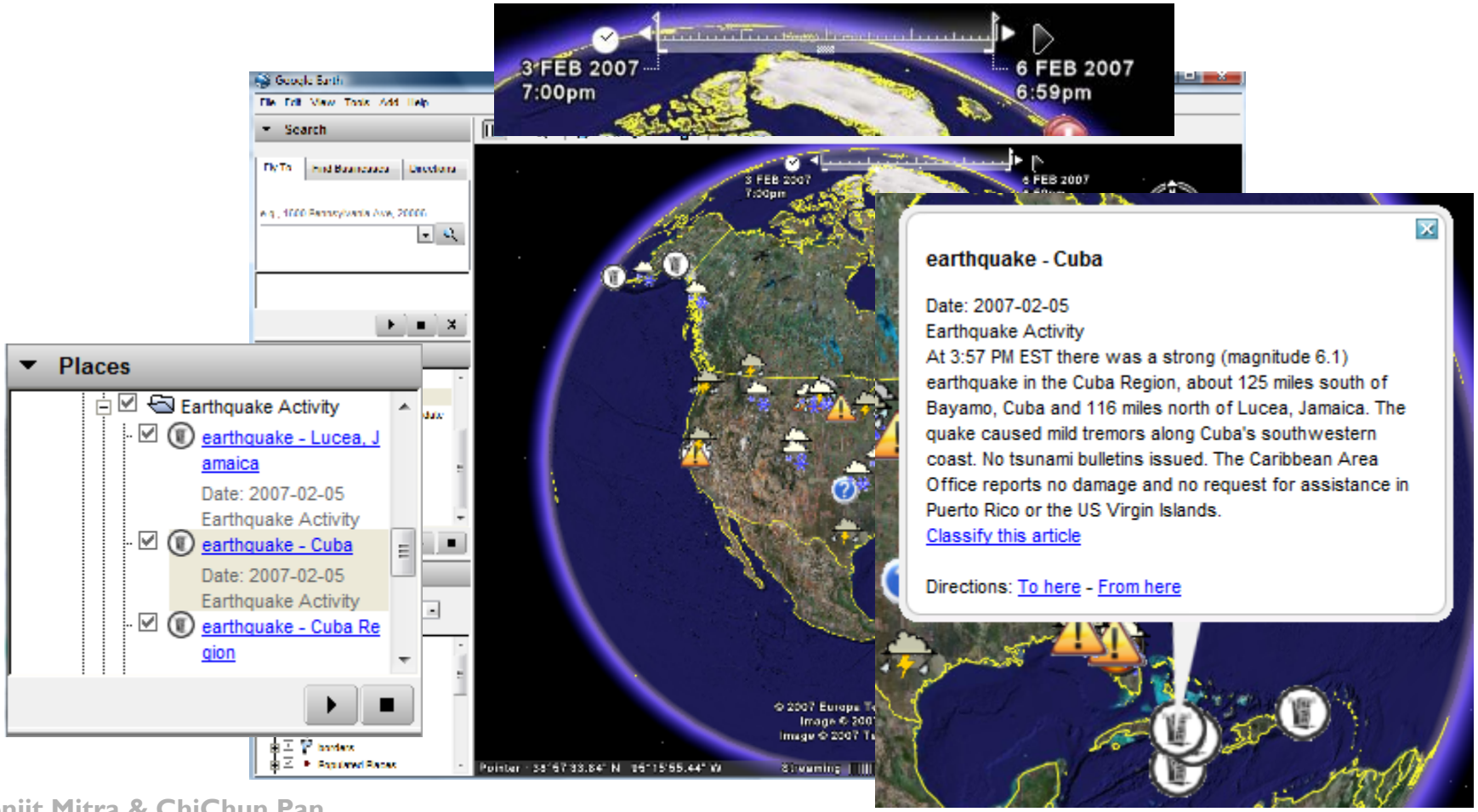
© Alan M. MacEachren, 2008

PENNSTATE

Approach: visual (group, select, filter) + computational analysis

FEMARepViz

- example info review session

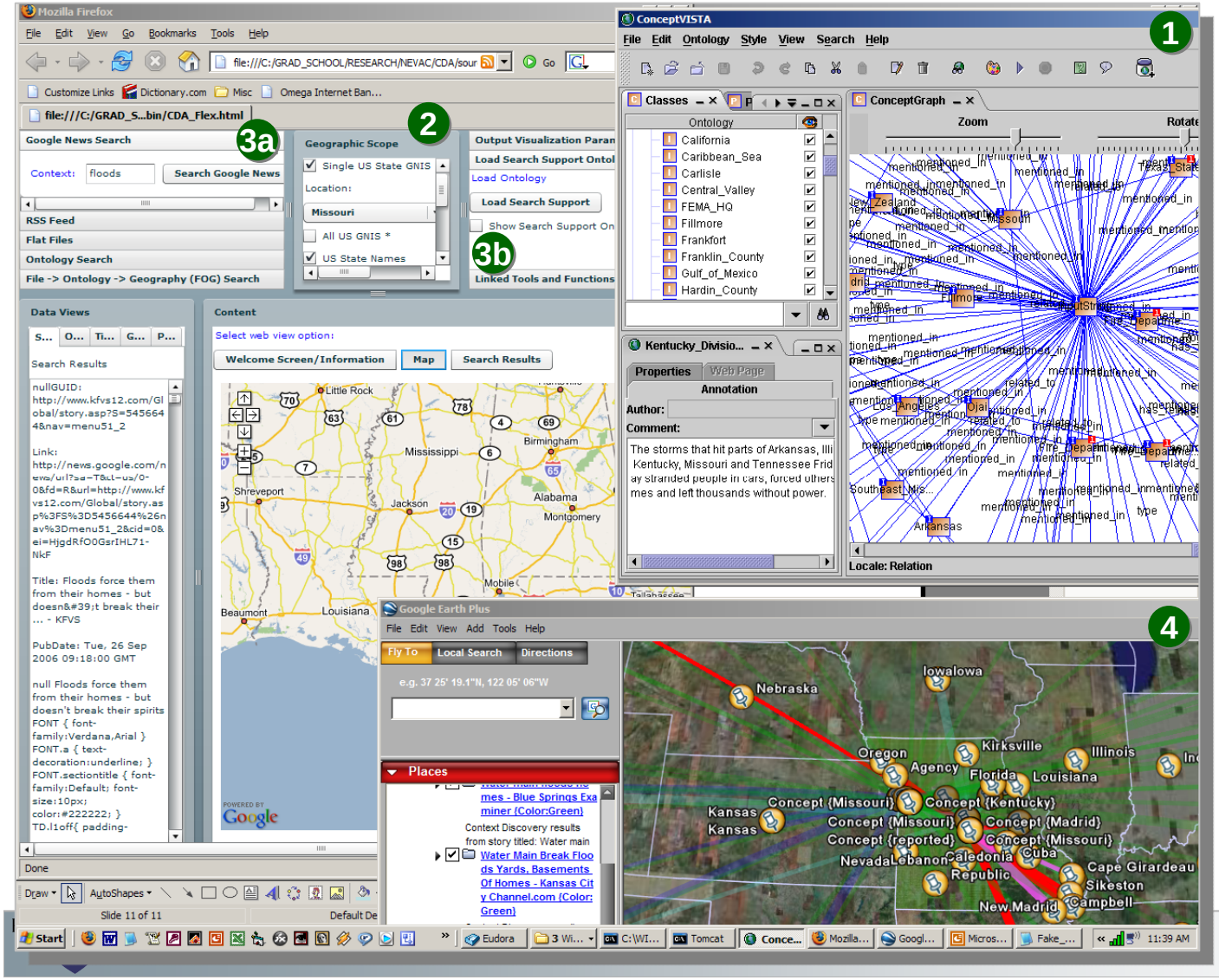


Presenjit Mitra & ChiChun Pan

© Alan M. MacEachren, 2008

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

Brian Tomaszewski, Alan MacEachren