

Geovisualization of News Stories

Julie Dillemath

PhD Candidate

Geography Dept.

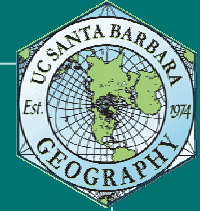
UC Santa Barbara

February 27, 2007

OSU Geovis Series

Collaborator:

**Alex Villacorta,
Statistics and Applied
Probability**



Interactive Digital Multimedia IGERT, UCSB

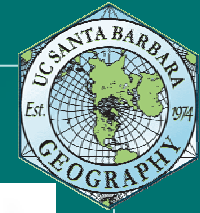
11 departments

Elec. & Comp. Engineering – Media Arts & Tech. – Geography –
Psychology – Music – Statistics – Computer Science –
Education – Film Studies – Art Studio – Envi. Science

30 students, 20 faculty

3 collaborative projects

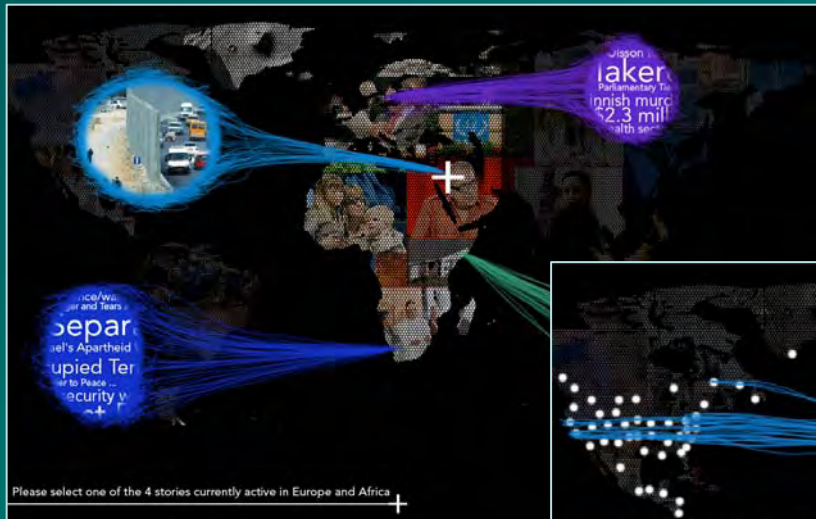
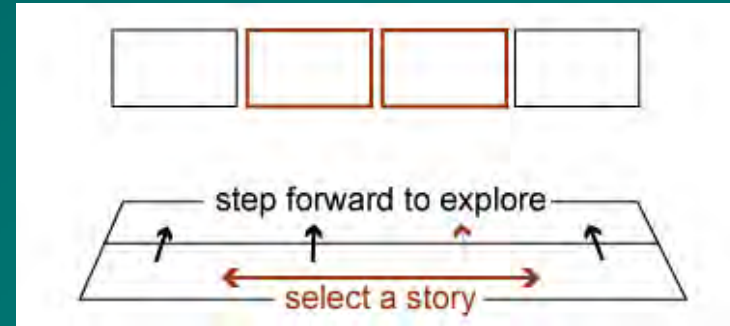
<http://media.igert.ucsb.edu/>



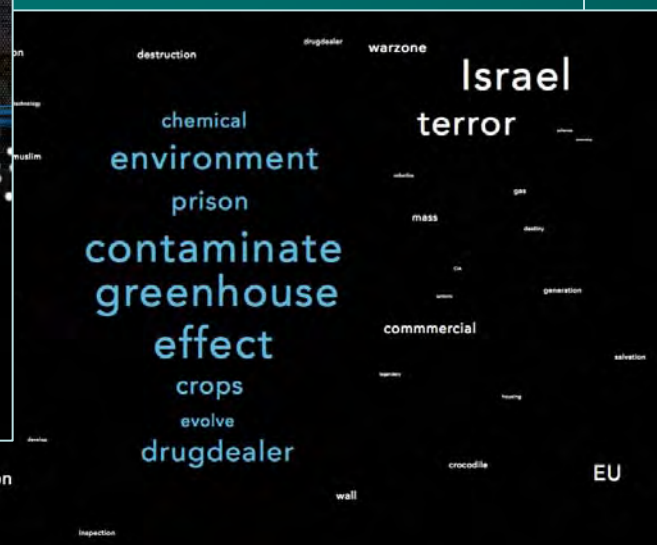
Projects

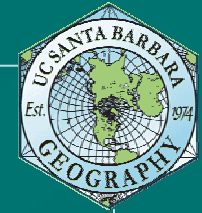
Spheres of Influence II

Geospatial and news data integration & visualization with a camera-tracking interface



MAT, ECE, Geography,
Statistics, Psych, CS, Art

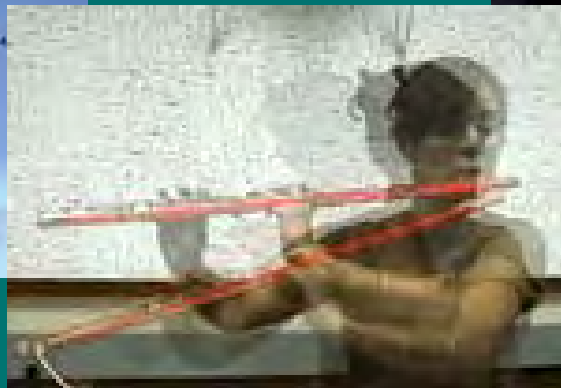




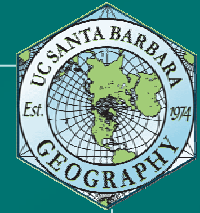
Projects - 2

Out of the Ether II

Audio, video, and electric field sensors for cueing control during live musical performance



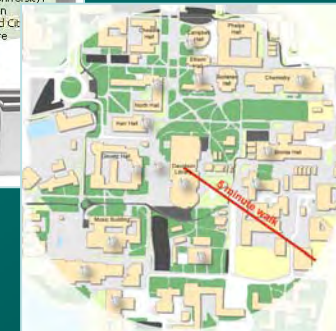
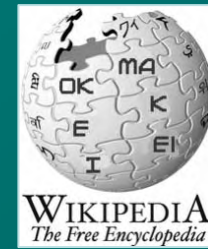
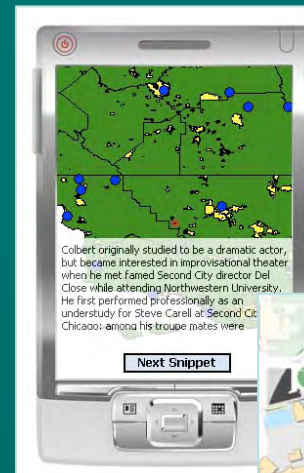
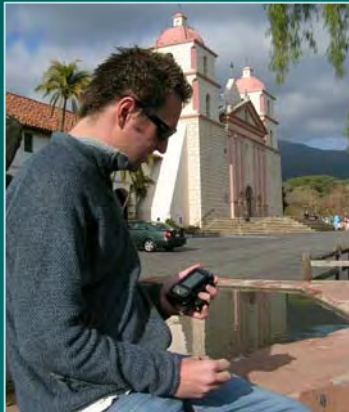
MAT, ECE, Music



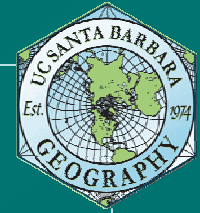
Projects - 3

Minotour: a mobile, spatially aware Wiki

A PDA-based tour guide that weaves a narrative for a user's journey based on Wikipedia article snippets



Geography, Psych, Film Studies, CS



Overview

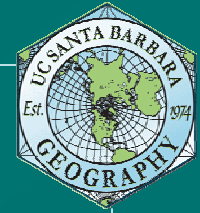
Driving Questions

Big Picture

Our approach: single story case study
national front-page stories

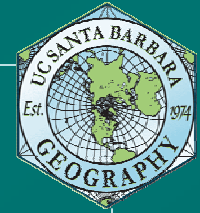
‘Spheres of Influence’ project

Driving Questions



How is news reported around the country; who has access to what news stories and when?

- How can semantic meaning and geography be related through visualization?
- Does news spread like a disease; can it be modeled?



Big Picture

3 'spaces' of news events

- geographic
- temporal
- semantic

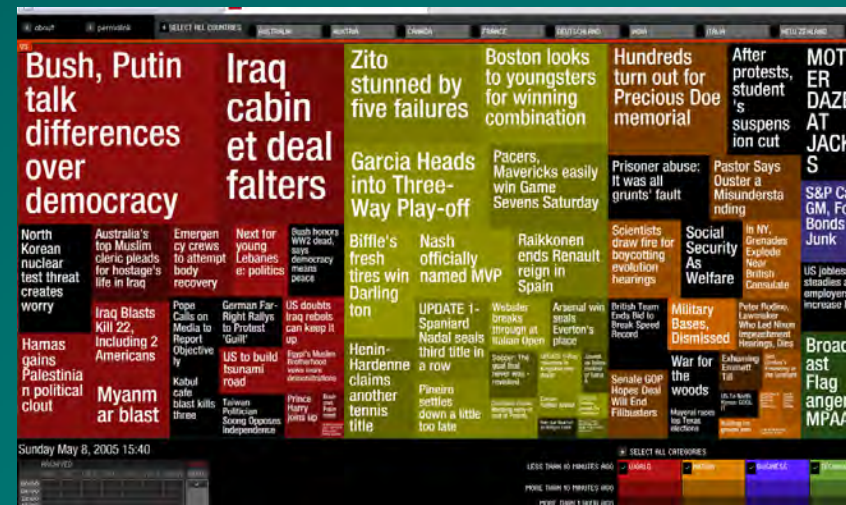
Geovisualization for:

- data exploration & pattern i.d.
- presentation to broader audience

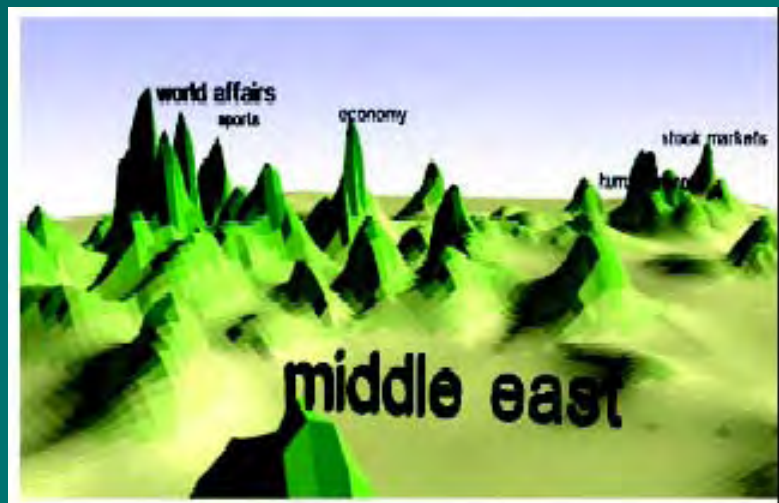
Spatial -- Temporal -- Semantic
analysis of news stories



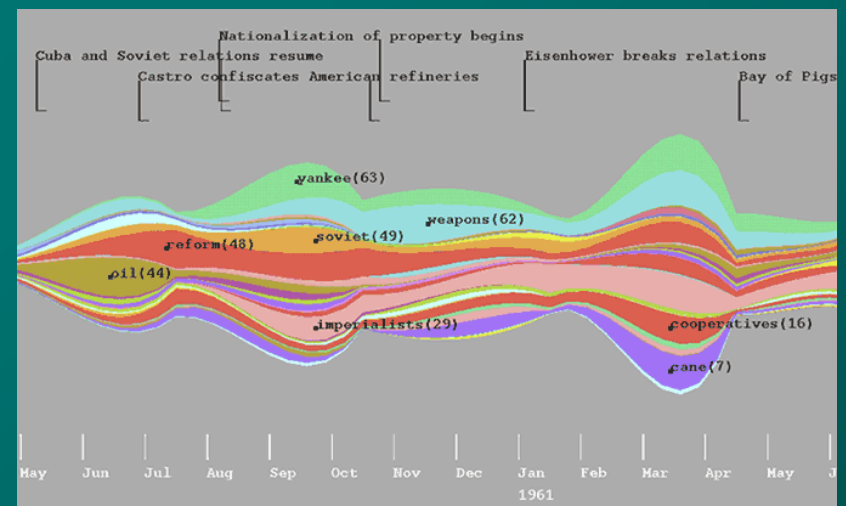
Buzztracker
www.buzztracker.org



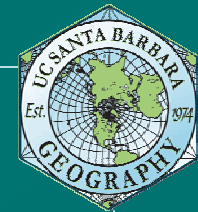
Newsmap
www.marumushi.com/apps/newsmap



Spacecast
geog.ucsb.edu/~sara



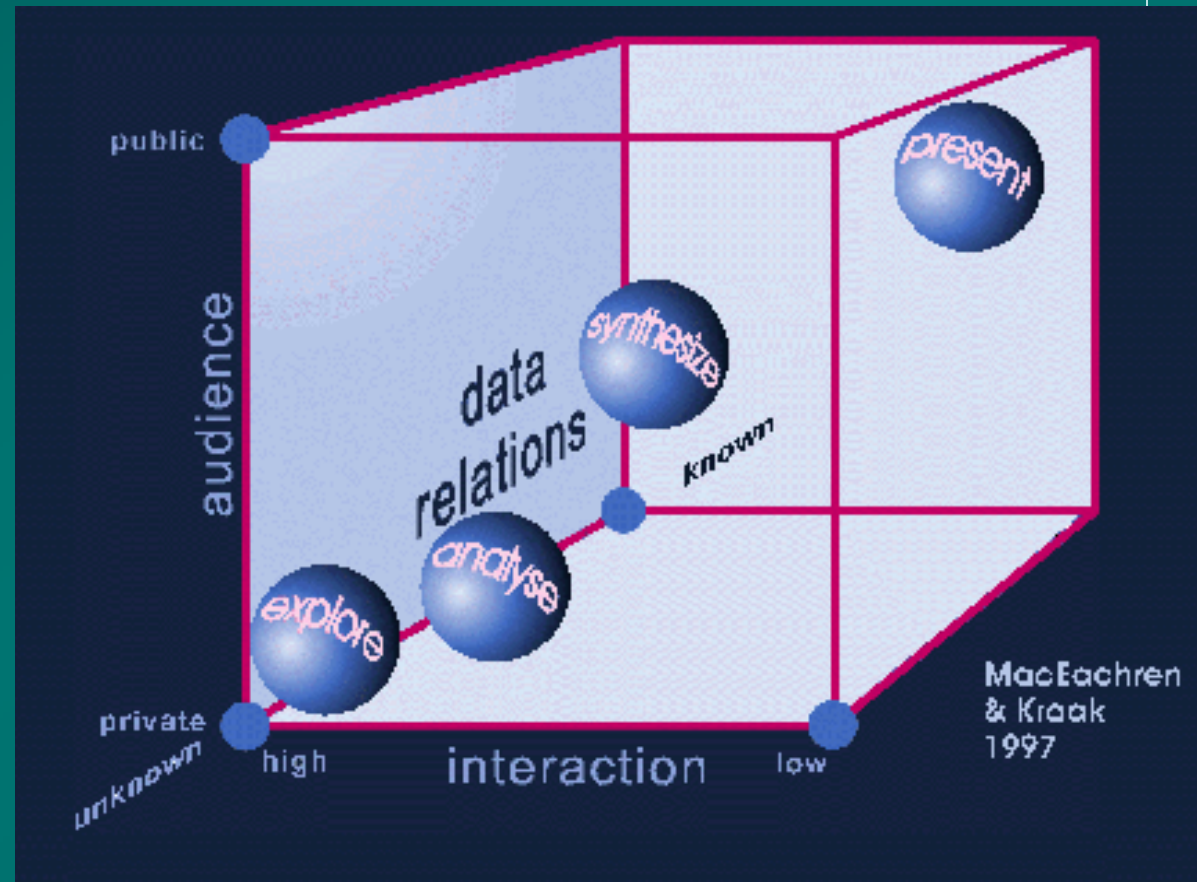
ThemeRiver
www.pnl.gov/infoviz/technologies.html



Geoviz

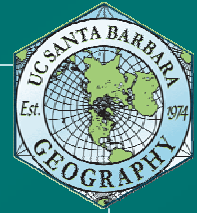
Graphic display
of complex
spatial data

- Audience
- Data relations
- Interaction

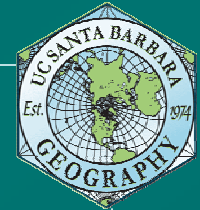


MacEachren, A. M. and Kraak, M.-J. 1997. Exploratory cartographic visualization: Advancing the agenda. *Computers & Geosciences* 23(4): 335-343

Our Approach



- case study – single story, compare 2 states
- what can geovisualization do
- feasibility of automation



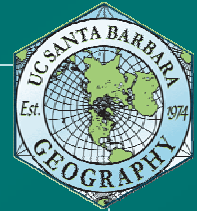
The Story

Semantic: Terri Schiavo (right-to-die case in FL)
Temporal: Feb. 05 - June 05
Geographic: reporting in Texas & California

Data source: America's Newsbank digital archive
Keyword search

Newspapers: 32 CA papers
13 TX papers

Constraints



- newspaper archive incomplete
- archives have inconsistent structure – impedes parsing
- newspaper circulation areas hard to determine



Interface lan



Access World News

[Skip navigation](#)

[Help / About | T](#)

SEARCH SHORTCUTS:

- [California Newspapers](#)
- [untitled bucket](#)
- [Major Metropolitan Newspapers](#)
- [Greater Los Angeles](#)
- [Acceda Noticias](#)
- [The World](#)
- [United States](#)
- [Create Custom List](#)

FOLDERS

- 📁 [Searches this Session](#)
- 📁 [Saved Articles](#)

MORE RESOURCES

- [America's Historical Newspapers](#)
- [American State Papers](#)
- [Archive of Americana](#)
- [Early American Imprints, Series I: Evans](#)
- [Early American Imprints, Series II: Shaw-Shoemaker](#)
- [U.S. Congressional Serial Set](#)
- [AccessUN](#)

SEARCHING: California selected title - [More Information](#)

for: in All Text

AND in All Text

return: Most recent matches first

timeframe: all years

custom date:

Results: 1 - 10 of 2158 [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#)

[Edit Search](#)

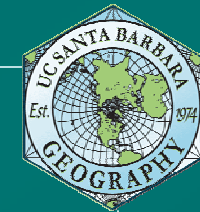
1. **San Diego Union-Tribune, The (CA)** - June 1, 2005
length: 2682 words
[Title: LETTERS TO THE EDITOR](#)
2. **Ukiah Daily Journal (CA)** - May 31, 2005 length: 327 words
[Berg again in national spotlight](#)
3. **Merced Sun-Star, The (CA)** - May 30, 2005
length: 490 words
[Around Town](#)
4. **Tribune, The (San Luis Obispo, CA)** - May 30, 2005
length: 412 words
[FREE HELP WITH END-OF-LIFE ISSUES - THE PUBLIC HAS THREE MORE CHANCES TO HEAR FROM EXPERTS ON HOW TO DRAFT ADVANCE DIRECTIVES](#)

Process

1) America's Newsbank keyword search

1367 articles CA
716 articles TX

Process - 2

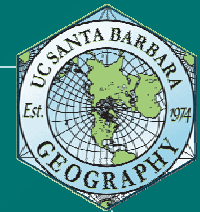


2) Exported text parsed to excel file

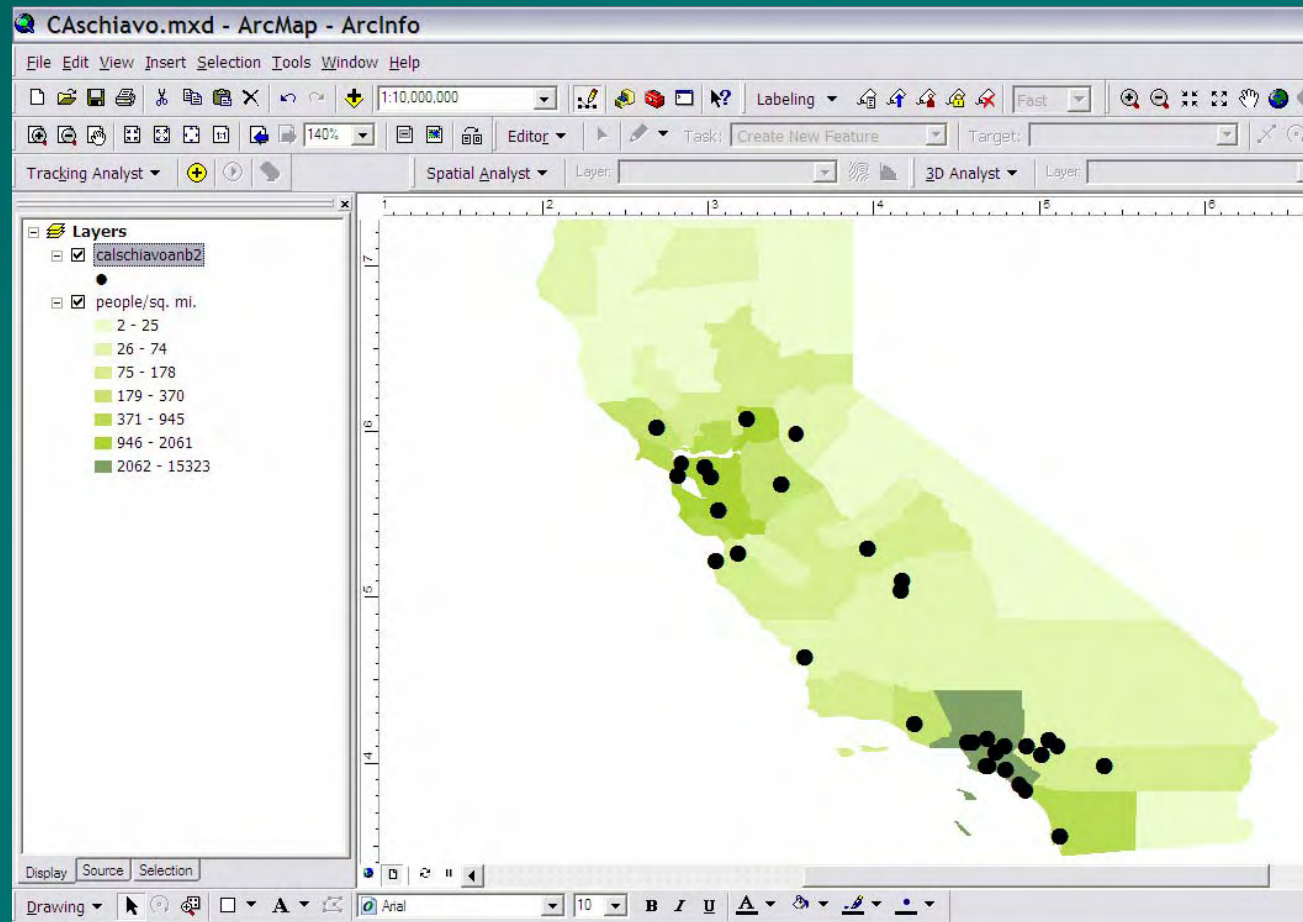
- Paper
- City
- Date
- Headline
- Word count

	A	B	C	D	E
1	Paper	City	Date	Headline	Words
2	Marin Independent Journal (Marin, CA)	Marin	22-Oct-2003	Gov. Bush orders feeding tube reinserted	311
3	Ventura County Star (California)	Ventura	23-Oct-2003	Bush's stay of execution;Florida case highlights value of	682
4	Sacramento Bee	Sacramento	25-Oct-2003	The Schiavo intrusion Governor, lawmakers tread on cou	500
5	The Daily Review (Hayward, CA)	Hayward	3-Nov-2003	Shocked by statements	572
6	The San Diego Union-Tribune	San Diego	6-Nov-2003	Guardian at issue in right-to-die fight	105
7	San Mateo County Times (San Mateo, CA)	San Mateo	10-Nov-2003	Right to live	368
8	Ventura County Star (California)	Ventura	25-Jan-2005	Court rules on Schiavo, asks for reviews of cases	575
9	Los Angeles Times	Los Angeles	24-Feb-2005	The Nation;;Schiavo Feeding Tube Stays for Now;;Judge	641
10	Los Angeles Times	Los Angeles	26-Feb-2005	Judge Orders Feeding Tube Removed;;A three-week de	722
11	The San Diego Union-Tribune	San Diego	11-Mar-2005	Schiavo's husband offered \$1 million;;Local businessma	380
12	Los Angeles Times	Los Angeles	13-Mar-2005	Passionate Debate Over the Right to Die	426
13	Los Angeles Times	Los Angeles	19-Mar-2005	Schiavo Taken Off Food Supply;;Congress tries to stop t	1987
14	Los Angeles Times	Los Angeles	20-Mar-2005	Weary Mother Begs for Her Daughter's Life;;As demonst	821
15	Ventura County Star (California)	Ventura	21-Mar-2005	Bush signs Schiavo legislation	893
16	Inland Valley Daily Bulletin (Ontario, CA)	Ontario	22-Mar-2005	Few California legislators on hand for Schiavo vote	297
17	Los Angeles Times	Los Angeles	22-Mar-2005	THE TERRI SCHIAVO CASE;;COST OF CARE;;Financi	237
18	Los Angeles Times	Los Angeles	22-Mar-2005	THE TERRI SCHIAVO CASE;;Judge Raises Doubts Ab	1013
19	San Gabriel Valley Tribune (San Gabriel Va	San Gabriel V	22-Mar-2005	If it's murder...	945
20	The Daily News of Los Angeles	Los Angeles	22-Mar-2005	FEW CALIFORNIA LAWMAKERS CAST SCHIAVO VO	309
21	The San Diego Union-Tribune	San Diego	22-Mar-2005	The Schiavo case;;Congress was wrong to intervene	494
22	Ventura County Star (California)	Ventura	22-Mar-2005	Judge defers decision in Schiavo case	989
23	The San Francisco Chronicle	San Francisco	22-Mar-2005	Life, death and politics	514
24	Los Angeles Times	Los Angeles	23-Mar-2005	THE TERRI SCHIAVO CASE;;MEDICAL PERSPECTIV	1084
25	Los Angeles Times	Los Angeles	23-Mar-2005	THE TERRI SCHIAVO CASE;;POLITICAL AFTERMATH	681
26	Los Angeles Times	Los Angeles	23-Mar-2005	THE TERRI SCHIAVO CASE;;Parents Lose Appeal in S	1866
27	San Gabriel Valley Tribune (San Gabriel Va	San Gabriel V	23-Mar-2005	Newspapers weigh in on Schiavo case	601
28	San Gabriel Valley Tribune (San Gabriel Va	San Gabriel V	23-Mar-2005	Disgusting intervention	1240

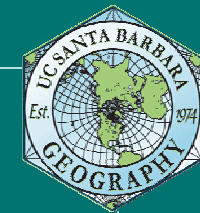
Process - 3



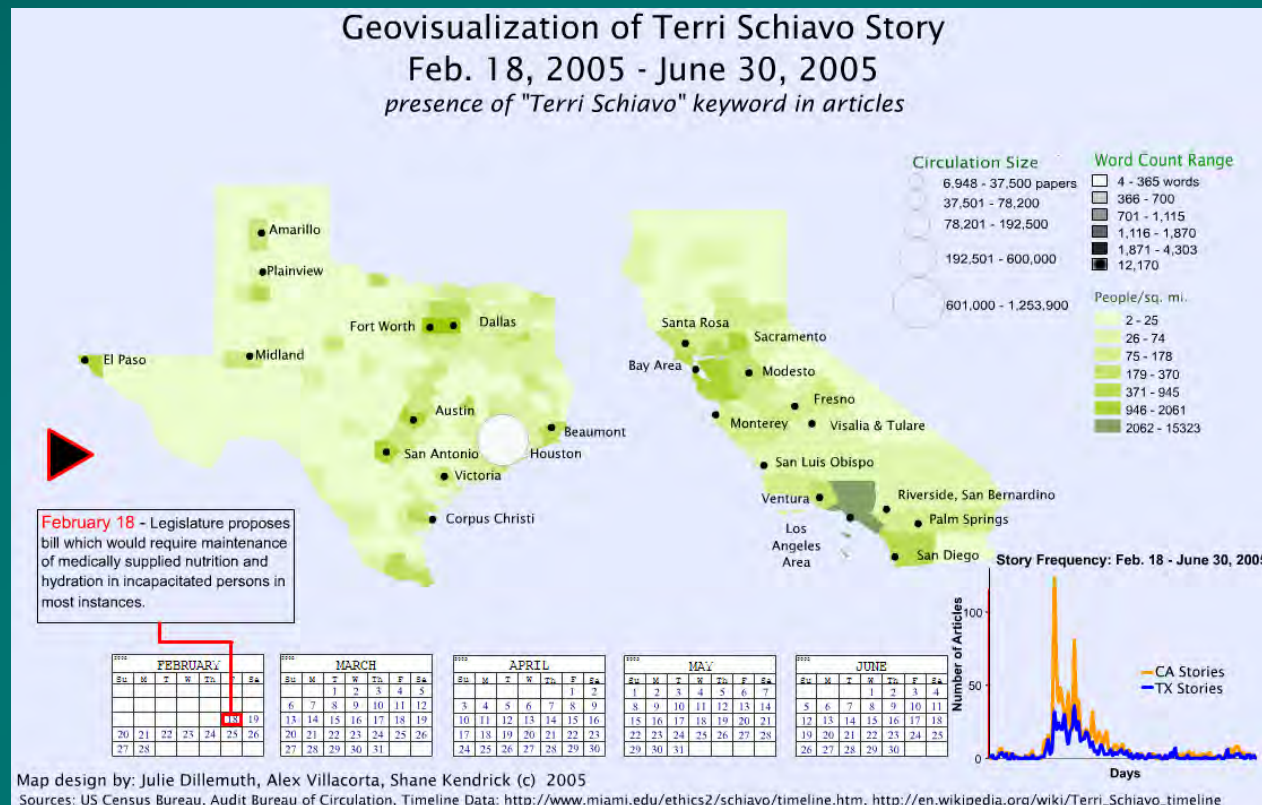
3) Geo-coded papers, mapped in ArcGIS



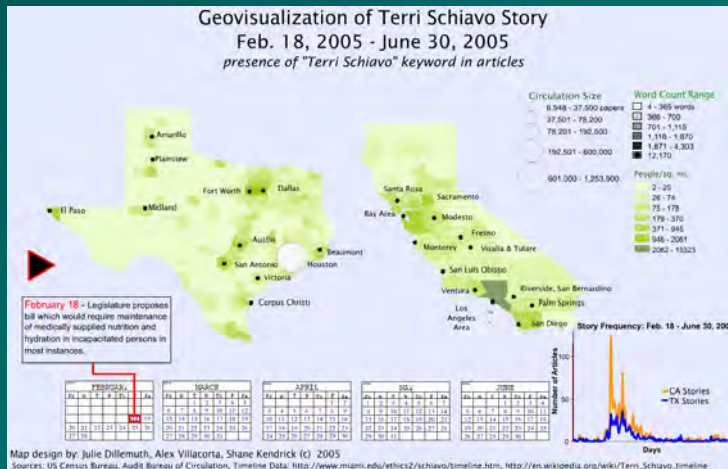
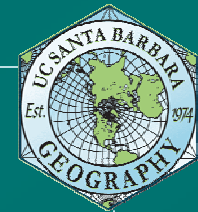
Process - 4



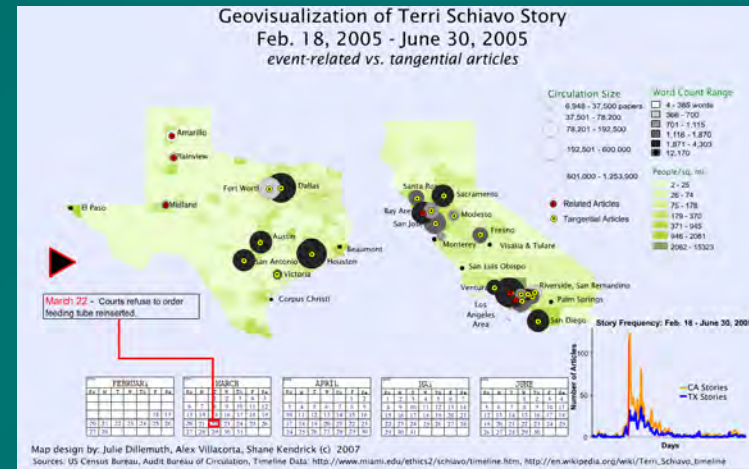
4) Annotated and animated in Flash



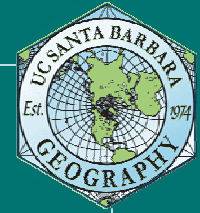
Visualizations



Viz 1: presence of 'Terri/y Schiavo' keyword



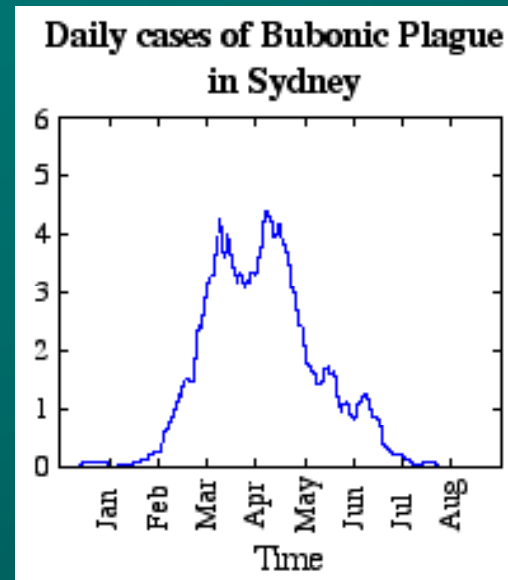
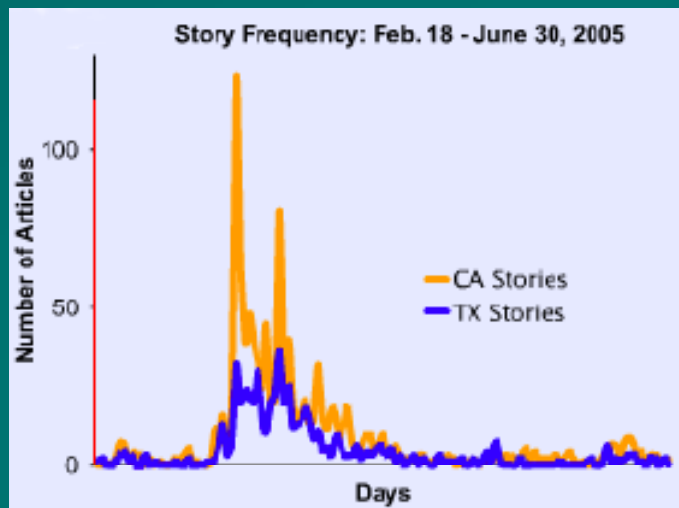
Viz 2: event-related vs. tangential articles



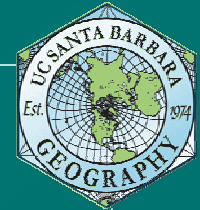
Viz 1

Viz 1: presence of 'Terri/y Schiavo' keyword

- First analysis: salient events
- Big cities first to carry story, smaller ones follow on
- Plot of articles-per-day similar to disease pattern
- Short attention span, or 'immunity' for news story



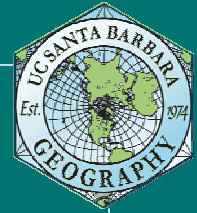
Viz 2



Viz 2: event-related vs. tangential articles

- Article classes:
 - events (176 CA, 184 TX) = fact-based
 - tangential (356 CA, 306 TX) = broader impact
 - unrelated (59 TX, 533 CA) = affected
- differences in two classes of article, 2 states
 - event-related until Congress & Bush involved (3/21)
 - western TX picks up story only then
 - tangential in TX before CA (3/21)
 - at TS death, more tangential than event

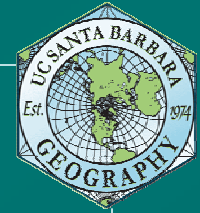
Viz 2



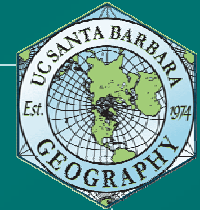
Evaluative questions, further analysis:

- what does the spatial component provide that graphs don't?
- how similar are the stories across papers and states?

Limitations

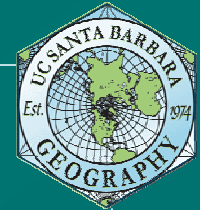


- potential for misinterpretation, overinterpretation
incomplete paper coverage
publication locations, not readership coverage
content type of article
- automation challenges: classifying articles



Observations

- Major cities 'early adopters' of story
- Uniformity in reporting only for major or unusual events
- Story growth and decay similar to disease
- Short attention span for story
- Event-related versus tangential stories have different patterns

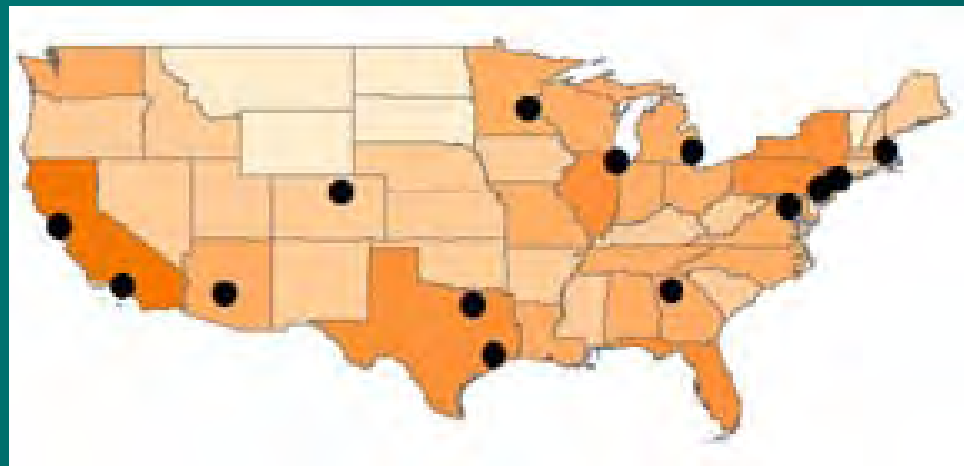


National Level

- Story: All front page news
- Duration: Aug. '99,'01,'03
- Location: Nationwide
- Newspapers: 15 papers
- Data: America's Newsbank online database

Filter Criteria

- Top 15 most circulated (excluding multiple cities)
- ≥ 3 papers carry story
- ≥ 5 publications total



Temporal



Top 15 Circulating US Newspapers

Aug. 2001 & 2003



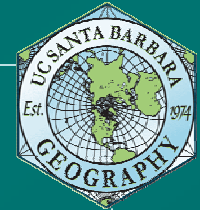
Avg. Circulation Size 2001 & 2003:

1,140,000 – 850,000: NYT, LA Times, Chicago Trib, WA Post

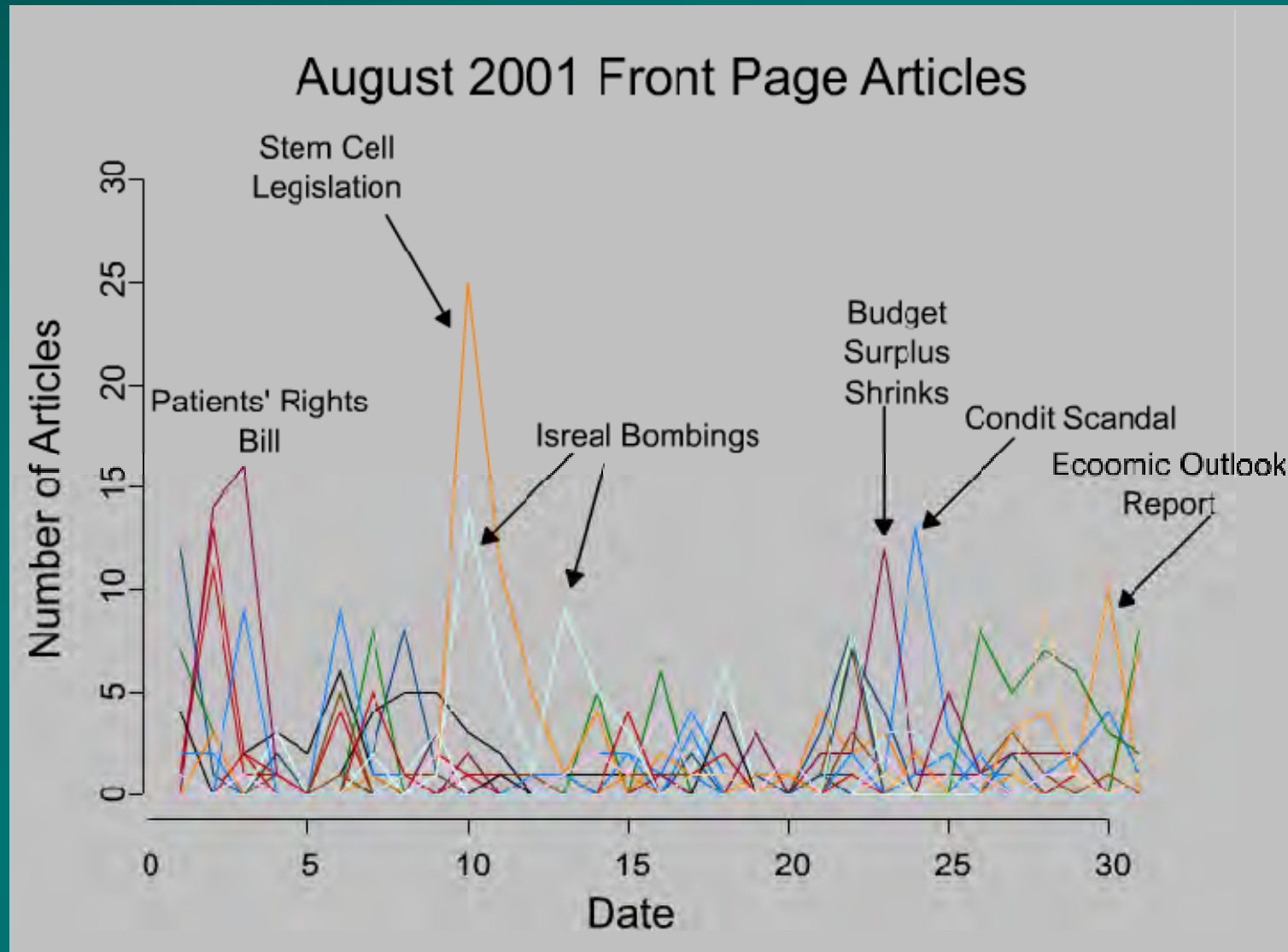
770,000 – 500,001: Denver, Detroit, Dallas, Houston, San Fran.

500,000 – 370,000: Boston, Atlanta, Ariz., Newark, Minneapolis, Philly

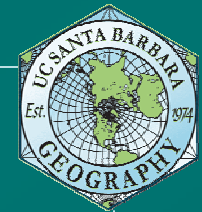
National News 2001



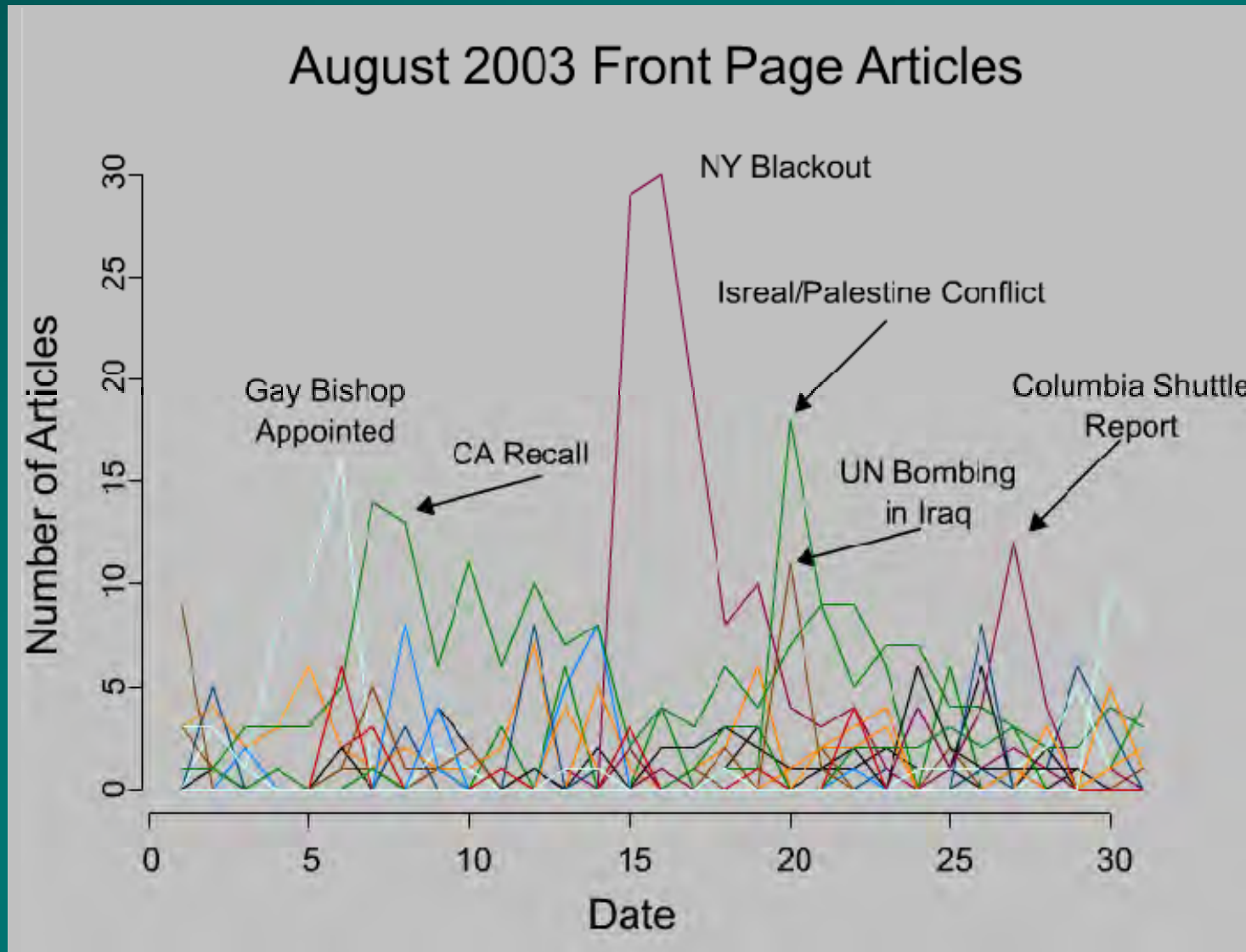
- 40 national-scale stories, 640 articles



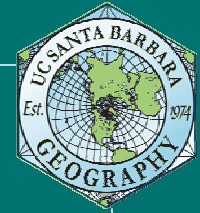
National News 2003



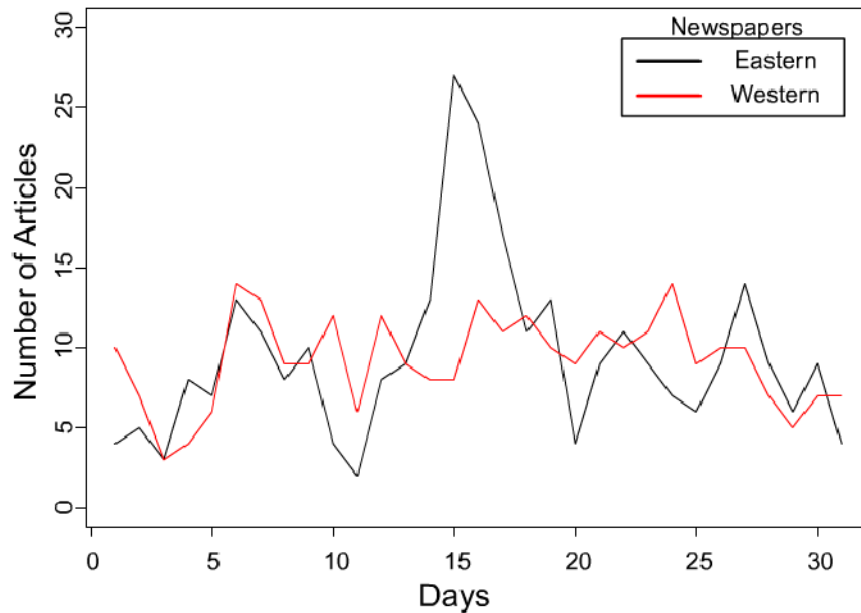
- 27 national-scale stories, 781 articles



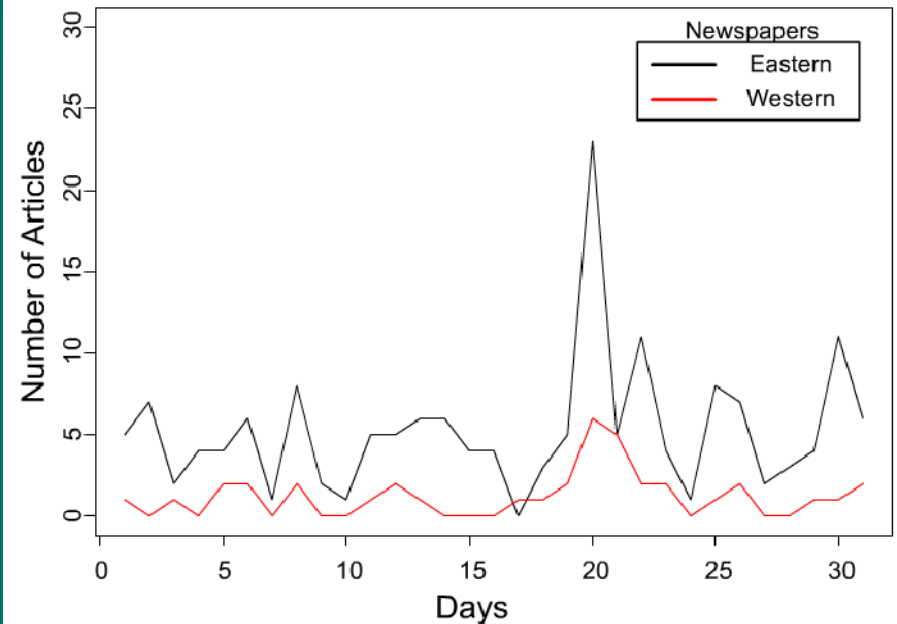
Spatial & Semantic



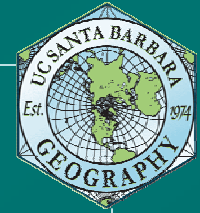
August 2003 Domestic Coverage by Region



August 2003 International Coverage by Region

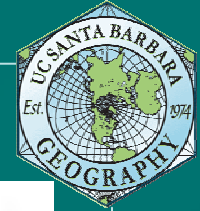


- On average, domestic stories are carried about the same
- International stories are carried more frequently in east papers



Implications

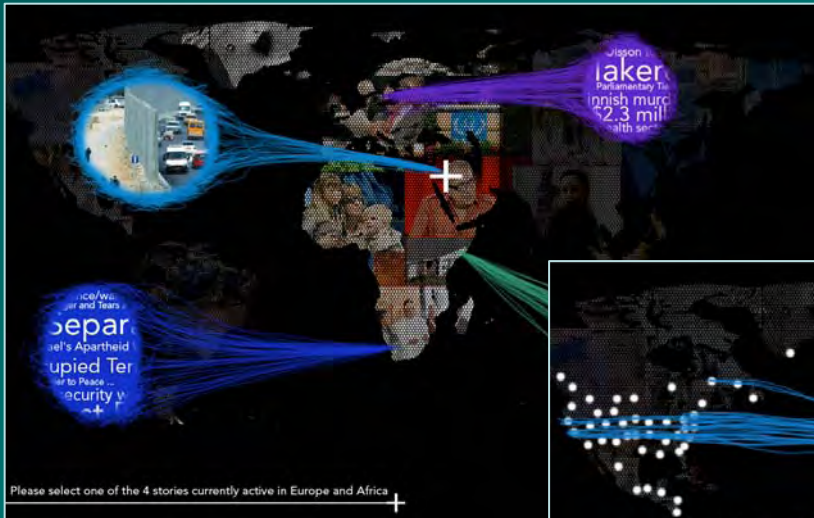
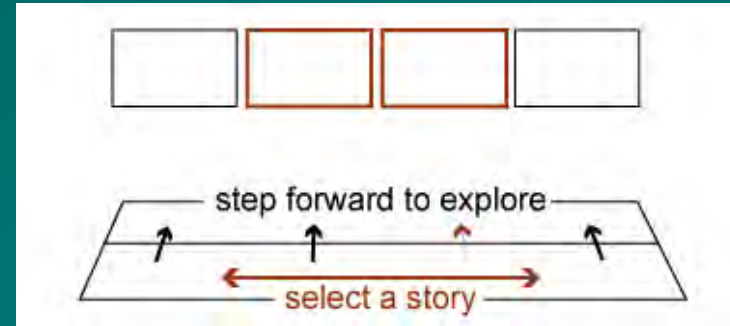
- contributes to understanding of communication channels
- Integrates semantic, spatial and temporal domains of dataset
- Towards an automated, interactive tool to visualize any text-based database query
- Inspired 'Spheres of Influence' concept & content



SOI

Spheres of Influence II

Geospatial and news data integration & visualization with a camera-tracking interface

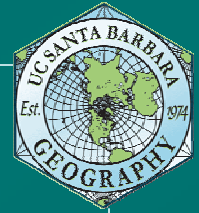


MAT, ECE, Geography,
Statistics, Psych, CS, Art

**Spatial,
Temporal &
Semantic
Dimensions
Spheres of Influence [SOI] 2007
Global
News
Reporting**

Alex Villacorta **statistics**
Anne-Marie Hansen **art**
Emily Moxley **electrical & computer engineering**
Jonathan Ventura **computer science**
Joriz De Guzman **computer science**
Karl Grossner **geography**
Matt Peterson **psychology**

SOI



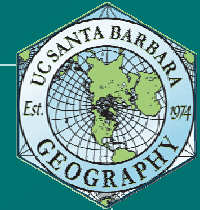
[Video demo](#)

Spheres of Influence 2007

Story Statistics:

Total number of reports: 211
Avg number of reports per day: 75
Total time in database: 2 weeks
Rank in database: 22nd
Highest rank: 3rd
Lowest rank: 131st
Max/Min reports: 249/6
Location w/ most reports: North America
% of all known locations: 2%
Number of keywords: 25

A c k n o w l e d g e m e n t s



Thanks to:

Keith C. Clarke, Advisor

S. Rao Jammalamadaka, Stats & Applied Prob.

Shane Kendrick, CNSI Apprentice Researcher Program

Sara Fabrikant

UCSB IGERT in Interactive Digital Multimedia

NSF Graduate Research Fellowship

Dr. Dawn Wright

OSU Geosciences, Eco-IGERT

Questions, comments?