

# Visual Analytics



CS 7450 - Information Visualization  
November 28, 2016  
John Stasko

## Agenda



- Overview of what the term means and how it relates to information visualization
- Specific example, Jigsaw, helping investigative analysis
- Related systems
- Some example VA research projects



## Acknowledgment



Slides looking like this provided  
courtesy of Jim Thomas

## Before there was VA



- Growing concern from some that infovis was straying from practical, real world analysis problems
  - Is it helping people enough?
- Infovis typically not applied to massive data sets
- Infovis “competes” with other computational approaches to data analysis
  - Statistics, data mining, machine learning

# Important Paper



- Shneiderman suggests combining computational analysis approaches such as data mining with infovis – Discovery tools
  - Too often viewed as competitors in past
  - Instead, can complement each other
- Each has something valuable to contribute

Shneiderman  
*Information Visualization '02*

Fall 2016

CS 7450

5

# Issues



- Issues influencing the design of discovery tools:
  - Statistical Algorithms vs. Visual data presentation
  - Hypothesis testing vs. exploratory data analysis
- Pro's and Con's?

Fall 2016

CS 7450

6

# Differing Views



- Hypothesis testing
  - Advocates:  
By stating hypotheses up front, limit variables and sharpens thinking, more precise measurement
  - Critics:  
Too far from reality, initial hypotheses bias toward finding evidence to support it
- Exploratory Data Analysis
  - Advocates:  
Find the interesting things this way, we now have computational capabilities to do them
  - Skeptics:  
Not generalizable, everything is a special case, detecting statistical relationships does not infer cause and effect

Fall 2016

CS 7450

7

# Recommendations



- Integrate data mining and information visualization
- Allow users to specify what they are seeking
- Recognize that users are situated in a social context
- Respect human responsibility

Fall 2016

CS 7450

8

## Further Questions



- Are information visualizations helping with exploratory analysis enough?
- Are they attempting to accomplish the right goals?

Fall 2016

CS 7450

9

## Another Important Paper



- Information visualization systems inadequately supported decision making:
  - Limited Affordances
  - Predetermined Representations
  - Decline of Determinism in Decision-Making
- “Representational primacy” versus “Analytic primacy”
  - Telling the truth about your data versus providing analytically useful visualizations

Covered earlier this term

Amar & Stasko  
InfoVis '04 Best Paper  
TVCG '05

Fall 2016

CS 7450

10

## Task Level



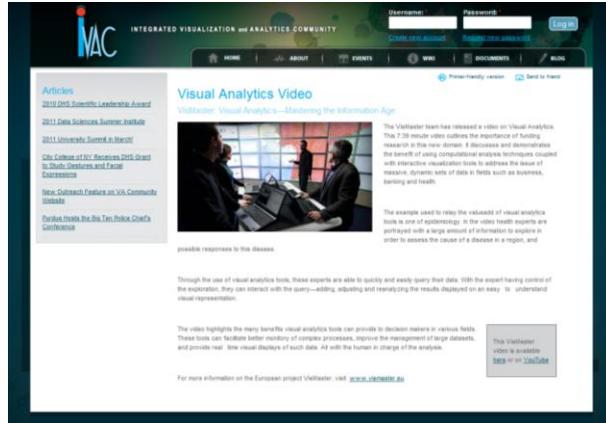
- Don't just help "low-level" tasks
  - Find, filter, correlate, etc.
- Facilitate analytical thinking
  - Complex decision-making, especially under uncertainty
  - Learning a domain
  - Identifying the nature of trends
  - Predicting the future

## More Motivation



- Increasing occurrences of situations and areas with large data needing better analysis
  - DNA, microarrays
  - 9/11 security
  - Business intelligence
  - ...

# Articulating the Motivation



Video

<http://videotheque.inria.fr/videotheque/doc/635>

Fall 2016

CS 7450

13

## Visual Analytics



- A new term for something that is familiar to all of us
- Informal description:
  - Using visual representations to help make decisions
  - Sounds like infovis, no?
  - Let's be more precise...

Fall 2016

CS 7450

14

# History



- 2003-04 Jim Thomas of PNNL, together with colleagues, develops notion of visual analytics
- Holds workshops at PNNL and at InfoVis '04 to help define a research agenda
- Agenda is formalized in book *Illuminating the Path*, shown on next slide

Fall 2016

CS 7450

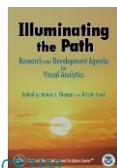
15



**Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces.**

**People use visual analytics tools and techniques to**

- Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
- Detect the expected and discover the unexpected
- Provide timely, defensible, and understandable assessments
- Communicate assessment effectively for action.



Thomas & Cook  
2005

**“The beginning of knowledge is the discovery of something we do not understand.”**  
~Frank Herbert (1920 - 1986)

# Visual Analytics



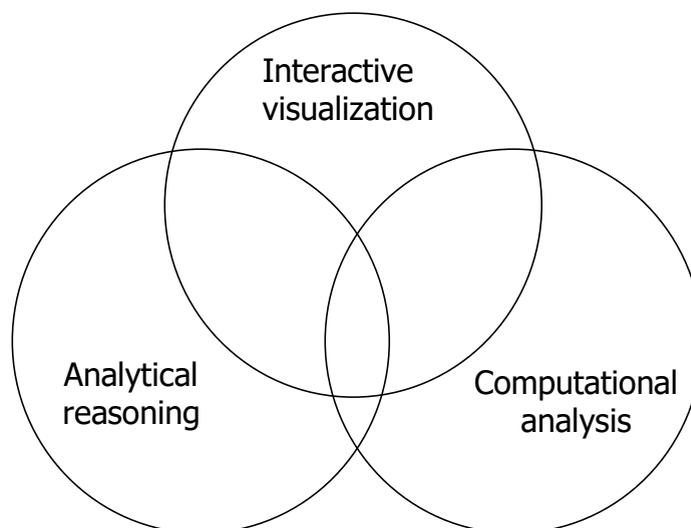
- Not really an “area” per se
  - More of an “umbrella” notion
- Combines multiple areas or disciplines
- Ultimately about using data to improve our knowledge and help make decisions

Fall 2016

CS 7450

17

# Main Components



Fall 2016

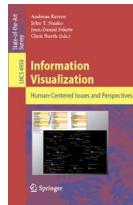
CS 7450

18

# Alternate Definition



- Visual analytics combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets

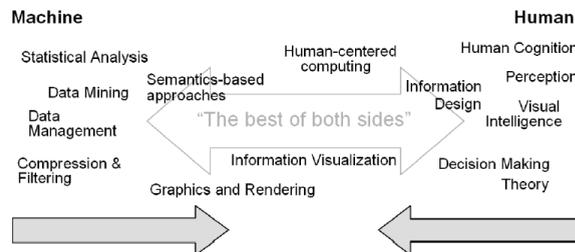


Keim et al, chapter in *Information Visualization: Human-Centered Issues and Perspectives*, 2008

# Synergy



- Combine strengths of both human and electronic data processing
  - Gives a semi-automated analytical process
  - Use strengths from each



From Keim

# InfoVis Comparison



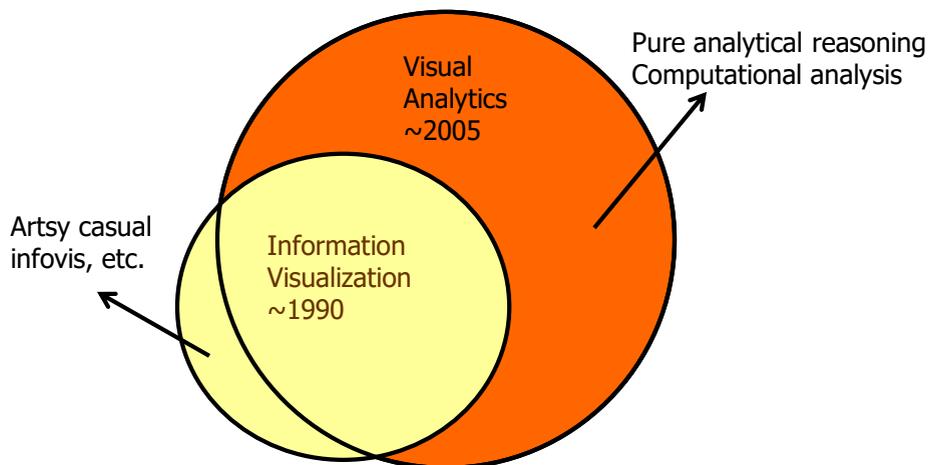
- Clearly much overlap
- Perhaps fair to say that infovis hasn't always focused on analysis tasks so much and that it doesn't always include advanced data analysis algorithms
  - Not a criticism, just not focus
  - InfoVis has a more narrow scope
  - (Some of us actually do believe that infovis has/should include those topics)

Fall 2016

CS 7450

21

# Academic Context



Fall 2016

CS 7450

My interpretation

22

# Visual Analytics



- Encompassing, integrated approach to data analysis
  - Use computational algorithms where helpful
  - Use human-directed visual exploration where helpful
  - Not just “Apply A, then apply B” though
  - Integrate the two tightly

Fall 2016

CS 7450

23

# Domain Roots



- Dept. of Homeland Security supported founding VA research
- Area has thus been connected with security, intelligence, law enforcement
- Should be domain-independent, however, as other areas need VA too
  - Business, science, biology, legal, etc.

Fall 2016

CS 7450

24

# VA-related Research Topics



- Visualization
  - InfoVis, SciVis, GIS
- Data management
  - Databases, information retrieval, natural language
- Data Analysis
  - Knowledge discovery, data mining, statistics
- Cognitive Science
  - Analytical reasoning, decision-making, perception
- Human-computer interaction
  - User interfaces, design, usability, evaluation

Fall 2016

CS 7450

25

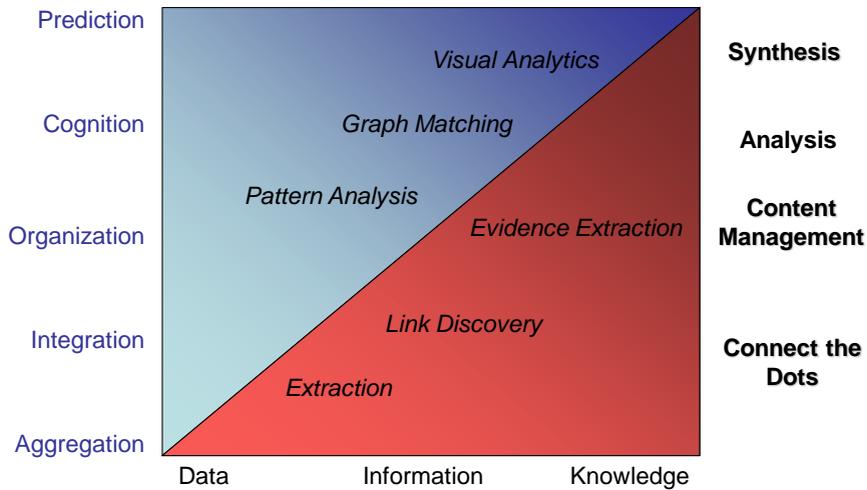


- **Statistics, data representation and statistical graphics**
- **Geospatial and Temporal Sciences**
- **Applied Mathematics**
- **Knowledge representation, management and discovery**
  - Ontology, semantics, NLP, extraction, synthesis, ...
- **Cognitive and Perceptual Sciences**
- **Communications: Capture, Illustrate and present a message**
- **Decision sciences**
- **Information and Scientific Visualization**

*And far more than homeland security*

CS 7450

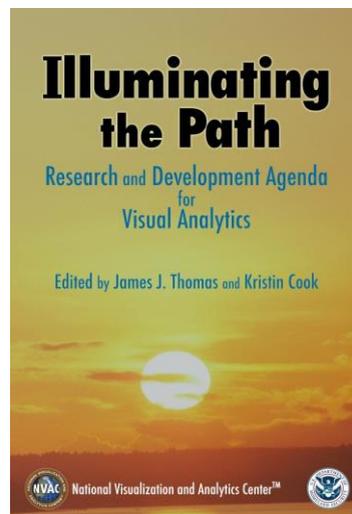
26



CS 7450

27

- Available at <http://nvac.pnl.gov/> in PDF form
- At IEEE Press in book form
- Special thanks to IEEE Technical Committee on Visualization and Graphics



CS 7450

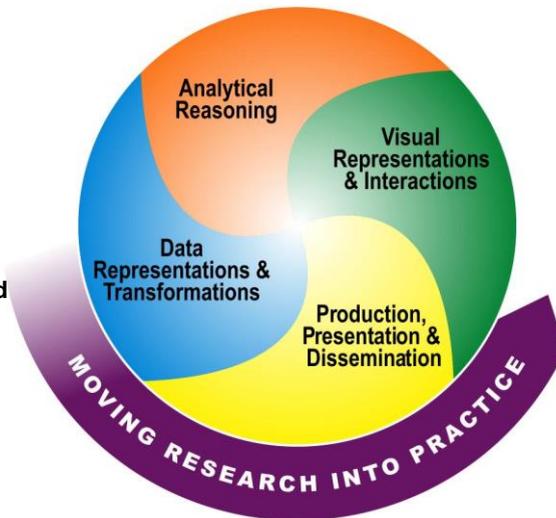
28



# Overview of the R&D Agenda



- Challenges
- Science of Analytical Reasoning
- Science of Visual Representations and Interactions
- Data Representations and Transformations
- Production, Presentation, and Dissemination
- Moving Research Into Practice
- Positioning for an Enduring Success



CS 7450

## More History



- European Union has become very active in visual analytics area
  - VisMaster project



Fall 2016



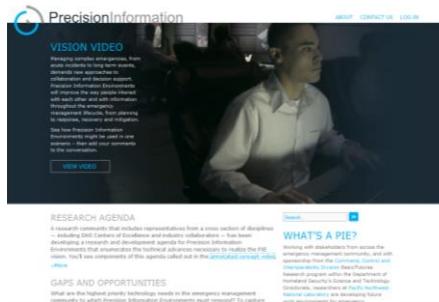
CS 7450

30

# Vision of the Future



- PNNL Precision Info Environments (PIE) video
- Emergency response scenario



<http://precisioninformation.org>

Fall 2016

CS 7450

31

# Application Area



- Investigative & Intelligence Analysis
  - Gather information from various sources then analyze and reason about what you find and know
  - Analyze situations, understand the particulars, anticipate what may happen

Fall 2016

CS 7450

32

# Intelligence Process

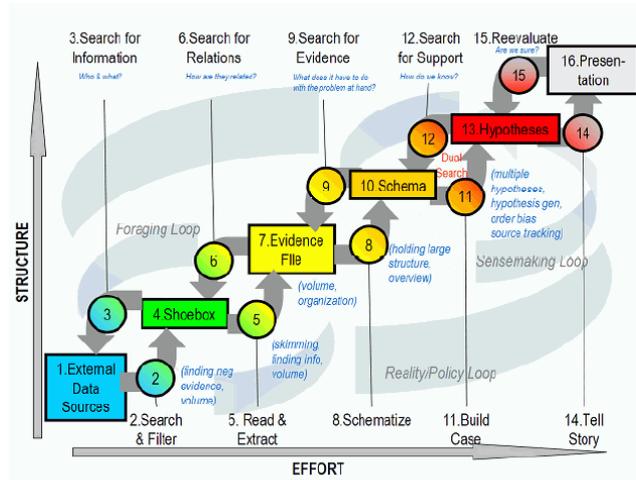


Figure 2.1. Notional model of sensemaking loop for intelligence analysis derived from CTA.

Pirolli & Card

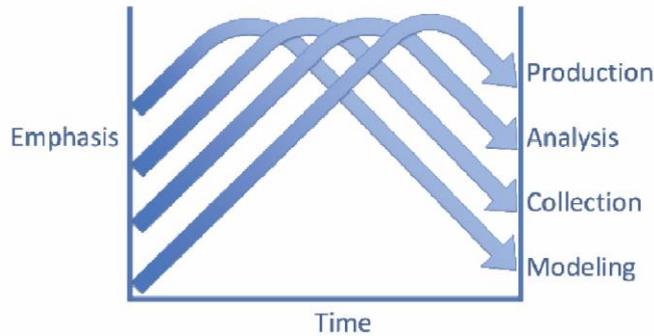
Intl Conf Intelligence Analysis '05

33

Fall 2016

CS 7450

# Intelligence Process



Wheaton  
In preparation

Fall 2016

CS 7450

34

# Pain Points



- Cost structure of scanning and selecting items for further attention
- Analysts' span of attention for evidence and hypotheses

Fall 2016

CS 7450

35

# Jigsaw

Stasko, Görg, Liu  
*Information Visualization '08*

Görg et al  
*TVCG '13*



## Visualization for Investigative Analysis across Document Collections

Law enforcement & intelligence community  
Fraud (finance, accounting, banking)  
Academic research  
Journalism & reporting  
Consumer research

**"Putting the pieces together"**



Fall 2016

CS 7450

36

# The Jigsaw Team



Carsten Görg  
Zhicheng Liu  
Youn-ah Kang and many others  
Jaeyeon Kihm  
Jaegul Choo  
Chad Stolper  
Anand Sainath  
Sakshi Pratap

Fall 2016

CS 7450

37

## Problem Addressed

*Analogy*



Help “investigators” explore, analyze and understand large document collections

Articles & reports

XML documents

Spreadsheets

Blogs

Fall 2016

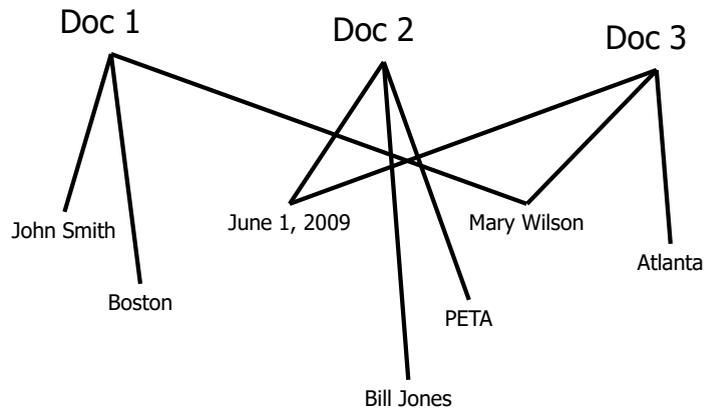
CS 7450

38

# Our Focus



- Entities within the documents
  - Person, place, organization, phone number, date, license plate, etc.
- Thesis: A story/narrative/plot/threat within the documents will involve a set of entities in coordination



# Entity Identification



- Must identify and extract entities from plain text documents
  - Crucial for our work
- Not our main research focus – We use tools from others

# Sample Document



Report: 20040510-4\_16  
May 14 2004

VANCOUVER, British Columbia - A Canadian immigration panel is considering whether accused environmental saboteur Tre Arrow can apply for refugee status in Canada.

Arrow, 30, who is wanted for fire bombing logging and cement trucks in Oregon, asked the Canadian authorities to remain in Canada as a political refugee at a hearing in Vancouver on Tuesday.

A key issue will be whether Arrow is affiliated with a terrorist group, which would immediately disqualify him from receiving refugee status in Canada, authorities said.

The Immigration and Refugee Board is scheduled to decide by May 31 whether Arrow is affiliated with the Earth Liberation Front, a group the FBI considers a terrorist organization responsible for scores of attacks on property over the past dozen years.

# Entities Identified



**Source:**  
**Date:** May 14, 2004

**VANCOUVER, British Columbia** - A Canadian immigration panel is considering whether accused environmental **saboteur Tre Arrow** can apply for refugee status in **Canada**.

**Arrow**, 30, who is wanted for fire bombing logging and cement trucks in **Oregon**, asked the Canadian authorities to remain in **Canada** as a political refugee at a hearing in **Vancouver** on **Tuesday**.

A key issue will be whether **Arrow** is affiliated with a terrorist group, which would immediately disqualify him from receiving refugee status in **Canada**, authorities said.

The **Immigration and Refugee Board** is scheduled to decide by **May 31** whether **Arrow** is affiliated with the **Earth Liberation Front**, a group the **FBI** considers a terrorist organization responsible for scores of attacks on property over the past dozen years.

Fall 2016

CS 7450

43

# Sample Document 2



Title: Proving Columbus was Wrong  
Abstract: In this work, we show the world is really flat. To do this, we build a bunch of ships. Then we...  
PI: Amerigo Vespucci  
Co-PI: Vasco de Gama, Ponce de Leon  
Organization: Northwest Central Univ.  
Amount: 123,456  
Program Mgr: Ephraim Glinert  
Division: IIS  
ProgramElementCode: 2860

Fall 2016

CS 7450

44

# Entities Already Identified



Title: Proving Columbus was Wrong

**Abstract:** In this work, we show the world is really flat. To do this, we build a bunch of ships. Then we...

---

PI: Amerigo Vespucci

Co-PI: Vasco de Gama, Ponce de Leon

Organization: Northwest Central Univ.

Amount: 123,456

Program Mgr: Ephraim Glinert

Division: IIS

ProgramElementCode: 2860

**Entities**

# Connections



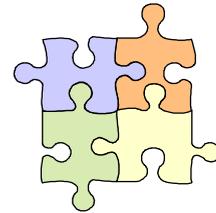
- Entities relate/connect to each other to make a larger “story”
- Connection definition:
  - Two entities are connected if they appear in a document together
  - The more documents they appear in together, the stronger the connection

# Jigsaw

“Putting the pieces together”



- Computational analysis of document text
  - Entity identification, document similarity, clustering, summarization, sentiment
- Multiple visualizations (views) of documents, analysis results, entities and their connections
- Views are highly interactive and coordinated



Fall 2016

CS 7450

47

## System Views

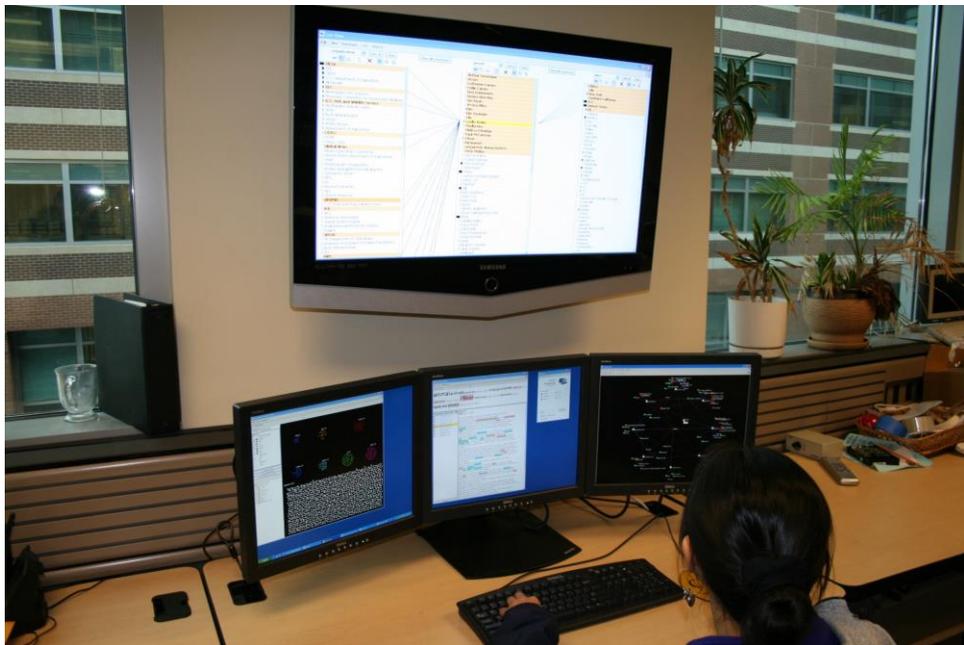
The collage displays several views from the Jigsaw system:

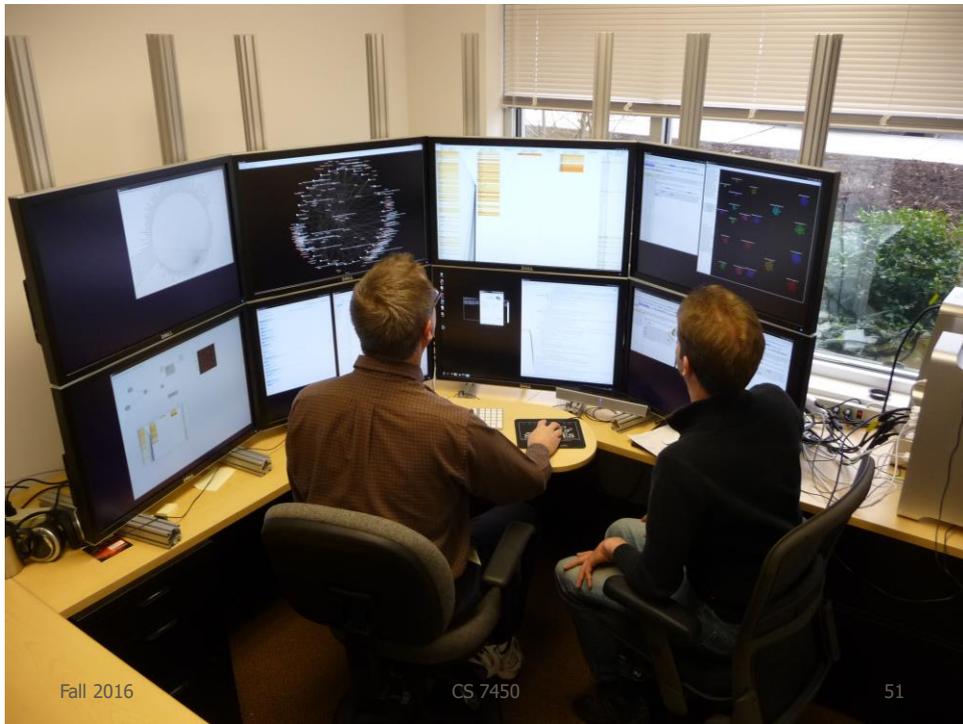
- Top Left:** A small overview window showing a document structure.
- Top Middle:** A large window with a hierarchical tree view on the left and a list of document segments on the right.
- Top Right:** A network graph visualization with nodes and edges on a dark background.
- Middle Left:** A document viewer window showing a list of document segments with a selected segment highlighted.
- Middle Center:** A window showing a horizontal bar chart or timeline visualization.
- Middle Right:** A window displaying a grid of data points or a matrix visualization.
- Bottom Left:** A window showing a hierarchical tree view of document segments.
- Bottom Center:** A window showing a network graph with nodes colored in various colors.
- Bottom Right:** A window showing a circular network graph visualization.

Fall 2016

CS 7450

48





# Input Data Formats



- Text, pdf, Word, html, Excel
- Jigsaw data file format
  - Our own xml
  
- DB?
  - Go to Excel
  - Go to text, transform to Jigsaw data file

Fall 2016

CS 7450

53

# Document Import



The screenshot shows two overlapping dialog boxes. The 'Import' dialog is in the background, showing a 'Source' dropdown menu with options like 'Files', 'Jigsaw Datafiles', 'Web Sites', 'Web Search', 'DHS Reports', and 'BibTex Files'. Below it is a 'Files:' text field and a 'Browse...' button. The 'Entity Identification' dialog is in the foreground. It has three sections: 'Statistical Entity Identification' with radio buttons for 'None', 'LingPipe', 'Calais', 'GATE', and 'Illinois-NER', and checkboxes for 'Person', 'Location', 'Organization', 'Date', and 'Money'; 'Rule-Based Entity Identification' with checkboxes for 'Date', 'Phone', 'Zip code', 'Email', 'URL', and 'IP address'; and 'Dictionary-Based Entity Identification' with 'Entity Type' and 'Dictionary File' labels, three input fields with 'Browse...' buttons, and 'Case sensitive' checkboxes. Both dialogs have 'Import' and 'Cancel' buttons at the bottom.

Various document formats with entity identification

Fall 2016

CS 7450

54

```

<award>
<awardnumber>0640291</awardnumber>
<title>SGER: Distributed Spatial Partitioning Algorithms for Scalable Processing of Mobile
<nsforganization>IIS </nsforganization>
<programs>DATA MANAGEMENT SYSTEMS</programs>
<startdate>September 1, 2006</startdate>
<lastamendmentdate>September 12, 2007</lastamendmentdate>
<principalinvestigator>Liu, Ling</principalinvestigator>
<state>GA</state>
<organization>GA Tech Research Corporation - GA Institute of Technology </organizatio
<awardinstrument>Standard Grant </awardinstrument>
<programmanager>Le Gruenwald </programmanager>
<expirationdate>February 29, 2008</expirationdate>
<awardedamounttodate>65502</awardedamounttodate>
<co_pinames></co_pinames>
<piemailaddress>lingliu@cc.gatech.edu
<organizationstreetaddress>Office of Sponsored Programs </organizationstreetaddress>
<organizationcity>Atlanta </organizationcity>
<organizationstate>GA</organizationstate>
<organizationzip>30332</organizationzip>
<organizationphone>4048944819</organizationphone>
<nsfdirectorates>CSE </nsfdirectorates>
<programmelementcodes>7485</programmelementcodes>
<programreferencecodes>HPCC|9218|7484</programreferencecodes>
<fieldofapplications>0104000 Information Systems </fieldofapplicati
<awardnumber>0640291</awardnumber>
<abstract>IIS-0640219 Ling Liu &lt;lingliu@cc.gatech.edu&gt; Georgia Institute of Instit
</award>

```

## Scraped XML

Fall 2016

CS 7450

55

```

<document>
<docID>0808863</docID>
<docDate>July 1, 2008</docDate>
<docSource></docSource>
<docText>FODAVA-Lead: Dimension Reduction and Data Reduction: Foundations for Visualization

FODAVA-Lead: Dimension Reduction and Data Reduction: Foundations for Visualization The FODAVA (Foundations of
Data Analysis and Visualization) Lead research team at the Georgia Institute of Technology provides unified
expertise in the critical areas for providing leadership of the FODAVA effort, including machine learning and
computational statistics, information visualization, massive-dataset algorithms and data structures, and
optimization theory. The team is focused on the fundamental theory and approaches to make breakthroughs in data
representations and transformations. The work is directed along the two main axes of scale reduction, data reductio
<directorate>CSE</directorate>
<award-instr>Continuing grant</award-instr>
<programreferencecode>HPCC</programreferencecode>
<programreferencecode>9218</programreferencecode>
<keyword>visualization</keyword>
<keyword>algorithms</keyword>
<fieldofapplication>0000912 Computer Science</fieldofapplication>
<state>GA</state>
<organization>GA Tech Research Corporation - GA Institute of Technology</organization>
<keyword>data analysis</keyword>
<keyword>information visualization</keyword>
<keyword>machine learning</keyword>
<amount>1200000</amount>
<pi>Park, Haesun</pi>
<copi>John Stasko</copi>
<copi>Alexander Gray</copi>
<copi>Renato D. C. Monteiro</copi>
<copi>Vladimir Koltchinskii</copi>
<progmgr>Lawrence Rosenblum</progmgr>
<division>CCF</division>
<keyword>visual analytics</keyword>
<programmelementcode>I114</programmelementcode>
<programmelementcode>H194</programmelementcode>
</document>

```

## Jigsaw Datafile Format

Fall 2016

CS 7450

56

# Computational Analyses



- Document summarization
- Document similarity
- Document clustering by content
  - Text or entities
- Sentiment analysis

Görg et al  
*TVCG*'13

Fall 2016

CS 7450

57

## Demo



- Amazon Samsung TV reviews
- Entities
  - Built-in:
    - Author
    - Rating
  - Extracted from text:
    - Feature (audio, picture, stand, delivery, ...)
    - Brand (Samsung, Sony, LG, Vizio, ...)

Fall 2016

CS 7450

58

# Console

**JIGSAW**  
infovis-vast  
612 documents

Color Legend:

- author (1017)
- concept (77)
- conference (2)
- indexterm (1790)
- journal (17)
- keyword (1202)
- year (16)

Search

Entities  Documents

Workspace: no active workspace

Fall 2016

CS 7450

59

# Document View

analysis analysts analytic animation based cognition design discuss display evaluation

framework information infovis interaction level localization paper research

systems tasks techniques video visual visualization visualizations

Documents

- 1 infovis00-885091
- 0 infovis01-963277
- 0 infovis03-1249027
- 1 infovis04-1382902
- 1 infovis05-1532136
- 1 infovis07-4376134
- 0 infovis07-4376144
- 2 infovis08-4658127
- 0 infovis08-4658139
- 2 infovis08-4658146
- 0 infovis09-5290706
- 0 infovis95-529685
- 0 vast07-4389006
- 0 vast07-4389013
- 0 vast09-5332596
- 1 vast09-5333878

**Summary:** Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study Despite the growing number of systems providing visual analytic support for investigative analysis, few empirical studies of the potential benefits of such systems have been conducted, particularly controlled, comparative

**Source:** Visual Analytics Science and Technology, 2009. VAST 2009.  
**Date:** Oct 12, 2009

Evaluating **visual analytics** systems for **investigative analysis**.  
Deriving design principles from a **case study**

Despite the growing number of systems providing visual analytic support for **investigative analysis**, few empirical studies of the potential benefits of such systems have been conducted, particularly controlled, comparative evaluations. Determining how such systems foster **insight** and sensemaking is important for their continued growth and study, however. Furthermore, studies that identify how people use such systems and why they benefit (or not) can help inform the design of new systems in this area. We conducted an **evaluation** of the **visual analytics system**, Jigsaw employed in a small **investigative sensemaking exercise**, and we compared its use to three other more traditional methods of analysis. Sixteen participants performed a **simulated intelligence analysis** task under one of the four conditions. Experimental results suggest that Jigsaw assisted participants to analyze the data and identify an embedded threat. We describe different analysis strategies used by study participants and how computational support (or the lack thereof) influenced the strategies. We then illustrate several

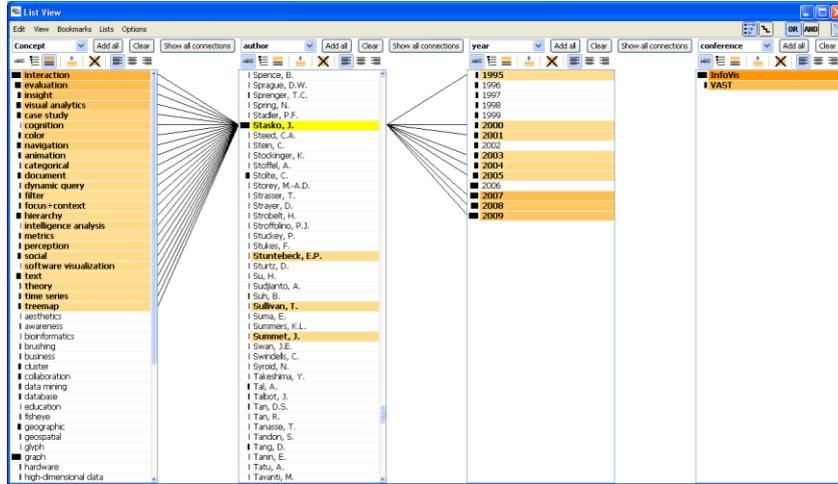
Fall 2016

CS 7450

60

# List View

Lists of entities by type  
Connections highlighted



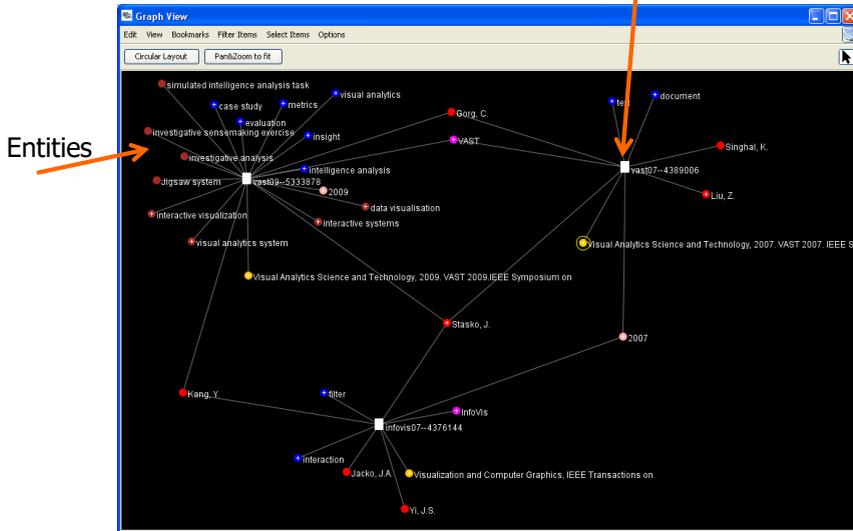
Fall 2016

CS 7450

61

# Graph View

Document



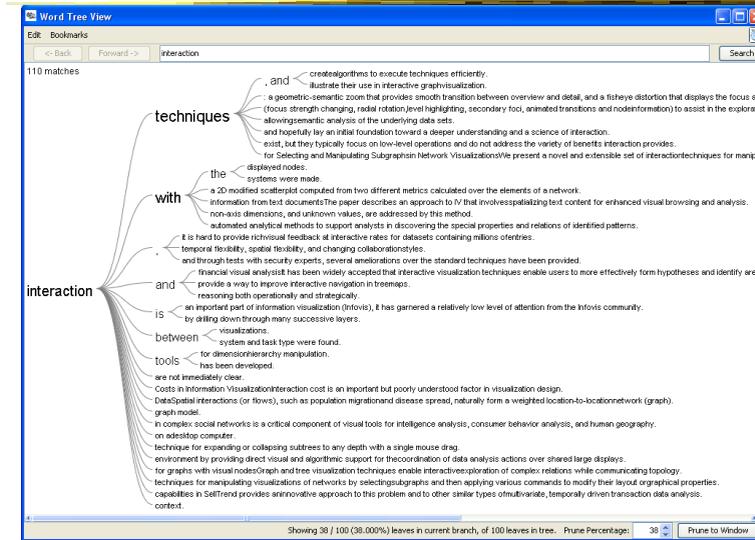
Fall 2016

CS 7450

62

# WordTree View

Context of a word in the collection



Fall 2016

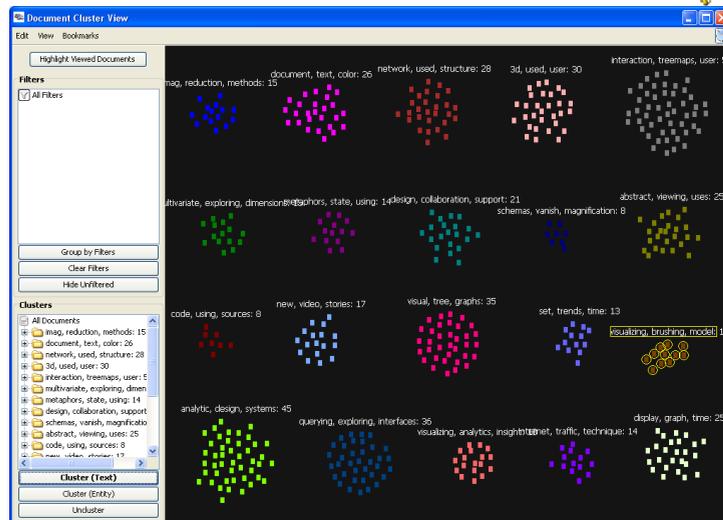
CS 7450

63

# Document Cluster View

Clustered by document text or by entities

Summarized by three words

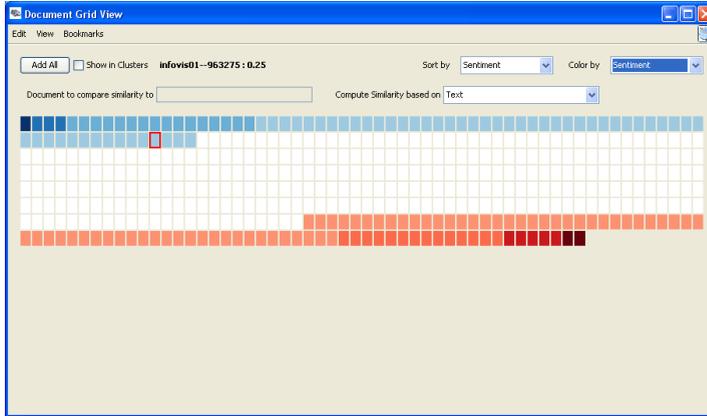


Fall 2016

CS 7450

64

# Document Grid View



User controls order and color

Sentiment analysis shown here

Fall 2016

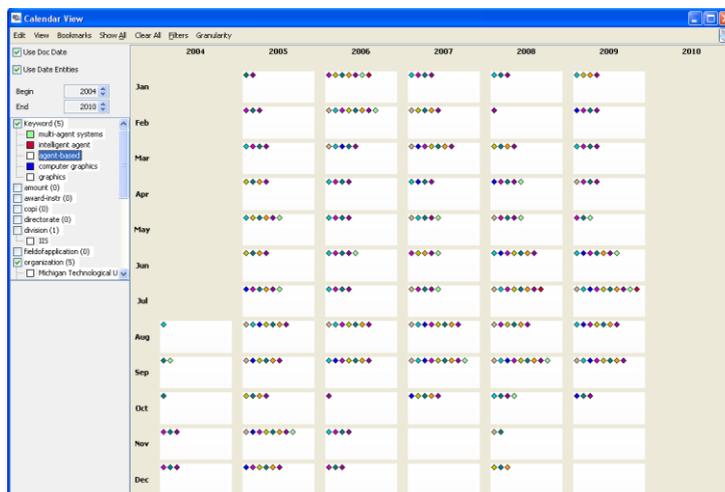
CS 7450

65

# Calendar View



Showing connections between entities and dates



Fall 2016

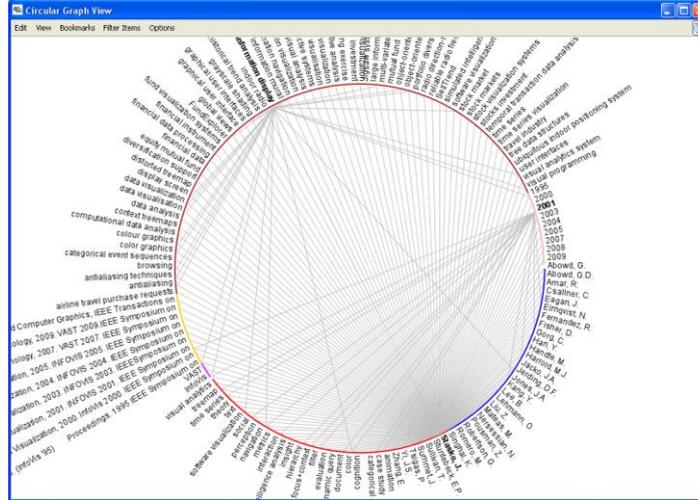
CS 7450

66

# Circular Graph View



Connections between entities



Fall 2016

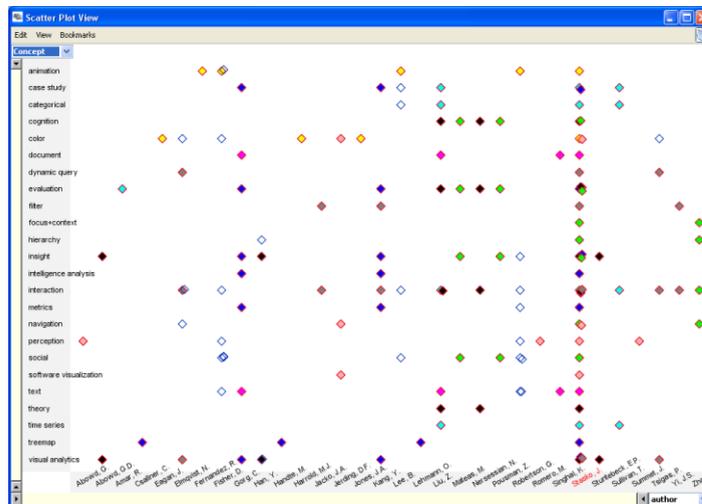
CS 7450

67

# Scatterplot View



Documents containing pairs of entities



Fall 2016

CS 7450

68

# EI Correction



The screenshot shows a 'Document View' window with a text document. The text contains several terms related to software testing and visualization, such as 'analysis', 'approach', 'based design', 'exploration', 'graph', 'information', 'interaction', 'interactive', 'large', 'layout method', 'network', 'paper', 'patterns', 'present', 'results', 'space', 'technique', 'techniques', 'users', 'using', 'visual', 'visualization', and 'visualizations'. A context menu is open over the word 'graph', offering options like 'Add as author entry', 'Add as conference entry', 'Add as journal entry', 'Add as keyword entry', 'Add as year entry', 'Add as new entity type', and 'Add Text to Tablet'. A 'Summary' section is visible below the text, and a list of documents is on the left.

Fall 2016

CS 7450

69

# Entity Aliasing



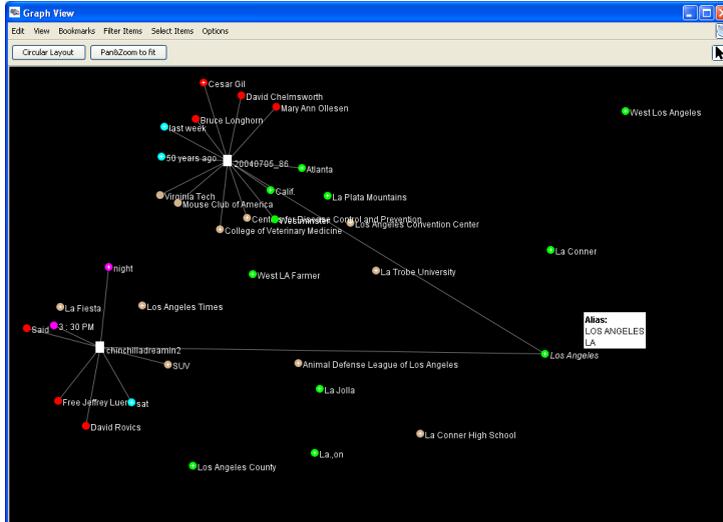
The screenshot shows a 'List View' window with a hierarchical list of concepts, authors, years, and keywords. The 'concept' column lists various terms like 'interaction', 'graph', 'visualization', etc. The 'author' column lists names like 'Kern, D.A.', 'Stasko, J.', etc. The 'year' column lists years from 1995 to 2009. The 'keyword' column lists terms like 'graph visualization', 'graphical user interface', etc. The list is organized into a tree structure with expandable nodes.

Fall 2016

CS 7450

70

# Alias Representation



Fall 2016

CS 7450

71

# Tablet

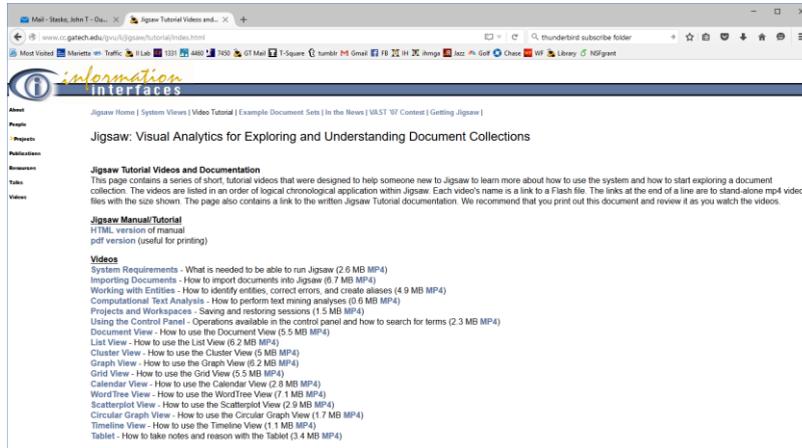


Fall 2016

CS 7450

72

# Help Using the System



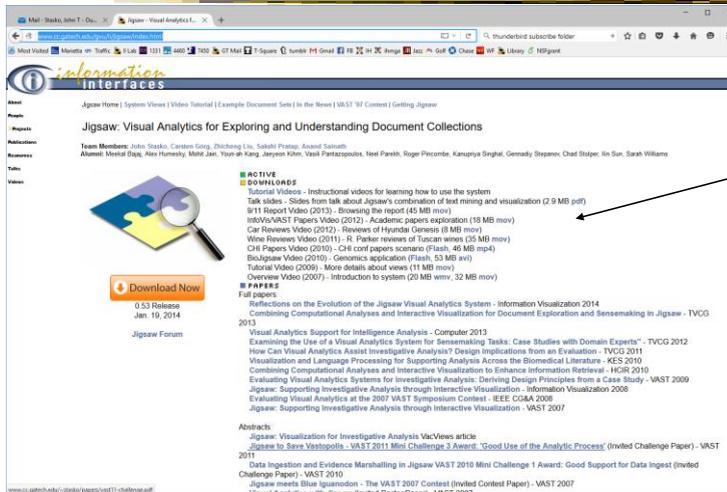
Manual, how-to videos

Fall 2016

CS 7450

73

# See Examples



Usage scenario videos

Fall 2016

CS 7450

74

# Application Domains



- Intelligence & law enforcement
  - Police cases
  - Won 2007 VAST Contest
  - Stasko et al, *Information Visualization* '08
- Academic papers, PubMed
  - All InfoVis & VAST papers
  - CHI papers
  - Görg et al, KES '10
- Investigative reporting
- Fraud
  - Finance, accounting, banking
- Grants
  - NSF CISE awards from 2000
- Topics on the web (medical condition)
  - Autism
- Consumer reviews
  - Amazon product reviews, edmunds.com, tripadvisor.com
  - Görg et al, HCIR '10
- Business Intelligence
  - Patents, press releases, corporate agreements, ...
- Emails
  - White House logs
- Software
  - Source code repositories
  - Ruan et al, SoftVis '10

Fall 2016

CS 7450

75

# Potential Jigsaw Future Work



- Collaborative capabilities
- Improved evidence marshalling
- Present/browse investigation history
- Scalability upward
- Web document ingest
- Implement network algorithms
- DB import
- Wikipedia & Intellipedia
- Geospatial view
- Better timeline capabilities
- Reliability/uncertainty
- Other types of data
- Active crawling/RSS ingest
- Try it on display wall
- Deployment to real clients

Fall 2016

CS 7450

76

## Room to Improve



- What Jigsaw doesn't do so well now
  - The end-part of the Pirolli-Card model
    - Helping the analyst take notes, organize evidence, generate hypotheses, etc.  
(The Tablet is a first step)
  - Sometimes called "evidence marshalling"
  - Others have focused more on that aspect...

Fall 2016

CS 7450

77

## PARC's Entity Workspace



- Tools for rapid ingest of entities from documents
- Can snap together entities into groups
- Can indicate level of interest in objects
- Four main view panels, with zooming UI

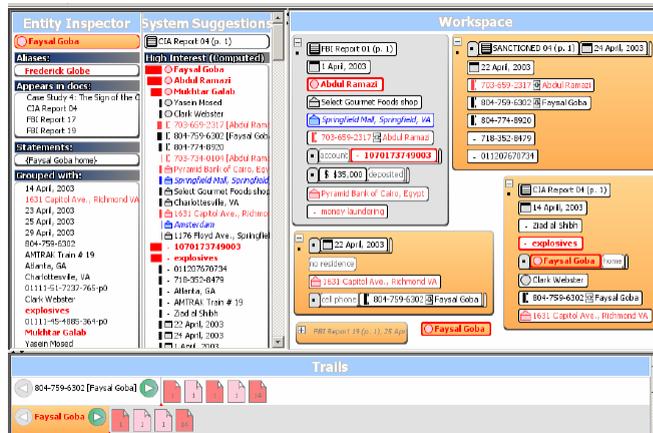
Bier, Card & Bodnar  
VAST '08

Fall 2016

CS 7450

78

# PARC's Entity Workspace



Video

Fall 2016

CS 7450

79

# VT's Analyst's Workspace



- Uses spatial affordances from a large display area for benefit in sensemaking
- Analysts move around and arrange items (documents, entities, search results) to externalize the thinking process
  - Like working with pieces of paper on a conference table, but with computational capabilities

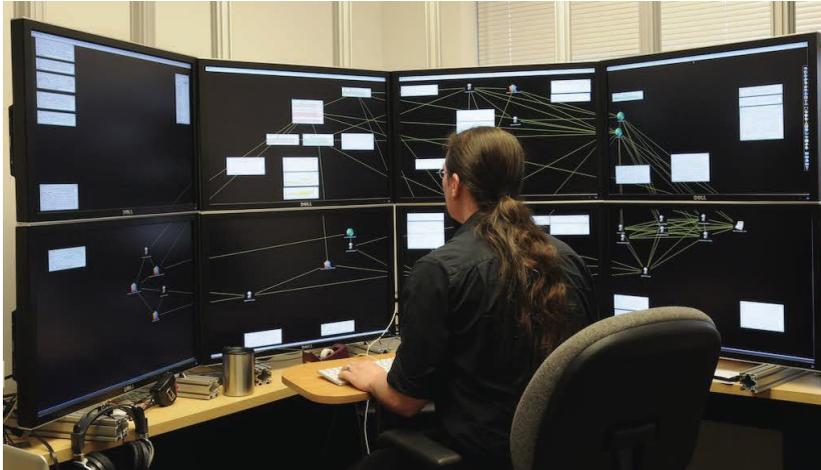
Andrews & North  
VAST '12

Fall 2016

CS 7450

80

# VT's Analyst's Workspace



Fall 2016

CS 7450

Video

81

## Related Area of Interest



- Sensemaking
- A general term that has been used in a number of different contexts
  - E.g., How large corporations make decisions
- To me, ultimately about people working with data and information to understand it better

Fall 2016

CS 7450

82

# Sensemaking



Nice definition:

“A motivated , continuous effort to understand connections (which can be among people, places, and events) in order to anticipate their trajectories and act effectively.”

– Klein, Moon and Hoffman  
*IEEE Intelligent Systems '06*

# Alternate Definition



“The process of creating situation awareness in situations of uncertainty”

– D. Leedom, '01 SM Symp. Report

Situation awareness:

“It’s knowing what’s going on so you know what to do”

– B. McGuinness, quoting an Air Force pilot

# Other VA Projects



- Just a few other nice examples of visual analytics...

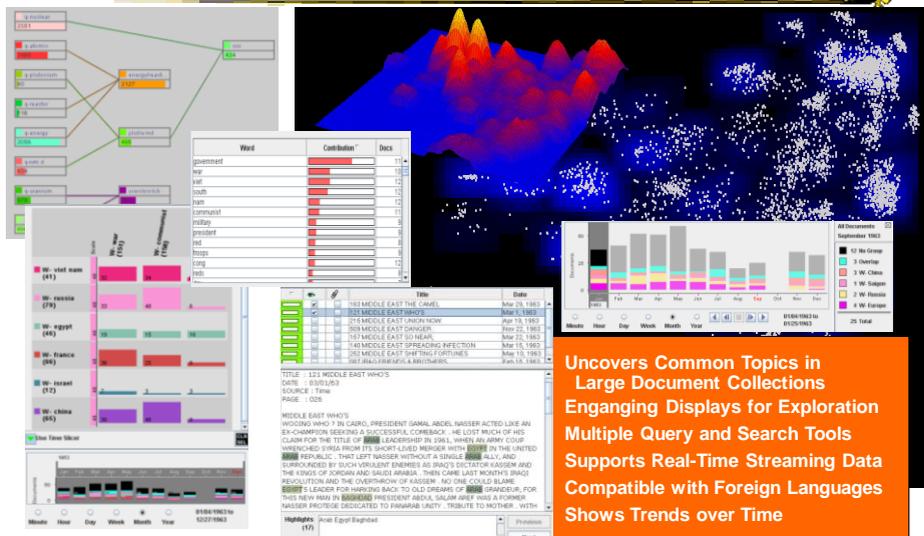
Fall 2016

CS 7450

85

## IN-SPIRE™ Visual Document Analysis

A "Thinking Aid" for advanced investigation of unstructured text



Fall 2016: <http://in-spire.pnl.gov>

CS 7450

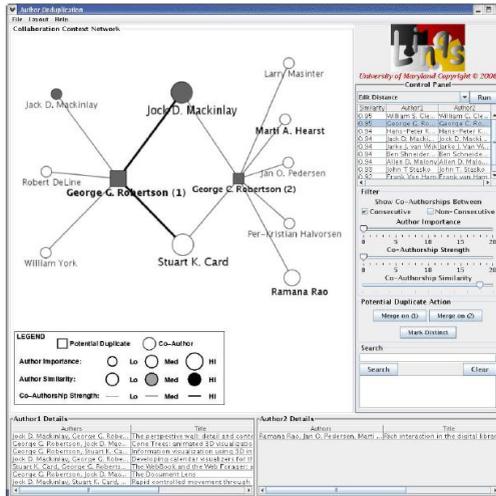
Video

86

86

# D-DUPE

Video



System for entity resolution in large networks such as bibliographic collections

System does computational analysis and provides suggestions and user can augment and correct

Fall 2016

CS 7450

Bilgic et al  
VAST '06

87

# WireVis

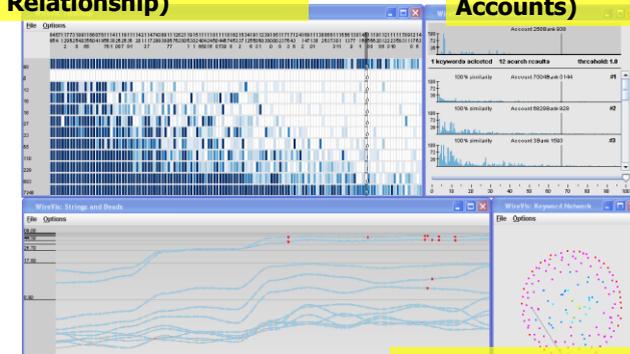
Video



Heatmap View  
(Accounts to Keywords Relationship)

Search by  
Example (Find Similar Accounts)

Helping Bank of America examine wire transfers of money to detect money laundering and fraud



Strings and Beads  
(Relationships over Time)

Keyword Network  
(Keyword Relationships)

Look for certain temporal patterns and keywords in descriptions

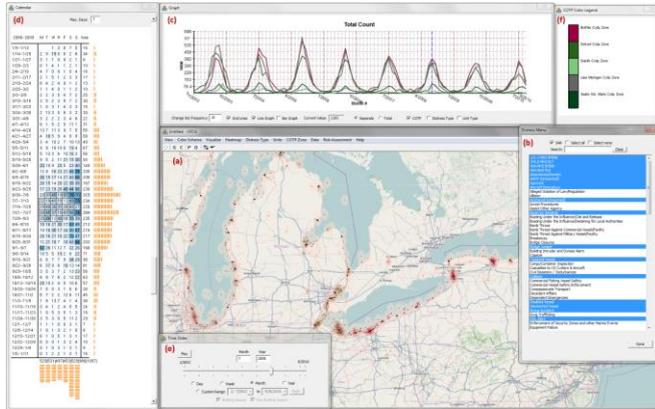
Fall 2016

CS 7450

Chang et al  
Information Visualization '08

88

# Coast Guard Search & Rescue



Shows stations, incidents, response times

Visualize historical data and support "what if" explorations

Calculate risk assessments and then communicate visually

Malik et al  
VAST '11

Fall 2016

CS 7450

89

## Many Others



- A number of nice examples shown earlier on Graph & Network visualization day
  - Perer: Social Action
  - etc.

Fall 2016

CS 7450

90

## Other Courses to Take



- CSE 6242 (Prof. Chau)
  - Data and Visual Analytics
- CS 8803 Special Topics (Prof. Endert)
  - Visual Data Analysis

Fall 2016

CS 7450

91

## Learning Objectives



- Describe motivation behind visual analytics
- Discuss differences between "statistical" and human-centered data analysis processes, including strengths of each
- Explain visual analytics
  - Define the term
  - List its components
  - Explain the differences between it and information visualization
- Define sensemaking
- List and describe some visual analytics applications

Fall 2016

CS 7450

92

# Reading



- Keim et al, '08

# Project Timeline



- Finish up system development this week and next
- Meet TAs/myself
- Prepare video
- Demo (20 minutes) next Thu-Fri 8<sup>th</sup> & 9<sup>th</sup>
  - Sign up on t-square
- Video showcase Fri 9<sup>th</sup> 2:50-5:40pm

## Video Advice



- Use Camtasia
- Process
  - 1. Develop script (rehearse timing)
  - 2. Record script
  - 3. Capture video of demo to script
  - 4. Add effects
- You've seen examples all semester
  - eg, <http://www.cc.gatech.edu/gvu/ii/videos.html>

Fall 2016

CS 7450

95

## Video Advice

< 5 minutes



- Script
  - Introduce problem
  - Describe visualization & system
  - Walk through usage scenario

Fall 2016

CS 7450

96

# Exam



- Evaluating the learning objectives
  - Lectures and reading material
- Short answer questions
  - Draw visualization techniques
  - Explain concepts, +/-, differences, ...
  - Identify technique and systems
  - Analyze and critique visualizations
  - Samples coming tonight or tomorrow morn

Fall 2016

CS 7450

97

# Upcoming



- Exam
- Evaluation in visualization

Fall 2016

CS 7450

98