



Examining the Use of a Visual Analytics System for Sensemaking Tasks: Case Studies with Domain Experts

Youn-ah Kang and John Stasko
Georgia Institute of Technology



Evaluation in Visual Analytics

We evaluate systems because..

- We hope our technologies are making an impact
- We hope they are helping people gain value from their information

Evaluation is challenging and not very common

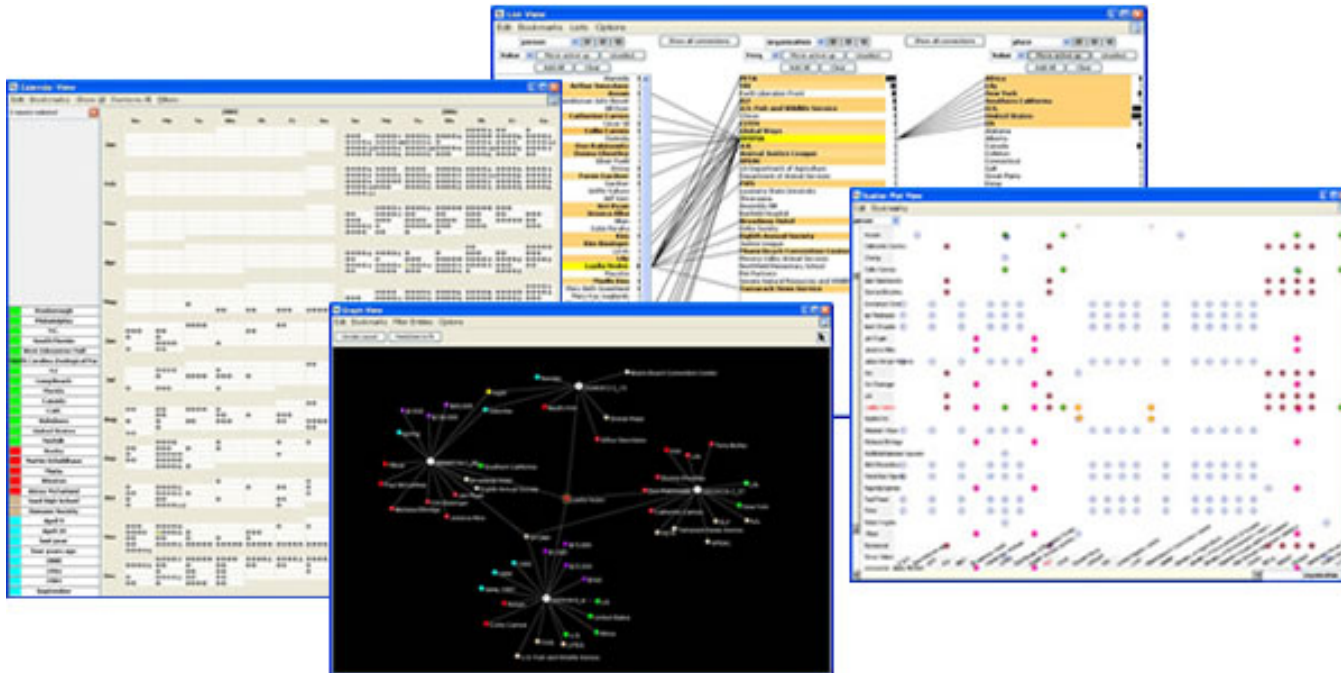
- Usability testing and controlled experiments remain crucial
- Actual case studies of prolonged system use by analysts working with their own data are still rare

Case Studies

- Case studies can **provide valuable findings and insights**
 - Yield a description of **how a tool was used and where users had problems**
 - Difficult to achieve through controlled lab studies
- Multi-dimensional in-depth long-term case studies (MILC)
 - B. Shneiderman and C. Plaisant (2006). Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies.

System of Study

- Jigsaw (<http://www.cc.gatech.edu/gvu/ii/jigsaw/>)



Motivation and Goals

- Is Jigsaw helping analysts with their tasks and problems?
- For what types of documents and analyses does Jigsaw help?
- What are particularly useful features/capabilities as well as missing or problematic ones?

Recruitment and Study Protocol

- 6 working/practicing investigators who were using the system
 - 3 intelligence analysts, 2 academic researchers, and 1 business analyst
- Used Jigsaw for a range of 2-14 months
- Semi-structured interviews
- Follow-up email conversations

Case Studies



P1: Aerospace Engineering Researcher

- ▶ **Goal:** Identify similarities/differences and **create a mapping** between the two programs
 - ▶ Does a concept or capability suggested in one program also appear in the other program?

Program A	Program B
	L07-02 TS-0201: Basic Departure Management (DMAN)
	L07-02 TS-0202: Departure Management Synchronized with Pre-Departure Sequencing
	L07-02 TS-0203: Integration of Surface Management Constraint into Departure Management
OI-0320: Initial Surface Traffic	L07-02 TS-0306: Optimized Departure Management in the Queue Management Process
	L10-02 AO-0205: Automated Assistance to Controller for Surface Movement Planning and Routing
	L10-03 AO-0501: Improved Operations in Adverse Conditions through Airport Collaborative Decision Making

P1: Aerospace Engineering Researcher

- ▶ **Originally done manually** using MS word and search
 - ▶ Search for descriptions of program A -> identify keywords -> review descriptions of program B containing matching keywords one by one
- ▶ **Jigsaw helped:** Review and compare the huge document collection and complete the mapping between the two
 - ▶ Published the work at *Aviation Technology Integration and Operations ATIO Conference*

P2: Business Analyst at an Accounting Firm

- **Task:** Analyze unstructured data and identify any linkages between people/ companies relevant to a financial fraud
- **Goal:** Find **evidence for a financial fraud**
- **Before Jigsaw:** Put all documents into an Excel spreadsheet, search for keywords, and read all returned documents
- **Jigsaw helped:** Reveal connections between people & companies that were not easily identifiable
 - Found evidence of a financial fraud after analyzing 100,000 emails



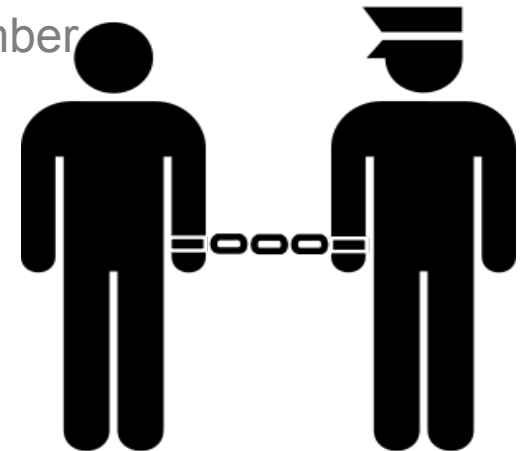
P3: Industrial & Systems Engineering Researcher

- **Task:** Validate her model about company transformation by combining historical company data (5,000+ announcements and news articles of 9 IT firms for 10 years)
- **Goal:** **Make sense of the documents** and extract keywords for the next step – data mining
- **Jigsaw helped:** Attain a clear understanding of the documents in a short amount of time



P4: Intelligence Analyst at a Police Department

- **Task:** Make sense of daily incident reports and identify patterns, trends, and any top issues in the city
- **Goal:** Find **connections between individuals, places, and other incidents** within accumulated crime reports
- Originally read all the reports individually and tried to remember different connections using printed copies of the documents
- **Jigsaw helped:** Develop a repository of important connections
 - Helped the police arrest a criminal by identifying where he might be



P5: Intelligence Analyst at a National Lab

- **Task:** Review resumes and find a good candidate with a certain specialty
- **Goal:** Examine connections in candidate info and **find an expert in a specialized area**
 - Skills, publications, co-authors, education, employment history.
- Performed using Analyst's Notebook but felt limited
- **Jigsaw helped:** Identify possible connections between people and technology



P6: Intelligence Analyst with the US Air Force

- **Task:** Examine budget summaries of R&D programs in the Department of Defense and identify common themes (10,000+ documents from 20+ agencies such as Air Force, Navy, and DARPA)
- **Goal:** Highlight **what programs/topics are similar**, what makes them similar, and who are working on similar topics
- **Jigsaw helped:** Effectively search for similar tools and technologies that required further investigation



Types of Tasks

- Relationship / connection between entities
 - Targeted investigation rather than seeing the big picture
- Search / comparison
 - If the documents contain specific keywords
- Understanding
 - Getting an overview of documents
- As a communication aid / shared understanding of data
 - Persuasive power in communication

Learning the System

- Learning curve existed
 - “How to better analyze my data using this tool”
- **Constructing a frame**
 - Which views are most appropriate for my data and task?
 - What entity types do I want to put in this column?

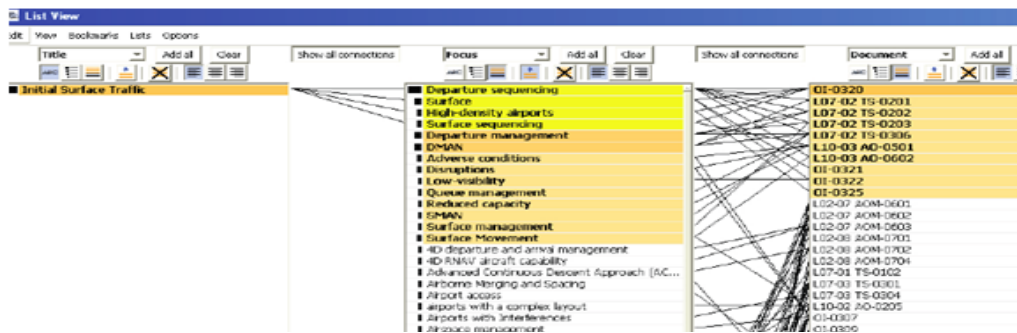
→ Finding the optimal approach
in their own way



Unexpected Use of the System

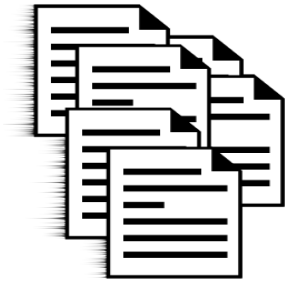
- Views for evidence/output generation

NextGen Operational Improvements	SESAR Operational Improvements and Steps
O1-0320: Initial Surface Traffic	L07-02 TS-0201: Basic Departure Management (DMAN)
	L07-02 TS-0202: Departure Management Synchronized with Pre-Departure Sequencing
	L07-02 TS-0203: Integration of Surface Management Constraint into Departure Management
	L07-02 TS-0306: Optimized Departure Management in the Queue Management Process
	L10-02 AO-0205: Automated Assistance to Controller for Surface Movement Planning and Routing
	L10-03 AO-0501: Improved Operations in Adverse Conditions through Airport Collaborative Decision Making
	L10-03 AO-0602: Collaborative Pre Departure Sequencing



A mapping created manually (top) and by Jigsaw (bottom)

Unexpected Use of the System



Information-dense documents

Separate docs into several projects

Merge new incoming documents with an existing Jigsaw project

Build a historical dataset

Issues and problems

- **Technical issues in the preparation stage**
 - Importing data into Jigsaw
 - Identifying entities
- **Limited filtering options**
 - Not being able to easily select a subset of data in the views



Design Implications

- **Supplement automatic entity identification**
- Allow flexible data (document) management
 - Provide an ability to easily select a subset of documents
- Empower with numbers
 - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
 - Limit user interaction vs. give more power
- Invest in tutorial
 - Break down into subtopics with use-cases and examples

Design Implications

- Supplement automatic entity identification
- **Allow flexible data (document) management**
 - Provide an ability to easily select a subset of documents
- Empower with numbers
 - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
 - Limit user interaction vs. give more power
- Invest in tutorial
 - Break down into subtopics with use-cases and examples

Design Implications

- Supplement automatic entity identification
- Allow flexible data (document) management
 - Provide an ability to easily select a subset of documents
- **Empower with numbers**
 - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
 - Limit user interaction vs. give more power
- Invest in tutorial
 - Break down into subtopics with use-cases and examples

Design Implications

- Supplement automatic entity identification
- Allow flexible data (document) management
 - Provide an ability to easily select a subset of documents
- Empower with numbers
 - Degree centrality, betweenness, closeness
- **Consider allowing visualization modification**
 - Limit user interaction vs. give more power
- Invest in tutorial
 - Break down into subtopics with use-cases and examples

Design Implications

- Supplement automatic entity identification
- Allow flexible data (document) management
 - Provide an ability to easily select a subset of documents
- Empower with numbers
 - Degree centrality, betweenness, closeness
- Consider allowing visualization modification
 - Limit user interaction vs. give more power
- **Invest in tutorial**
 - Break down into subtopics with use-cases and examples

Contributions

- Identified real-world cases of how an interactive visual system for investigative analysis assisted document sensemaking in various domains and tasks
- Discussed issues and findings that emerged upon the use of the visual analytic system
- Provided design recommendations for the system and future visual analytics tools.

Acknowledgements

- We thank our six professionals for sharing their experience with Jigsaw.
- This work was supported by the National Science Foundation under awards IIS-0915788, CCF-0808863, and the VACCINE Center, a Department of Homeland Security's Center of Excellence in Command, Control and Interoperability.

