VISUAL TEXT ANALYTICS IN JIGSAW

Nadia Boukhelifa nb@lri.fr

VISUAL ANALYTICS 8 OCT 2014

BEFORE WE START

1. **Import a dataset**

   InfovisVast-papers.jig

2. **Run Computational Analysis**

   Tools -> compute all *5 min*

OVERVIEW

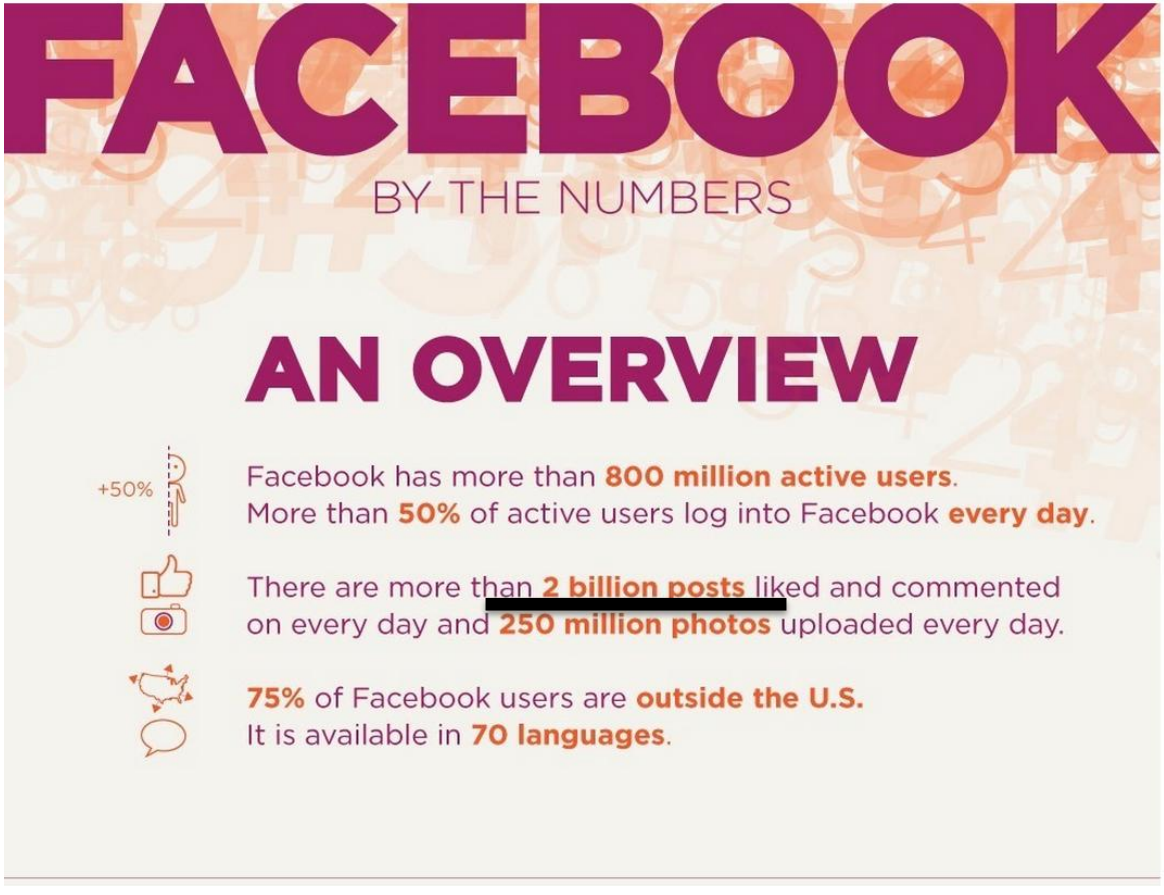
Introduction to Jigsaw

Visual Text Analytics using Jigsaw

UBIQUITOUS IMAGE

UBIQUITOUS IMAGE

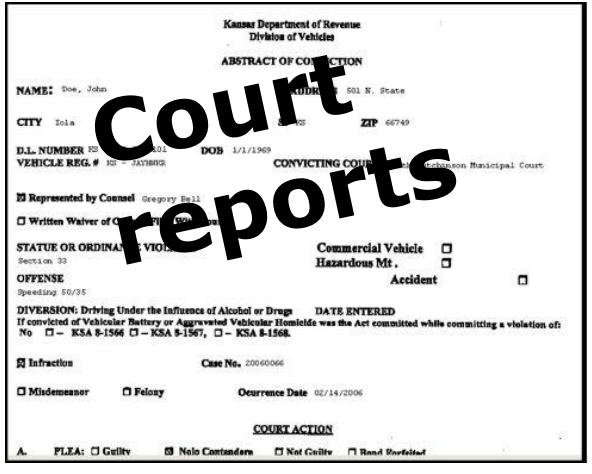
UBIQUITOUS IMAGE



[url0]

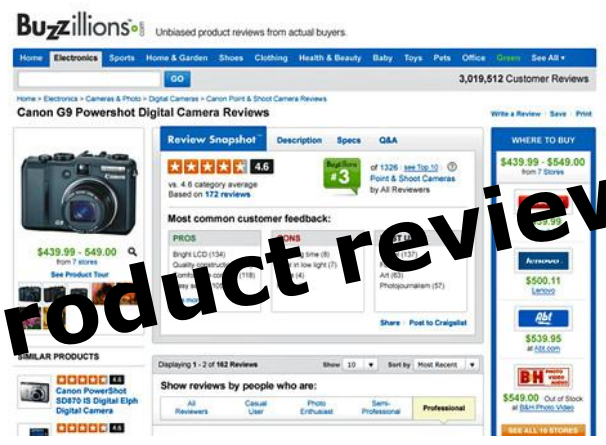
TEXT EVERYWHERE

News reports



Court reports

Academic articles



Product reviews

Online forums

E-mail

TEXT SEARCH

TEXT SEARCH

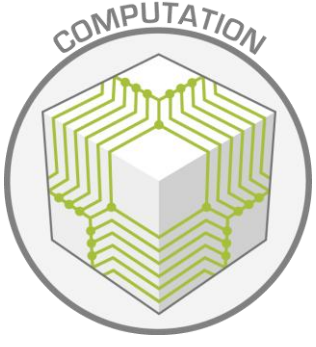
VISUAL TEXT ANALYTICS

- Assist analysts to overcome **information-overload**
  - Information foraging
  - Sense-making

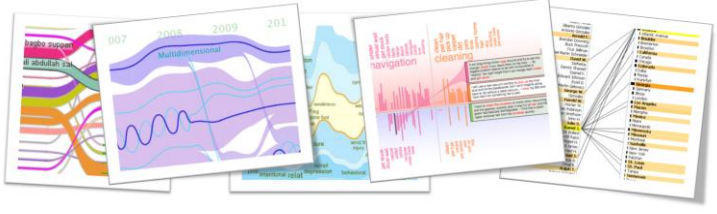
[url1]

VISUAL TEXT ANALYTICS

- Text mining & Computational Analysis
- Interactive text visualization



+



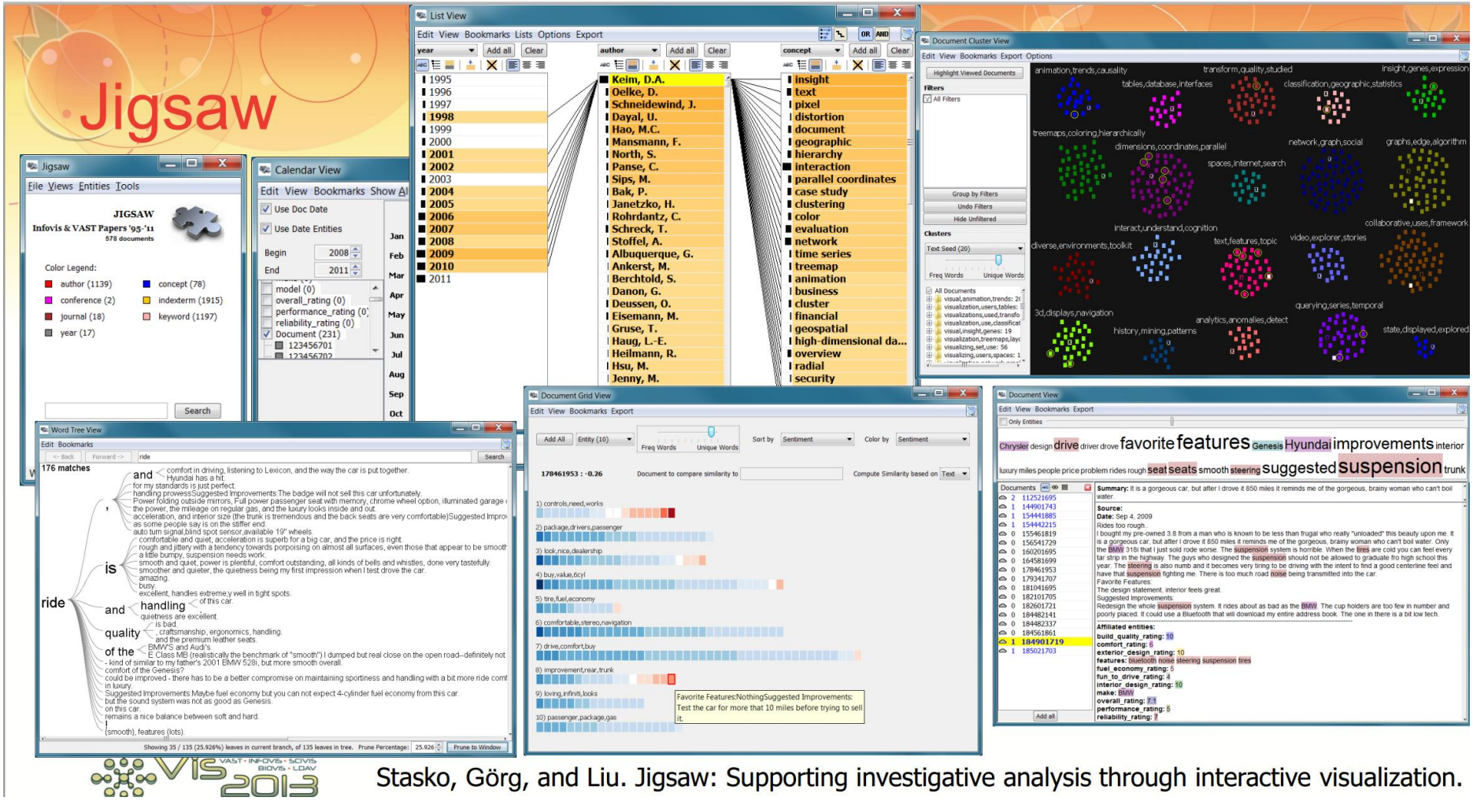
[url2,url3]

CHALLENGES

- How to *extract & visualize* concepts and relations from large noisy corpora

- How to design new visualization / interaction techniques to *browse* and *refine* content without becoming data mining expert

[url4]



Stasko, Görg, and Liu. Jigsaw: Supporting investigative analysis through interactive visualization.

Jigsaw: Supporting Investigative Analysis through Interactive Visualization

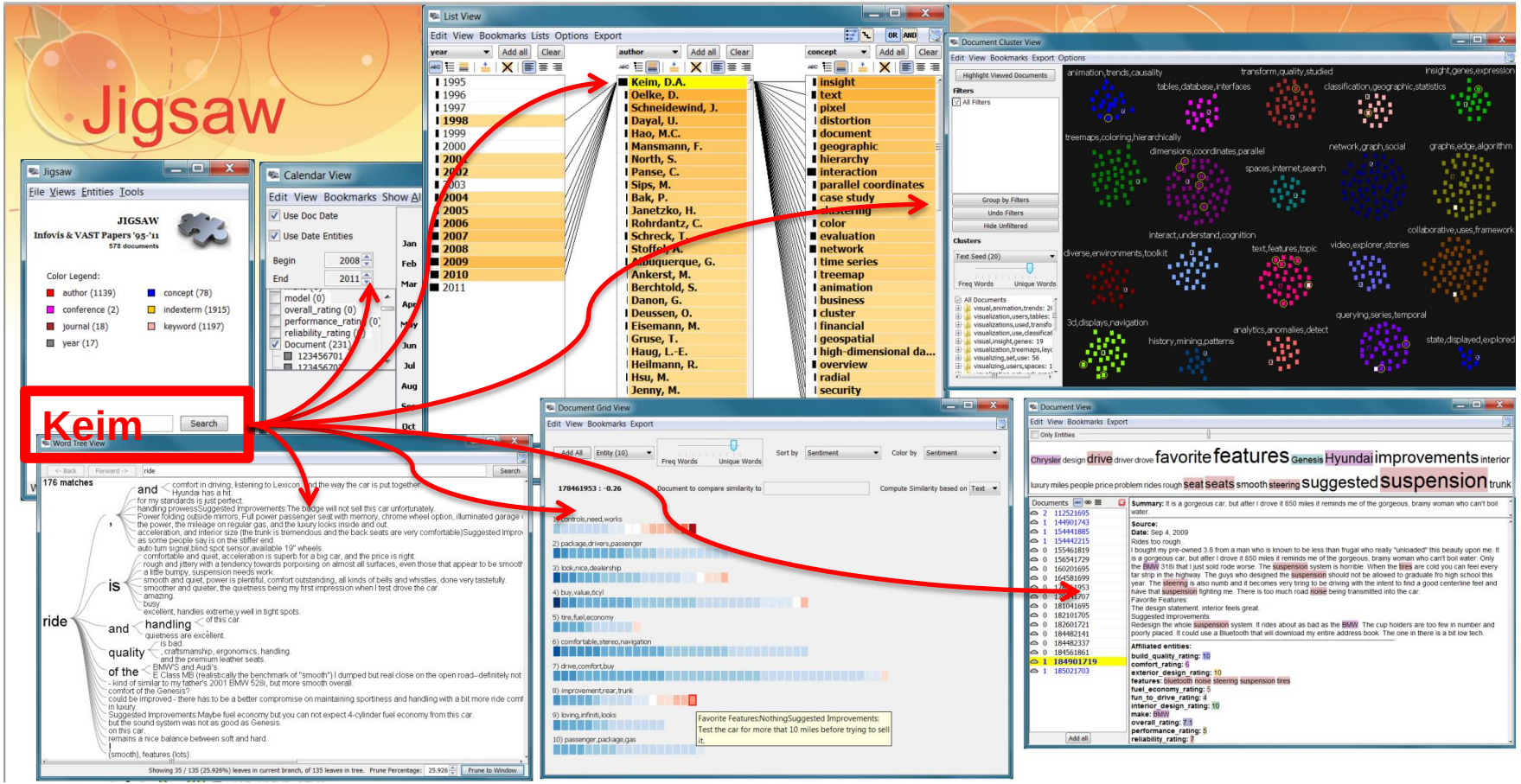
John Stasko, Carsten Görg,
Zhicheng Liu, Kanupriya Singhal

School of Interactive Computing & GVU Center
Georgia Institute of Technology

NAMED ENTITIES

Entities are connected if they appear in the same document

CORDINATED MULTIPLE VIEWS



STRATEGIES

- Overview, filter & detail

Summarisation, doc metrics

- Build from detail

Doc similarity, recommended related docs

- Hit the keyword

Keyword search

- Find a clue and follow the trail

Document clustering

DEMO

Exploring academic publications: Infovis and VAST 1995-2013

ASSIGNMENT

1. **Import a dataset**

   articles_enriched.jig

2. **Extract entities** ***3 minutes***

   entities -> identify entities -> Ilinois-NER

   entities -> clean up entities

3. **Run Computational Analysis** ***10 minutess***

   Tools -> compute all

ASSIGNMENT

- Don’t forget to save your work!
- Send me PDF file with answers
- Deadline: Mon13th Octobre 2014, 23:00

ACKNOWLEDGEMENTS

- John Stasko, Georgia Tech
- Alex Kachkaev, City University London
- Aidan Slingsby, City University London

REFERENCES

- John Stasko, Carsten Görg, and Zhicheng Liu, **"Jigsaw: Supporting Investigative Analysis through Interactive Visualization"**, *Information Visualization*, Vol. 7, No. 2, Summer 2008, pp. 118-132.

- Carsten Görg, Zhicheng Liu, Jaeyeon Kihm, Jaegul Choo, Haesun Park, John T. Stasko, **"Combining Computational Analyses and Interactive Visualization for Document Exploration and Sensemaking in Jigsaw"**, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 10, October 2013, pp. 1646-1663

- http://www.cc.gatech.edu/gvu/ii/jigsaw/

URLS

1. http://mashable.com/2011/10/21/facebook-infographic/
2. http://markjowen.wordpress.com/2012/12/19/a-very-brief-history-of-search/
3. http://msr-waypoint.com/en-us/um/people/shliu/tasmc/
4. http://www.cloudsoftwareprogram.org/superior-user-experience
5. https://projects.cs.dal.ca/visualtextanalytics/