Session 7: Early Case Assessment

LBSC 708X/INFM 718X
Seminar on E-Discovery

Jason R. Baron
Adjunct Faculty
University of Maryland
March 8, 2012

Hradcany Castle and Charles Bridge, Prague

How much does counsel know?

- About the nature and scope of the relevant evidence existing in custody and control of the opposing party?
- Same question re the relevant evidence existing in one's own client's custody and control – how much is known?
- How much is know-able?
- Have the answers changed in past decade of practice?

Early Case Assessment

- **Early case assessment** refers to estimating risk (cost of time and money) to prosecute or defend a legal case.
- Human expertise
- Technology

Why ECA?

--Most cases settle prior to trial

-- Exposure (likelihood of negative outcome in litigation)

--Discovery and document review in particular are costly, burdensome

--Protecting against sanctions (spoliation should preservation obligations not be carried out with due diligence)

--Speed

--Insight into case (readings this week)

--Establishing a budget

ECA lifecycle

- Risk/benefit analysis
- Putting into place legal hold
- Collecting/culling info for attorney review
- Processing relevant info using filtering, search terms, predictive coding, etc.
- Producing info (in discovery)
- Re-using info, KM, BI

ECA Downsides

Up front costs

False negatives (example case)

Software approaches to ECA

- Keywords
- Relevance ranking
- Clustering
- Predictive Coding
- Sentiment detection
- Social network analysis
- Visual analytics

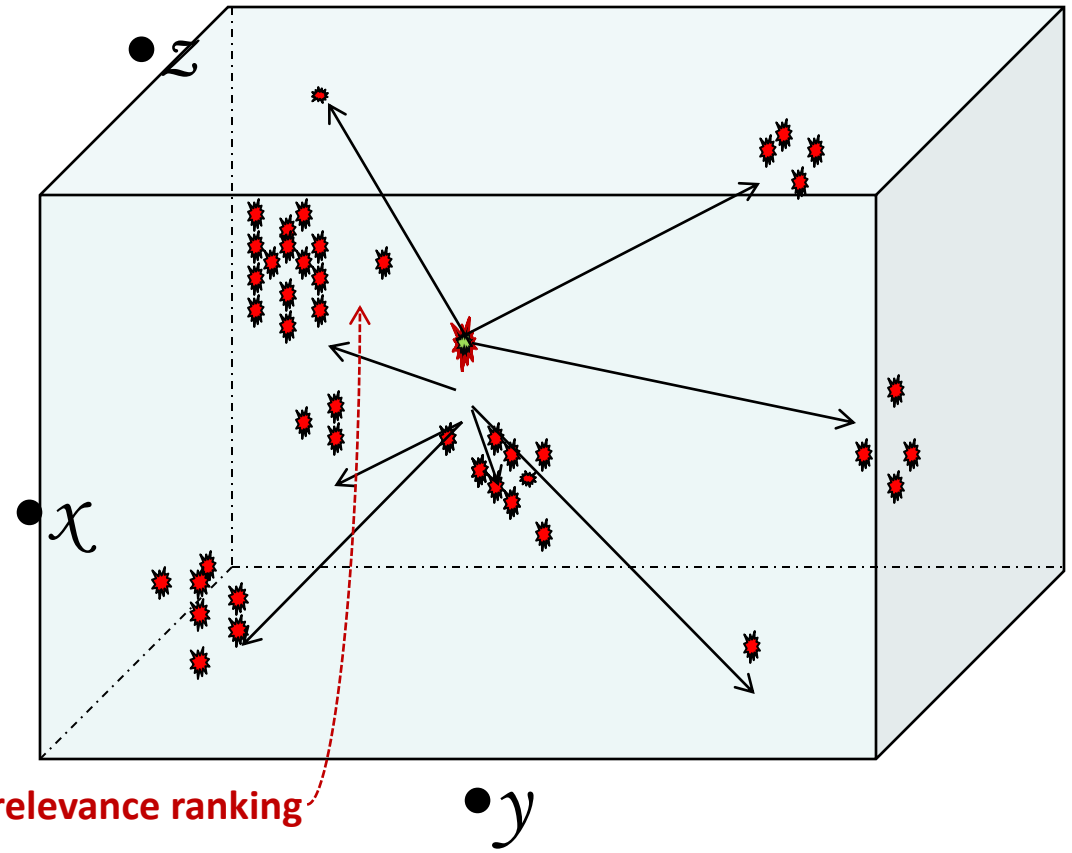
Bayesian Statistical Models

•Based on mathematical models of Statistical Probability to recognize documents of similar content.

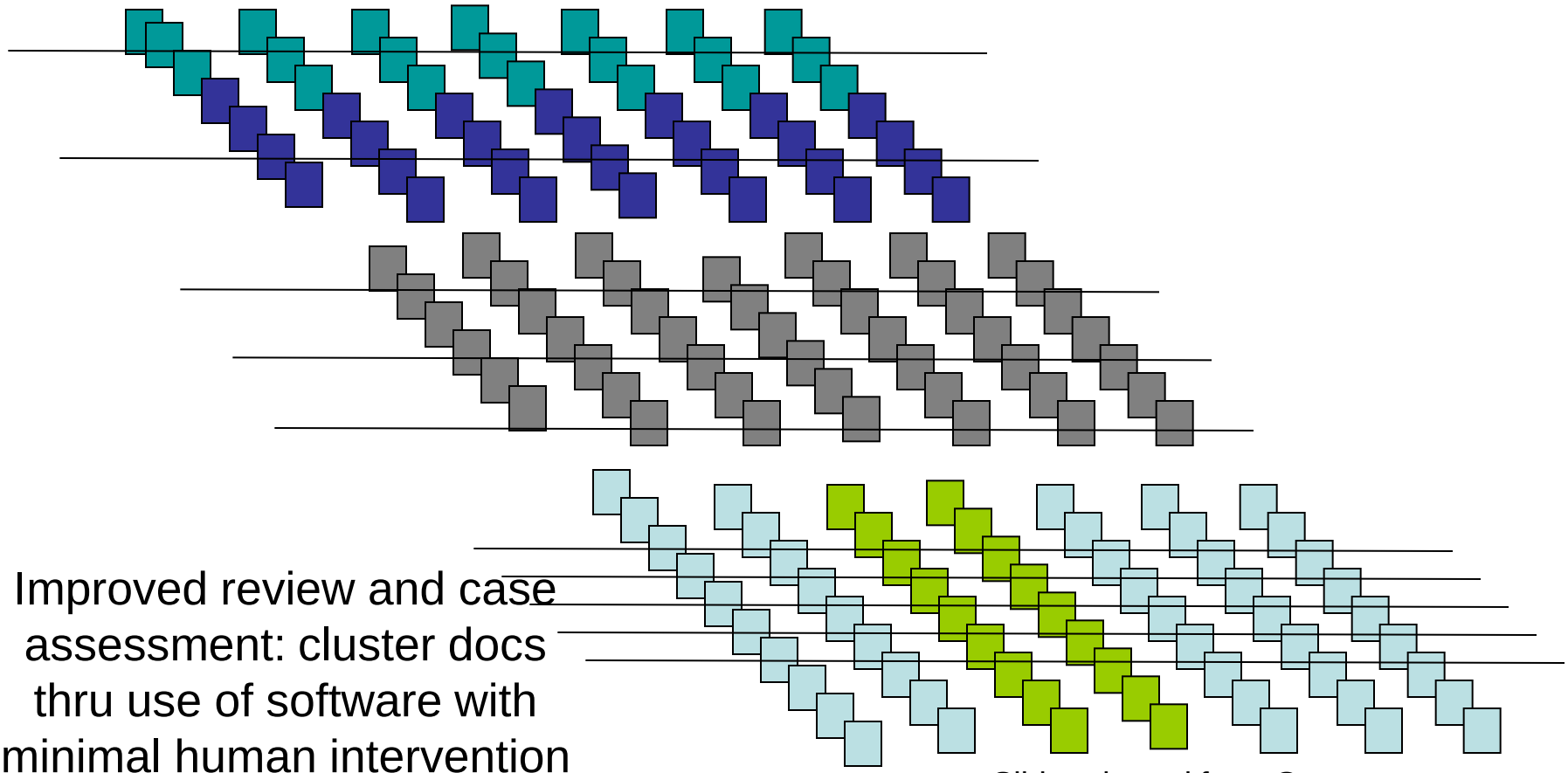
- Learns passively from the document content
- Position, frequency and proximity of terms (language independent) combine to create a mathematical “thumbprint” of concepts contained in documents.
- Useful to “cluster” documents by content
- Can “learn” to build clusters from exemplar sets
- Requires re-indexing and assessment can change

Latent Semantic Indexing (LSI)

1. SVD (Singular Value Decomposition) assigns each record to a place creating "clusters"
2. "Query" documents are SVD analyzed and placed in the matrix
3. "Hits" and rankings are determined by the distance from clusters

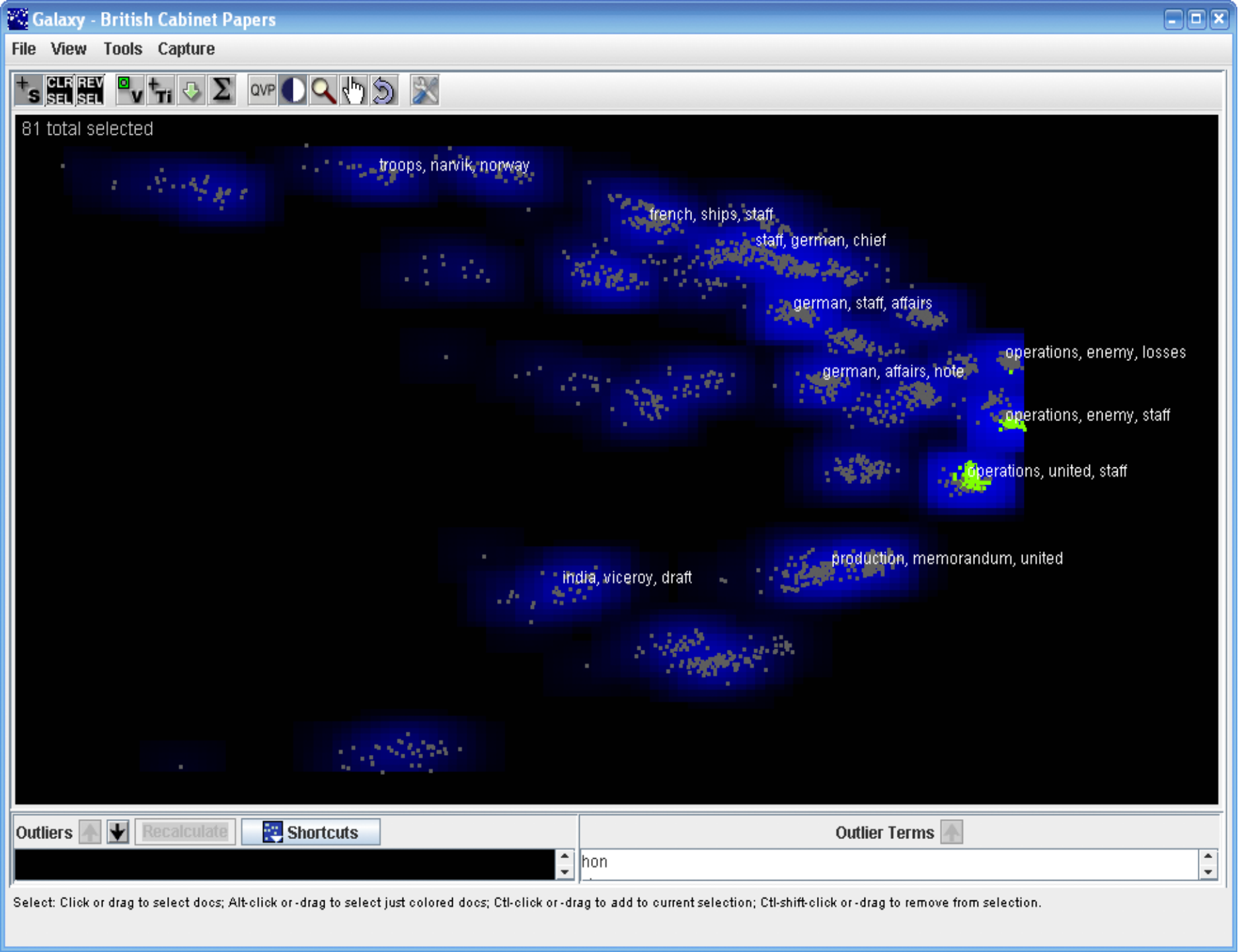


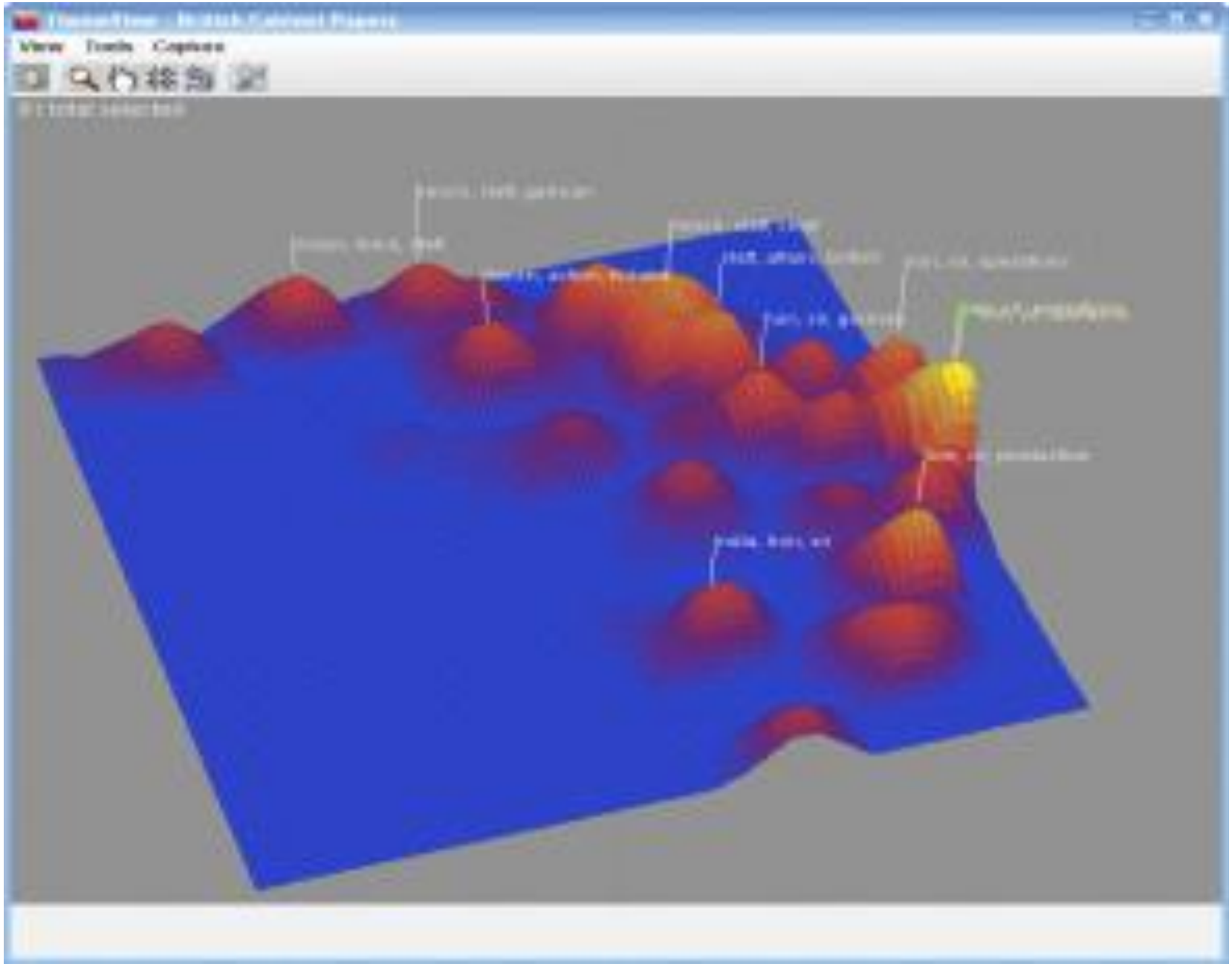
Emerging New Strategies: "Predictive Analytics"



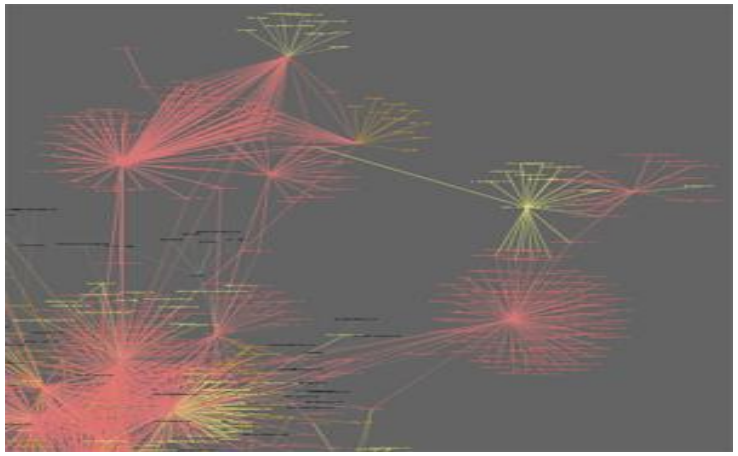
Improved review and case assessment: cluster docs thru use of software with minimal human intervention at front end to code "seeded" data set

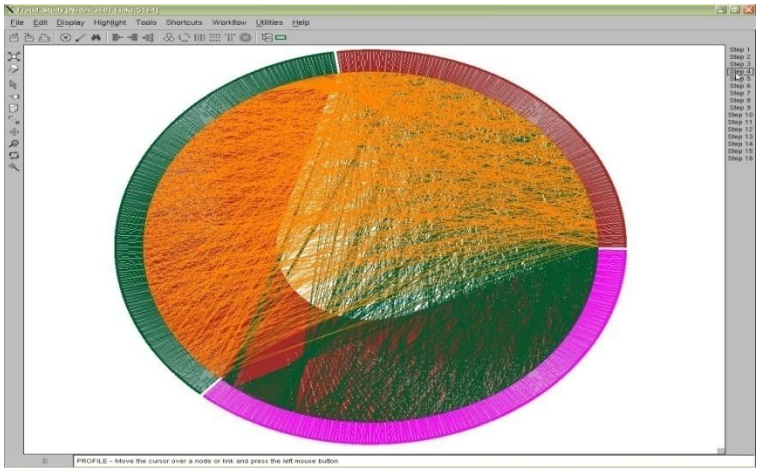
- Slide adapted from Gartner Conference
- June 23, 2010 Washington, D.C.

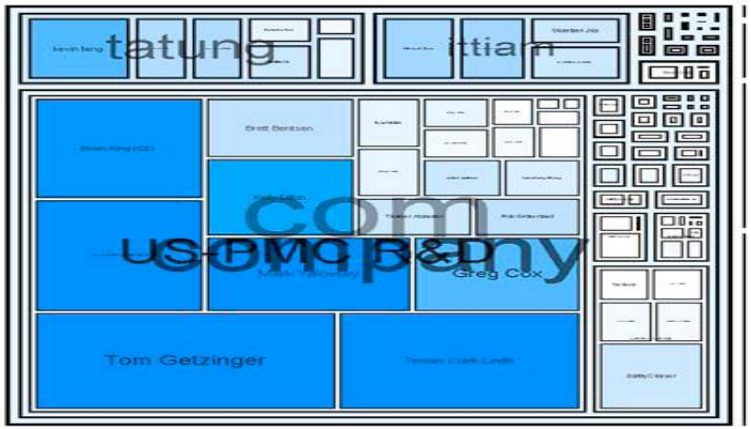

Galaxy - British Cabinet Papers
File View Tools Capture
81 total selected
troops, narvik, norway
french, ships, staff
staff, german, chief
german, staff, affairs
operations, enemy, losses
german, affairs, note
operations, enemy, staff
operations, united, staff
production, memorandum, united
india, viceroy, draft
Outliers
Recalculate
Shortcuts
Outlier Terms
hon
Select: Click or drag to select docs; Alt-click or -drag to select just colored docs; Ctl-click or -drag to add to current selection; Ctl-shift-click or -drag to remove from selection.


View
Tools
Capture

Visual Analysis Examples
(Presentation by Dr. Victoria Lemieux, Univ. British Columbia, at Society of American Archivist Annual Mtg. 2010, Washington, D.C.)





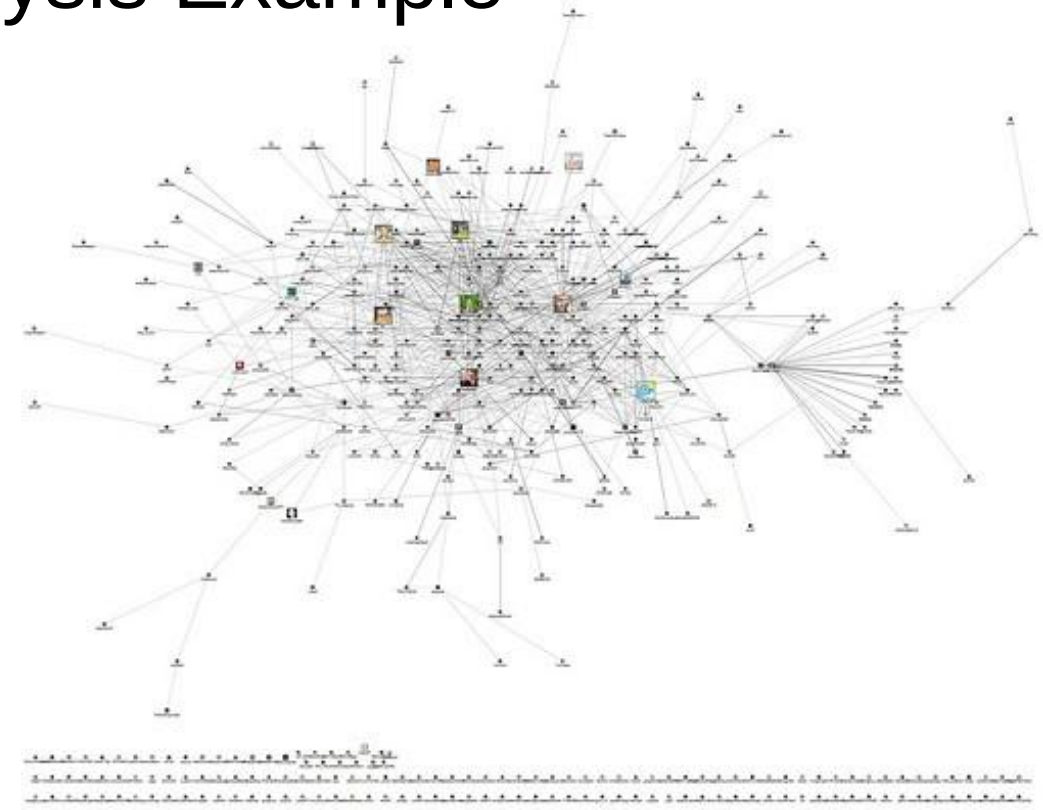


•With acknowledgments to Jeffrey Heer, *Exploring Enron*, http://hci.stanford.edu/jheer/projects/enron/,

•Adam Perer, *Contrasting Portraits*, http://hcil.cs.umd.edu/trs/2006-08/2006-08.pdf,

•and Fernanda Viegas, *Email Conversations*, http://fernandaviegas.com/email.html

•Social Networking/Links Analysis Example



•From Marc Smith
•Posted on Flickr
•Under Creative Commons License