

# CIS 4930/6930-002

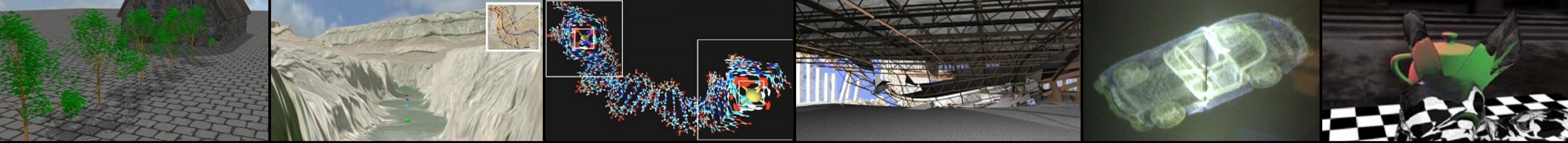
## DATA VISUALIZATION



### VISUALIZING TEXT

Paul Rosen  
Assistant Professor  
University of South Florida

slides credits Miriah Meyer (U of Utah)



## REMINDERS

3/19/2018 – Paper Review Presentations Begin

3/26/2018 – Paper Review 3 Due

3/28/2018 – Project 6 Due



Tables

Items

Attributes

Networks &  
Trees

Items (nodes)

Links

Attributes

Fields

Grids

Positions

Attributes

Geometry

Items

Positions

Clusters,  
Sets, Lists

Items

Text

?

## WHAT DOES IT MEAN TO BE AN “ITEM”?



## TEXT DATA TYPE

no numbers (implicitly)  
characters (ASCII)  
strings

b <sub>7</sub> b <sub>6</sub> b <sub>5</sub>				USASCII code chart								
b <sub>4</sub>	b <sub>3</sub>	b <sub>2</sub>	b <sub>1</sub>	Column	0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
↑	↑	↑	↑	Row ↓	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	\$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	8	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	
1	1	0	1	13	CR	GS	-	=	M	]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
1	1	1	1	15	S1	US	/	?	O	—	o	DEL



## TEXT DATA TYPE

no numbers (implicitly)  
characters (ASCII)  
strings

b <sub>7</sub> b <sub>6</sub> b <sub>5</sub>				USASCII code chart								
b <sub>4</sub>	b <sub>3</sub>	b <sub>2</sub>	b <sub>1</sub>	Column	0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
0	0	0	0	Row	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	\$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	8	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	
1	1	0	1	13	CR	GS	-	=	M	]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
												DEL

[6:10]

0	1	2	3	4	5	6	7	8	9	10	11
M	o	n	t	y		P	y	t	h	o	n
-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1

[-12:-7]



# TEXT DATA SEMANTICS

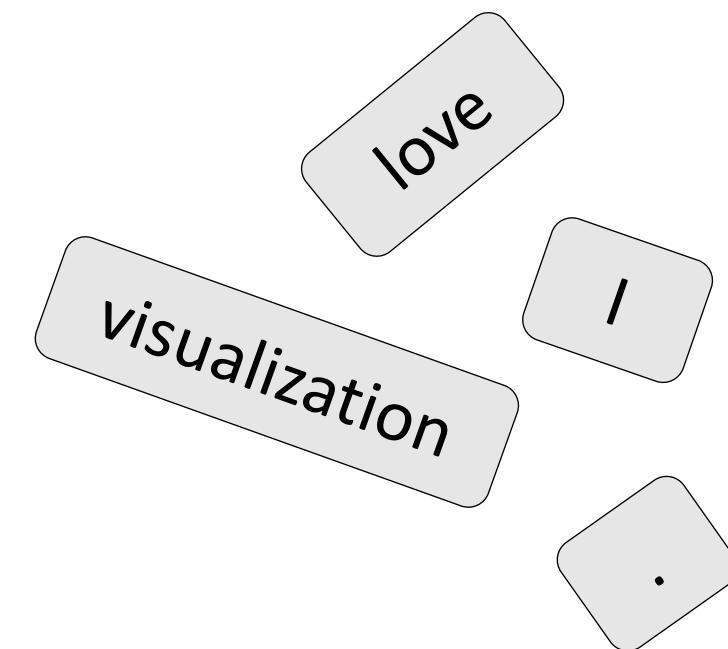
words

lines

sentences

paragraphs

chapters



I love visualization.





**SINGLE DOCUMENT**



# TAG CLOUD / WORD CLOUD

abstract accepted analogue applications applying attuned bar burgeoning challenging chapters chart collections combine communicate conducted convert **data** date difficult discussed earlier effectively end evaluation evocative familiar field focus focused form **general goal** graph highly **human** hundreds ideas images improve

**information** innovative insight kinds line makes m

meta-analysis nature new numbers order ost perceive perceptual points positive problems providing purpose range rapidly read reading reasons representations results retrieval robust search shortciten{chen2000esi} shortcite{larkin1987dsw} shown space studies successful system table task tasks text textual time translate underlying

# usability vibrant visual visualization



## Visualizations : definitions of visualization word tree

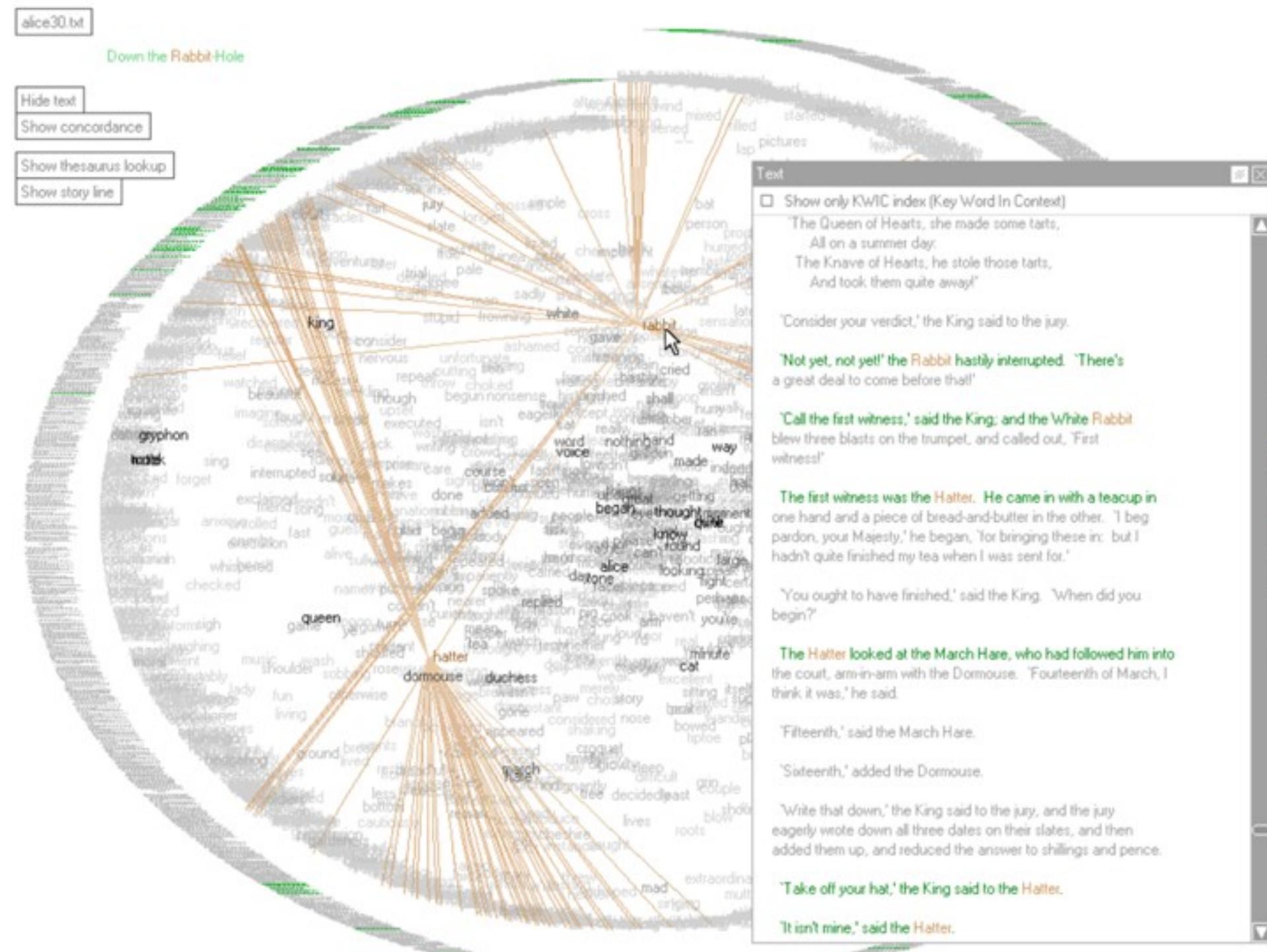
Uploaded by: mhalle

Created at: Wednesday May 21 2008, 11:37 PM

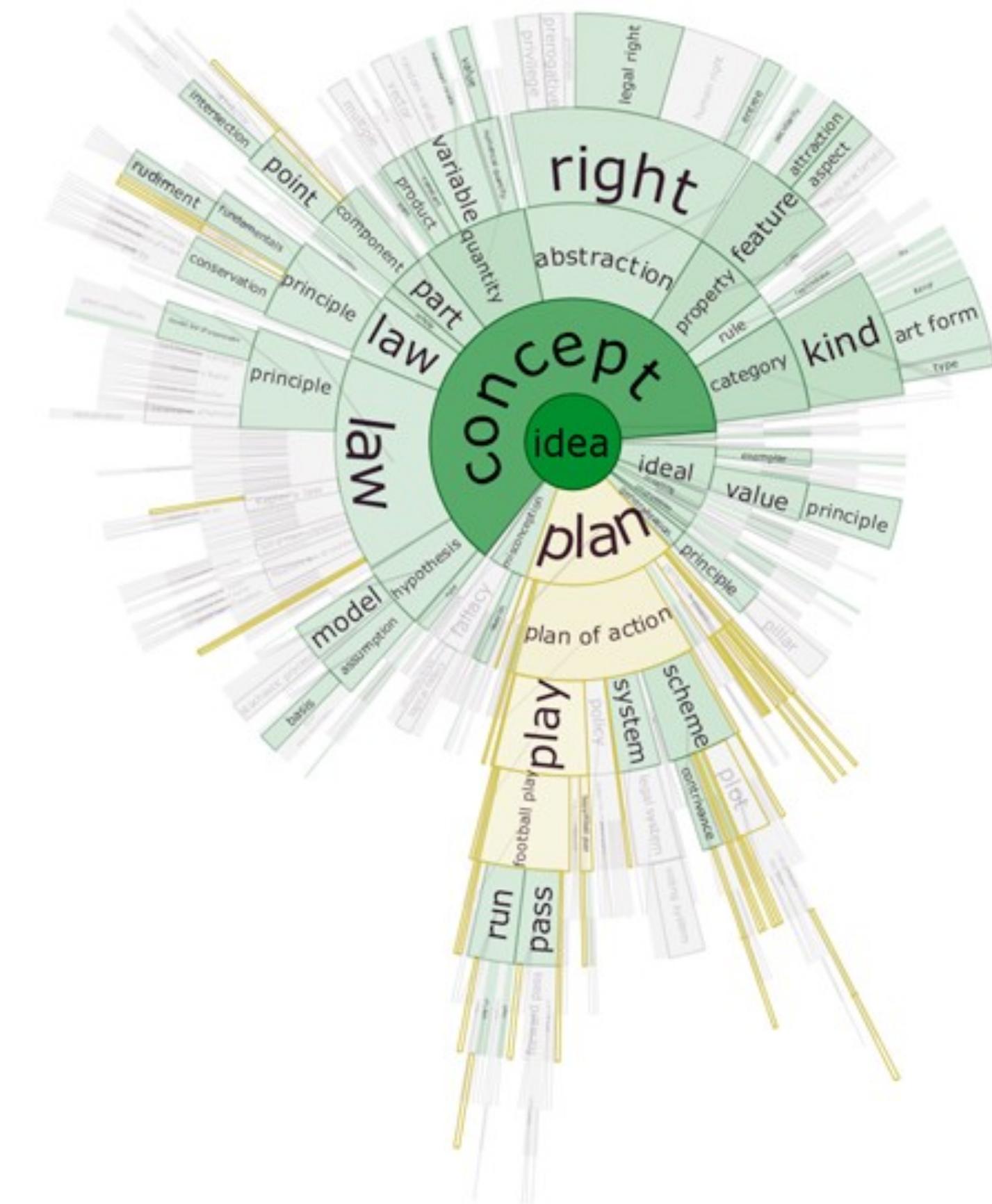
Tags: text



# TEXT ARC



# DOCUBURST

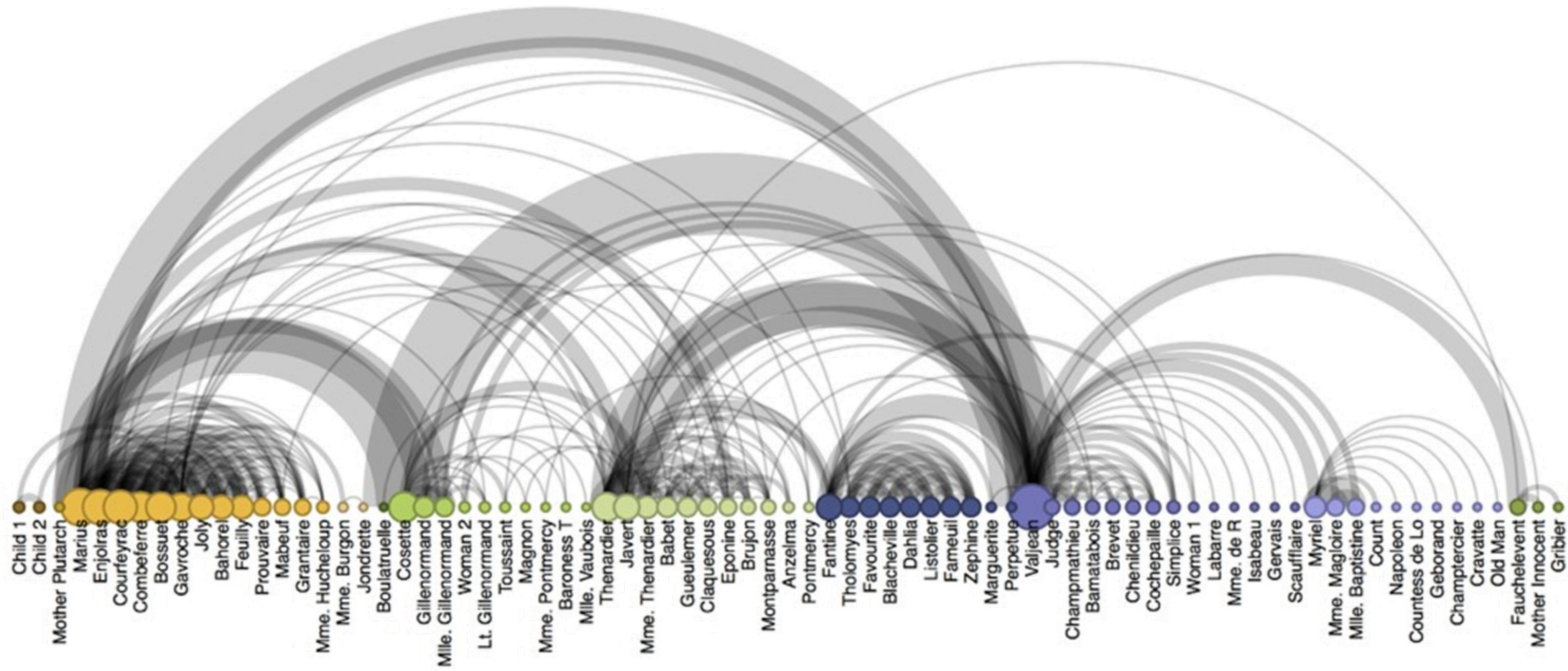


COLLINS, CARPENDALE, PENN 2008



# ARC DIAGRAM

abcd111110000011111abcd



# COLLECTION OF DOCUMENTS





# Parallel Tag Clouds to Explore and Analyze Faceted Text Corpora

Christopher Collins  
Fernanda B. Viégas  
Martin Wattenberg

▶ ▶◀ 0:01 / 7:40

CC HD □ □ [ ]

## Parallel Tag Clouds to Explore and Analyze Faceted Text Corpora



Christopher Collins

Subscribe 26

2,768

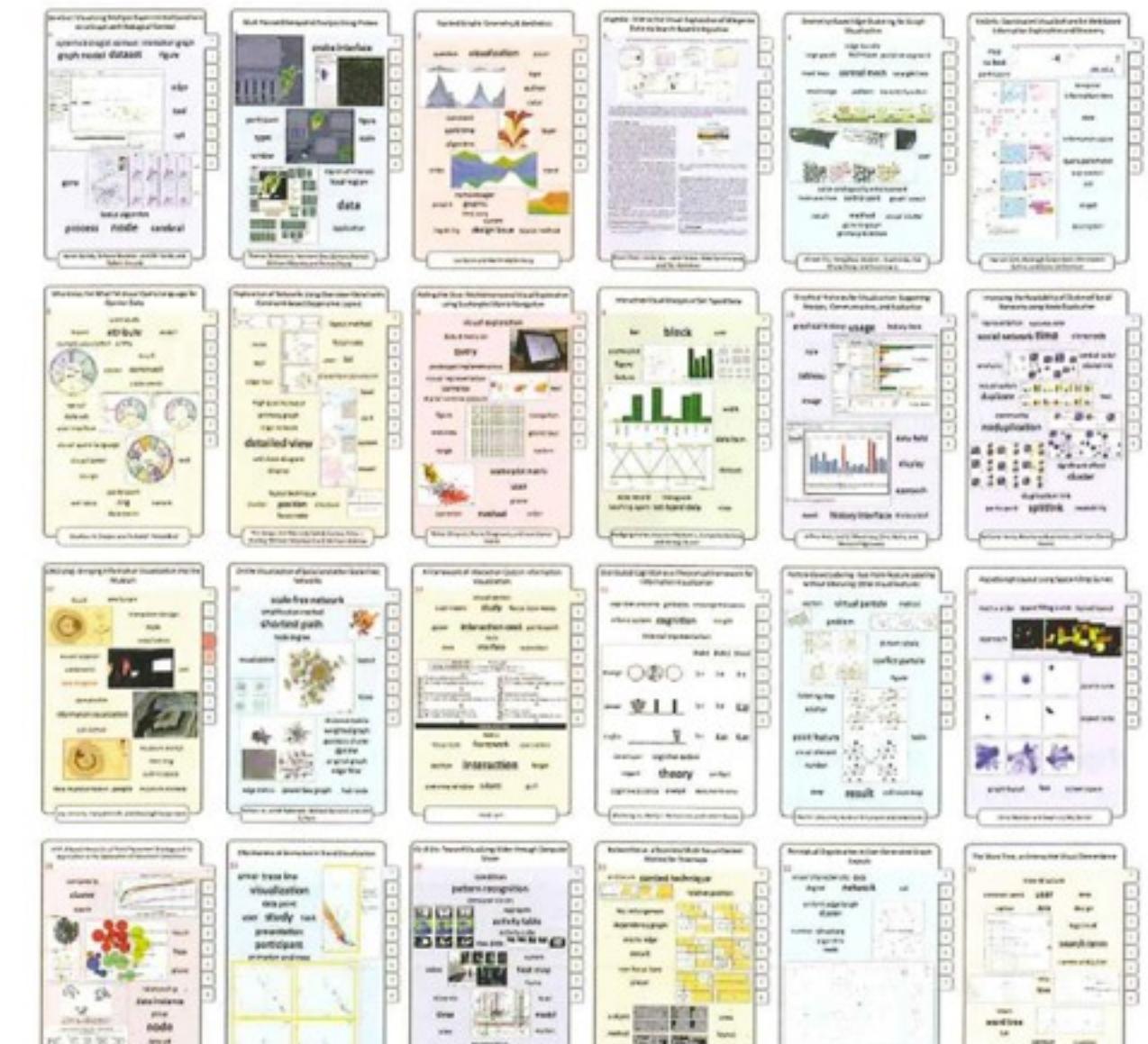
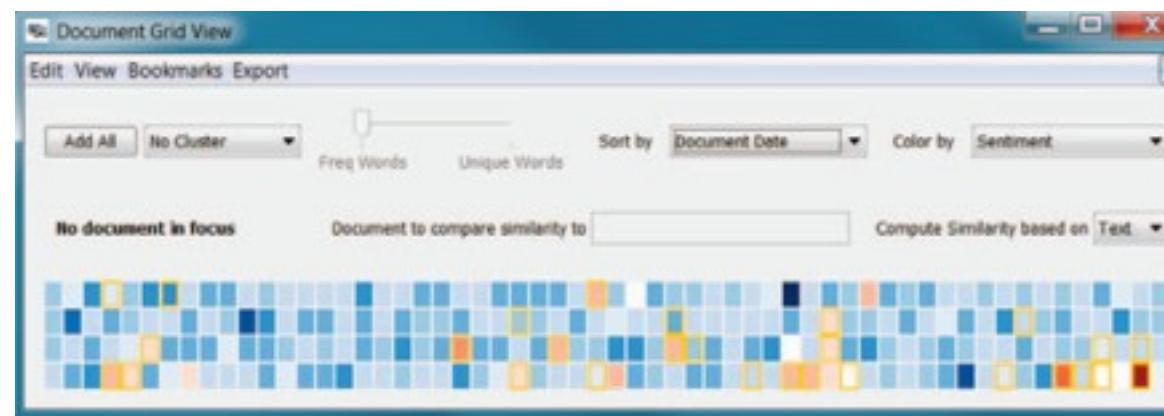
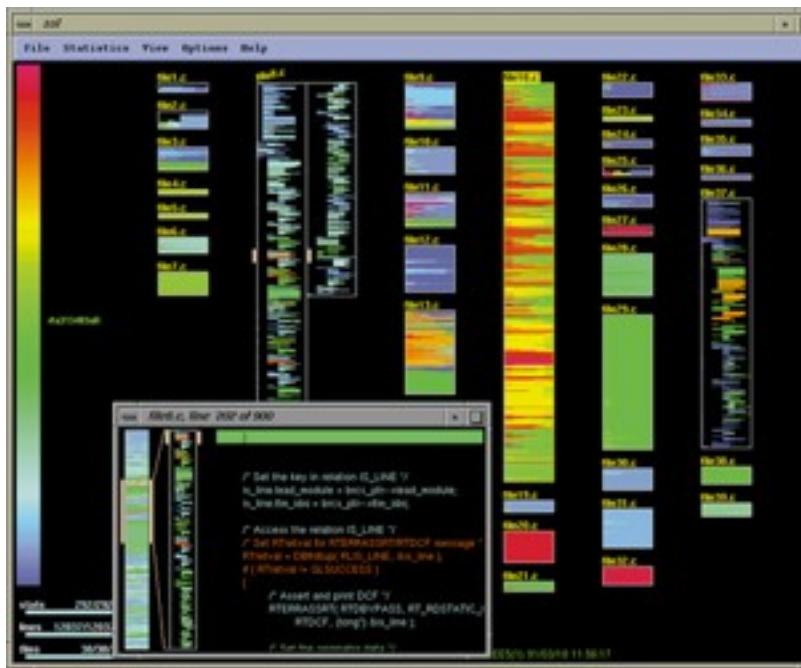
+ Add to Share ... More

12 0

Uploaded on Aug 19, 2009

This video accompanies a full publication which appeared in the Proceedings of the 2009 IEEE Symposium on Visual Analytics Science and

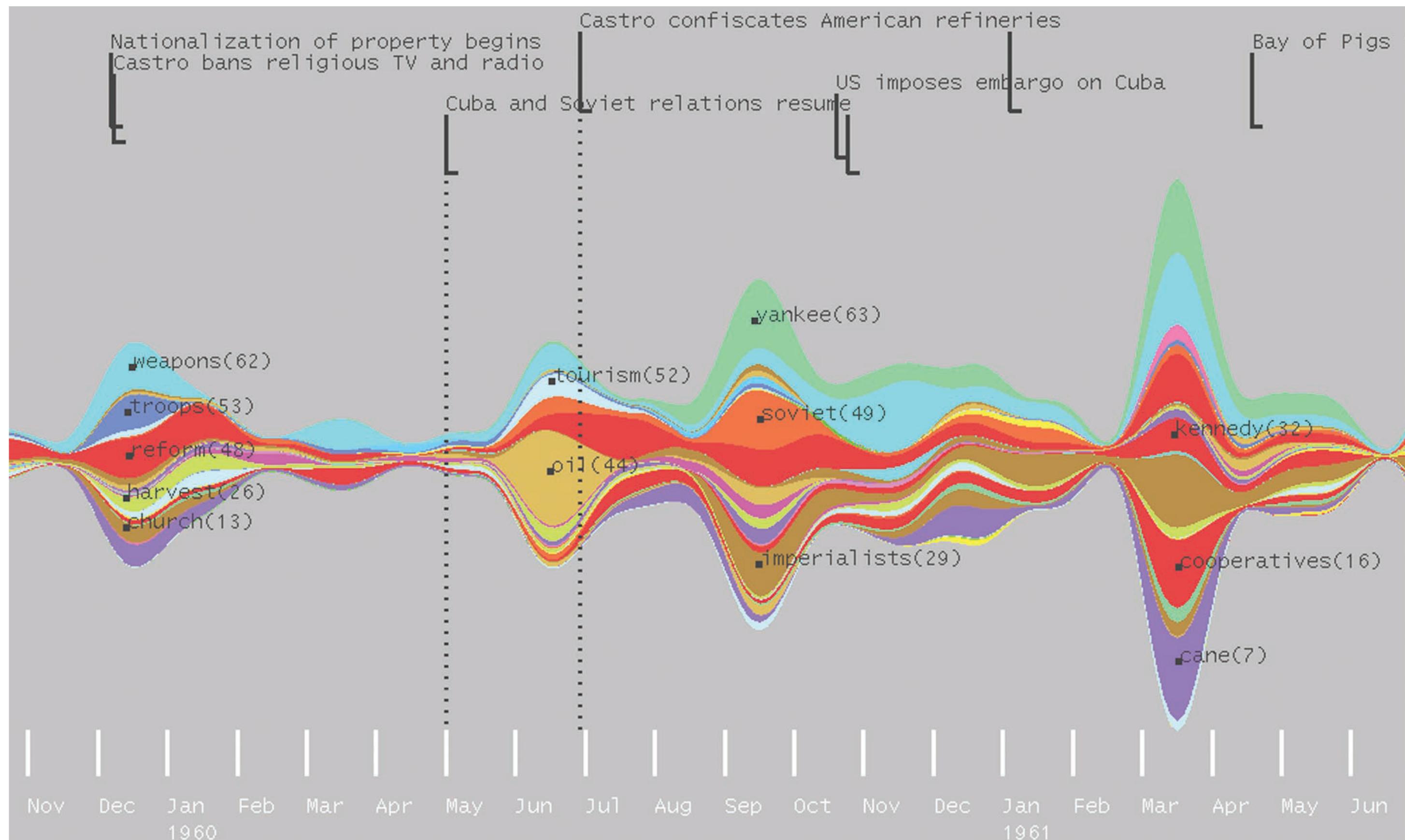




# DOCUMENT CARDS

(small multiples)





# JIGSAW: MANY LINKED VIEWS

**Visual Analytics Support for Intelligence Analysis  
Case Study: The 9/11 Report**

Carsten Görg

Youn-ah Kang

Zhicheng Liu

John Stasko



Information Interfaces Group  
Georgia Institute of Technology



## AFTER BREAK

Visualizing – sets & maps

Importance of aggregation in visualizing large data – including statistics, data mining, and topology



