

CIS 4930/6930-002

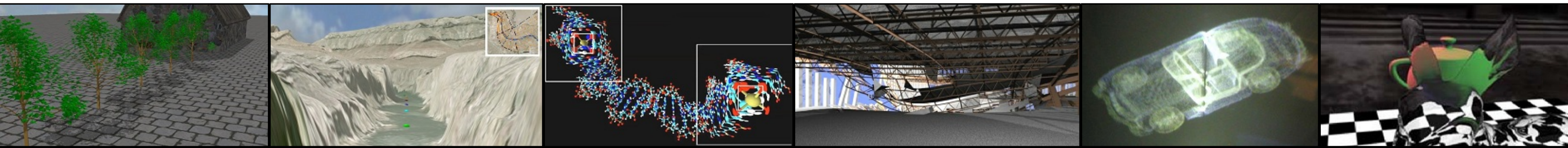
DATA VISUALIZATION

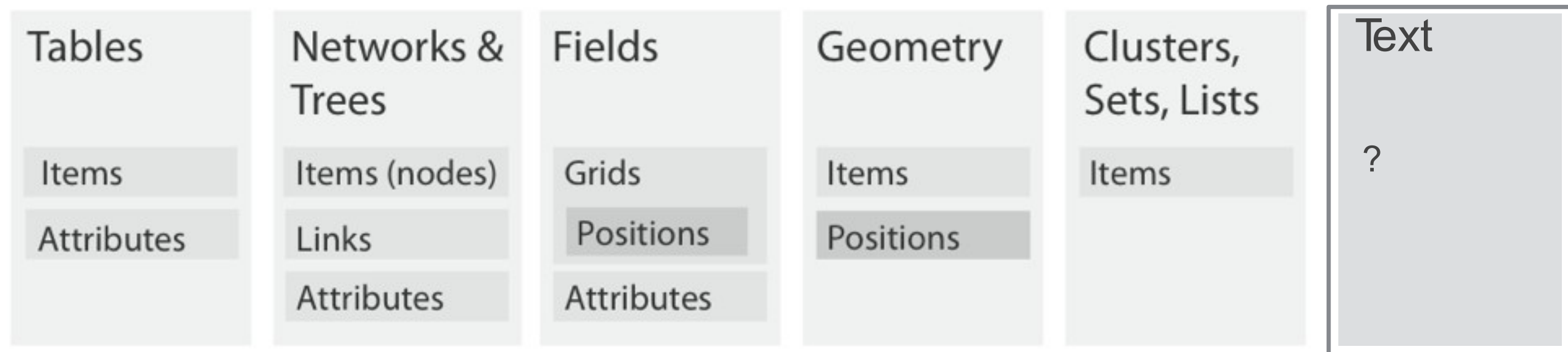


VISUALIZING TEXT

Paul Rosen
Assistant Professor
University of South Florida

slides credits Miriah Meyer (U of Utah)





WHAT DOES IT MEAN TO BE AN “ITEM”?



TEXT DATA TYPE

no numbers (implicitly)

characters (ASCII)

strings

USASCII code chart

<div> <div> b7b6b5 Bits </div> <div> b4b3b2b1 Column Row </div> </div>					000	001	010	011	100	101	110	111
					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	o	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	&	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	SI	US	/	?	O	_	o	DEL



TEXT DATA TYPE
no numbers (implicitly)
characters (ASCII)
strings

USASCII code chart

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

-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
-----	-----	-----	----	----	----	----	----	----	----	----	----



TEXT DATA SEMANTICS

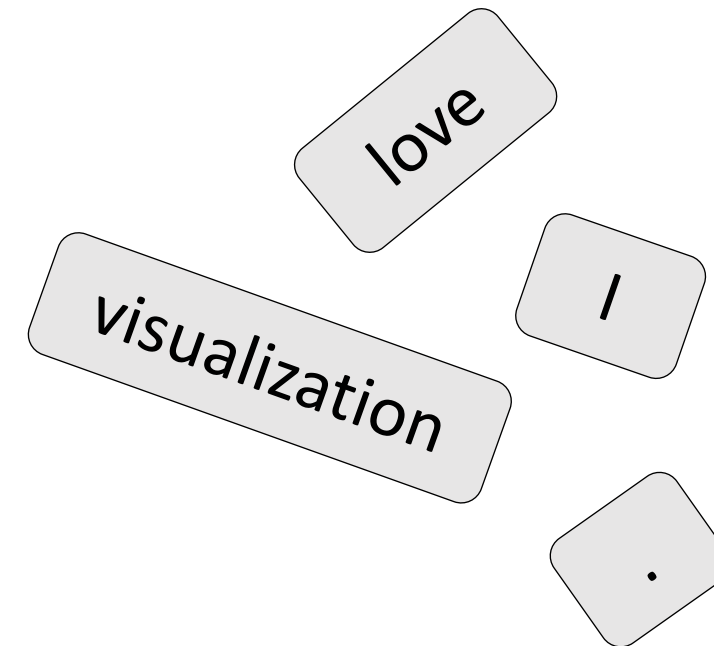
words

lines

sentences

paragraphs

chapters



I love visualization.



TEXT DATA SEMANTICS

documents: books, papers, webpages, e-mails, twitter post corpus: collection of documents

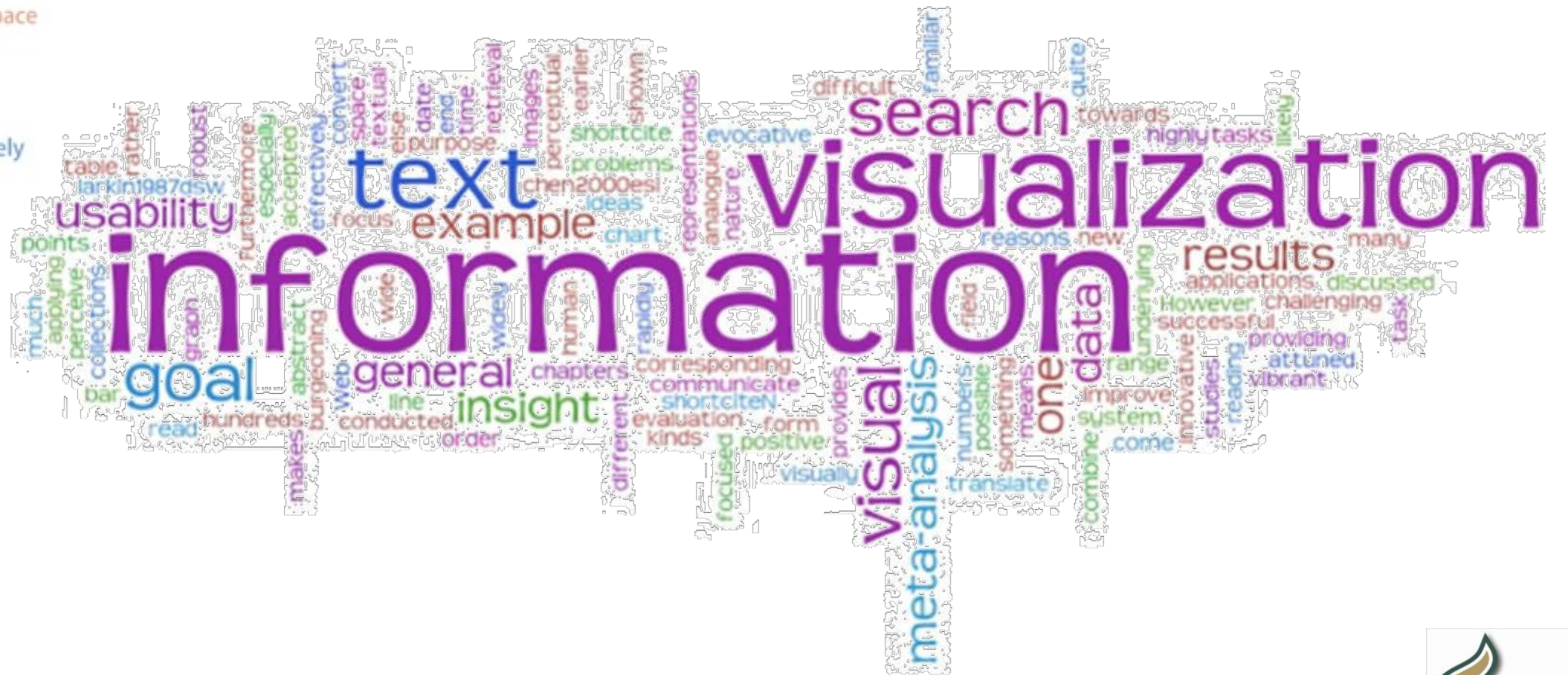


SINGLE DOCUMENT



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 means
 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** visually web wide widely



Visualizations : definitions of visualization word tree

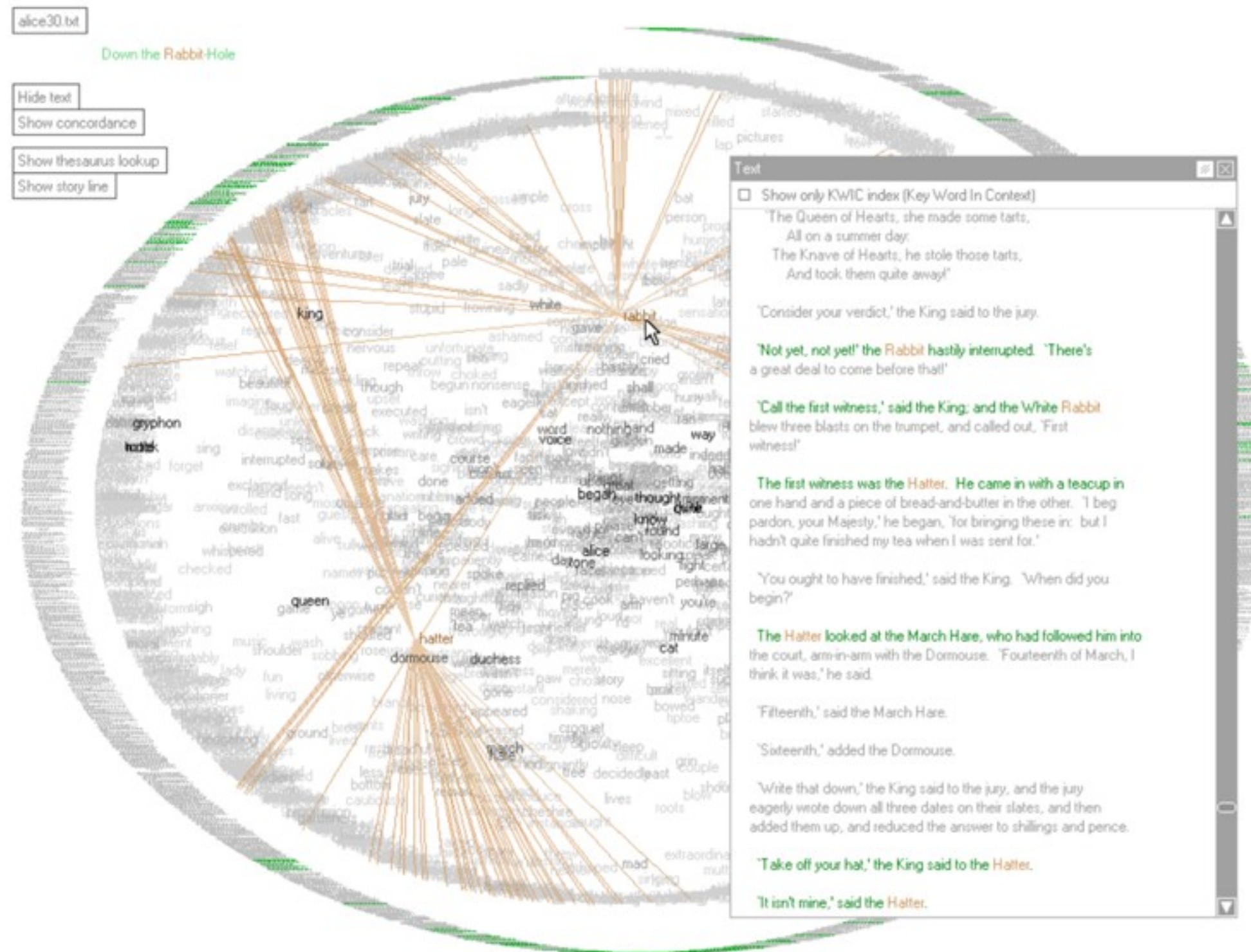
Uploaded by: mhalie

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

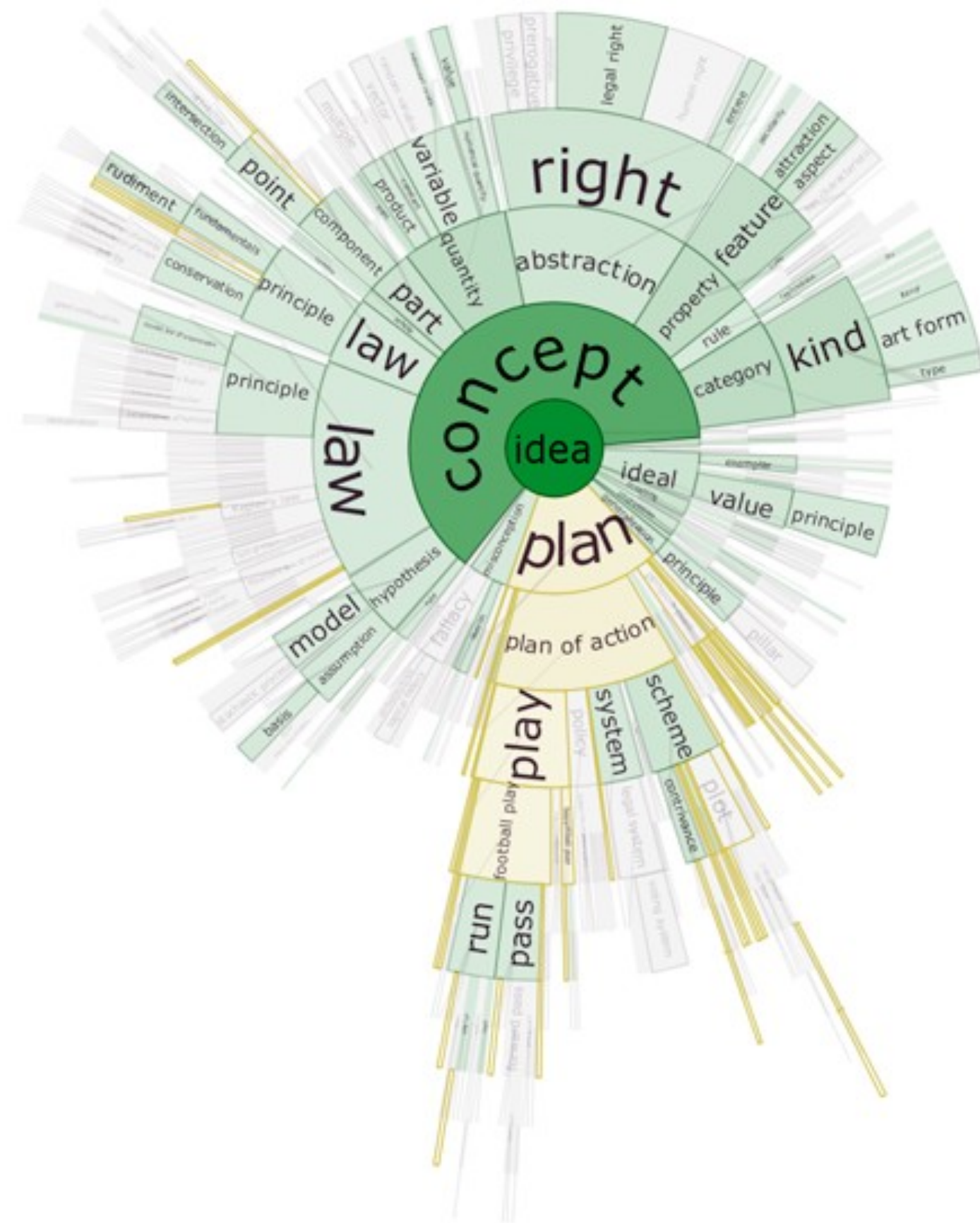
Tags: text



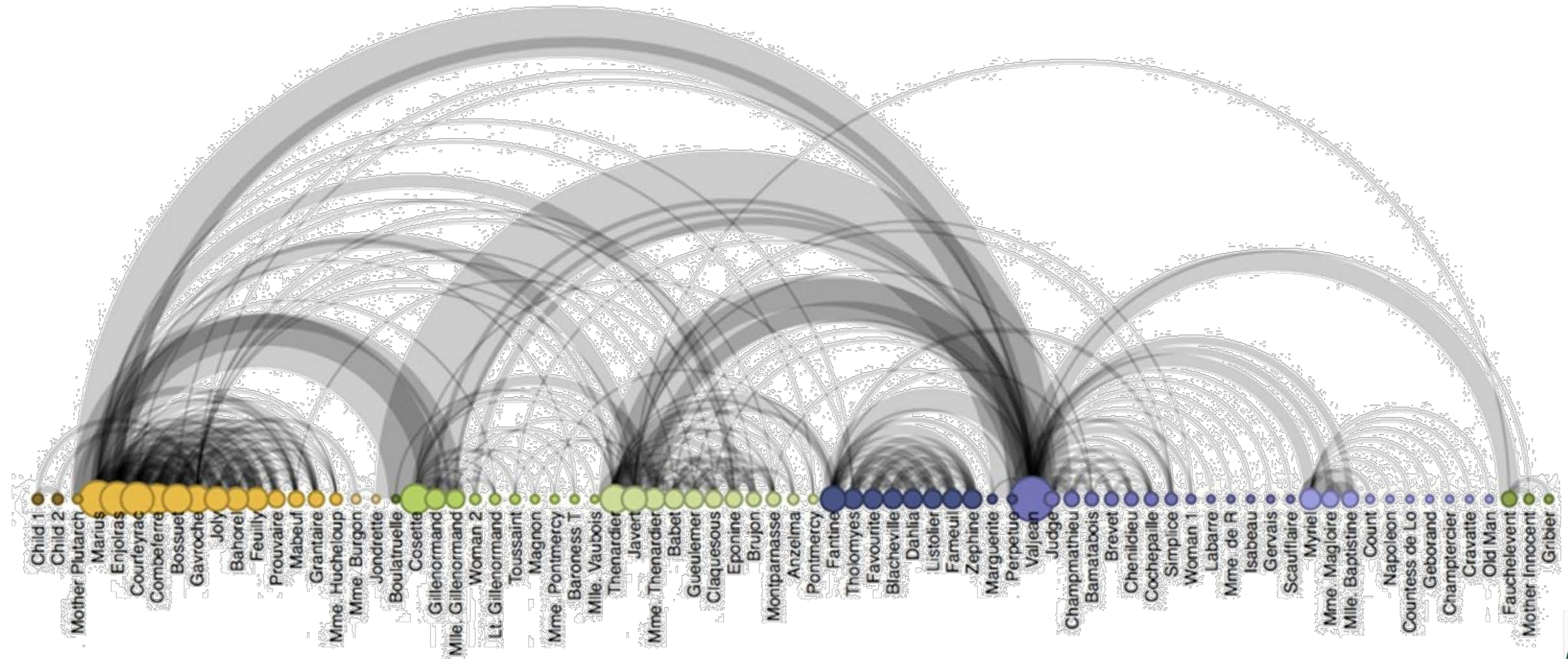
TEXT ARC



DOCUBURST



ARC DIAGRAM



COLLECTION OF DOCUMENTS



Parallel Tag Clouds to Explore and Analyze Faceted Text Corpora

Christopher Collins
Fernanda B. Viégas
Martin Wattenberg

0:01 / 7:40



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

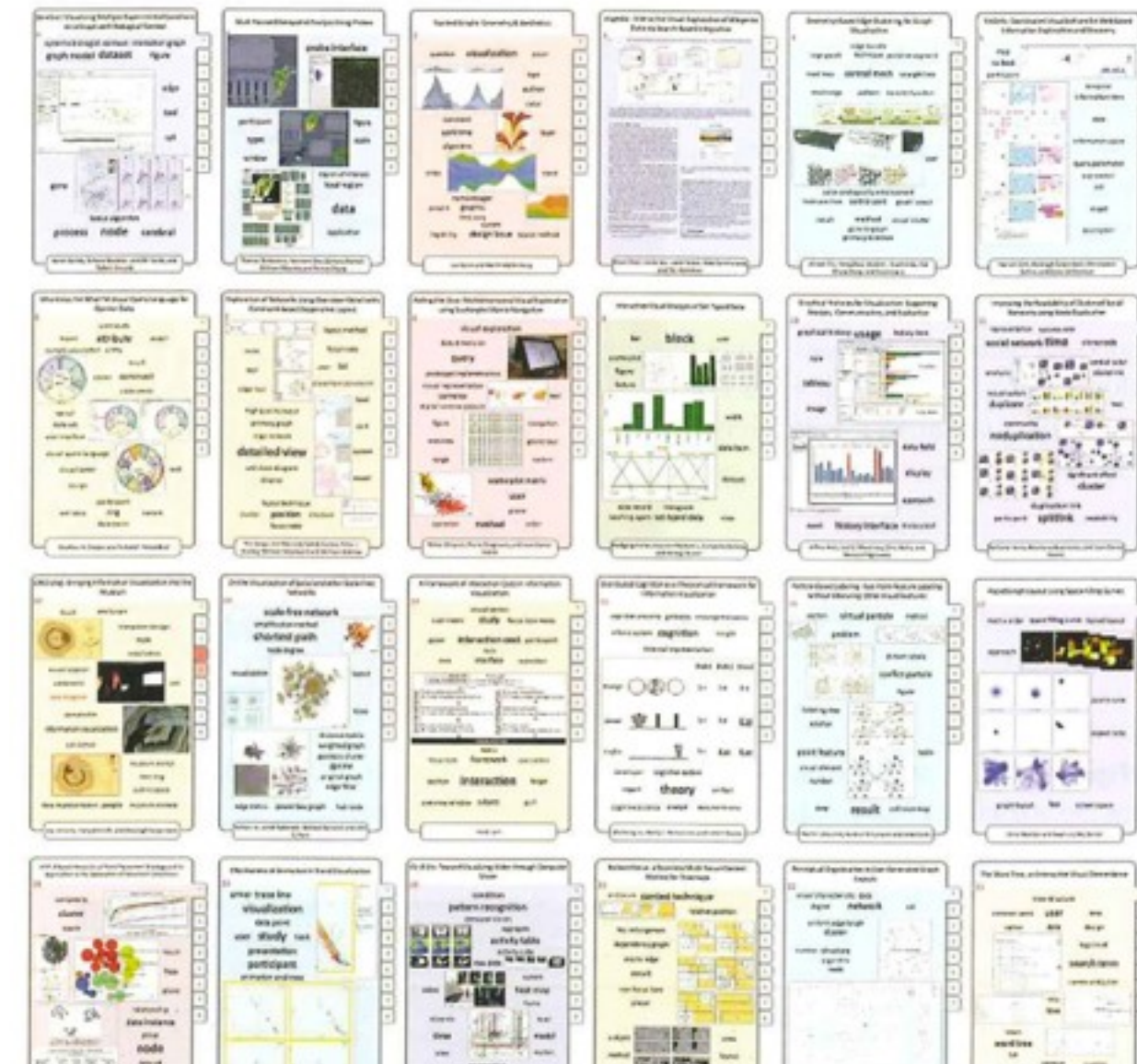
This video accompanies a full publication which appeared in the Proceedings of the 2009 IEEE Symposium on Visual Analytics Science and

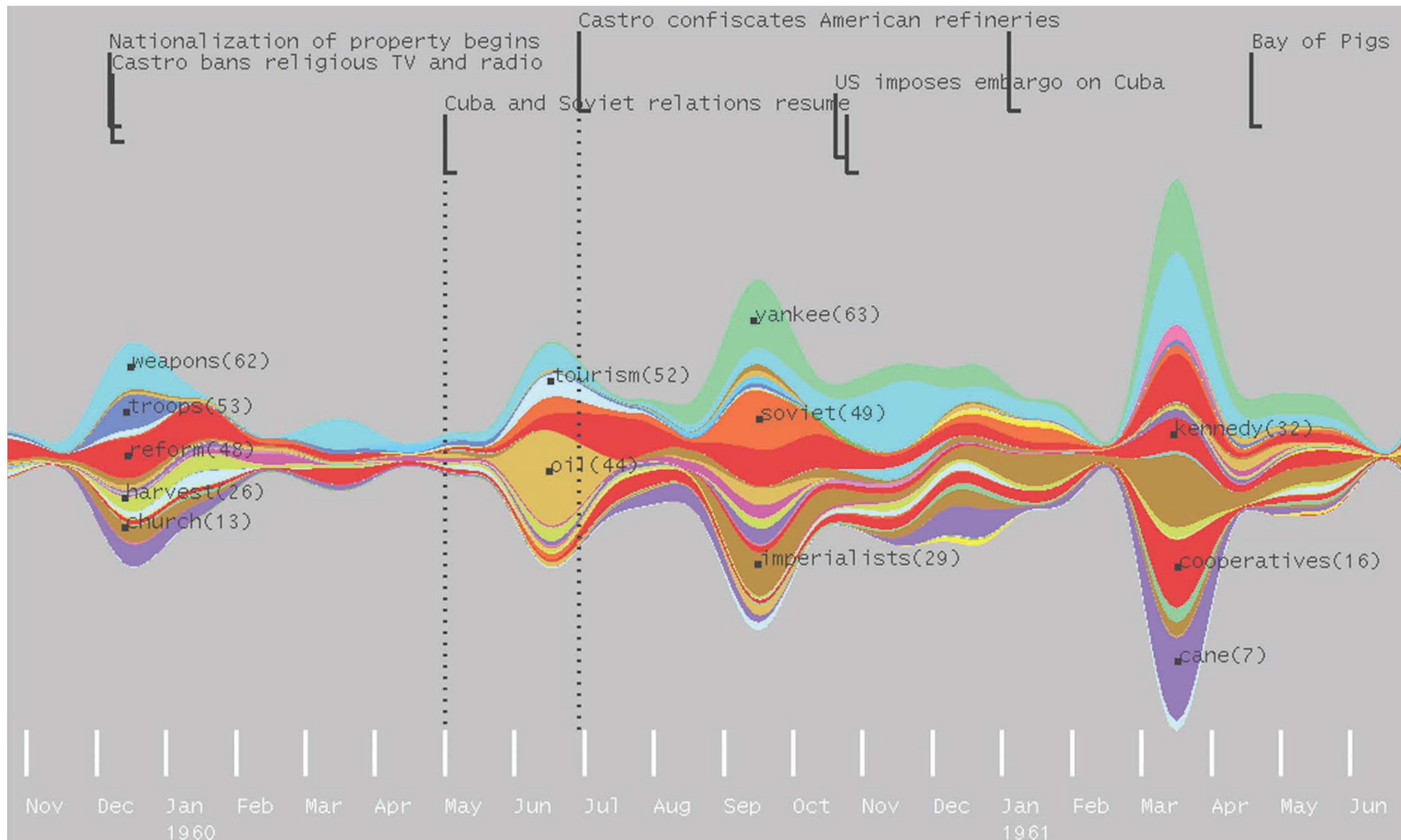




DOCUMENT CARDS

(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



