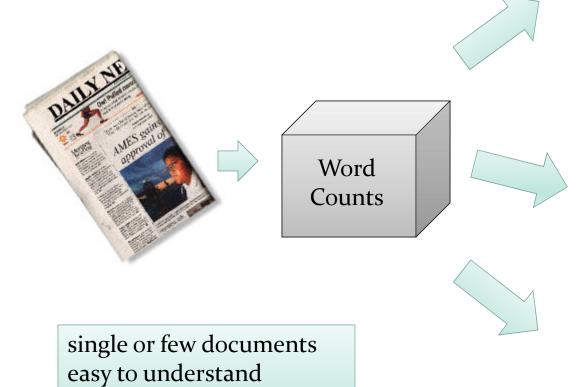
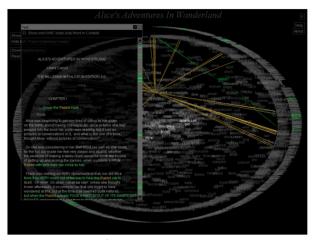
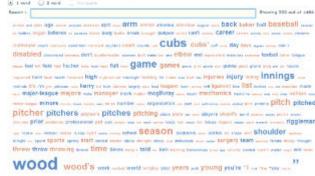
# Parallel Tag Clouds to Explore and Analyze Faceted Text Corpora





Text Arc, Paley, 2002

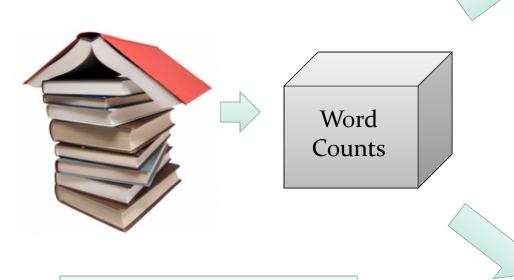


Many Eyes Tag Cloud, Viégas et al., 2007

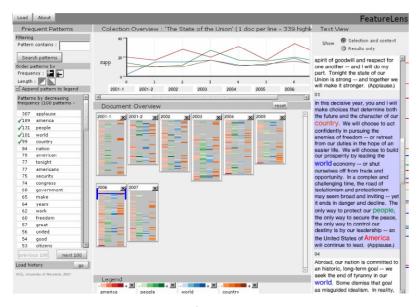


DocuBurst, Collins et al., 2006, 2009

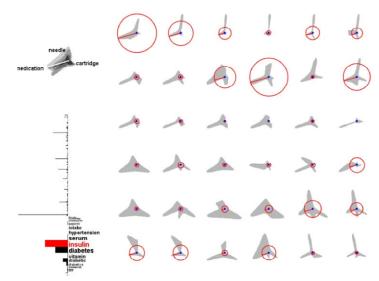
less analytic power



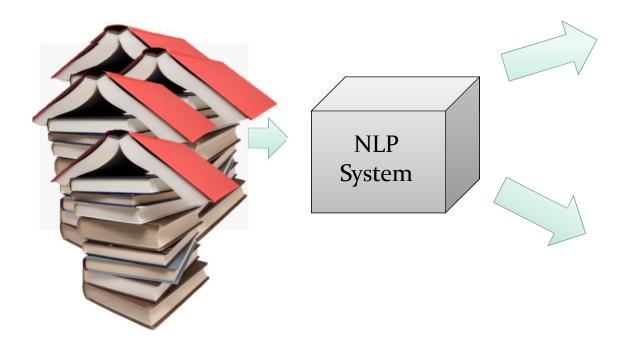
medium collections (100's) document comparison information retrieval entity relations



Feature Lens, Don et al, 2007



Gist Icons, De Camp et al., 2005

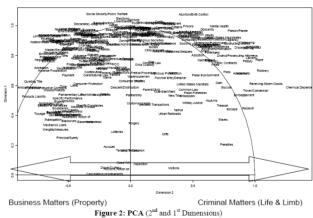


large collections (>> 1000)

data experts long term use document clustering



Attenex Patterns, 2008



Topics Maps, Hook 2007

# Keep it simple, but smart!

# Motivating Problem

### Forum Shopping:

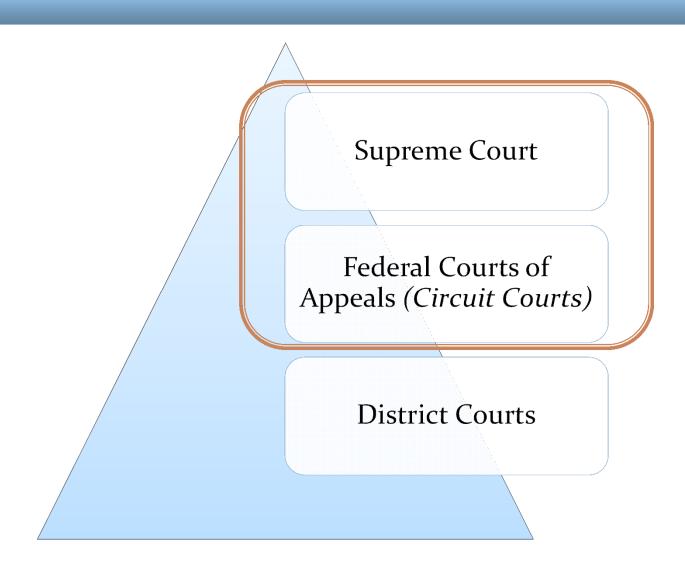
Bringing a case before a specific court based on higher likelihood of favorable judgement

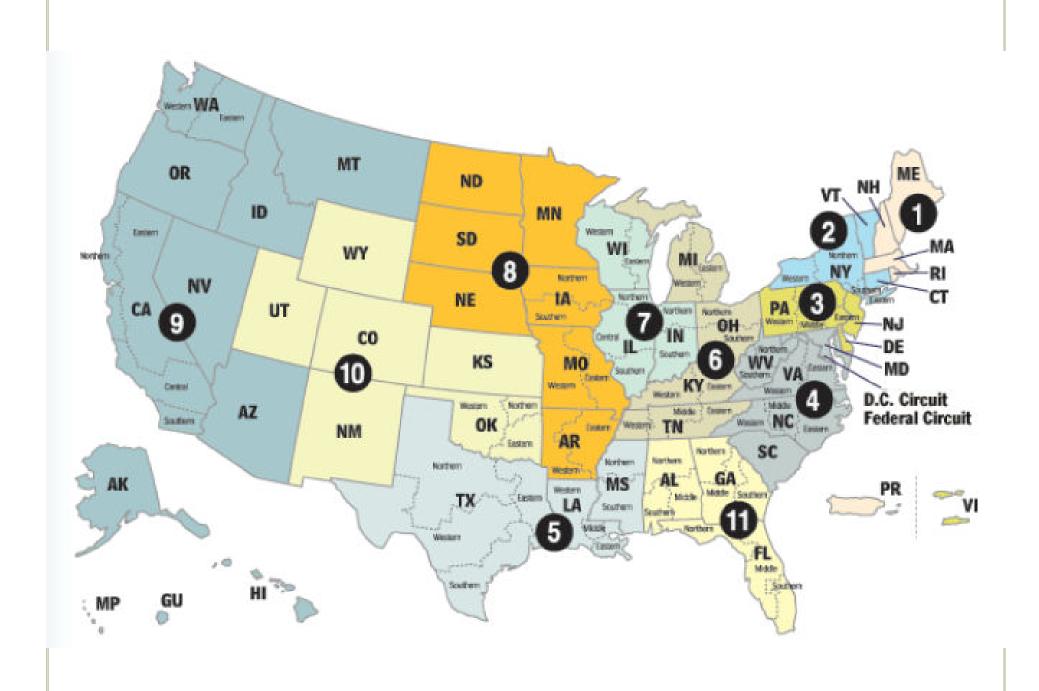
# Mini-Law Scl

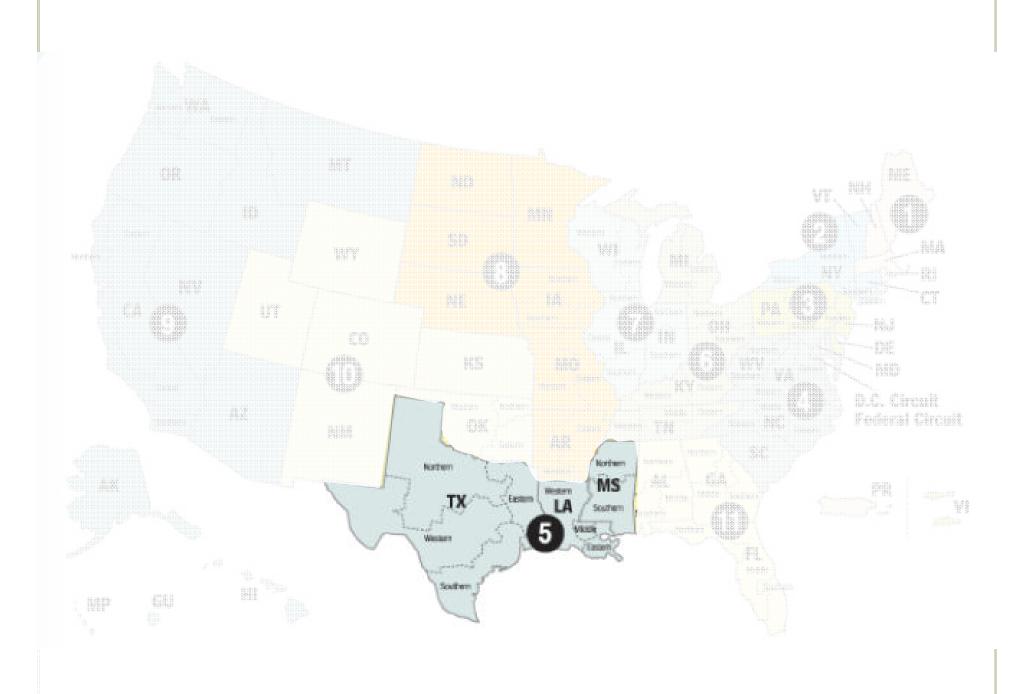
Data Backgrounder



# U.S. Court System

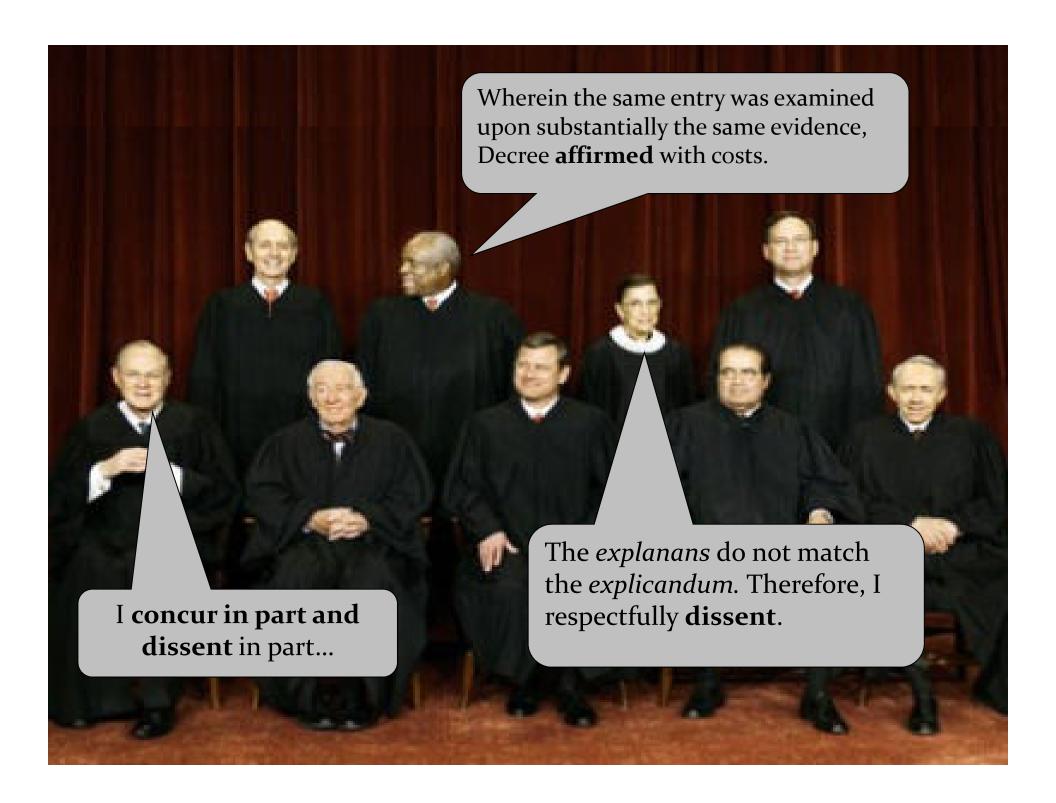


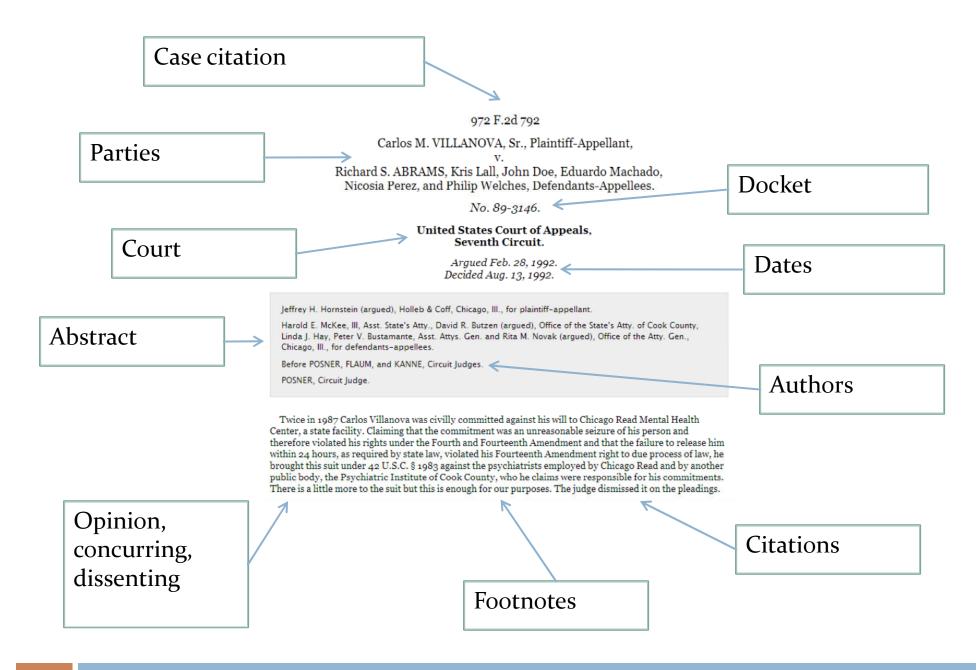




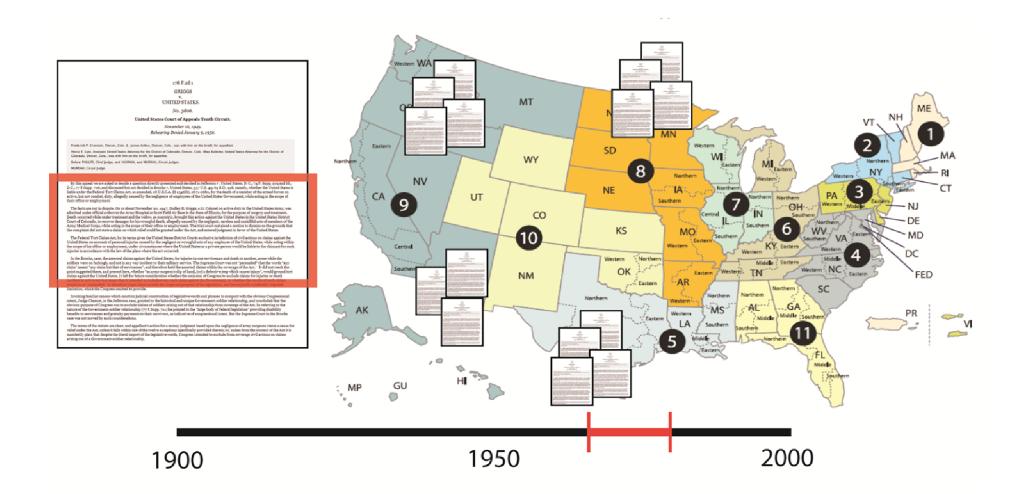
## U.S. Federal Reporters

- Written decisions of:
  - US Supreme Court (1750-2007)
  - Circuit Courts of Appeal (1950-2007)
- Issues of scale relevant part of the problem
  - □ 628,000 unique documents (12 GB text)!
  - 48,086 *Dante's Inferno*





### **Data Facets**



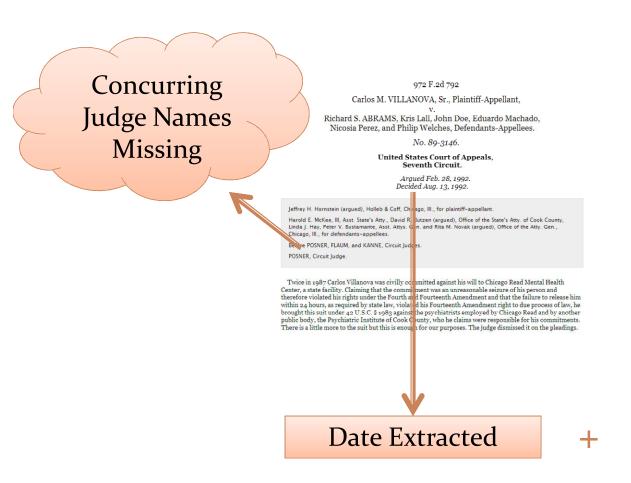
# Data Processing

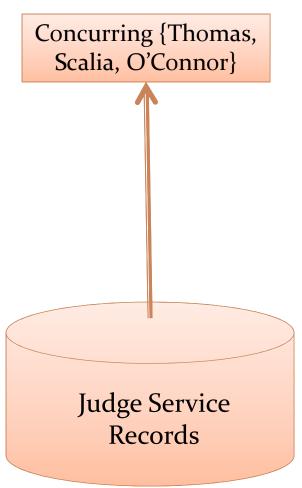
Federal Reporter Database

### Large Text Analysis Infrastructure

- Detect and separate parts of each text
- Apply simple NLP techniques
  - $\blacksquare$  reversible stemming (going  $\rightarrow$  go)
  - stop words (remove a, and, the)
  - dynamic stop words (remove court, judge, and hapax legomena)
- Pre-process expectation statistics
- Retain quick access to original text (search)

### Augmenting Authorship Information





# Finding Significant Words

E1 = c	(a+b)
	$\overline{(c+d)}$

$$E2 = d\frac{(a+b)}{(c+d)}$$

		Reference Text	Total
Count of word	a	Ь	a + b
Count of other words	c - a	d - b	c + d - a - b
Total	С	d	c + d

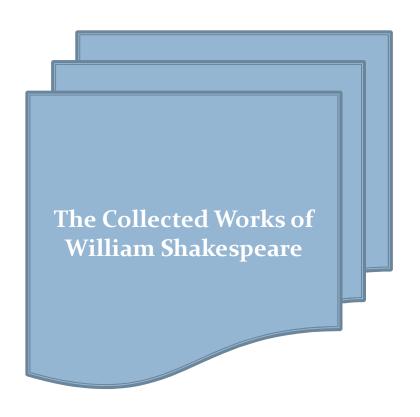
$$G^2 = 2((a \cdot \ln(a/E1)) + (b \cdot \ln(b/E2)))$$

Translates approximately to significance, e.g. 6.63 = p < 0.01

Dunning, Ted. 1993. *Computational Linguistics*, Volume 19, number 1, pp. 61-74. Rayson, P. and Garside, R. ACL 2000.

### Two Forms of Measurement

- Cases
  - Number of cases containing a word at least once
  - Good for finding words characteristic to many cases
- Occurrences
  - Number of occurrences overall
  - Good for finding outliers (long cases), repeated jargon



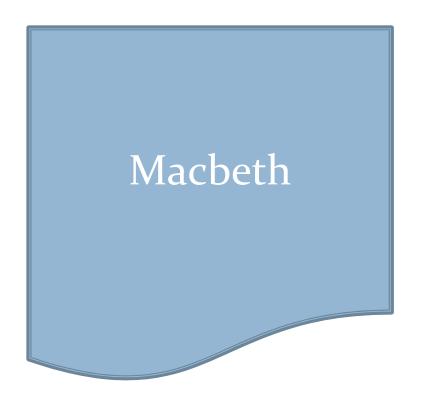
### **Word Counts**

begotten

hast

love

thou



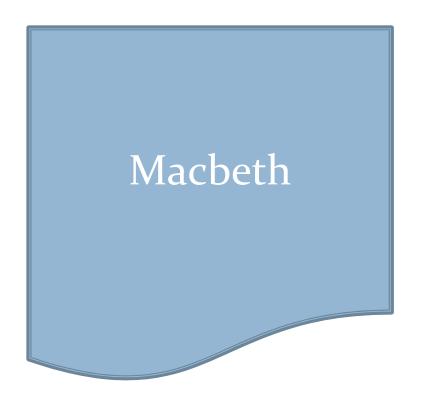
### **Word Counts**

begotten

hast

love

thou



### Log Likelihood

wrath

blood

lady

witch

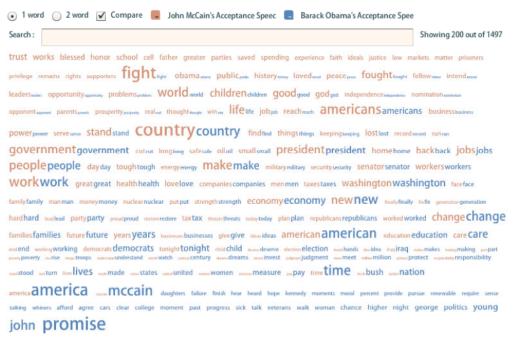
# Parallel Tag Clouds

Visualization Design

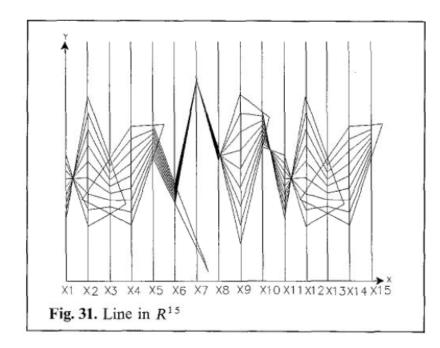
Rank	Unigram	(-2λ)	Bigram	(-2λ)	Trigram	(-2λ)
1	Slovenia	319.48	federal army	21.27	Slovenia central bank	5.80
2	Yugoslavia	159.55	Slovenia Croatia	19.33	minister foreign affairs	5.80
3	Slovene	87.27	Milan Kucan	17.40	unallocated federal debt	5.80
4	Croatia	79.48	European Community	13.53	Drnovsek prime minister	3.86
5	Slovenian	67.82	foreign exchange	13.53	European Community countries	3.86

Figure 2. Top 5 unigram, bigram, and trigram concepts for topic "Slovenia Secession from Yugoslavia".

Lin and Hovy, 2002



Comparison Word Cloud, Wattenberg et al., 2008

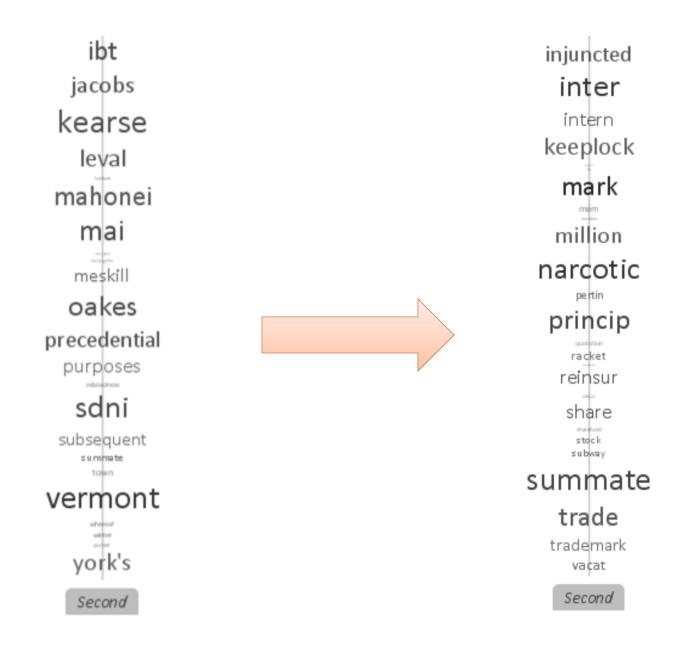


Parallel Coordinate Plots, Inselberg, 1985

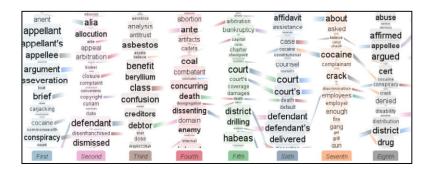
## Visual Encoding

- □ One column per *Circuit*
- □ Top *N* words only (scrollable)
- □ Text Size → significance score
- □ Order → alphabetic

```
ibt
  jacobs
 kearse
   leval
mahonei
   mai
  oakes
precedential
 purposes
   sdni
 subsequent
   s u mima te
vermont
   Second
```



Words that occur in multiple columns?





ippellant's appellee	*ante appeal *arbitration	asbestos  assets ebelieve	artifacts cadets	*capital cargo charter checkpoint	*Case *cocaine constitutional	*believe called cfleck
argument sseveration	broker *closure	benefit beryllium	coal combatant *concluded	*court	counsel counsel's *Court	complainar *crack
*brief	conveniens *copyright curiam	·class confusion	concurring  *death  *desegregation	coverage damages edeath edear	*court's  *death	•discriminatio •employee employer
carjacking castroore *cocaine	*defendant	creditors •debtor	<ul><li>dissenting</li><li>domain</li></ul>	<ul><li>district</li><li>drilling</li></ul>	•defendant	enough fire
conspiracy	•disenfranchised dismissed	diet dose exercise	enemy •internal	*habeas	*defendant's delivered *dissenting	gang get grill gun

anent appellant appellant's appellee argument sseveration but but cocaine cocaine coconspiracy count	alia allocution are appeal arbitration brider closure complaint complaint converses copyright curiam date defendant disentranchised dismissed	analysis anutrust asbestos benefit beryllium class confusion creditors debtor	abortion ante artifacts cadets coal combetant concurring death desayspain dissenting domain enemy internal	arbitration bankruptcy capital capital court courts coverage damages district drilling	affidavit assistance case coccine contributional countsel count's death default defendant's defivered dissenting	about asked before cocaline complainant crack discemination employees employees employees employees gang get guil gun	affirmed appellee argued cert cosaine conspiracy crack denied disability distribution district drug
First	Second	Third	Fourth	Fifth	Sixth	Seventh	Eighth

appellant's	ante appeal arbitration	asbestos assets believe	artifacts	capital charter checkpoint	case cocaine constitutional	believe called check
argument sseveration brief	broker closure complaint conveniens copyright curiam	benefit beryllium class confusion	coal combatant concluded concurring death desegrecation	court's coverage damages death	counsel court court's death	complainant  crack  discrimination employees employer
carjacking cashmero  COCAINE commonwealth  CONSpiracy count	date  defendant disenfranchised dismissed	creditors debtor diet dose exercise	dissenting domain enemy internal	district drilling habeas	defendant defendant's delivered dissenting	enough fire gang <sup>get</sup> grill gun

anent appellant appellant's appellee	alia allocution appeal arbitration	aerobics analysis antitrust asbestos assets beleve	abortion ante artifacts cadets coal	arbitration bankruptcy capital cargo charter checkpoint	affidavit assistance case cocaine constitutional	about asked befeve called check	abuse  affirmed  appellee  argued
argument sseveration brief carjacking	broker closure complaint conveniens copyright curiam date	benefit beryllium class confusion	combatant concluded  concurring death desegregation dissenting	court's coverage damages death district	counsel count's court's death default	discrimination employees employer enough	cert cecaine conspiracy crack denied disability
cocaine commonwealth conspiracy count	defendant disenfranchised dismissed	debtor dest dose exercise	enemy	drilling	defendant's delivered	fire gang get grill gun	distribution district drug

anto	asbestos	antifacts	capital	case	oder check	appellee	asylum	atrocious	larels .
appeal	assata	cadets	o harter	cocaine continue	cocaine		banc	brief	case
arbitration broker	benefit	coal	court	counsel	complainant	argued	bankruptcy bordering	cattle competent concluded	circuit
closure	beryllium	combatant	court's	court	crack	Cert	categorical	court	class
complaint	class	concur	coverage		decorrentor	conspiracy	country	courts	commerce
copyright curiam	confusion	death	damages	court's	employed employees	denied	crime dba	degree	conspiracy
date	creditors	dissent	district	default	enough	disability	deportation	district	county
efendant senfranchised	debtor	domain enemy	drilling	defendant's	fire gang	distribution	dissert	doc	court
ismissed	dose exercise	internal joined	habeas	delivered	grill	drug	extradition	eagles	disenfranchised district
foreign fraud	fiduciary	loss	indemnity	district	had harassing	evidence	fisheries forest	evidence	enhanced festivals
ground	have	maintained	insurance	exigencies	have	grams	habitat	gas	graduated
inter	here inasmuch	majority	tonder.	fact fam	her	guidelines	harvesting harvesting	grazing	indictness
iones judiament	insurance	majority's marks	judgment jurists	prievance habeas	his	her	hold	heinous her	interstate jail
marks	liability market	mineral	law	ineffective	job	his mair karv	land	instructions	lading lan mitigation

anent appellant appellant's appellee argument sseveration	alia allocution apreal arbitration broker closure complaint	aerobiss analysis antitrust asbestos assets believe benefit beryllium class	abortion ante artifacts cadets coal combatant	arbitration bankruptcy capital cape checkpoint contest court's coverage	case cocaine constitutional counsel coursel coursel	about asked believe condine complainant crack	abuse  affirmed appellee argued cert cocaine conspiracy
commonwealth //	copyright curlam date  defendant disenfranchised dismissed	confusion creditors debtor diet dose exercise	death desegregation dissenting domain enemy internal	damages death district drilling habeas	default defendant defendant's delivered	employees employer enough fire gang get grill gun	denied disability distribution district drug

# Edge Stubs

mem members millions narcotics payment pertinent plaintiff principal quotation racketeering reinsurance respect security See shareholders shares stock subway summation trade trade trademark	liability majority market non notes  our parents plaintiff plan plenary predicate prediction provision radiation recognized required section settlement syrup	from his interlocutory legal lung magistrate material merits mining opinion oral order pagers pneumoconiosis present pro process pulmonary recommendation relief respiratory recond	marihuana maritime negligence notes nre oil parish parish policy recovery ref'd removed rigging school stated suit usury vessel vote	magistrate's marijuana medical motion  office pain  panel paupers  plaintiff  plaintiff's  pneumoconiosis  pursuant  recommendation  recommendation  report search sentence sitting summary unanimous	lawyer might more more notified one ostrich ought aures out para police statement suit tentative than thought told too want
vacated	which would	would	white	union work	what would

# Edges

II VALL liability lawyer from magistrate's mem marihuana members marijuana his might millions majority maritime interlocutory medical more narcotics market motion negligence much revisions legal non payment notes notified notes office pertinent nre one magistrate our panel plaintiff ostrich offshore parents material ought ources oil paupers principal plaintiff parish out merits plaintiff quotation рага platform miner's plan racketeering point plaintiff's mining police policy reinsurance plenary opinion recovery pneumoconiosis respect 500 ref'd oral predicate security she pulmonary prediction order removed principle see pursuant statement pauperts and product rigging shareholders pneumoconiosis suit schools recommendation provision shares seaman present referred state d radiation report tentative suit stock pro search recognized subway than required ..... process sentence usury thought summation section sitting pulmonary = told vessel recommendation summary settlement record p 500 trade relief vote penetened unanimous want respiratory trademark review p syrup voters sentence p union what white vacated which would spon writ view

work

■would

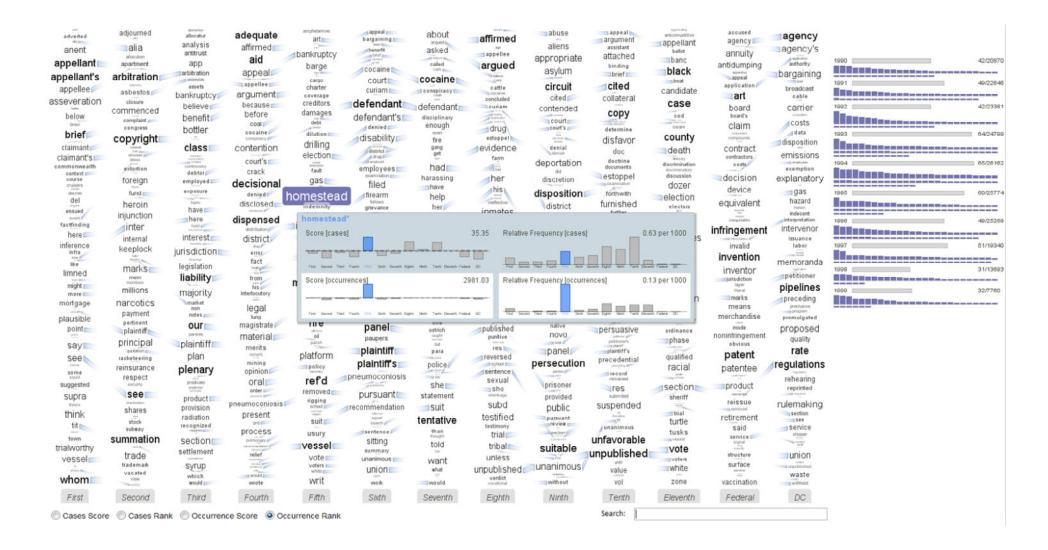
wrote

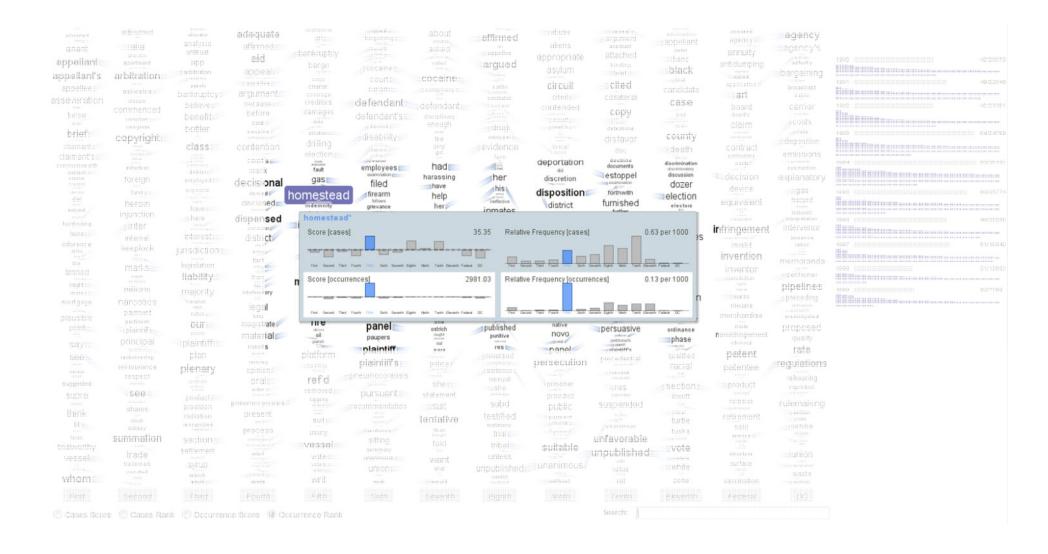
would

a where

adverted	adjourned	allocatur	adequate	amphetamine art	appealbargaining	about	affirmed	abuse	appeal	anticompetitive	accused	agency
anent	alia	analysis	affirmed	CIL	benefit	argued	- Kips	aliens	assistant	appellant	agency	agency's
	allocution	antitrust	aid	bankruptcy	brief	asked	appellee	appropriate	attached	ballot	annuity	at application
appellant	apartment	app		barge	cocaine	called	argued		binding	black	antidumping	authority
appellant's	arbitration	arbitration	appeal	cargo	court	cocaine	balance c-are	asylum	brief	boat	appeal	bargaining
appellee	articità	assets	appellee	charter	curiam	conspiracy	cattle	circuit	cited	candidate	application 🚩	broadcast
asseveration	asbestos	bankruptcy	argument	coverage	444.5	could closed	concluded	cited	collateral		art	cable
Italian	commenced	believe	because	creditors	defendant	defendant	curiam	contended		case	board	carrier
below	complaint =	benefit	before	damages	defendant's	disciplinary	streeting streetings at classical k	court	copy	cod	board's	⊿ costs
brovi	congress	bottler	coal	dilution	denied	enough	drug	court's	determine	coram	claim	data
brief	copyright	Care City Control Cont	conspiracy ==	drilling	disability	fire	estoppel	chon checks born \$6 declared	disfavor	county	compounds	disposition
claimant	Contraced Code gas deforeders	class	contention	13.000000000000000000000000000000000000	district	gang get	evidence	denial	doc	death	contract	-
claimant's	droeses g	content content	court's	election	drug s emptyyet	gar.	farm	deportation	doctrine	discrimination	contractors	emissions
commonwealth context	extortion	debtor	crack	fault	employees	had	Fred	dd	documents	discriminatory		exemption
course	foreign	employed	decisional	gas	filed	harassing	her	discretion	estoppel	dozer	decision	explanatory
decree	fund	exposure	denied	homestead	⊿firearm	have	his	disposition	forthwith		device	gas
del	heroin	nam	disclosed	indemnity	follows	help	ineffective	district	furnished	election	equivalent	hazard
ensued	injunction	have	dermant	instant	grievance	0.000	inmates	err	further	firefighters	estings	indecent
factfinding	⊿inter	here	dispensed	insurance	hereby	him	netsclore	except	greating	hold	inequitable	interpretation
here		interest	distribution	interest	his	his	judicata	failed	Amend .	immunities	infringement	intervenor
inference	internal keeplock	(description) Management	district	pick Naw B	job	judge	jury	fear	judgment	insanity	invalid	labor
infra	later .	jurisdiction	error	liability	judgment	jaage		finding	judicata	ivory	invention	Tomas Company
like	marks	legislation	fact	loon	magistrate	kilograms '''	ethamphetamii	hardship	material	jul Jury		memoranda
limned	mem	liability	from	loan	magistrate's	lawyer	murder	his	nevertheless	law	Inventor	petitioner
might	members millions	otto:	his	marihuana	marijuana	might	opinion	immigration	opinion	migrant	layer	pipelines
more ===		majority	interlocutory	maritime	medical	more	pain	jurisdiction		mitigation	marks	preceding
mortgage	narcotics	market	legal	negligence	motion	men	persuasive	may	oral	avenue On extens	means	preclusive program
plausible	payment	notes	lung	notes	caffices a	notified	postconviction	methamphetamine	order	nobis nonstatutory	merchandise	promulgated
point	pertinent	our	magistrate	nre	panel	one ostrich	published	native	persuasive	ordinance	mode	proposed
bayon.		parents	material	oil partsh	paupers	ought ources	punitive	novo	petioner's	phase	noninfringement	quality
say	principal	plaintiff	merits	10	plaintiff	out para	res	panel	plaintiff's	pice pice	obvious	rate
see	racketeering	plan	miners	platform		The same of	reversed	persecution	precedential	qualified	patent	
some	reinsurance	plenary	opinion	policy	plaintiff's	police	sentence	persecution	record	racial	patentee	regulations
squid	respect	predicate	oral	ref'd	pneumoconiosis	sho -	sexual	pricence	remained	across	parameter and the second	rehearing
suggested	see	prediction principle	order	removed	pursuant	she statement	she	prisoner	res	section	product	reprinted
supra	sheelstles	product	pneumoconiosis	rigging			subd	provided	suspended	sheriff	reissue	rulemaking
think	shares	provision	present	school states	recommendation	suit	testified	public	suspended	trial	removed	section
tit	stock	radiation recognized	pro	suit	report	tentati∨e	testimony	pursuant	51	turtle	retirement	see
Selection	subway	rangament ger	process	usury	sentence	th an thought	trial	section	unanimous	tusks	Said service	service
trialworthy	summation	section	pulmonary	vessel	sitting	told	tribal	suitable	unfavorable	vossel	signal signal state	ayadera Brill
	trade	settlement	relief	vote	summary	ton	unless	Sultable	unpublished	vote	structure	union
vessel	trademark	syrup	respiratory, respiratory, sectorico	voters	unanimous	want		unanimous	until	white	surface	unpublished
whom	vacated view	which	would	white	union	what	unpublished =	violation	value	you	SUPPLIES.	waste
whom	WANT STATE	would ===	wrote	writ	work	would	vocational	without	vol	zone	vaccination	withhold
First	Second	Third	Fourth	Fifth	Sixth	Seventh	Eighth	Ninth	Tenth	Eleventh	Federal	DC

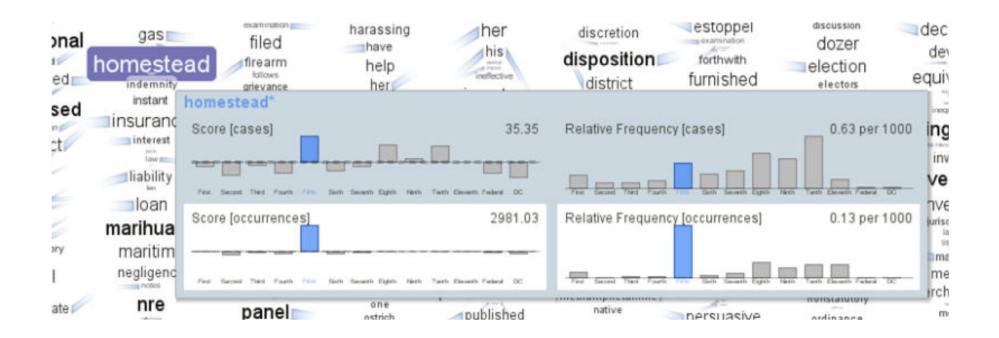
# Coordinated Views



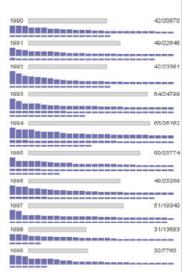


## Tooltip Graphs

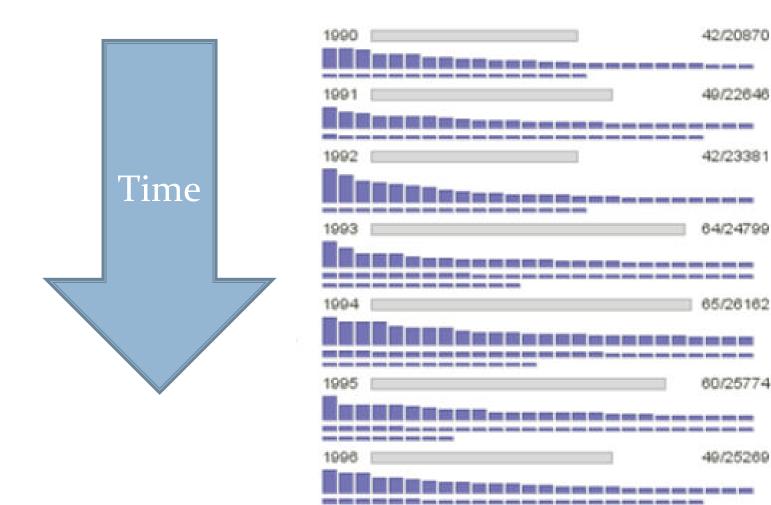
- Compare several measures
- □ Significant *absence*
- See scores for all columns at once



	adjourned		adequate	angladanina BIŠLB	nag apprairies Transportation	shout	adVirread	ma albuma	11 App 0 of 16.		acoused agency:	agency
smont	alia	emally sis erenaties	affirmed		printer fill	965494		aliens	assistant	m appeare	annuity	magerkty's
appellant	nimuion ne artment	2000	214				en nappallea	арргоргіайа	attached	mbanc		
	arbitration	a serial anion	EEE EEE				argued	asylum		black	anlidumping	mbargaining
appellent's	GHARLEMAN I		Manapelleo III	6.26(24)	COURT	cocaine		circuit	cited	mpheut	appdiontions <sup>pp</sup>	
appellee#		bankruptcys	angumentiiii	0.50/9/0886 0.50/9/08.86	custarn ma	gilos sepiratoritos	1121110	citedal		candidate	Mart	broadcast cable
asseveration			becausem	cre ditors	defendant	udefendant.			collateral	case	board	
		benefits	before		defendant's	distiplinary			capy	0.43-8	becentu	
	ocemphant possi	boliler	12006 (61		_gdenieds!	eacugh	raruğ.	in courting			clains	
brief	copyright				cisability		estropolito		disfavor	county		
		class	contention				mendence			death	contract	Hitisposition
claimant's iii			coust's all	election			fam	Accessed to the	destrine		constructions.	ernissions
commonwealth			crack	1995		hadiiii		deportation	spinorausmispii			
		emiployed III			film of	hamasing				\$1100 PE \$110PE		esplanishary
	fundimi	aspanire	decisional		filed	minove		disposition	teritoriti	dozer	device	
	heroin		declored -						furnished		equivalent	
ensued					grievance	heriP		Tidistrict	nan nan rawa	eleston mannamillana		
			dispensed									
herefill			distribution of T		8]		-3535 - Raidhi				infringement	intervenor
			districti								invalid	interes :
Print .	keeplock	jurisdiction	# HKHL								invention	
liko	marks	legislation										memoranda
limned	DISCH P.D.	Hability										
majahi m	millions			n Beore lacer	18181							pipelines
mane IIII		mayority	меньили									aproceding
	narcolics		103301								2016/2015	
gilausiidie	payment putinent										menchandise	
grointage	ili plaintifis.	DUF		100 M 1 M	2.07141				apersuasive		mede	proposed
1811	principal	plaintiff			şhan şperisi		partire			mighase	noninmingement atrices	quality
sayı	Princeduce		menite		plaintiff	0.808			nigitalistiff's			1000
566h	ta alicia seja B			platform	olaintiff's			persecution	precedential	qualified	patent	aregulations
3417914 1-3417	10/15/20/10	plemany		Hipshey			- spentence k	(artisti gri		racial	patentee	
	respect			red"d	g presentoconiosis						maroduct	
supra	500	initia.	180140 (81-	removedim	pursuanti				1105	giSOCTION ::		reprinted
THE STATE OF		production		all during	anecommendation	115.111	subd		suspended		P0:03:03U8	rulemeking
18111151	shares		protections				bestifi ed				retirement	This entire
	stack subwey					tentative	181869009	nevine	gunanimous		9310	
town	summation	n n n di n n u u	process	namk	and sentence in		tralii		unfevorable	tusks	servicerii	
brahvorthy	Dunnanun	section		vessei	siting	811111	tribalii	aultable		17010		
	treide	is entillerenterent	relef	votes	CHIBINITEDIA CHIBINITEDIA		uniess		unpublished	and the second	strature	HAMICH'I
	tradientade	syrrup			union	minut	unpublished	Luntariimous		anwinite)	surface	
whom	With the second	ushish muhis		veril	ma ši		191603					waste
									Teach			
									Seach: I			
THE CHISES SCORE	e 🕕 Cases Ran	a de cacuman	ce acore le Cc	carience Hank					Annual Indiana			

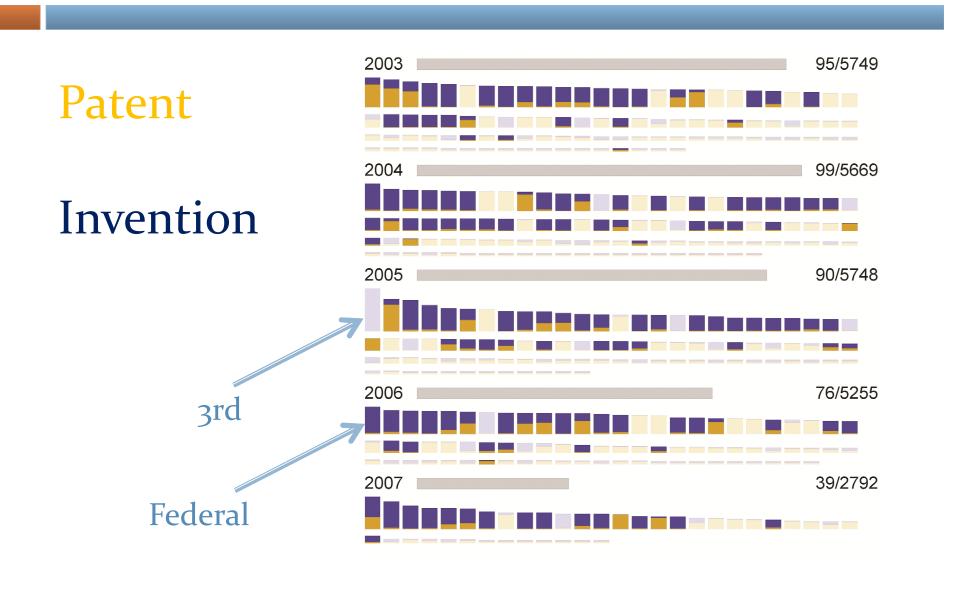


#### Document Browser



51/19340

### Document Browser



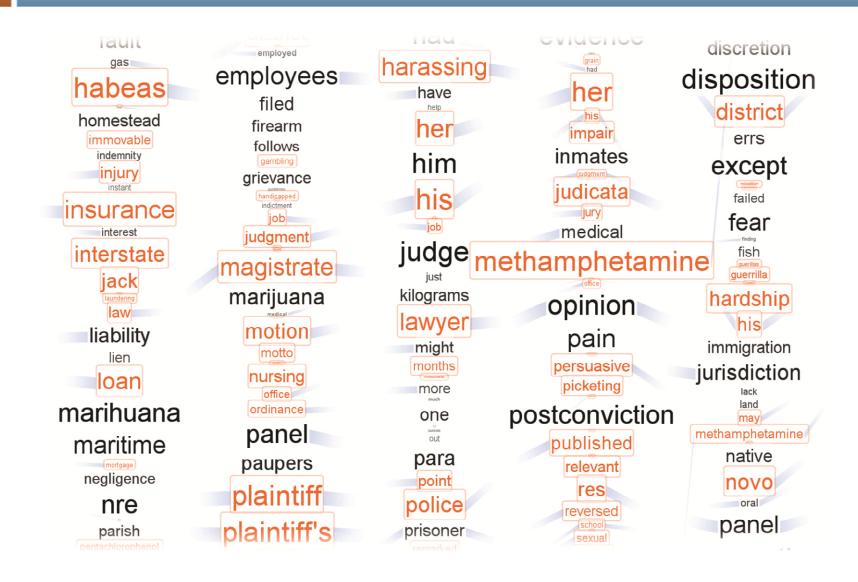
# Interaction

#### Selection of Data

Cases Score Cases Rank Coccurrence Score Coccurrence Rank Third (15660) Sixth (20581) Ninth (39875) Federal (10610) Supreme (33736) Options Fourth (23432) Seventh (17557) Tenth (15635) First (8311) DC (6454) Second (14623) Fifth (32224) Eighth (21096) Eleventh (26321) Unknown (518)

<Video>

# Change Over Time



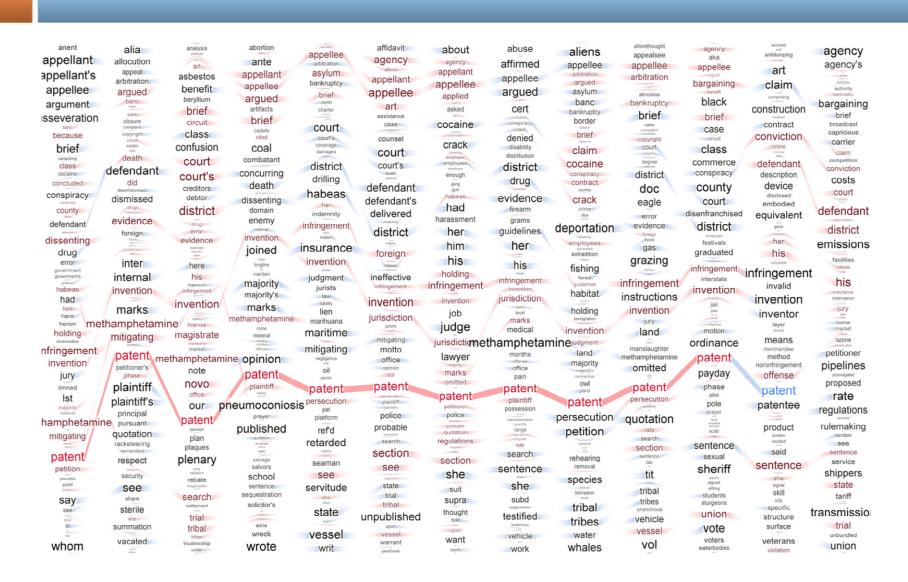
# **Exploring Circuit Relations**

add	adjourned aff'd	abstention allocatur	adequate	aboard admirally	appeal	about	The state of the s	abuse	argument	aggravating	agency	agency
anent	alia	analysis	affirmed	art	beganny	asked	affirmed	aliens	assistance	appellant		agency's
ante	apartment	app	aid	erticle	benefit claimant		appellee	appropriate	attached	n/relies	annuity	Application
appellant	appelleet	arbitration	appeal	barge	claimants	Certiofall	argued	asylum	brief	banc	antidumping	authority
appellant's	arbitration	bankruptcy	appellee	black	coal	cocaine	District in	11.800	cited	black	appeal	broadcast
isseveration	books	bargaining	argument	eargo	court	Test april 1	4 calle	circuit	collateral	boat	application	cable
below	tharge every	believe	authoritative	charter	and the same of the same of	defendan	cocaine	cited		capital	art	carrier
berzidine brow	commenced	benefit	because	CIBM	defendant	discolvey	Novice to the second	contended	сору	case	board	competition congressional
brief	congress	bottler	before	damages	defendant's	enough	timens.	court	defendant	confession	board's	costs
_stet claimant	copyright	class	coal	death	disability	lice	district	court's	determine	conspiracy	claim	data
claimant's		CHARACTER	cocaine	drilling	BARTANIA	gang	drug	denial	disfavored	county	compounds	
commonwealth	counsel		contention	fault	dismissed	get	evidence	deportation	doc	court	contract	emissions
	event	employee	court's	gas	employees	had	farm		doctrine	death	contractors	exemption explanatory
defendant del	films	forum	crack	homestead	examination	harassment	farmers	discretion	estoppel	defense	decision	bil
	foreign	holding	400	indemnity	filed	have	guilty	disposition	forthwith	dozer	device	gas
here	had		decisional denied	100	1	her	her	district	furnished	Tormer	declared	interpretation
inference	heroin	interest	disclosed	insurance	hereby	him	his	917	gas	guilt	equivalent	intervenor
	injunction	jurisdiction	dismissed	liability	his	his	ineffective	exception	-	habeas	inequitable	legislative
like	inter	legislative	dispense	loan	impair		inmates	fear	adgment	indictment	infringement	Science Sciences
limned	internal	liability	distribution	marihuana	judgment	judge	instructions	fishing	judicata	insanity	invalid	neroneste
might	keeplock	majority	district	maritime	magistrate	kilograms	interdistrict	his	material	insurance	invention	petitioner
more mortgage	marks	majority's	error	Thispalier	medical	lawyer	jury	minigration	nevertheless	Jurors	inventor	pipelines
plaintiff		non	fact	negligence	mine	might	medical	jurisdiction	opinion	jury	jurisdiction tayor	program
plaintiff's plausible	million	OUL	ton	nre	motion	more	nethamphetam	ine land	oral	marijuana	Spylor Stocial market	promulgated
point	narcotice	pension	his	offshore	pain	The same of the sa	murder	may may		methaqualone	marks means	proposed
	narcotics	plaintiff	legal	oil	panel	ostrich 2	opinion	inetribilipitetamilie	order	mitigating	merchandise	public
said		plan	lung	parish	politicos/g	plaintiff	pain	novo	persuasive	murder	noninfringement	rate
say	principal	plenary	magistrate	pre		police	postconviction	panel	plaintiff's	nonstatutory	obvious	regulations
see	racketeering	policy	material	platform	plaintiff	prison	punitive	persecution		penalty	patent	regulations
seems	reinsurance respect	precept	merits	recovery	plaintiff's	PRET DE	school	prison	record	petitioner's	patentee	reprinted
some suggests		predicate	opinion	ref'd	oneumoconiosis	rel	SCHOOL	provides	res	phase		rulemaking
sordy tenura	see	product	oral	The second second	pulmonary	she	sexual	public	sambid	precedent	product.	section
think	share	provision	The state of the s	rigging	pursuant	SHIT	she	review	suspended	racial	reexamine	
tit	shareholders Since		pneumoconiosis	seaman	recommendation	tentative	subd	section	tit trial	sentence	reissue	service
	stock	section	present		ricot	than	testified	sentence	unanimous	sir	removal	standard station
town	summation		pro	SLIP	sitting	thought	testimony	suitable		trial	retirement	statutory
trialworthy vessel		syrup	process	Inseaworthiness	Sitting	told	trial		unfavorable		said	supra
vis	supra	tit	pulmonary	Sing	unanimous	too	tribal	tribal tribes	unpublished	vessel	same a spec skill specific	tariff technology
	trade trademark	typescript	relief	vessel	union	want	unpublished	unanimous	value	vote white	structure surface	
whom	view	which	Saltania (	writ	work	what	verdict	ulolation	vol	you	use	waste whom
	-	would	wrote		11102111	Would	//orginal	WEYN.	1.01	100		

# **Exploring Circuit Relations**

add	adjourned aff'd	abstention allocatur	adequate	aboard admiralty	appeal	about	affirmed	abuse	argument	aggravating	agency	agency
anent	alia	analysis	affirmed	art	targaning	asked	appellee	aliens	assistance	appellant	advice, a grand	agency's
ante	apartment	app	aid	a ticle	benefit	1 4 4 4		appropriate	attached	100	annuity	Ampleaton
appellant	apoles.	arbitration		barge	claimant	certiorare	argued	asylum	brief	banc	antidumping	authority
appellant's	arbitration	bankruptcy	appeal	black	coal	cocaine	balava Not Care	1675		black	appeal	broadcast
isseveration	books	bargaining		cargo	court		CALL	circuit	cited	boat	application	cable
below	1000	believe	argument	charter	curiem	defendant	cattle	cited	collateral	capital	art	carrier
benzidine	commenced	benefit	authoritative because		defendant	dispinory		contended	copy	case	board	competition
brief	congress	bottler	before	CLGM	defendant's	enough			defendant		board's	congressional
brief	contempt	and the second	a professional and a second	damages	denied	even		court	defendant's	confession	claim	costs
claiment	copyright	class	coal		disability	fire gang	district	denial	determine	conspiracy	temporter	data
claimant's commonwealth	counsel		cocaine	drilling	dismissed	get	drug	denial	disfavored	county	compounds	emissions
	ettelete	9194	contention	fault	district		evidence	deportation	doc	court	contract	exemption
crusers elefonedment	event	employer	court's	gas	employees	had	farm	discretion		death	contractors	explanatory
defendant del	films	forum	crack	homestead	examination	harassment	farmers		estoppel	defense dozer	decision	lai
	foreign	holding		indemnity	filed	help	guilty	disposition	forthwith		device	gas
here		instance	decisional			her	her	district	furnished	Tormer	disclosed	interpretation
inference	had	interest	distribution .	insurance	hereby	nim	hio	err	020		equivalent	intervenor
IIII CI CI II C	heroin	jurisdiction	disclosed	liability	his		neffective	exception	1-11-1	guilt- habeas	inequitable	Selection .
like	injunction	legislative	dispense	Hability		his	inmates	failed fear	judgment		infringement	legislative
limned	inter	liability	distribution	loan	job	judge	instructions	fishing	judicata	insanity	invalid	
	keeplock	notionity	district	marihuana	judgment			Name of School	lent	insurance		note
might more	neeproon	majority	drug	maritime	magistrate	kilograms	interdistrict	his	material	Justin	invention	petitioner
more mortgage	marks	majority majority's		maritime navigation	magistrate magistrate medical	lawyer	jury	his Immigration	nevertheless	Jurous Her	inventor	petitioner pipelines
more mortgage plaintiff	marks	majority's	drug error fact			lawyer might	jury	his Immigration jurisdiction	contraction (even		inventor jurisdiction	
more mortgage	marks		drug	navigation		lawyer might	jury medical thamphetan	his immigration jurisdiction	nevertheless opinion	jury jury siling marijuana	inventor jurisdiction layer Person wantee	pipelines
more mortgage plaintiff plaintiff's plausible	marks	majority's non order OUI	drug error facts help help help help help	negligence	medical	lawyer might more	jury medical thamphetam murder	his immigration jurisdiction line land	nevertheless opinion oral	jury jury Miling marijuana methequelone	inventor  jurisdiction  layin  merits  marks	pipelines
more mortgage plaintiff plaintiff's plausible point	marks million narcotics	majority's non erder	drug error facts the the the the the the the	negligence nre	medical mine motion pain	lawyer might more	jury medical thamphetan	his immigration jurisdiction ine land may methamphetamine	opinion oral order	jury jury siling marijuana	inventor jurisdiction layer Person wantee	pipelines program promulgated proposed
more mortgage plaintiff plaintiffs plausible point said	marks million narcotics payment	majority's	drug error fact 25 25 15 15 15 15 15 15 15	negligence nre offshore oil	medical mine motion	might me more ostrich para	jury medical thamphetam murder murder	his immigration jurisdiction line land	opinion oral order persuasive	jury jury Miling marijuana methequelone	inventor  jurisdition  jurisdit	pipelines program promulgated proposed
more mortgage plaintiff plaintiffs plautiffs plautiffs plautiffs said said say	marks million narcotics payment principal	majority's non erder OUF pension plaintff plan	eug error fried haben his his legal lung magistrate	navigation negligence nre offshore oil	medical mine motion pain panel	lawyer might more ostrich para plaintiff	jury medical thamphetam murder printing opinion pain	his immigration jurisdiction ine land may methamphetamine	opinion oral order	jury jury sarg marijuana methequitore mitigating	inventor  aparisdation  aparis  marks  means  merchandise  noninfringement	pipelines program promulgated proposed public rate
more mortgage plaintiff plaintiffs plausible point said	marks million narcotics payment	majority's norr our our pension plaintff plan plenary	eug error frot han his his legal lung magistrate material	navigation negligence nre offshore oil owner parish	medical mina motion pain panel panel panel panel panel panel	might more estimates pora plaintiff police	jury medical thamphetam murder murder opinion	his immigration jurisdiction including land may methamphetamine	nevertheless opinion oral order persuasive	jury  Jigg  Marijuana  methogoalone  mitigating  murder	inventor spiris deten spiris s	pipelines program promulgated proposed public rate regulations
more mortgage plaintiff plaintiffs plausible point said say see seems	marks million narcotics payment principal racketeering reinsurance	majority's roon erdess OUF pension plaintff plan plenary policy	dug error fred his his legal lung magistrate material	navigation negligence nre offshore oil owner parish parish parish parish parish parish parish parish	medical mine motion pain panel	lawyer might more ostrich para plaintiff	iury medical thamphetam murder murder popinion pain postconviction punitive reversed	his immigration jurisdiction ine land may methamphetamine novo panel persecution	opinion oral order persuasive	jury jury marijuana restrupulstre mitigating murder nonstatutory	inventor  Juris diction  Styre  Juris diction  Styre  Juris diction  Styre  Juris diction  Juris	pipelines program promulgated proposed public rate regulations
more mortgage plaintiff plaintiffs plausible point said say see	marks million narcotics payment principal racketeering	majority's non ender OUT pension plaintiff plan plenary policy precept	eug error frect fr	navigation negligence nre offshore oil owner parish pre platform recovery	medical mina motion pain panel panel panel panel panel panel	might more para plaintiff police prison rider	iury readical thamphetam murder am opinion pain postconviction puntive	his immigration jurisdiction ince land may methamphetamine novo panel	opinion oral order persuasive plainerrs procedential	jury  states jury  states marijuana merijuana mitigating murder nonstatutory penalty petitioner's phase	inventor  Juris de ton  Juris de ton  Juris	pipelines program program proposed public rate regulations regulations
more mortgage plaintiff plaintiff plaintiffs plausible point said say see seems some suggests	marks million narcotics payment principal racketeering reinsurance respect	majority's non ender OUF pension plaintiff plan plenary policy precept predicate	eug error frect fr	navigation negligence nre offshore oil owner parish pps platform	medical motion pain panel panel plaintiff plaintiff's	might more para plaintiff police prison redure	iury medical thamphetam murder murder popinion pain postconviction punitive reversed	his immigration jurisdiction ine land may methamphetamine novo panel persecution prison	opinion oral order persuasive procedential	marijuana methopulore mitigating murder nonstatutory penalty petitioner's phase	inventor  Juris diction  Styre  Juris diction  Styre  Juris diction  Styre  Juris diction  Juris	pipelines program promulgated proposed public rate regulations
more mortgage plaintiff plaintiff plaintiff plaintiffs plausible point sald say see seems some suggests sarely lanus	marks million narcotics payment principal racketeering reinsurance respect	majority's non sider OUT pension plaintiff plan plenary policy precept predicate product	legal lung magistrate material merits mine opinion	navigation negligence nre offshore oil owner parish pre platform recovery ref'd rigging	medical motion pain panel panel plaintiff plaintiff's	might me more control para plaintiff police prison rather rel	iury redical thamphetam murder pointon pain postconviction portive reversed school	his immigration jurisdiction ine land may methamphetamine novo panel persecution prison provides public	opinion oral order persuasive parametrs precedential	marijuana retriegueire mitigating murder nonstatutory penalty petitioner's phace	inventor jurisdiction by the state of the st	pipelines program program proposed public rate regulations regulations
more mortgage plaintiff plaintiffs plautiffs plautiffs plautiffs saide say see seems some suggests says	marks million narcotics payment payment principal racketeering reinsurance respect	majority's non sider OUT pension plaintiff plan plenary policy precept predicate product provision	eug error frec. fr	navigation  negligence nre offshore oil owner parish platform recovery ref'd rigging school	medical minita motion pain panel	might more para plaintiff police prison redure	thamphetam murder opinion pain postconviction puritive reversed school sentence savual	his immigration jurisdiction ine land may methamphetamine novo panel persecution prison provides	opinion oral order persuasive persuasive precedential res	marijuana restruepustore mitigating murder nonstatutory penalty petitioner's phase precedent	inventor  Juris diction  Styring  marks  means merchandise  mole mole mole mole mole mole mole mole	pipelines program promulgated proposed public rate regulations regulations reprinted rulemaking
more mortgage plaintiff plaintiff plaintiff plaintiffs plausible point sald say see seems some suggests sarely lanus	marks million narcotics payment principal racketeering reinsurance respect shareholders since	majority's non ender Our pension plaintiff plan plenary policy precept predicate product provision	eug error frec. fr	navigation negligence nre offshore oil owner parish par platform recovery ref'd rigging school seaman	medical mine motion pain panel panel plaintiff plaintiff's pulmonary pursuant ray recommendation	might me more control para plaintiff police prison rather rel	thamphetam murder opinion pain postconviction puritive reversed school sentance sexual	his immigration jurisdiction jurisdiction in and may methamphetamine novo panel persecution provides public review section section.	opinion oral order persuasive parametrs precedential	marijuana restructuring marijuana restructuring mitigating murder nonstatutory penalty petitioner's phace procedent procedent facial sentence	inventor jurisdiction by the state of the st	pipelines program promulgated proposed public rate regulations reprinted rulemaking section service
more mortgage plaintiff plaintiff plaintiff's plausible point said say see seems some suggests surgests think tit	marks million narcotics payment principal racketeering reinsurance respect see share shareholders	majority's non sider Our pension plaintir plan plenary policy precept predicate product provision maggies section	eug error frot frot frot frot frot frot frot	navigation negligence nre offshore oil owner parish ppe platform recovery ref'd rigging school seaman	medical mine motion pain panel panel plaintiff plaintiff's poursuant ray recommendation recorr	might more more para plaintiff police prison reduce she suit tentative	iury medical thamphetam murder pointion pain postconviction puntive reversed school sentence sownal she subd testified	his immigration jurisdiction jurisdiction in and may methamphetamine novo panel persecution prison provides public review team section section section is a constant of the co	opinion oral order persuasive reserve plaintre precedential res suspended	marigating murder nonstatutory penalty petitioner's phase	inventor  jurisdiction  superior  marks means merchandise noninfringement obvious  patent patentee  product reexamine reissue removal	pipelines program promulgated proposed public rate regulations reprinted rulemaking section service standard station
more mortgage plaintiff plaintiff plaintiffs plausible point said say see seems some suggests says think tit town	marks million narcotics payment principal racketeering reinsurance respect shareholders since	majority's non ender Our pension plaintiff plan plenary policy precept predicate product provision	eug error froct fr	navigation negligence nre offshore oil owner parish pre platform recovery ref'd rigging school seaman	medical mine motion pain panel	might more more para plaintiff police prison redure she suit tentative than thought	postconviction postconviction puntive reversed school sentence sexual she subd testified testimony	his immigration jurisdiction.  jurisdiction.  ine land may methamphetamine novo panel persecution prison provides public review team section section.	opinion oral order persuasive reserve plaintre precedential res suspended trial unanimous	marijuana restructuring marijuana restructuring mitigating murder nonstatutory penalty petitioner's phace procedent procedent facial sentence	inventor  jurisdiction  light in the control of the	pipelines program promulgated proposed public rate regulations reprinted rulemaking section service standard statutory
more mortgage plaintiff plaintiff plaintiff plaintiffs plausible point sald say see seems some suggests sarely tanus think tit town trialworthy	marks million narcotics payment principal racketeering retrisurance respect share share share stock summation	majority's read read read read read read read read	eug error frect fr	navigation negligence nre offshore oil owner parish parish recovery ref'd rigging school seaman	medical minis motion pain panel panel plaintiff plaintiff's pulmonary pursuant ray recommendation record selectory	might more para plaintiff police prison reder she suit tentative than thought told	postconviction postco	his immigration jurisdiction jurisdiction in and may methamphetamine novo panel persecution prison provides public review team section section section section in and may misser in section section section in a sect	nevertheless opinion oral order persuasive persuasive persuasive persuasive persuasive persuasive persuasive persuasive precedential res suspended trial unanimous unfavorable	marigating murder nonstatutory penalty petitioner's phase precedent formation and the state of t	inventor  jurisdiction  jurisdiction  jurisdiction  marks means merchandise moninfringement obvious  patent  patentee  patentee  product reexamine reissue removal retirement said	pipelines program promulgated proposed public rate regulations reprinted rulemaking section service standard statulory supra
more mortgage plaintiff plaintiff plaintiffs plausible point said say see seems some suggests surely tanks tit town trialworthy vessel	marks million narcotics payment principal racketeering reinsurance respect share share shareholders since stock summation supra	majority's room room plantiff plan plenary policy precept predicate producti provision section statuta	eug error friot his legal lung magistrate material merits mine opinion oral pneumoconiosis present pro process putnomary	navigation negligence nre offshore oil owner parish ppe platform recovary ref'd rigging school seaman	medical mine motion pain panel panel plaintiff plaintiff's plaintiff's pulmonary pursuant ray recommendation report sodentary sitting unanimous	more more more more more more more more	postconviction portive reversed School sentence sexual she settined testimony trial tribal	his immigration jurisdiction ine land may methamphetamine novo panel persecution pison pison provides public review section section section section terance suitable	opinion oral order persuasive reserve plaintre precedential res suspended trial unanimous	marijuana marijuana mitigating murder nonstatutory penalty petitioner's phase precedent racial sentence strial vessel	inventor  jurisdeten  jurisdeten  jurisdeten  jurisdeten  jurisdeten  jurisdeten  jurisdeten  marks  means merchandise  noninfringement obvious  patent  patentee  jurisdeten  patentee  jurisdeten  product reexamine reissue removal retirement  said  said  said	pipelines program promulgated proposed public rate regulations regulations reprinted rulemaking section service standard station statutory supra tanff
more mortgage plaintiff plaintiff plaintiffs plausible point said say see seems some suggests surely tanks tit town trialworthy vessel	marks million narcotics payment principal racketeering reinsurance respect share share share share stock summation supra trade	majority's non rotation plantiff plan plenary policy precept predicate product provision section statute syrup	eug error frec. fr	navigation negligence nre offshore oil owner parish par patform recovary ref'd rigging school seaman suit unseaworthiness usey vessel	medical mine motion pain panel panel panel plaintiff plaintiff's plaintiff's preumoconiosis pulmonary pursuant ray recommendation recommendation sedentary sitting unanimous union	more more para plaintiff police prison rether suit tentative than told too want	postconviction postco	his immigration jurisdiction ine land may methamphetamine novo panel persecution provides public public sertence suitable imser tribal tribes	opinion oral order persuasive parametrs precedential results or the trial unanimous unfavorable unpublished value	particular and processing and proces	inventor jurisdiction jurisdiction jurisdiction jurisdiction jurisdiction jurisdiction marks means merchandise moinfringement obvious patent patentee patent	pipelines program promulgated proposed public rate regulations reprinted rulemaking section service standard statulory supra
more mortgage plaintiff plaintiff plaintiff plaintiffs plausible point said say see seems some suggests say think tit town trialworthy vessel	marks million narcotics payment principal racketeering reinsurance respect share share shareholders since stock summation supra	majority's non order our pension plaintiff plan plenary policy precept predicate product provision section statuta syrup tit	eug error friot his legal lung magistrate material merits mine opinion oral pneumoconiosis present pro process putnomary	navigation negligence nre offshore oil owner parish ppe platform recovary ref'd rigging school seaman	medical mine motion pain panel panel plaintiff plaintiff's plaintiff's pulmonary pursuant ray recommendation report sodentary sitting unanimous	more more more more more more more more	thamphetam murder opinion pain postconviction puntive reversed school sentence sexual she subd- testified testimony trial tribal	his immigration jurisdiction ine land may methamphetamine novo panel persecution pison pison provides public review section section section section terance suitable	opinion oral order persuasive plantiffs procedential resord res suspended trial unanimous unfavorable unpublished	marijuana marijuana mitigating murder nonstatutory penalty petitioner's phase precedent racial sentence strial vessel	inventor  jurisdeten  jurisdeten  jurisdeten  jurisdeten  jurisdeten  jurisdeten  jurisdeten  marks  means merchandise  noninfringement obvious  patent  patentee  jurisdeten  patentee  jurisdeten  product reexamine reissue removal retirement  said  said  said	pipelines program promulgated proposed public rate regulations reprinted rulemaking section service standard station statutory supra taniff techology

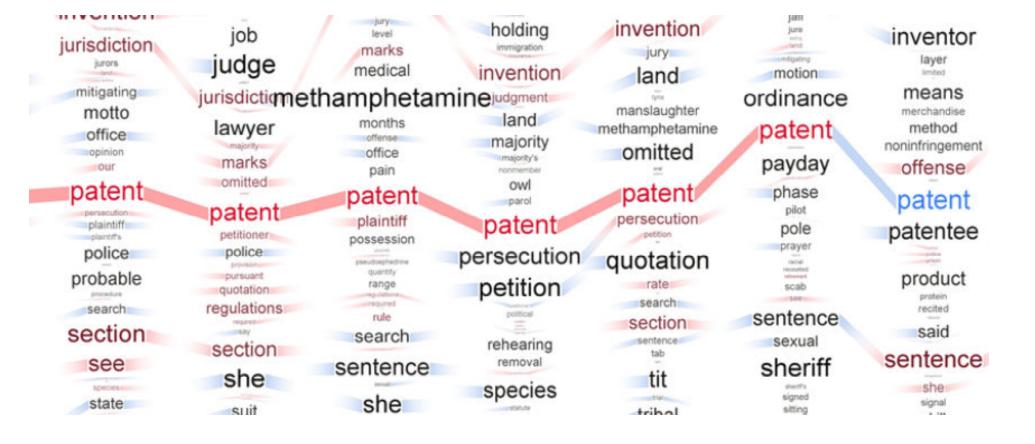
# Significant Absence



# Significant Absence



# Significant Absence



# Size by Score

accused niin affirmaxi and perional 85000 agency adequate DEBUX ....ivivi 9000 argued annuity antidumping and color aid Dinding art 1911119 181 - brief defendant e argument board Ciled alastias claim (2) (2) 3191 construction defendard graduated harri contract deportation CODY fisteign incresumities. decision his district device methamphetam@hsp.osition.disfavor emissions judge inter equivalent 11121111 decisional imayer 112 entrant est 19880 Staff exports pand inequitable PER BEREIT 031120120 чевародия (21971) dispensed nfringeme galainia?"a. furnished invalid principal amor ...... invention (817)19 888 100000 1/01/11/11/11 inventor laysit. 11111111 layer literal whorn engeleind navarthalasa means merchandise DSISCUROT opinion mode noninfringement photemaconiosis patent persuasive patentee patentee's product reexamination reissue retirement said unfavorabie signal structure subheading surface

veterans

# Analysis

# Ostriches in the 7<sup>th</sup> Circuit

#### ostrich\*

2000 6/5888

2001 7/5910

2002 6/5674

2003 7/5749

2004 6/5669

2005 2/5748

2006 6/5255

#### THE WALL STREET JOURNAL.

WSJ.com

JUNE 25, 2008, 5:39 PM ET

#### Conrad Black's Sentence Upheld; 7th Circuit OKs Ostrich Instruction



Former newspaper tycoon Conrad Black enters the Chicago federal courthouse, June 30, 2006 (Credit: Associated Press/Spencer Green)

Media baron and <u>author</u> Conrad Black inched closer to the end of his legal road today when the opinion, penned by Judge Posner — whom Black recently called "part of the prosecution" — as

Black, 63, and three other former Hollinger International executives — Peter Atkinson, John Bout newspaper company out of millions of dollars. The defendants allegedly caused Hollinger, whice for "private gain."

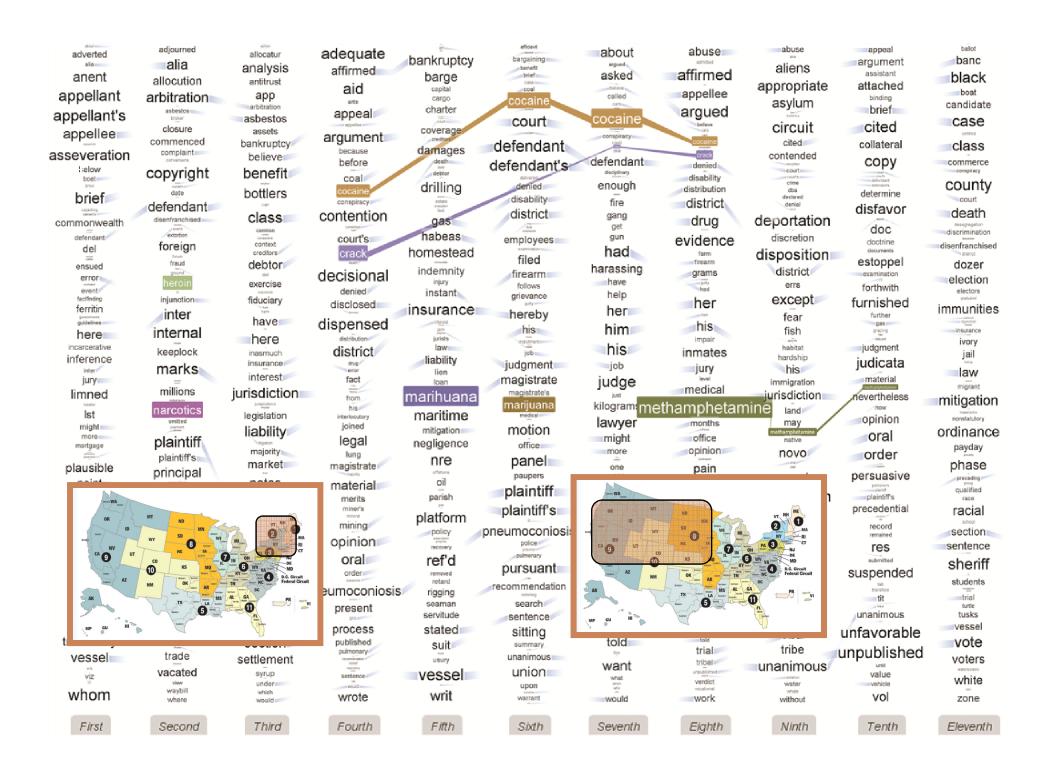
"If you receive a check in the mail for \$1 million that you have no reason to think you're entitled know you weren't entitled to the money," Posner wrote, rejecting the defense's argument that t instruction"? An "ostrich instruction," Posner explains, "tells the jury that to suspect that you is the equivalent of intending to commit the crime." Posner also clears up a fallacy: The legend legend and a canard on a very distinguished bird.")

Patrick J. Fitzgerald, the U.S. Attorney for the Northern District of Illinois, said in a statement, multimillion dollar corporate fraud scheme and deprived the public shareholders of Hollinger Inte

Black's defense attorney, Mayer Brown's Andrew Frey, told the Sun Times that, the defense to straightened out." Black could still try appealing to the entire 7th Circuit Court and the U.S. Su

# Highfalutin Judge Selya

- First Circuit shows unusual words ... "furculum", "impuissant", "immurement", ...
- Tooltips in document browser reveal all associated cases written by Judge Selya



# Conclusion

## Related Work



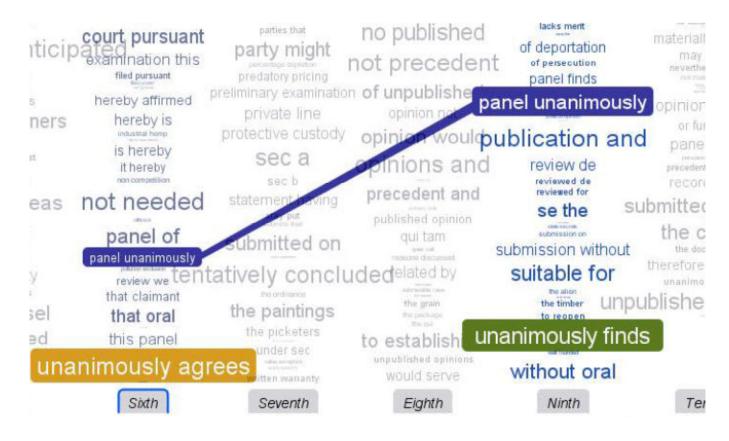
Jigsaw, Stasko et al., 2007



Themail, Viégas et al., 2006

# Challenges

- Scale of data
- Scale of visualization



### Conclusion

- Analytics generalizable:
  - News articles about different companies
  - Academic articles
  - Collections of product reviews
  - Literary analysis
- Ideas for extension:
  - Semantic zoom
  - Hybrid case-occurrence scoring metric



#### **Acknowledgements:**

Irene Ros and the VCL team, helpful reviewers, iLab Calgary folks!

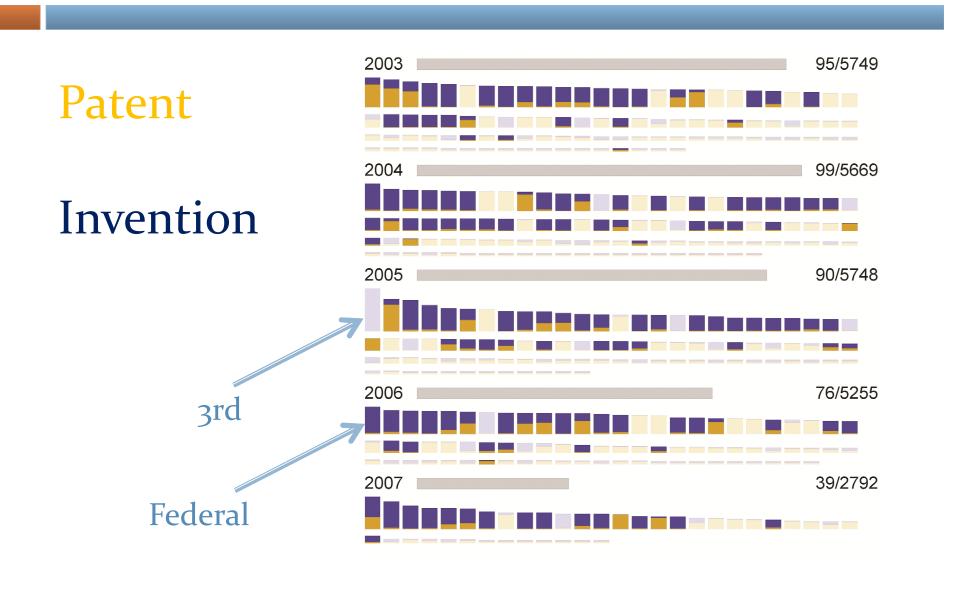
Video:

http://www.christophercollins.ca

Christopher Collins, Fernanda B. Viégas, and Martin Wattenberg

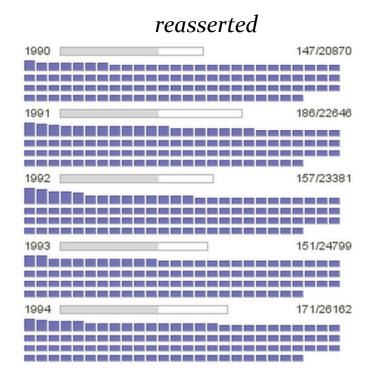
# Extra Slides

### Document Browser



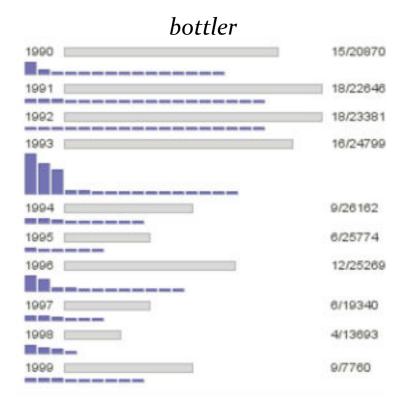
# Scoring by Case

- Discovers linguistic nuance: capricious, reasserted
- Geographic variances: drilling, oil, cocaine
- Circuit term preferences: typescript vs. transcript



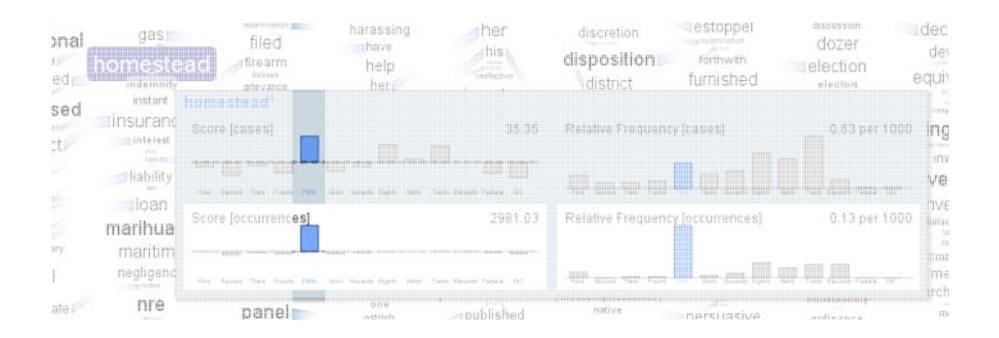
# Scoring by Occurrence

- Common words in long cases:
   asbestos, discrimination,
   bottler
- Highly skewed distribution: she, police, damages



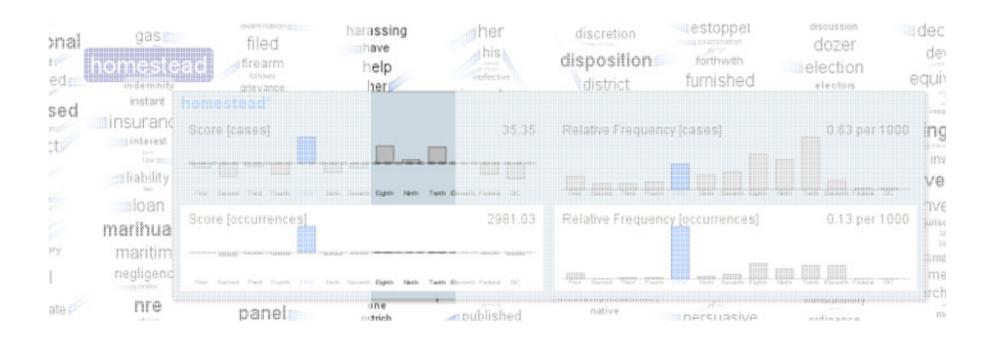
## Tooltip Graphs

- Compare measures
- Significant absence
- □ See scores for all columns at once



## Tooltip Graphs

- Compare measures
- Significant absence
- □ See scores for all columns at once



## Research Process

### Meet with legal experts

 Basic grasp of data and questions Design initial visualization(s)









#### Acquire the data

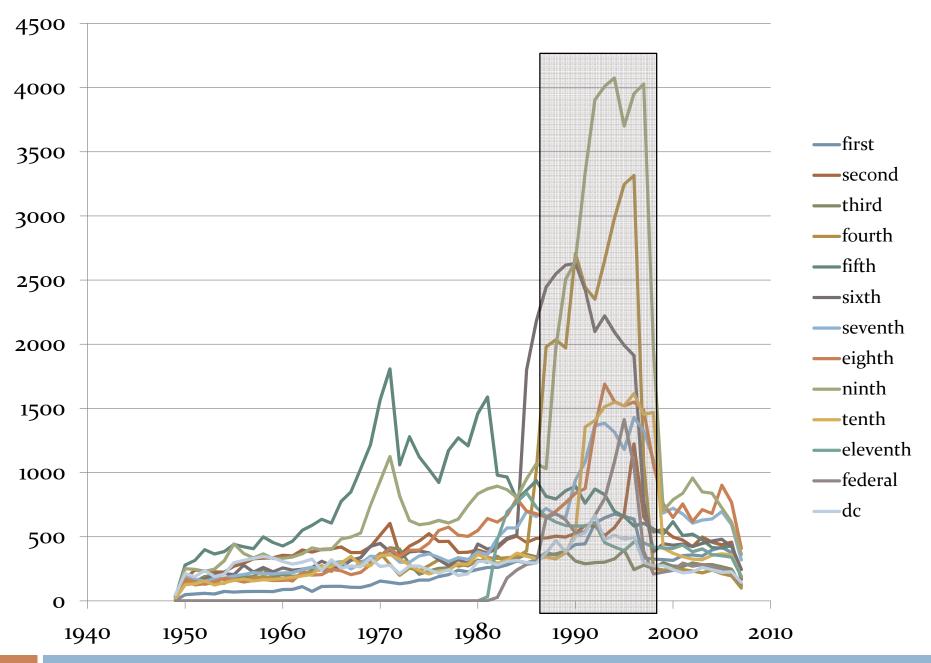
• Develop analysis infrastructure

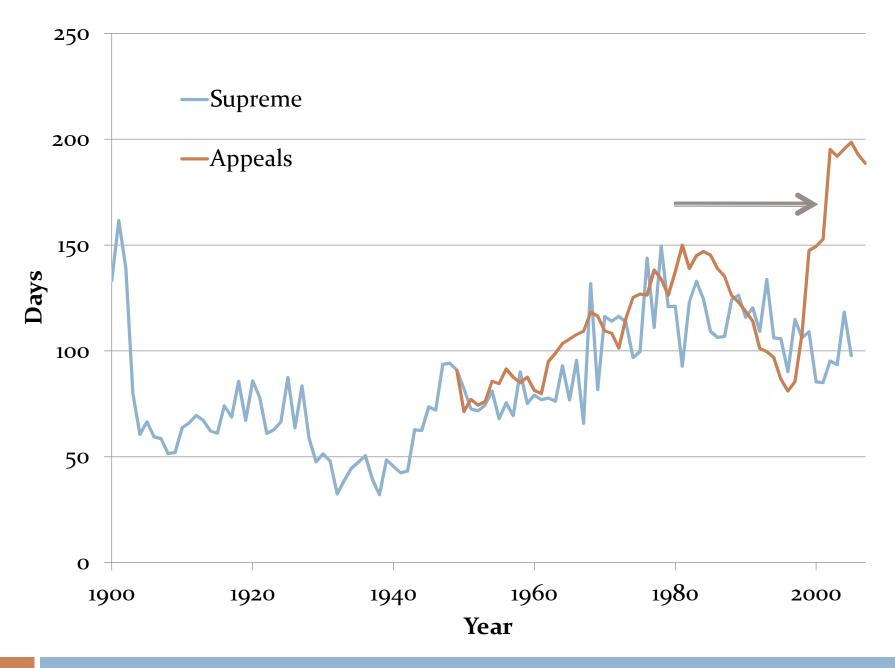
#### Revise visualization

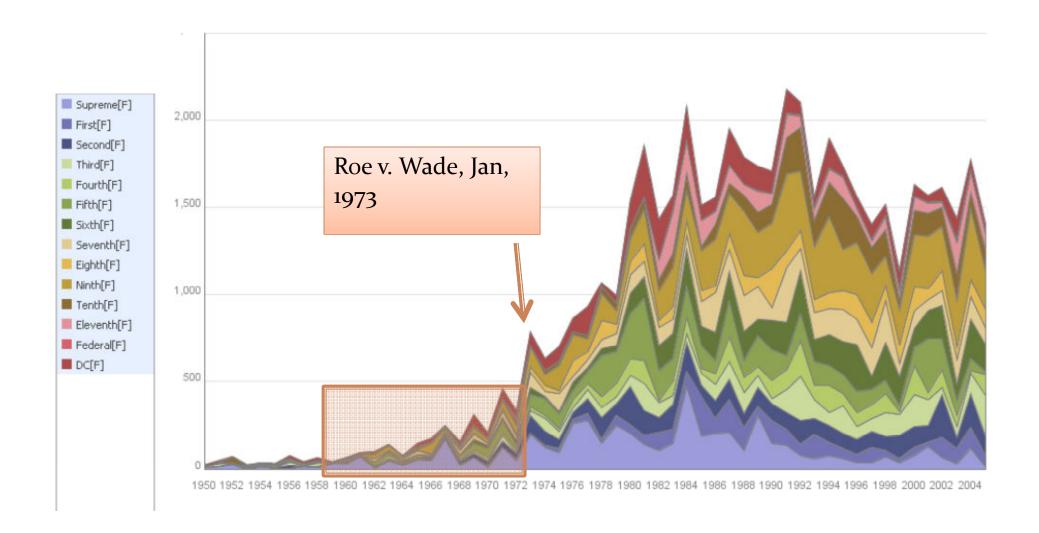
• Meet with data experts for review

## Characteristics of Data

Federal Reporter Database







# Dynamic Stop Words

- □ Top ½ percentile by frequency
  - Common words, often do not carry meaning
  - Better than stop words; captures context
  - [court, and, his, judge, circuit, first, usapp, ...]
- Bottom percentile by frequency
  - Spelling errors
  - Hapax legomena (words occurring once)

# Data Parsing Tools

- Xerces XML Parser
  - Mark-up provides hints to data content, but often incorrect
- Regular Expression Processing
  - recognize authorship statements
  - concurring, dissenting opinions
  - citations
  - remove copyright notices
- Date Parsing
  - 15 different formats!

# Indexing and Search

#### Lucene:

Industry standard open-source search engine.

#### □ Extensions:

- Full text retrieval with more than one field of same name per document (e.g. multiple dissenting opinions).
- Capture Analyzer to capture and retain original words during stemming.

# Storage and Data Processing

- Postgres Database
  - Stores aggregate term vectors and statistics per document
- Multi-threaded
  - Parsing, indexing, and storage concurrent
  - 10-12 hours to index and process all documents
- Result
  - Code infrastructure for handling large text corpora