

A proof of concept for automated discourse analysis in support of identification of relationship building in blogs

David Phillips and Bruno Amaral

- Overview
 - Large scale empirical evidence
 - Advanced search and content analysis tools
 - Identifying how relationships are formed, sustained and changed in public discourse
 - Very far reaching for PR, but also all soc sciences, psychology, evolutionary science, political and economic theory.
 - Some early evidence about the dynamic of relationships
 - Exciting that PR has a fundamental breakthrough in societal research

- Reputation, trust, brand loyalty, lobbying, corporate license to operate depend on relationships
- PR is a relationship discipline.
- What do we know about the dynamics of relationships?
- Formation, Sustaining & Change

- Towards Relationships Management
 - Widely understood tokens
 - Values make up tokens
 - Actors have different perception of tokens
 - Common beliefs in values support relationships
 - Actors with different understanding of value invoke dissonance.

- There is considerable research about relationships.
- Very few large scale observations to identify the actual building blocks in the formation sustaining and change in relationships
- Empirical research on a very large scale needed – especially in discourse.

- Personal blogs offer large scale discourse for social observation
- Identifying actors
- Over time
- Tokens & values
- Relationships

- Triangulated testing:
 - Hyperlink network analysis
 - Word map networks
 - Semantic concept networks

Values and tokens as integral parts of relationships

Hyperlinks can be more than just resource locators — tokens

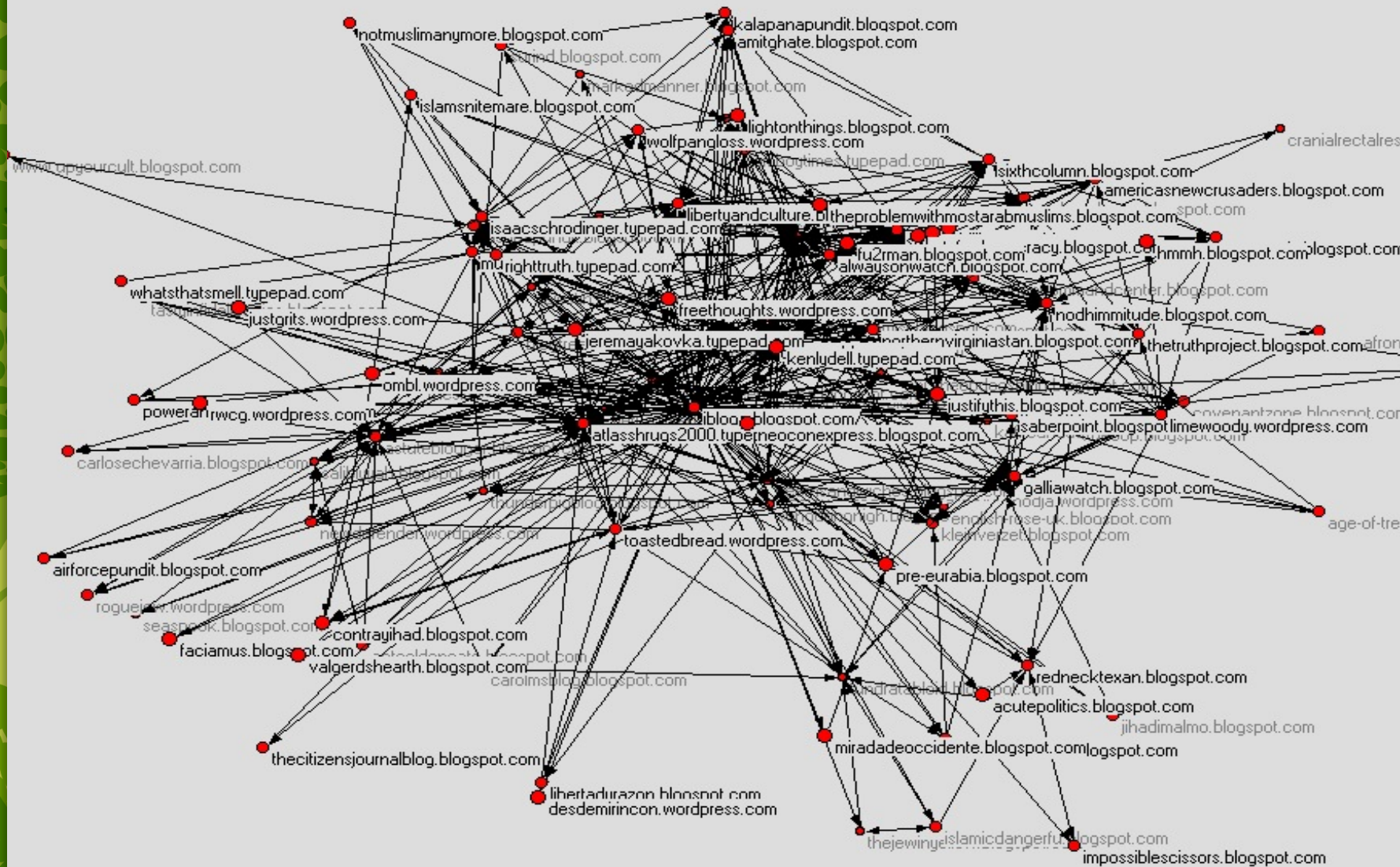
The subjects actors talk about in blogs reflect actor tokens & values

- Our goals
 - To explore the possibility of mapping both blogs and relationships between actors
 - To draft a methodology for future research in mapping relationships online

- The tools
 - Urlnet
 - Pajek (large network analysis)
 - Latent Semantic Analysis
 - Word Density Analysis

Looking at networks

The next graph will show us the network around “the infidel bloggers’ alliance” blog

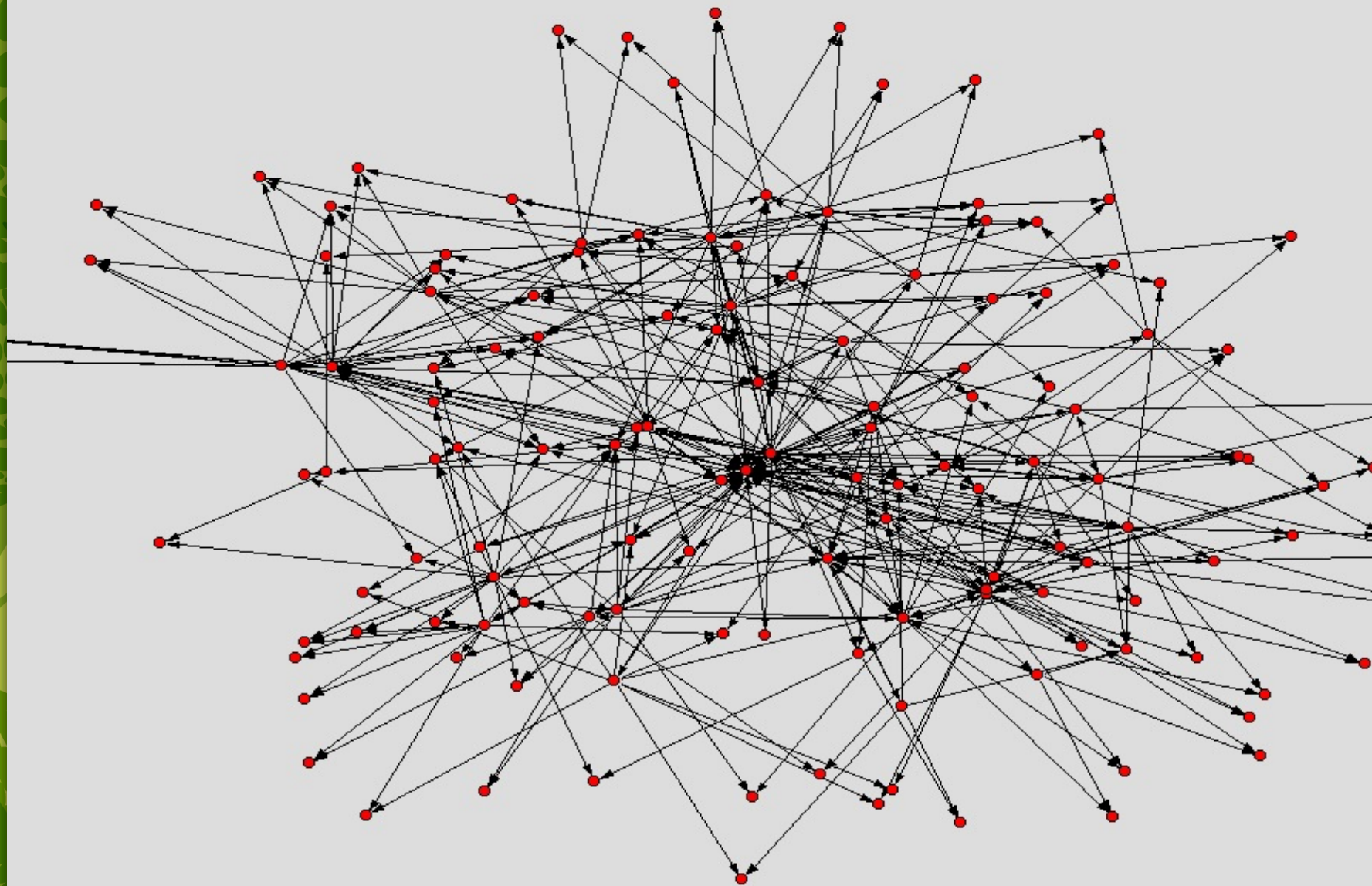


Word frequencies in the LSA Results

word	iraq	islam	gaza	palest	arab	muslim
number of occurrences in LSA	9	46	6	12	33	12

Looking at networks

The next graph will show us the network around Jeremiah Grossman's blog



A	B	C	D	E	F	G	H	I	J
jeremiahgrossman.blogspot.com/	web applica	new silicon	logic flaws	figure sums	ip statistics	silver tail	business an	vulnerability	bad act
agguvenligi.blogspot.com	devlet planlar	ulakbim	lamak	bilgi toplumu	csirt	kilat	embe	ulak	amba
andyitguy.blogspot.com	budget crunch	traffic nightm	arnold palmer	key line	upper 20s	sink hole	city of atlanta	pennzoil	line bre
anil-identity.blogspot.com	personal onlin	chief bottle w	city of london	networking	data wareho	computer fo	personal imag	w3c workshop	enisa
armorize-cht.blogspot.com	carnegie mell	chenxi wang	tim bass	web proxy	bass web	cylab	google	dhruv	black ha
carnal0wnage.blogspot.com	script kiddie	risk identifi	penetration	mundane task	automated to	nessus	subnet	shells	exploit
christ1an.blogspot.com	cryptic error r	chris shiflett	php program	valid attribute	content mode	java port	html spec	line script	matthie
cleartext.wordpress.com	john mcmanu	heartland pay	legal sanction	ben edelman	insightful ana	ecommerce tr	writing secu	payment sec	interne
codeczar.blogspot.com	decimal numb	number twelv	negative num	bcd	fingers	ba	labels	long time	
coding-insecurity.blogspot.com	double byte c	double byte c	special filter	pression	application s	level applicati	manico	sample applic	google
datasecurityblog.wordpress.com	malicious so	malicious so	open source s	data loss pr	intrusion pr	security pos	laptop hard d	information	inbound
ddanchev.blogspot.com	courier delive	english voices	credit card de	multilingual c	female voices	life crime	cybercrimin	cultural divers	commen
dutcherstiles.blogspot.com	narrative para	information	collective wis	corporate sto	stock value	dutcher	e mail	bar charts	pie char
ethicalhackernet.blogspot.com	computer es	electronic fr	dns protocol	computer cr	due authoriz	article talks	technology ar	falsification	el delito
extremesecurity.blogspot.com	windows sec	technet micro	exe windows	windows erro	privileged a	wsus	domain adm	tool windows	reportin
fr3dc3rv.blogspot.com	quick brown f	message dige	hash algorithm	hash algorithm	hash function	secure hash	downgrade i	hash function	cryptog
hackathology.blogspot.com	risk manage	fight identit	flag activities	manual check	red flags	fraudulent act	risk factor	response plan	financia
holisticinfosec.blogspot.com	russ mcree	sql	trustmark	pierini	xss	forgery	php web	csrf	insertion
hollistalkimontheblog.blogspot.com	lufthansa stev	lufthansa flight	jeremy jacob	ladies and ge	lufthansa	emergency ex	lsq	air crash	15 janu
ilkeryoldas.blogspot.com	image society	quick quiz	health society	ignorance is b	geographic kr	age of inform	fifth grader	2 year olds	map rea
infosecramblings.wordpress.com	business cor	disaster rec	flash develop	script functi	administrati	action script	twits	bragging righ	differen
jdm-tech.blogspot.com	beta nvidia	irc www	nvidia drivers	disk encrypti	package man	boot time	lirc	26 december	tree str
kuza55.blogspot.com	javascript stri	david ross	active code	meta tag	security me	le8	sensitive da	server side	whole lo
kyran.wordpress.com	csrf	rapid success	ckers	logout button	malicious us	captcha	nonce	9 out of 10	rundown
labnol.blogspot.com	microsoft offic	lonely planet	dell laptops	wall street jou	popular search	free addon	html webpage	wall street jou	docume
log0.wordpress.com	web applica	hex numbers	stack size	stack trace	createinstanc	message loop	system defau	msdn	workar
loquens-caesu.blogspot.com	computer se	tcp dos	security met	robert e lee	management	couple memb	smart person	fyodor	local cor
michaelcoates.wordpress.com	logon credent	url bar	passerby	security risk	attacker	guise	page content	url address	worksta
muchikon.blogspot.com	linux x86	google	debian etch	proyeccion	virtual machi	kvm	gnome	x2	kernel
myappsecurity.blogspot.com	information	web applica	india confer	shadow bar	summit event	owasp	security tre	leading minds	measur
pentesterconfessions.blogspot.com	arp spoofing	arp spoof	visual cues	javas	mitm	recomendatio	financial instit	co worker	subnet
preachsecurity.blogspot.com	grudge match	epic proporti	gas pipeline	deep winter	russian sour	gas crisis	political situat	google	val kiln
rationalsecurity.typepad.com	chris hoff	human eyeba	technorati tag	cross pollinati	logic analysis	storage availa	fuzzy logic	diverse netw	securit
schmoil.blogspot.com	domain valida	validation pro	full disclosure	security exp	ca domain	registered ag	application s	certificate a	ssl cert
securitycoin.blogspot.com	bank pins	encryption k	registration a	web apps	crown jewels	access cards	security mo	magstripe	dual cor
securityninja.blogspot.com	ninja website	security for	new email ad	application s	rook	new security	dss	ramblings	informa
shoaibyouusuf.blogspot.com	western unior	western uni	persistent my	union mone	malicious vi	hacking gam	evil softwar	game account	softwar

- Special thanks to software developers who contributed to this research project:
 - Girish Lakshminarayana and his team of programmers in India – who developed a special Latent Semantic Analysis tool
 - Dale Hunscher Business Systems Analyst University of Michigan Medical School and Owner of South Wind Press – who provided the software and online capability to explore blog networks.

- Bruno offers an approach for discourse research that goes to the heart of many areas of research in PR and beyond.
- We are beginning to see evidence that common values create relationships and probably publics
- Indications that values expressed in discourse on and offline are a dynamic in formation of relationships, networks and social groups.
- Potentially powerful theoretical proof for PR research and practice in relationship management

Thank You