

DATA EXTRACTION, MODELING AND APPLICATIONS IN E-SOCIAL SCIENCES




M.R. Martínez Torres
University of Seville
Seville, Spain
rmtorres@us.es



2

1. e-Social Sciences

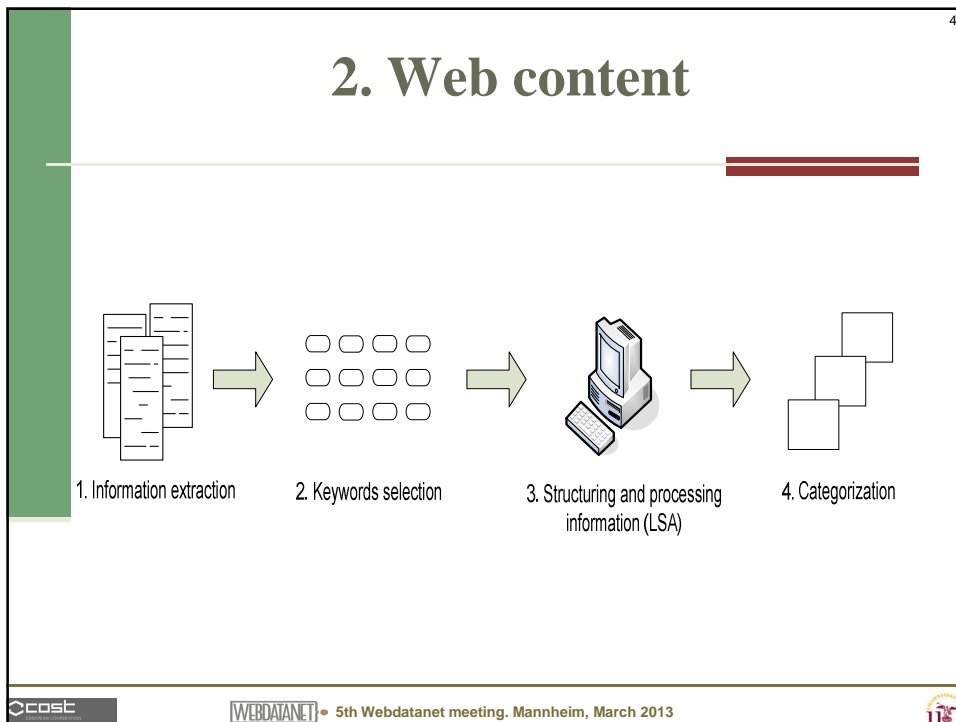
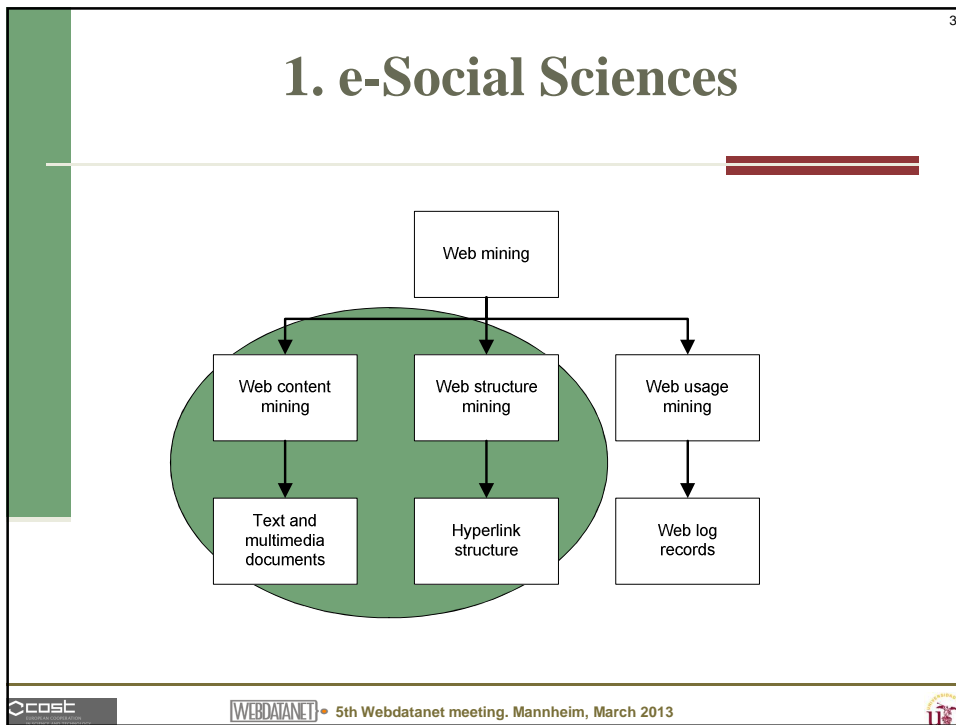
Extraction of data from the web and their analysis through the use of Social Network Analysis and Semantic Analysis techniques

Virtual communities  Open Source Software communities
Open innovation communities

Website structure analysis

Knowledge networks





2. Web content

Oss projects mailing list

debian-arm May 2012 by thread

Discussions / Leadership Qualifies

Page 1 of 1

Search: [Date Index] [Date Desc] [Thread Desc] [Thread Index] [Date Index] [Thread Index]

Report as spam

Number of messages Views

GCC 4.7 is now the default for x86 architectures 2 208

- To: debian-ports@lists.debian.org
- Cc: Debian Release <debian-release@lists.debian.org>, Debian GCC Maintainers <debian-gcc@lists.debian.org>
- Subject: GCC 4.7 is now the default for x86 architectures
- From: Matthias Klose <klo@debian.org>
- Date: Mon, 07 May 2012 12:11:22+0200
- Message-ID: <133637133.901000@lists.debian.org>

GCC 4.7 is now the default for x86 architectures for all architectures except the 68000series, including FPE6800 and the Busc.

There are still some build failures which need to be addressed. Out of the 350 bugs filed, more than the half are fixed, another quarter has patches available, and the remaining quarter isn't blocking any other 4.7 build failures. Busy people to the arms-architects and others, including Cyril Brulebois, Roger Steiner, Paul Pillay, for the fixes.

This will add one more transition for x86 (lib64of-? lib64of1), which needs extracting with updates of some software base packages.

The 3.9L transition is likely to be updated to 4.7, before the release, no build dependencies.

Best regards,

Matthias

CEOs with HR background
kerson@gmail.com - 08/05/2011, 5:13 am

CEOs with HR background
kiewow@gmail.com - 08/05/2011, 5:13 am

The Qualities of A Successful Leader
Tobias, Thane - 04/02/2011, 1:02 pm

Re: The Qualities of A Successful Leader
SSON - 07/12/2011, 6:51 am

Is it realistic or desirable to focus on individual leadership skills, or should we in fact be looking for different forms of leadership?
SSON - 07/12/2011, 6:07 am

Re: Is it realistic or desirable to focus on individual leadership skills, or should we in fact be looking for different forms of leadership?
SSON - 07/12/2011, 6:07 am

How Sharing Feelings Often Leads to Flame
meseup - 06/24/2011, 2:29 pm

Re: Work can make you happy
Chloef - 05/24/2011, 9:30 am



5th Webdatanet meeting, Mannheim, March 2013



2. Web content

Oss projects mailing list

Distribution of topics per year

	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7	Topic8	Topic9	Topic10	Topic11	Topic12	Topic13	Topic14	Topic15	Topic16	Topic17	Topic18	Topic19
1999	kernel	debian	Problem	struct	ram	build	time	king	machine	arm	file								
	linux	image	Ccc	instruction	system	start	time	coorig	compile	binary	server								
	gnu	ftp	Built	support	program	running	patch	version	main	write	standard								
	cross	install	error	status	install	script	glibc	fix	Documentation	basic	debian								
	compiler	stuff	Direct	memory	archive	trace	share	even	upload	value									
2000	kernel	install	Arms	image	coorig	build	fix	debian											
	problem	time	Machine	ftp	ram	gcc	patch	main	arm	binary									
	hosting	group	Linux	server	error	version	source	king	power	embedded									
	start	support	Compiled	system	build	software	gcc	king	glibc	power									
	shane	arm	Memory	files	off	direct	off	direct	glibc	power									
2001	kernel	build	debian	time	install	linux	compile	ram	arm	error	problem								
	king	arm	ftp	start	version	include	arm	source	files	report	machine								
	coorig	glibc	main	king	process	fix	patch	program	system	direct	gcc								
	support	cpu	open	write	status	stable	built	stuff	float	ports	gnu								
	image	default	basic	server	rmk	power	compiler	block	flash	struct	even								
2002	block	kernel	king	arm	ram	debian	build	system	machine	report	off	problem	linux	debian					
	server	coorig	patch	arm	fix	include	install	file	family	gnu	time	start	arm	main					
	flash	source	even	debian	program	running	stable	image	support	software	write	error	gcc	ftp					
	stuff	script	available	reference	project	device	target	user	king	number	open	compile	source						
	install	version	struct	traffic	module	process	upload	random	pas	system	memory	debug	cross	stable					
2003	stuff	gcc	debian	address	arm	support	script	arm	block	build	error	time	compile	problem	machine	image	arm	linux	kernel
	king	linux	return	include	main	cpu	report	built	start	running	memory	files	king	version	source	process	development	embedded	install
	software	error	flash	number	build	debug	subpod	patch	include	pas	function	console	compiler	nodechain	power	stable	project	main	ftp
	direct	arm	mailing	family	registered	struct	problem	driver	client	executable	main	failure	install	off	issue	gnu	Device	main	abi
2004	build	include	debian	install	problem	linux	sub	arm	off										
	arm	version	main	ram	built	arm	system	compile	support										
	upload	coorig	stable	machine	time	gcc	board	struct	king										
	available	error	kernel	patch	king	gnu	time	issue	even										
	machine	report	ftp	driver	fix	program	ram	instruction	family										
2005	gcc	install	system	debian	fix	machine	arm	time	kernel	arm	build	main	linux						
	compile	debian	ram	arm	struct	off	block	king	family	support	built	source	linux						
	patch	problem	board	report	software	access	stable	start	sub	user	upload	version	gnu						
	error	running	embedded	network	abi	available	issue	pas	image	high-dian	archive	process	coorig						
	cross	platform	files	cpu	result	hardware	flash	bp	script	flow	accept	open	direct						
2006	arm	kernel	debian	family	system	arm	build	ram	block	king	coorig	kernel							
	machine	partition	install	user	process	linux	problem	flash	abi	time	device	image							
	man	loader	toolchain	address	compile	source	upload	direct	arm	start	sub	support							
	ports	off	patch	archive	even	gnu	issue	driver	struct	script	support	version							
	target	randm	Developers	server	cross	gcc	fix	memory	float	available	sound	install							
2007	arm	error	gcc	patch	install	system	kernel	build	block	time	debian	config							
	abi	king	gnu	king	main	mount	fix	source	struct	off	issue	usb							
	ports	server	arm	glibc	ftp	running	linux	compile	partition	problem	report	device							
	applied	process	linux	debian	write	problem	user	ram	cpu	download	compiled	family							
	available	share	abi	support	problem	user	ram	cpu	download	version	archive	register							



5th Webdatanet meeting, Mannheim, March 2013



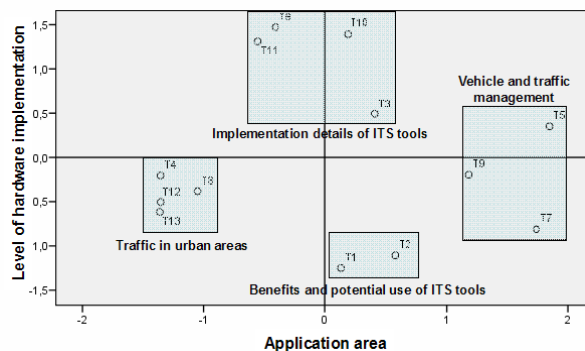
2. Web content

Research trends in ITS

The screenshot shows a search results page from Web of Knowledge. The search term is 'Intelligent Transportation Systems', yielding 1,600 results. The page lists several articles, with the first one highlighted: 'Schedule-driven Intersection control' by Xia, XF, Xiao-Feng, Smith, SF, et al. The abstract of this article is visible, discussing model-based intersection optimization strategies for distributed traffic signal control in road networks. The page also includes a 'Refine Results' sidebar with filters for databases, categories, and subject areas.

2. Web content

Research trends in ITS



2. Web content

Oss projects mailing list

debian-arm May 2012 by thread

[\[Previous thread\]](#)
 [\[Next thread\]](#)
 [\[Previous page\]](#)
 Page 1 of 1
 [\[Next page\]](#)
 [\[First page\]](#)
 [\[Last page\]](#)

[\[Date Index\]](#)
[\[Subject Index\]](#)
[\[Author Index\]](#)
[\[Other Debian Lists\]](#)
[\[Debian Home\]](#)

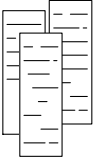
- [Re: arand 4.1.1.15 Register Bug:](#)
 - [Re: arand 4.1.1.15 Paul Shyne](#)
 - [Re: arand 4.1.1.15 open group](#)
- [GCC 4.7 in arm the default for ARM architectures Martin Kletz](#)
 - [Re: GCC 4.7 in arm the default for ARM architectures Thomas Gleixner](#)
 - [Re: GCC 4.7 in arm the default for ARM architectures Martin Kline](#)
- [Shreevatsa has Debian 5.0.3 pre-installed /: 616677](#)
 - [Re: Shreevatsa has Debian 5.0.3 pre-installed Mark Morgan-Lloyd](#)
 - [Re: Shreevatsa has Debian 5.0.3 pre-installed Ryan Lee](#)
 - [Re: Shreevatsa has Debian 5.0.3 pre-installed Ettore Watterhorn](#)
 - [Re: Shreevatsa has Debian 5.0.3 pre-installed Andrew H. C. Carter](#)
 - [Re: Shreevatsa has Debian 5.0.3 pre-installed Martin Alschlager](#)
 - [Re: Shreevatsa has Debian 5.0.3 pre-installed Lee Fisher](#)
 - [Re: Shreevatsa has Debian 5.0.3 pre-installed Jaf](#)
- [QNAP TS119 ported to problem /: 616678](#)
 - [Re: QNAP TS119 ported to problem Martin Michlmayr](#)
 - [Re: QNAP TS119 ported to problem Paul Oromas](#)
 - [Re: QNAP TS119 ported to problem Martin Michlmayr](#)
- [Earrinate or After](#)
 - [Re: Making an arand kernel package Paul Keroch](#)
- [Making an arand kernel package: Rp](#)
 - [Re: Making an arand kernel package: Tzafir Cohen](#)
 - [Re: Making an arand kernel package: Hector Oyon](#)
 - [Re: Making an arand kernel package: Paul Keroch](#)
- [Financemont per imposte a priori Zolt](#)
- [IPV6 on shreevatsa with new support kernel Rick Thomas](#)
 - [Re: IPV6 on shreevatsa with new support kernel Rp](#)
 - [Re: IPV6 on shreevatsa with new support kernel John Winters](#)
 - [Re: IPV6 on shreevatsa with new support kernel Steve McQuinn](#)
 - [Re: IPV6 on shreevatsa with new support kernel Paul Keroch](#)
 - [Re: IPV6 on shreevatsa with new support kernel Peter J. Roschberg](#)
 - [Re: IPV6 on shreevatsa with new support kernel Peter J. Roschberg](#)
- [Disable loading ipv6 on shreev NSU13 Alan Swigrove](#)
 - [Re: Disable loading ipv6 on shreev NSU13 John Winters](#)
 - [Re: Disable loading ipv6 on shreev NSU13 Alan Swigrove](#)
 - [Re: Disable loading ipv6 on shreev NSU13 Alan Swigrove](#)





5th Webdatanet meeting. Mannheim, March 2013

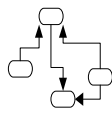


2. Web content





1. Information extraction






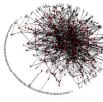
2. Nodes & Arcs







3. Structuring and processing information (SNA)






4. Social Network

5th Webdatanet meeting. Mannheim, March 2013

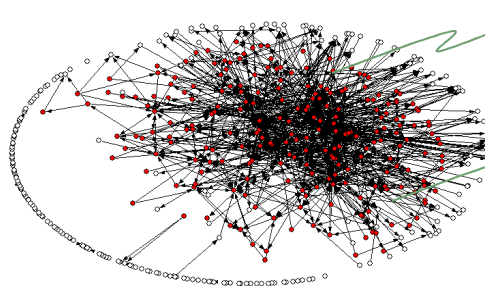


2. Web content

Oss projects mailing list

Vertex ⇒ Users and Developers

Arcs ⇒ Interactions in leads of discussions



Members:

- Alias – e-mail
- Analysis of the headings

Arcs:

- Valued graphs

Community of ARM Debian Linux (2007)



2. Web content

Dell IdeaStorm



13

2. Web content

Dell IdeaStorm

Vote Discuss

-bill_b

Ideas

Your Key to Action

118

cost WEBDATANET • 5th Webdatanet meeting. Mannheim, March 2013

14

2. Web content

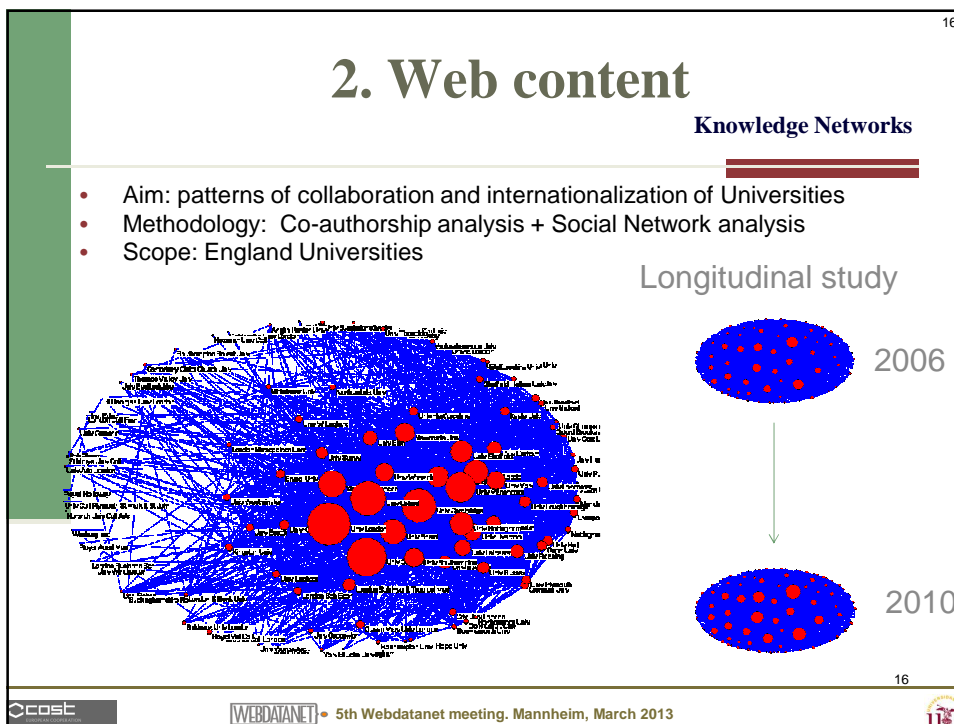
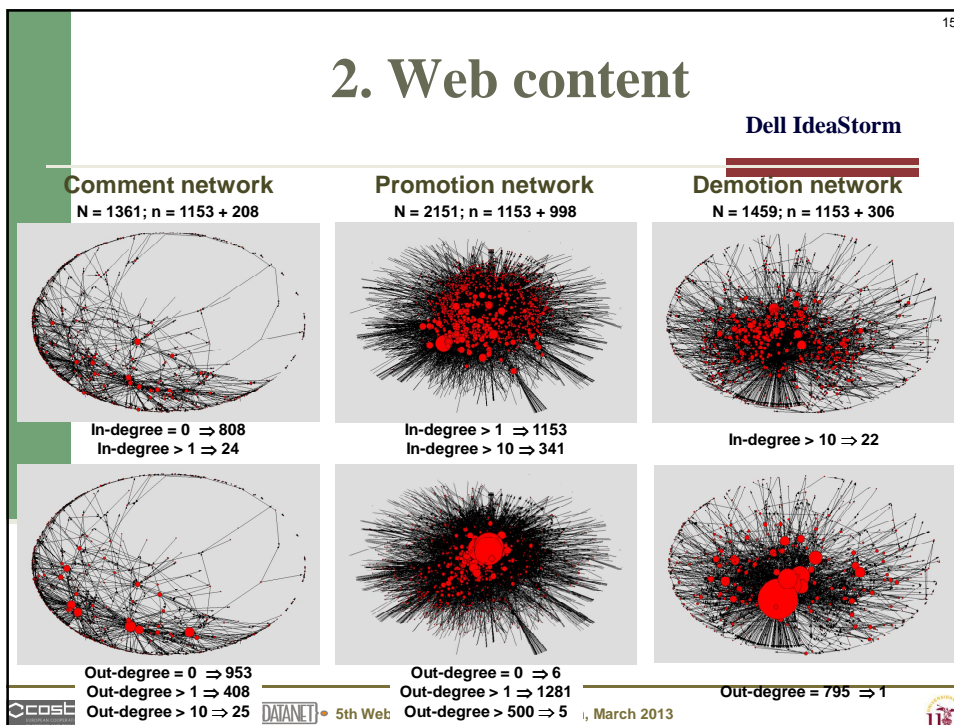
Dell IdeaStorm

Data extracted from IdeaStorm

• Idea name
• Author
• Date
• Comments <ul style="list-style-type: none"> ○ Number of comments ○ Authors who posted these comments
• Promotions <ul style="list-style-type: none"> ○ Number of received promotions ○ Authors who suggested promotions of the idea
• Demotions <ul style="list-style-type: none"> ○ Number of received demotions ○ Authors who suggested demotions of the idea

1482 ideas

cost WEBDATANET • 5th Webdatanet meeting. Mannheim, March 2013



3. Web structure

University website structure

	Total	Average	SD
Subdomains	2438	30,47	38,10
External domains	30500	381,25	580,32
Web pages	718272	8978,40	15334,01
Out-links	4429231	55365,38	73290,17

www.urjc.es www.ucam.edu

80 corporate Spanish Universities

www.webometrics.org

www.ucjc.edu www.fe.edu

www.deusto.es

www.udc.es

www.udima.es

www.urv.net

www.uma.es

www.uclm.es

www.uvic.es

www.ulpgc.es

www.umh.es

www.upc.edu

Spanish Universities web sites			
http://www.ucm.es/	http://portal.uned.es/	http://www.ual.es/	http://www.cef.es/
http://www.upc.edu/	http://www.uva.es/	http://www.uall.es/	http://www.uclm.edu.es/
http://www.upm.es/	http://www.upf.edu/	http://www.ujaen.es/	http://www.nebrija.com/
http://www.uab.es/	http://www.unav.es/	http://www.umh.es/	http://www.uic.es/
http://www.ehu.es/	http://www.uc3m.es/	http://www.deusto.es/	http://www.ur1.es/
http://www.ub.edu/	http://www.uniovi.es/	http://www.unavarra.es/	http://www.esdi.es/
http://www.us.es/	http://www.uma.es/	http://www.upct.es/	http://www.uax.es/
http://www.upv.es/	http://www.uco.es/	http://www.upo.es/	http://www.vives.org/
http://www.um.es/	http://www.uil.es/	http://www.ie.edu/	http://www.uimp.es/
http://www.ugr.es/	http://www.udc.es/	http://www.uncomillas.es/	http://www.ucjc.edu/
http://www.ua.es/	http://www.unex.es/	http://www.ceu.es/	https://www.ucv.es/
http://www.uvigo.es/	http://www.uah.es/	http://www.iese.edu/	http://www.uspcu.com/
http://www.uv.es/	http://www.uoc.edu/	http://www.ubu.es/	http://www.esdonbosco.com/
http://www.uam.es/	http://www.udg.edu/	http://www.urv.net/	http://www.ufv.es/
http://www.usal.es/	http://www.ulpgc.es/	http://www.unirioja.es/	http://www.esic.es/
http://www.uji.es/	http://www.unican.es/	http://www.uem.es/	http://www.cepade.es/
http://www.unizar.es/	http://www.unileon.es/	http://www.esade.edu/	http://www.eoi.es/portal/
http://www.usc.es/	http://www.urjc.es/	http://www.ucam.edu/	http://www.esmac.net/
http://www.uib.es/ca/	http://www.uca.es/	http://www.monragon.edu/	http://www.udima.es/
http://www.uclm.es/	http://www.uhu.es/	http://www.uv.es/	http://www.eupnt.es/



www.uab.es

www.uax.es

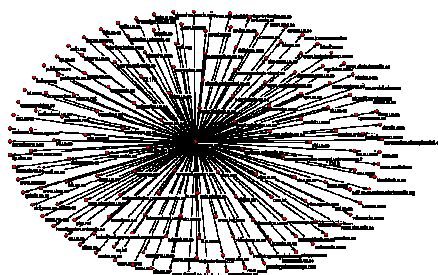
5th Webdatanet meeting. Mannheim, March 2013



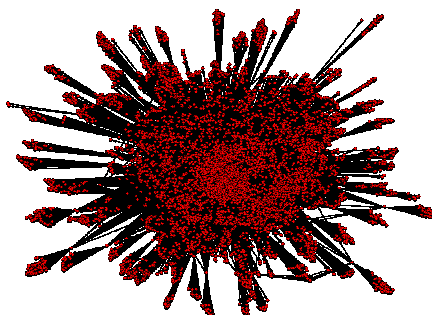
3. Web structure

University website structure

Representation of the University of Seville website



Domain network



Page network



5th Webdatanet meeting. Mannheim, March 2013





19


3. Web structure

University website structure

- Factor 1: Distributed structure
 - Information finding requires browsing through several web pages

5th Webdatanet meeting. Mannheim, March 2013





20


3. Web structure

University website structure

- Factor 2: Centralized structure
 - Accessibility of information

5th Webdatanet meeting. Mannheim, March 2013

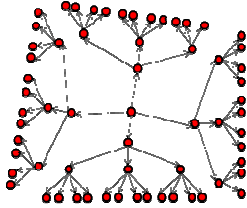


21




3. Web structure

University website structure

- Factor 3: Egocentric structure
 - Sum of more or less independent subnetworks



The network graph illustrates an egocentric structure, characterized by several distinct, interconnected subnetworks. Each subnetwork consists of a central node (red) connected to multiple peripheral nodes (black), forming a star-like pattern. These subnetworks are linked to each other, creating a larger, fragmented network.

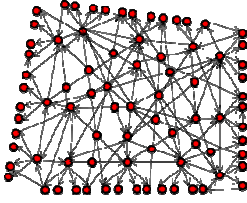
  • 5th Webdatanet meeting. Mannheim, March 2013 

22




3. Web structure

University website structure

- Factor 4: High sized websites
 - Long navigation process to achieve the information



The network graph illustrates a highly interconnected, dense network structure. It features a large number of nodes (red and black) connected by numerous edges, forming a complex, web-like structure. This represents high-sized websites with many internal links, leading to a long navigation process to reach specific information.

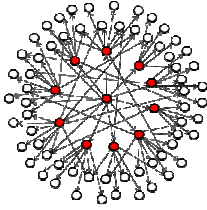
  • 5th Webdatanet meeting. Mannheim, March 2013 



23

3. Web structure


University website structure

- Factor 5: Small sized websites
 - Information provided through external references



5th Webdatanet meeting. Mannheim, March 2013

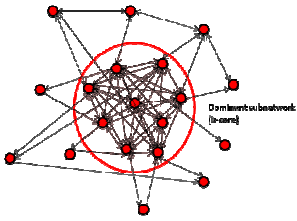




24

3. Web structure


University website structure

- Factor 6: Subnetwork structure
 - One subnetwork which contains the most relevant information



5th Webdatanet meeting. Mannheim, March 2013



DATA EXTRACTION, MODELING AND APPLICATIONS IN E-SOCIAL SCIENCES



M.R. Martínez Torres
University of Seville
Seville, Spain
rmtorres@us.es

