



AICA



UNIVERSITÀ DEGLI STUDI DI SALERNO



CONGRESSO NAZIONALE AICA 2013

Frontiere Digitali:
dal Digital Divide alla Smart Society



Big Data: strumenti e opportunità per affrontare la crisi

«Big Data: limitazioni e opportunità geopolitiche e geoeconomiche»

UNIVERSITÀ DEGLI STUDI DI SALERNO

CAMPUS DI FISCIANO

20 settembre 2013

<http://www.aica2013.it>

Carlo Muzzi
muzzi@acm.org

Abstract

Big Data: limitazioni e opportunità geopolitiche e geoeconomiche

Carlo Muzzi

AICA - Associazione Italiana per l'Informatica ed il Calcolo Automatico
muzzi@acm.org

Abstract. *The idea that Big Data can constitute a new frontier for innovation, competition and productivity in the global economy is now generally accepted; but there are many geo-political and geo-economic conditions that influence them in the global context. These conditions will be analyzed in this study, along with the opportunities offered.*

Keywords: Big Data, conditions, opportunities, geo-political, geo-economic.

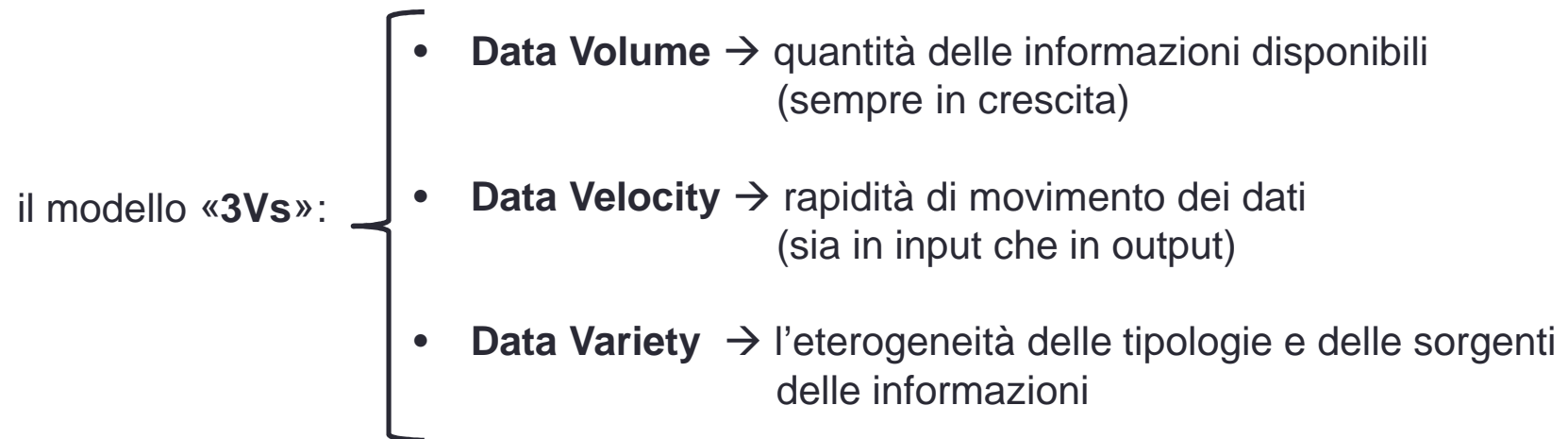
«Big Data»: alcuni concetti significativi

- Il termine Big Data [Manyika et al, 2011] è utilizzato per riferirsi a quegli insieme di dati dal volume talmente grande da non poter più essere catturati, conservati, gestiti e analizzati con gli strumenti tipici dei database tradizionali
- È una definizione soggettiva poiché, non utilizzando metriche precise per specificare quanti dati sono necessari per poterli definire “grandi”, preserva dal rischio di stabilire soglie che verrebbero certamente rese obsolete dal costante incremento delle informazioni digitali
- I Big Data sono già tra noi:
 - ❑ nella capacità della rete Internet di aggregare e rendere fruibili documenti, testi, indici di ricerca, immagini, filmati, ...
 - ❑ nell'accumulazione senza precedenti di dati personali derivanti dalla social data revolution [Weigend, 2009] innescata dalla diffusione dei social networks
 - ❑ dal continuo incremento delle informazioni digitalizzate su prodotti, merci, servizi, ... indotta dalla crescita del mercato dell'e-commerce
 - ❑ dalle raccolte di grandi masse di dati compiute da diverse discipline scientifiche (astronomia, chimica, biologia, meteorologia, genomica, ...)
 - ❑ dai nuovi dati creati dalla sempre più pervasiva diffusione delle reti di sensori (RFID, Internet of Things)

«Big Data»: alcuni concetti significativi

- L'equazione “Big Data = Grandi Quantità di Dati” è fuorviante → quando ci si riferisce ai Big Data si pone più attenzione alla complessità [Rindler, 2009] piuttosto che alla quantità dei dati da trattare

- Le 3 dimensioni dei Big Data:



Leve di analisi per i Big Data

Alcune leve possono influenzare significativamente i Big Data agendo come vettori di stimolo o di freno. Esistono limitazioni e condizioni geopolitiche e geoeconomiche che possono indurre:

- sia impatti negativi



- sia opportunità



Leve

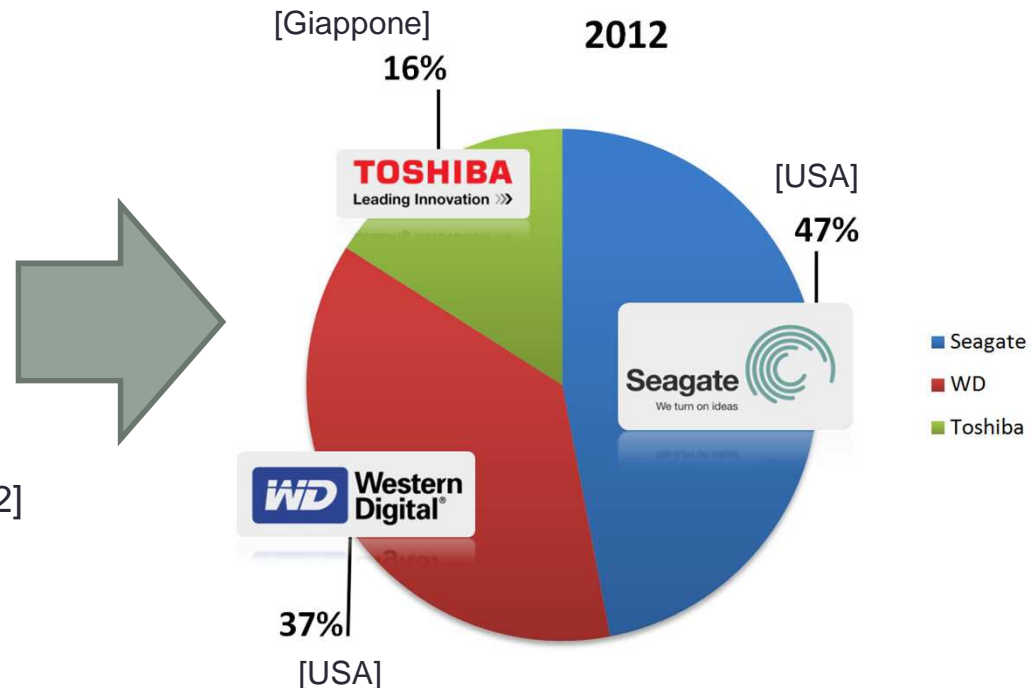
- **Storage**
 - **Internet**
 - **I grandi accumulatori di informazioni**
 - **I giganti del calcolo**
 - **Le influenze dei contesti giuridici**

Storage

Non si possono conservare informazioni senza disporre degli opportuni strumenti tecnologici di memorizzazione: strumenti che variano in funzione dei soggetti che accumulano i Big Data (individui, organizzazioni profit e non profit, governi e altri organismi istituzionali, ...) e della loro capacità di investimento.

- **Mercato degli hard disk:** elemento atomico per la costruzione dei vari livelli di storage tradizionale (SAN, NAS, DAS, ecc.)

Fonte: [Taylor, 2012]

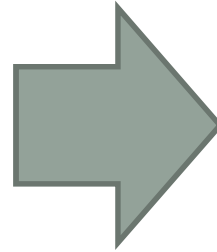


NOTA: nelle varie slides sono indicati gli stati in cui ha sede l'Head Quarter delle organizzazioni citate

Storage

- **storage USB** (prevalente mondo consumer)

Fonte: [Santa Clara Consulting Group
Executive Summary, 2012]

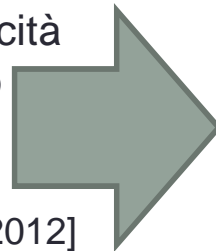


SanDisk (USA)
Kingston (USA)
Transcend (Taiwan)
Verbatim (Giappone)
HP (USA)
Sony (Giappone)
Buffalo (Giappone)
Imation (USA)
PNY (USA)

circa il 71%
del mercato
globale

- **storage cloud** (comparazione di 10 soluzioni rispetto a 6 capacità ritenute critiche per il mercato corporate o professionale)

Fonte: [Ruth e Chandrasekaran, 2012]



Amazon S3 (USA)
AT&T Synaptic Storage as a Service (USA)
Google Cloud Storage (USA)
HP Cloud Object Storage (USA)
IBM SmartCloud Enterprise Object Storage (USA)
Internap AgileFILES (USA)
Microsoft Windows Azure Blob Storage (USA)
Nirvanix Public Cloud Storage (USA)
Rackspace Cloud Files (USA)
SoftLayer Object Storage (USA)

Storage



Opportunità per l'Italia:

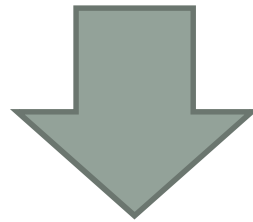
- mercato dello storage USB → dove l'aspetto del design, così tipico del nostro paese, permetterebbe anche alle FabLab di operare in settori di nicchia
- servizi ad alto valore aggiunto nello storage cloud → soluzioni appetibili anche da chi ha minori capacità di investimento e competenze ICT (le piccole imprese italiane).

Internet

Il modello “3Vs” ha enfatizzato nella velocità di trasferimento dei dati uno dei fattori qualificanti per la diffusione dei Big Data

→ oggi ciò avviene attraverso internet

→ quindi l'ampia disponibilità di connettività pubblica adeguata, veloce e pervasiva, diviene un elemento infrastrutturale essenziale per il progresso dei Big Data.



Recenti studi hanno evidenziato come, pur in presenza di sensibili divari tra i diversi continenti, la rete è oggi utilizzata dal 34,3% della popolazione mondiale (stimata in poco più di 7 miliardi di individui)

Fonte: [Internet World Stats, 2012]

Internet

Anche l'accesso veloce alla rete è una realtà per una significativa parte dei 2,4 miliardi di individui già oggi connessi.

Analisi effettuate sui 34 stati aderenti all'OCSE evidenziano l'esistenza di:

- 321.317.973 accessi sottoscritti alla banda larga su rete fissa
- 698.656.918 su rete wireless

Rank	Country	Fixed	Rank	Country	Wireless
34	Iceland	109.359	34	Iceland	206.359
33	Luxembourg	166.630	33	Luxembourg	334.800
32	Estonia	334.655	32	Slovenia	566.793
31	Slovenia	495.953	31	Estonia	850.601
30	Slovak Republic	771.638	30	Hungary	1.342.196
29	Ireland	1.011.730	29	Slovak Republic	1.972.551
28	New Zealand	1.239.569	28	Belgium	2.844.187
27	Finland	1.600.400	27	Ireland	2.845.357
26	Czech Republic	1.715.700	26	New Zealand	2.942.301
25	Norway	1.789.279	25	Israel	3.831.000
24	Israel	1.910.000	24	Chile	3.866.878
23	Austria	2.063.612	23	Portugal	3.960.947
22	Chile	2.101.524	22	Norway	3.987.776
21	Hungary	2.114.832	21	Switzerland	4.245.550
20	Denmark	2.131.167	20	Austria	4.334.468
19	Portugal	2.309.702	19	Greece	4.480.265
18	Greece	2.558.830	18	Denmark	5.125.855
17	Sweden	3.053.000	17	Finland	5.160.000
16	Switzerland	3.276.700	16	Czech Republic	6.134.301
15	Belgium	3.586.116	15	Sweden	9.616.000
14	Australia	5.601.000	14	Netherlands	9.760.000
13	Poland	5.602.280	13	Turkey	10.661.171
12	Netherlands	6.573.000	12	Mexico	10.725.623
11	Turkey	7.666.609	11	Canada	14.279.850
10	Canada	11.013.711	10	Italy	20.991.956
9	Spain	11.384.588	9	Poland	22.049.733
8	Mexico	12.108.478	8	Australia	22.178.000
7	Italy	13.444.129	7	Spain	22.277.983
6	Korea	18.010.256	6	France	30.562.000
5	United Kingdom	21.066.624	5	Germany	31.666.400
4	France	23.167.000	4	United Kingdom	37.629.000
3	Germany	27.677.770	3	Korea	51.894.381
2	Japan	35.142.132	2	Japan	107.508.636
1	United States	88.520.000	1	United States	237.824.000
	OECD	321.317.973		OECD	698.656.918

Fonti: [OECD Broadband statistics, 2012a] + [OECD Broadband statistics, 2012a]

Internet

Ovviamente i dati vanno analizzati nel contesto specifico:

- In rapporto al numero di abitanti → l'Islanda è in vetta anche se in fondo alla classifica !
- L'utente medio africano ha una banda media di 2.000 bps contro i 90.000 bps del corrispondente utente europeo [ICT Data and Statistics Division, 2011] → ma l'esigenza di diffondere le tecnologie della comunicazione ha creato in Africa il boom del mercato mobile (connettività di tipo Broadband Wireless Access o BWA)
- La diffusione del modello Ajax [Garret, 2005] per lo sviluppo di applicazioni web interattive ha ridotto sensibilmente i flussi dati http scambiati tra i nodi client e server in una comunicazione internet
- Piano Nazionale Banda Larga → Dal 2008 lo stato italiano ha lanciato il Piano Nazionale Banda Larga con l'obiettivo di offrire ampia connettività anche agli 8 milioni di cittadini residenti in aree ancora scoperte dalle normale offerta del mercato commerciale.

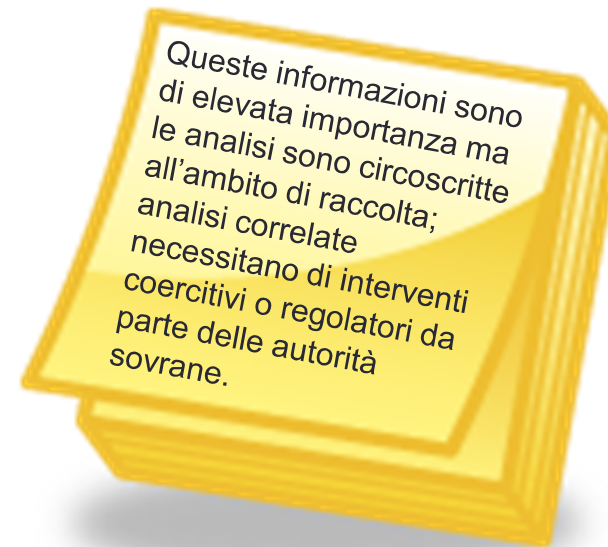


Gli ultimi investimenti definiti [Piano Nazionale Banda Larga, 2013] prevedono il raggiungimento di tale obiettivo nel breve periodo (anche se la crisi non è condizione favorevole) → è verosimile che nuove opportunità per i Big Data potranno aprirsi anche in Italia nel momento in cui internet diverrà effettivamente un servizio universale per tutta la popolazione

I grandi accumulatori di informazioni

Chi accumula grandi quantità di informazioni è un soggetto naturalmente avvantaggiato nel mondo dei Big Data:

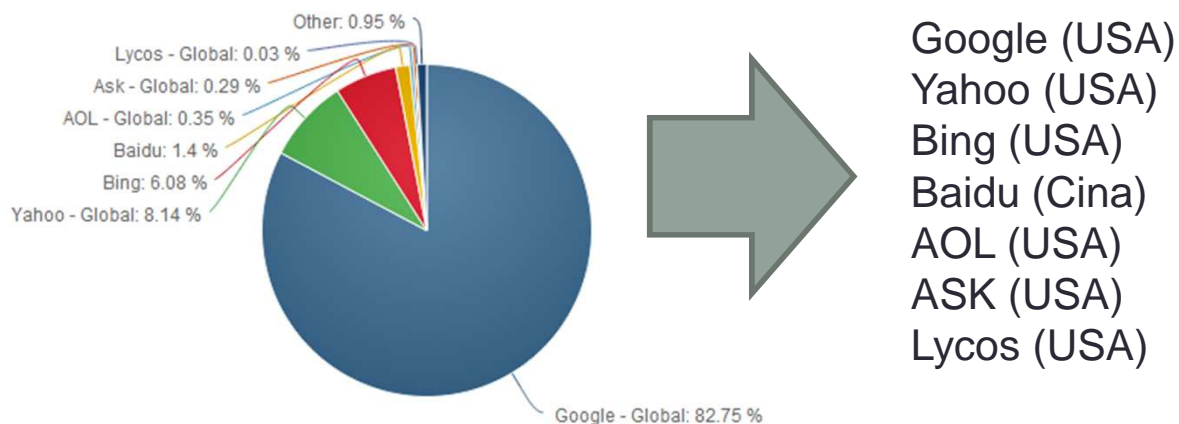
- le banche dati anagrafiche, fiscali, assistenziali, sanitarie, bancarie, assicurative, militari, ambientali, scientifiche, ... dei soggetti istituzionali



- i grandi operatori di internet → dove i dati divengono pubblici talvolta senza neanche la consapevolezza dei legittimi titolari

I grandi accumulatori di informazioni

- **Motori di ricerca**



Fonte: [Market Share Statistics for Internet Technologies, 2013]

- **Social network**
prime 10 posizioni

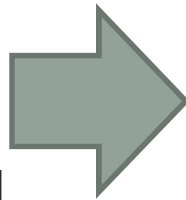
Fonte: [FindTheBest, 2013]

-
1. Facebook (USA) 960.000.000
 2. Twitter (USA), 170.000.000
 3. Tumblr (USA), 120.000.000
 4. LinkedIn (USA), 97.000.000
 5. Mobikade (Gran Bretagna), 70.200.000
 6. Multiply (USA), 56.000.000
 7. Badoo (Gran Bretagna) 56.000.000
 8. Flickr (USA), 51.000.000
 9. Orkut (USA), 45.000.000
 10. MySpace (USA), 31.000.000

I grandi accumulatori di informazioni

- **Siti internet**
i primi 20 siti più visitati al mondo

Fonte: [FindTheBest, 2013]



1. Facebook (USA)
2. Google (USA)
3. YouTube (USA)
4. Yahoo! (USA)
5. Amazon (USA)
6. Baidu (Cina)
7. Wikipedia (USA)
8. Windows Live (USA)
9. QQ.COM (Cina)
10. Taobao (Cina),
11. Google India (USA)
12. Twitter (USA)
13. Blogspot (USA)
14. Yahoo! Japan (USA)
15. LinkedIn (USA)
16. Bing (USA)
17. 新浪新闻中心/ sina.com.cn (Cina)
18. Яндекс yandex.ru (Russia)
19. ВКонтакте/vk.com (Russia)
20. Ask (USA)



I giganti del calcolo

I supercomputer sono uno degli strumenti coi quali disporre di grandi capacità di calcolo per trattare i Big Data.

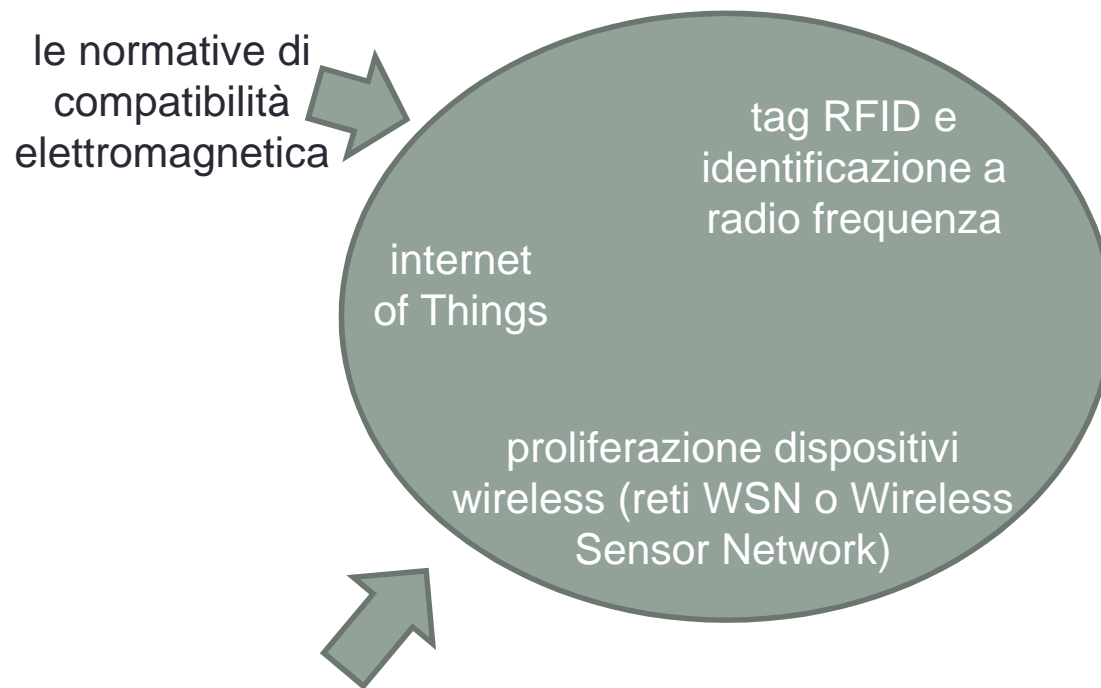
Stati / Continenti	Africa	Americhe	Asia	Europa	Oceania	Totale
USA		252				252
Cina			66			66
Giappone			30			30
Gran Bretagna				29		29
Francia				23		23
Germania				19		19
India			11			11
Canada		9				9
Russia				8		8
Svezia				7		7
Italia				6		6
Australia					5	5
Corea del Sud			4			4
Arabia Saudita			4			4
Svizzera				4		4
Brasile		3				3
Norvegia				3		3
Polonia				3		3
Spagna				3		3
Finlandia				2		2
Israele			2			2
Olanda				2		2
Austria				1		1
Belgio				1		1
Danimarca				1		1
Hong Kong			1			1
Taiwan			1			1
Totale	0	264	119	112	5	500

Fonte: [TOP500 Authors, 2013]

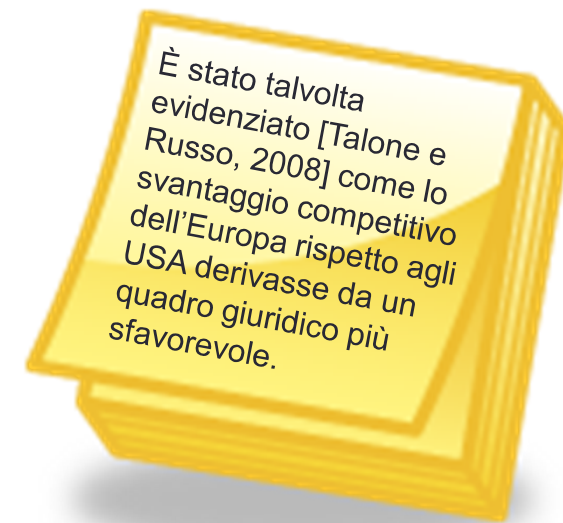
- il 50% sono negli USA
- la Cina, che è il secondo, ne ha solo ¼ degli USA
- la Cina ha quello più potente (il *Tianhe-2 MilkyWay-2*)
- l'Asia cresce
- l'Africa me è priva
- il primo italiano occupa il 12° posto (il *Fermi* del consorzio interuniversitario Cineca)

Le influenze dei contesti giuridici

Gli operatori dei Big Data sono fortemente influenzati dai contesti giuridici all'interno dei quali operano; talvolta tale influenza condiziona in misura ben maggiore che aspetti tecnologici o infrastrutturali.



l'attenzione crescente al tema controverso dell'elettrosensibilità [Knowledge note, 2012] e sui sintomi fisici e/o psicologici che sembrano indursi verso taluni soggetti esposti



Le influenze dei contesti giuridici



Conclusioni

L'esame delle leve sopra esposte evidenzia alcuni aspetti:

- la grande importanza strategica del tema dei Big Data
- la prevalente superiorità degli USA a livello mondiale e, in subordine, una significativa posizione degli attori asiatici
- i paesi europei subiscono più competere in questo mercato: quindi dovrebbero agire congiuntamente con politiche ed azioni a livello comunitario



the guardian
theguardian.com

Exclusive: how US and Britain unlock privacy on the internet

- Elaborate safeguards broken by NSA and GCHQ
- Encryption meant to protect emails, bank and medical records
- New Snowden revelations certain to cause political row



L'Italia dovrebbe promuovendo una maggiore sinergia tra:

- i distretti industriali (la tipica anima produttiva del nostro paese)
- le università e i centri di ricerca di cui ancora si dispone
- l'attore pubblico (governo e regioni) per sfruttare i vantaggi competitivi dei Big Data

Bibliografia

[Manyika et al, 2011] Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Hung Byers A., 2011, Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute.

[Weigend, 2009] Weigend A., 2009, The Social Data Revolution(s), Harvard Business Review.

[Rindler, 2009] Rindler A., 2012, Big Data Definition, MIKE2.0 The open source standard for Information Management.

[Taylor, 2012] Taylor P., 2012, HDD business to become Mexican standoff, TechEYE.net.

[Santa Clara Consulting Group Executive Summary, 2012] Santa Clara Consulting Group, Executive Summary–Q4/Year-End 2012, 2012, USB Flash Tracker™ - Trends for USB Flash Drive Markets, Santa Clara Consulting Group (Vol. 9 No. 4).

[Ruth e Chandrasekaran, 2012] Ruth G., Chandrasekaran A., 2012, Critical Capabilities for Public Cloud Storage Services, Gartner.

[Internet World Stats, 2012] Internet World Stats, 2012, Internet users on June 30, 2012, Miniwatts Marketing Group.

[OECD Broadband statistics, 2012] OECD Broadband statistics, 2012, Total number of fixed (wired)/wireless broadband subscriptions, by country, millions, June 2012, Organisation for Economic Co-operation and Development (OECD/OCSE).

[ICT Data and Statistics Division, 2011] ICT Data and Statistics Division- Telecommunication Development Bureau, 2011, The World in 2011 — ICT Facts and Figures, International Telecommunication Union.

Bibliografia

[Garret, 2005] Garret J.J., 2005, Ajax: A New Approach to Web Applications, Adaptive Path.

[Piano Nazionale Banda Larga, 2013] Piano Nazionale Banda Larga, 2013, Al via bandi per 900 mln di euro per azzerare digital divide e banda ultralarga, Agenda Digitale Italiana.

[Taylor, 2012] Taylor P., 2012, HDD business to become Mexican standoff, TechEYE.net.

[Market Share Statistics for Internet Technologies, 2013] Market Share Statistics for Internet Technologies, 2013, Device Types Search Engine Market Share, Netmarketshare.

[FindTheBest, 2013] FindTheBest unbiased, data-driven comparisons, 2013, Compare social networking, monthly visitors, FindTheBest.

[TOP500 Authors, 2013] TOP500 Authors, 2013, Top500 List - June 2013, TOP500 supecomputer sites.

[Knowledge note, 2012] Knowledge note 2012-001, 2012, Electrosensitivity. Health complaints near sources of electromagnetic fields, Dutch Knowledge Platform on Electromagnetic Fields and Health.

[Talone e Russo, 2008] Talone P., Russo G., 2008, RFId, fondamenti di una tecnologia silenziosamente pervasiva, Fondazione Ugo Bordoni.

[Mantelero, 2012] Mantelero A., 2012, Big Data: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo, Il diritto dell'informazione e dell'informatica (Fasc. 1, 2012).