



Istituto di Ricerche sulla Popolazione
e le Politiche Sociali - CNR

IRPPS Working Papers

ISSN 2240-7332

www.irpps.cnr.it: visite nel triennio 2005-2007 e osservazioni recenti

Maria Girolama Caruso,
Loredana Cerbara

What is IRPPS?

IRPPS is an Interdisciplinary Research Institute that conducts studies on demographic and migration issues, welfare systems and social policies, on policies regarding science, technology and higher education, on the relations between science and society, as well as on the creation of, access to and dissemination of knowledge and information technology.

www.irpps.cnr.it

IRPPS WPs n. 23 (2008)

www.irpps.cnr.it: visite nel triennio 2005-2007 e osservazioni recenti

Maria Girolama Caruso, Loredana Cerbara

Abstract

The IRPPS's processing and databases, among other tasks, has to design and realize the web pages making them effective from a communication to a varied audience point of view. Indeed, the potential users of the site can be both specialists in the Institute's fields, who may seek scientific content (e.g. research outputs), or just curious students, who approach the Institute issues often without a specific skill. In that sense, an observation site to study the visits to web pages for monitoring the use of published information, has been created since 2005. Furthermore, an observatory report is periodically published with on-line technical reports, a publication type historically performed by the Institute. As to the complexity of the site itself, the report presented here is made up of two parts: the first relating to a comprehensive and historical analysis of visits, the second more specific and focused on the so-called 'mini sites' that are related to specific projects. Some of these are very important in order to know the traffic generated as well as possible actions for improvement of the pages published again for the purpose of communication through the web.

Keywords: Website, Network, Statistics, Users, Visit, Visitor, File log, Focus

Riassunto

Il servizio elaborazioni e banche dati dell'IRPPS ha, tra gli altri compiti, quello di progettare e realizzare le pagine web dell'Istituto in modo che esse risultino efficaci da un punto di vista della comunicazione rivolta ad un pubblico variegato. Infatti, i possibili fruitori del sito possono essere sia specialisti della materia, che potrebbero cercare contenuti scientifici (ad esempio output della ricerca), sia semplici curiosi oppure studenti, che si avvicinano alle tematiche trattate in Istituto spesso senza competenze specifiche. In tal senso, è stato istituito, fin dal 2005, un osservatorio delle visite alle pagine web per il monitoraggio continuo dell'utilizzo delle informazioni pubblicate; periodicamente viene allora elaborato un resoconto dell'attività dell'osservatorio pubblicato attraverso rapporti tecnici on-line, una tipologia editoriale storicamente presente in Istituto. Per la complessità del sito stesso, il rapporto qui presentato si compone di due parti: la prima, relativa ad una analisi complessiva e storica delle visite, la seconda più specifica ed incentrata sui cosiddetti 'mini siti' che sono relativi a specifici progetti, per alcuni dei quali è molto importante conoscere il traffico generato oltre che le possibili azioni di miglioramento delle pagine pubblicate sempre ai fini dell'efficacia della comunicazione effettuata via web.

Parole chiave: Sito web, Rete, Statistiche, Utenti, Visite, Visitatori, File log, Focus

Citazione consigliata:

Caruso, Maria Girolama, Cerbara, Loredana. www.irpps.cnr.it: visite nel triennio 2005-2007 e osservazioni recenti. *IRPPS Working Papers*, n. 23, 2008.

Marisa Girolama Caruso è tecnologo presso l'Istituto di Ricerche sulla Popolazione e le Politiche Sociali del CNR (e-mail: mariagirolama.caruso@irpps.cnr.it).

Loredana Cerbara è ricercatrice presso l'Istituto di Ricerche sulla Popolazione e le Politiche Sociali del CNR (e-mail: loredana.cerbara@irpps.cnr.it).



Istituto di Ricerche sulla Popolazione e le Politiche Sociali - CNR
Via Palestro, 32 - 00185 Roma
<http://www.irpps.cnr.it/it>

Indice

1. Introduzione	1
2. Statistiche generali	1
2.1 Confronto su dati di sintesi tra intervalli temporali	1
2.2 Profili temporali	3
3. Statistiche di accesso	6
4. Utenti visitatori	10
5. Motori e frasi di ricerca	12
6. Focus su alcune aree del sito	14
Bibliografia	20

1. Introduzione

Ormai da diversi anni il 'servizio elaborazioni e banche dati dell'IRPPS' dedica una parte delle sue attività all'analisi degli accessi al sito dell'Istituto di Ricerche sulla Popolazione e le Politiche Sociali attraverso l'elaborazione statistica dei file di log (file che tengono traccia delle visite alle pagine web del sito) contenuti nel server e messi a disposizione dal Servizio Reti del CNR. L'intento è quello di monitorare periodicamente le visite alle pagine pubblicate in rete sia per comprenderne l'utilizzo reale, sia per poter formulare ipotesi circa le caratteristiche dell'utenza finale e le richieste di informazione dall'esterno. Come conseguenza di tali attività, viene monitorata nel tempo anche l'efficacia comunicativa del sito, che, come è ovvio, è uno strumento molto potente e particolarmente adatto alla comunicazione della scienza delle fasce di popolazione più giovanili.

Una prima analisi riguarderà il confronto tra le statistiche in una serie di intervalli temporali a partire dal 2003 al 2007, seguirà un'analisi di approfondimento di alcuni temi focalizzata solo sugli ultimi tre anni (dal 21/12/2004 al 23/12/2007) in quanto molte considerazioni sul primo anno sono già state fatte nel precedente lavoro (Working Paper 02/04). Infine, saranno riportate analisi ad hoc fatte su alcuni mini siti legati ad eventi speciali; analisi che chiameremo *focus* per distinguerle dalle altre.

2. Statistiche generali

2.1 Confronto su dati di sintesi tra intervalli temporali

La tabella 1 esprime in modo massimamente sintetico l'informazione sulle visite al sito web dell'IRPPS che qui si concentra su un arco temporale di 4 anni a partire dal 22/05/2003 al 23/12/2007 per raccordare le analisi del presente rapporto con quelle del rapporto precedente (Working Paper 02/2004). Le statistiche del primo anno qui considerato (dal 22/05/2003 al 16/06/2004) si riferiscono alla versione precedente del sito che aveva una progettazione completamente diversa dall'attuale versione che è in linea dalla fine del 2004. Sono proprio le strutture di base delle due versioni del sito ad essere completamente differenti più da un punto di vista tecnico, che di contenuto: la vecchia versione del sito era stata progettata a frame¹ e, di fatto, accedendo alla home page (cioè la prima pagina del sito, quella da cui si può navigare in tutto il resto delle pagine pubblicate e a cui si accede scrivendo l'indirizzo base nel programma di navigazione) si visitavano ben 3 pagine, quelle di cui, appunto, era composta la home page; l'ultima versione, invece, è stata ideata a pagina singola, con la conseguente e del tutto fittizia, riduzione del numero di pagine visitate.

Ciononostante, nell'arco temporale dei quattro anni qui considerati, le pagine visitate si sono più che raddoppiate passando da 179.222 nel primo anno a 318.874 nel quarto anno anche se in quest'ultimo si evidenzia una diminuzione di pagine visitate rispetto all'anno precedente (Grafico 1). Anche la stessa media giornaliera si è più che raddoppiata passando da 457 accessi nel primo anno a 871 nel quarto anno, anch'essa, coerentemente con quanto è accaduto per le pagine visitate, ha subito un decremento nell'ultimo anno rispetto all'anno precedente. Questo dato induce una prima riflessione sul fatto che lo strumento del web, viaggiando ad una velocità senza pari nel mondo della comunicazione, presenta allo stesso modo una tendenza all'invecchiamento superiore ad ogni altro *media*

¹ Un frame è una suddivisione di una pagina htm, cioè scritta in un linguaggio di programmazione adatto alla pubblicazione sul web, in più riquadri, in ognuno dei quali è possibile far apparire un documento diverso.

ed abbisogna, quindi, di aggiornamenti nei contenuti ed adeguamenti tecnici continui. Ma su questo tema torneremo anche in seguito, al momento di discutere delle possibili evoluzioni del nostro sito.

Ogni visitatore ha visto in media quasi quattro pagine nel primo anno; questo valore decresce negli anni successivi, fino ad arrivare nell'ultimo anno a 1,25. Questo dato però, non va letto in senso assoluto ma va considerato innanzi tutto che nel primo anno la struttura del sito era organizzata a frame (una pagina era composta da tre pagine), e poi che l'informazione presente in ciascuna pagina è andata sensibilmente aumentando nel tempo. Possiamo dire che dunque che la media è aumentata nel corso degli anni, ad accezione del quarto anno in cui il suo valore risulta essere equiparabile al valore "reale" (cioè calcolato come se le pagine non fossero a frame e dunque dividendo il numero di pagine visitate per il numero di frame) del primo anno.

Sempre nell'ultimo anno si sono più che quintuplicati i visitatori diversi, passando da 46.298 nel primo anno a 254.360 nell'ultimo anno, anche la stessa media giornaliera del numero di visitatori è cresciuta nel corso degli anni, passando da 118 nel primo anno a 694 nell'ultimo anno. Anche gli indirizzi IP², hanno mantenuto una crescita costante nel corso degli anni, passando da 22.558 nel primo anno a 73.001 nell'ultimo anno.

Tabella 1 – Quanti visitatori? Dettaglio di alcuni periodi tra il 2003 e il 2007

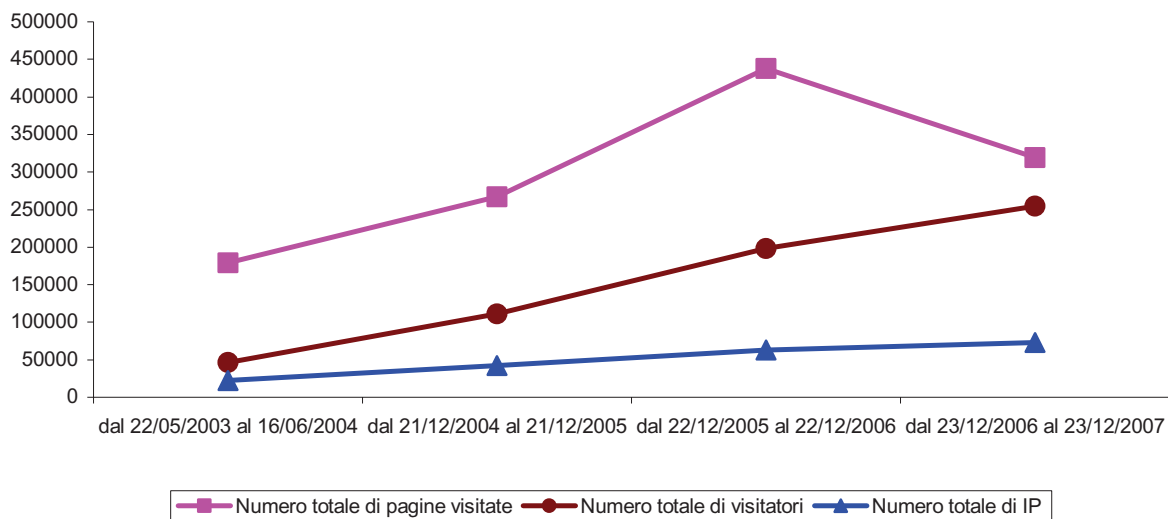
	dal 22/05/2003 al 16/06/2004	dal 21/12/2004 al 21/12/2005	dal 22/12/2005 al 22/12/2006	dal 23/12/2006 al 23/12/2007
Pagine visitate				
Numero totale di pagine visitate	179.222	266.874	438.079	318.874
Numero medio di pagine visitate al giorno	457	729	1.196	871
Numero medio di pagine visitate per visitatore	3,87	2,39	2,21	1,25
Utenti visitatori				
Numero totale di visitatori	46.298	111.511	198.520	254.360
Numero medio di visitatori al giorno	118	304	542	694
Numero totale di IP	22.558	42.283	62.740	73.001

È dunque soprattutto la diminuzione del numero di pagine visitate, mostrata nella tabella 1, nell'ultimo anno di osservazione a preoccuparci, benché il numero di visitatori diversi sia comunque in costante aumento; questo dato può rappresentare l'indicazione di un calo di interesse dovuto al fatto che molte delle le pagine pubblicate sono presenti sul sito da molto tempo. Ciò può determinare un fenomeno legato all'abitudine che i visitatori possono aver adottato di visitare periodicamente solo le pagine di maggior interesse per rilevare esclusivamente le novità. Questo, che di per se non è un fatto negativo, dall'altra parte può generare a lungo andare un certo disinteresse per il sito, soprattutto in presenza di una

² IP: Internet Protocol, un numero a 32 bit composto da 4 cifre comprese tra 0 e 255 separate da punti. È fornito dal proprio amministratore di rete o dal provider dei servizi di rete. Si tratta di un particolare numero che ci viene assegnato ogni qual volta ci connettiamo in rete e che permette molto spesso di identificare alcune caratteristiche dell'utente visitatore.

cronica mancanza di aggiornamenti continui in alcune delle aree (ad esempio quelle statiche o quelle di alcune aree della ricerca non aggiornate di recente).

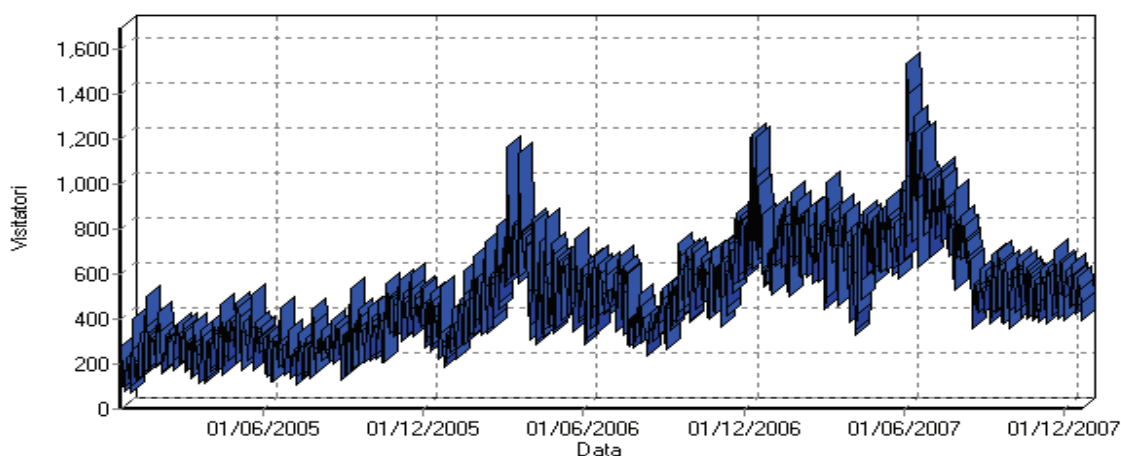
Grafico 1 - Differenze numeriche tra intervalli temporali



2.2 Profili temporali

I dati sulle visite giornaliere del triennio 2005-2007, così come anche i dati sull'attività mensile, ci forniscono informazioni interessanti per capire se in tutto il periodo di osservazione hanno agito dei fattori di stagionalità, o periodi in cui l'attività del web è stata particolarmente rilevante.

Grafico 2 – Panorama delle visite giornaliere tra il 2005 e il 2007

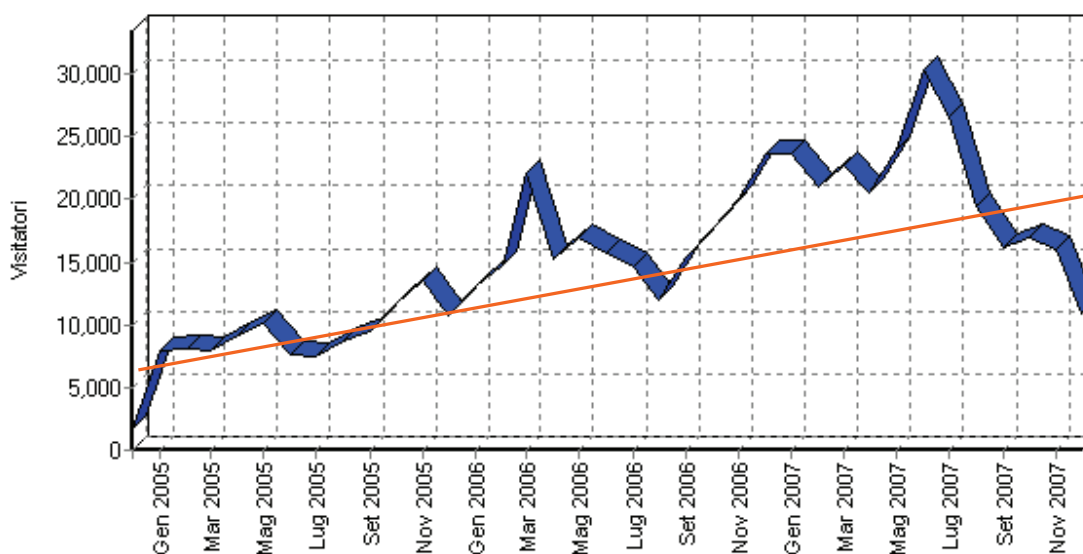


Dal grafico 2 si evidenzia un crescente aumento di visitatori fino alla metà dell'anno del 2007, il cui numero poi decresce nella seconda metà dello stesso dell'anno. In tutto l'arco del tempo di osservazione si evidenziano picchi in particolari periodi, marzo 2006, dicembre 2006 e giugno 2007, con buona probabilità dovuti alle visite della sezione "eventi e notiziario", che è anche quella più dinamica e in continuo aggiornamento, dunque di maggiore attrattiva per i visitatori.

Un'analisi più approfondita, in quanto il dato mensile è maggiormente leggibile, si può fare osservando il grafico 3. Possiamo individuare esattamente i mesi in cui alcune pagine web, senza altro legate ad eventi particolari e a noi, ovviamente, ben noti, sono state intensamente visitate. Gli avvenimenti che hanno visto il nostro Istituto coinvolto in questi particolari momenti sono stati: l'uscita in rete del "1st International Workshop on Multimodal And Pervasive Services" nel marzo 2006; una tavola rotonda "Il nuovo mercato del lavoro scientifico: ricercatori più flessibili o più precari?", "I risultati dell'inchiesta IRPPS - Ambrosetti - The European House sulla qualità della vita degli anziani" nel dicembre 2006 e la presentazione di DIVA (Science in a Different Voice) un progetto finanziato dalla Commissione Europea nel giugno 2007.

Disegnando una linea che interpola questa curva, possiamo percepire il fatto che i visitatori nel tempo sono aumentati anche se gli ultimi dati qui considerati sono al livello della fine del 2005. Si tratta cioè di un fenomeno che può essere descritto come attività residuale indipendente dagli eventi speciali che determinano i picchi di visitatori. In altre parole, in corrispondenza di un evento aumentano molto i visitatori diversi, ma, dopo tale evento, l'attività delle visite non torna al punto di partenza, ma rimane a livelli leggermente superiori rispetto al passato. Quindi, oltre al fatto che durante gli eventi speciali il sito è utilizzato in modo efficiente (efficienza testimoniata dai particolari picchi di visite) per la comunicazione in quelle particolari occasioni, esse veicolano la conoscenza delle nostre pagine anche a nuovi visitatori che rimangono fedeli ancora per qualche tempo. È nostro compito garantire loro una buona attrattiva del sito (attraverso la manutenzione tecnica e l'aggiornamento continuo delle pagine) perché essi rimangano a lungo tra i nostri visitatori usuali.

Grafico 3 – I visitatori di ogni mese tra il 2005 e il 2007

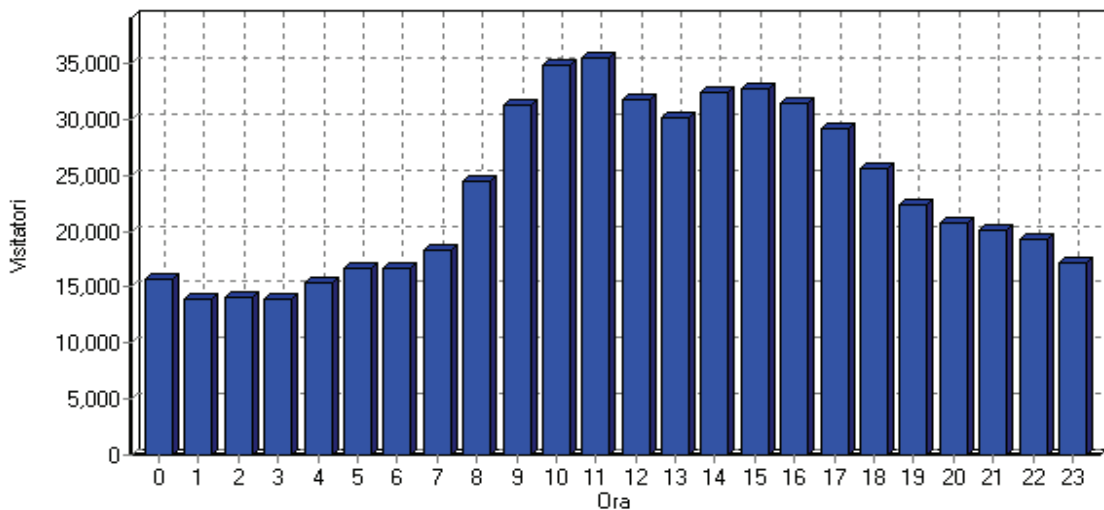


L'andamento della distribuzione oraria delle visite al sito e la distribuzione per giorno della settimana si presenta simile alla precedente analisi (Working Paper 02/04). Dal grafico 4 si

nota subito che la maggiore attività si svolge nelle ore diurne (dalle 8 alle 18), corrispondenti alle ore lavorative per l'Italia (dato ovviamente rilevato per la predominanza di pagine in italiano), mentre un'attività marginale si riscontra nelle restanti ore, molto probabilmente frutto di visite da collegamenti internazionali o di programmi automatici di ricerca, senza dubbio molto attivi sul nostro sito visti i buoni livelli di indicizzazione³ che possiamo rilevare.

L'indicizzazione del sito è il risultato di anni di lavoro che hanno portato il sito a buoni livelli di popolarità in rete dato l'ottimo *posizionamento* di esso a fronte di una ricerca (cioè l'inserimento di una parola su un motore di ricerca della rete) per numerose parole chiave di cui parleremo in seguito, correlate ai temi oggetto delle ricerche dell'istituto, soprattutto quelle storiche, che sono cioè presenti come linee di ricerca da molto tempo e che quindi hanno avuto tempo sufficiente per essere considerate dai principali spider internazionali. Allo stesso modo il *Page Rank*⁴ mette il sito in grande evidenza e ne facilita il reperimento delle pagine in rete.

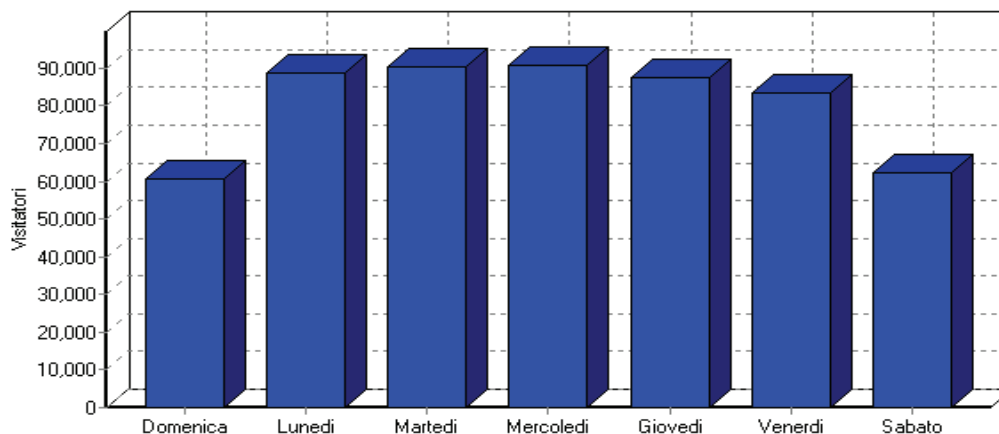
Grafico 4 – I visitatori a ciascuna ora del giorno (medie tra il 2005 e il 2007)



³ Per indicizzazione si intende l'inserimento di un sito web nel database di un motore di ricerca, mediante l'uso di apposite parole-chiave. L'indicizzazione di un sito internet, in altre parole, è il momento in cui il sito viene riconosciuto dai motori di ricerca e quindi compare nelle loro pagine di risposta alle interrogazioni degli utenti web. Diversa cosa è il posizionamento, che è invece un'azione volta a determinare che un sito compaia in una *specifica posizione* nelle pagine di risposta dei motori (cioè nelle prime posizioni).

⁴ Il PageRank è un algoritmo di analisi che assegna un peso numerico ad ogni elemento di un collegamento ipertestuale di un insieme di documenti con lo scopo di quantificare la sua importanza relativa all'interno della serie. L'algoritmo può essere applicato a tutti gli insiemi di oggetti collegati da citazioni e riferimenti reciproci. Il peso numerico assegnato ad un dato elemento E è chiamato anche "il PageRank di E", siglato in PR (E). L'algoritmo di PageRank è stato brevettato (brevetto US 6285999) dalla Stanford University; è inoltre un termine ormai entrato di fatto nel lessico dei fruitori dei servizi offerti dai motori di ricerca. Il nome PageRank è un marchio di Google, uno dei principali motori di ricerca.

Grafico 5 – I visitatori per giorno della settimana (medie tra il 2005 e il 2007)



3. Statistiche di accesso

Il grafico 6 mostra le distribuzioni in valore assoluto delle visite alle pagine più popolari del nostro sito. Si tratta di pagine che riguardano in particolare informazioni sull'attività di ricerca, come ad esempio la descrizione delle attività di alcune sezioni, sulla presentazione dell'istituto, sui risultati della ricerca, come ad esempio la lista delle pubblicazioni on line, che nel nostro sito contiene una intera sezione di pubblicazioni interne open source, la presentazione di alcune indagini svolte dall'Istituto, i servizi offerti alla ricerca, come ad esempio le informazioni sulla biblioteca di Roma. Una buona visibilità, anche se in misura minore, è stata registrata anche per alcuni mini siti come "Mobidis" e "Diva" legati a particolari progetti⁵.

Anche da questa lettura dei dati appare chiaro che il sito dell'IRPPS ha principalmente una funzione strumentale di comunicazione mirata alla soddisfazione di specifiche richieste degli utenti. Inoltre, oltre a cercare informazioni sui contenuti delle ricerche, gli utenti del web cercano anche informazioni sulle offerte di lavoro, oppure come raggiungere l'Istituto.

Come abbiamo già accennato, nel sito dell'Irpps, sezione "prodotti", sono stati messi a disposizione degli utenti web pubblicazioni on line prodotte dal nostro Istituto in formato pdf⁶ che possono essere scaricate gratuitamente.

Nel grafico 7 sono riportati in valori assoluti quelle pubblicazioni che nel periodo di osservazione di tre anni qui considerato sono state maggiormente scaricate dai visitatori. Tra essi si distinguono: un rapporto bilingue sulle indagini dell'Istituto "6 miliardi di abitanti: opinione pubblica e media" del 2002 e la presentazione di un progetto sulla Telemedicina

⁵ MobiDis è un progetto finalizzato in particolare all'analisi dell'uso della telefonia mobile all'interno di un processo di cura di pazienti non localizzati in ospedale. DIVA è un progetto finanziato dalla Commissione Europea nell'ambito dell'FP6- Programma "Strutturare lo spazio europeo della ricerca" - "Scienza e società: Donne Scienza".

⁶ Il formato Adobe PDF (Portable Document Format) è di fatto lo standard per la distribuzione e lo scambio di documenti su Internet. È un formato molto diffuso perché i documenti in questo formato possono essere aperti da tutti con l'ausilio di software gratuiti facilmente reperibili.

tramite Telefonia Mobile del 2004. Seguono in questa graduatoria una monografia dal titolo "Le scienze sociali e la nuova crisi balcanica" del 2001, un working paper che contiene l'elenco delle pubblicazioni edite dalla Sede di Roma dell'Istituto a partire dal 1981 al 2005. Sempre in questa graduatoria si evidenziano lavori relativi al progetto sulla Telefonia mobile e una serie di working paper che sono presenti ormai da molti anni su queste pagine web, ma che ancora riscuotono un notevole successo.

Il grafico 8 mostra le sezioni più visitate del sito. Al primo posto si posiziona la home page in cui sono riportate, in modo molto sintetico, tutte le informazioni che riguardano l'Istituto. Infatti, si presenta come un vero "portale", dove ogni utente grazie a queste prime informazioni può decidere se approfondire o meno l'argomento andando avanti nella visione delle pagine collegate ad esse. Seguono le pubblicazioni, i link dedicati alle migrazioni, le linee di ricerca, le informazioni sull'Istituto (struttura, servizi e staff) e le aree di due progetti MobiDis e DIVA che abbiamo già in precedenza citato. Per questi due progetti sono stati costruiti due mini siti che contengono una visione completa di tutto il progetto, partendo dall'organizzazione, ai risultati della ricerca e a tutti i prodotti scientifici. È evidente dunque che gli utenti cercano nel nostro sito principalmente informazioni sui possibili prodotti dell'Istituto come le pubblicazioni on line e le competenze nei vari settori della ricerca.

Grafico 6 – Le pagine più popolari

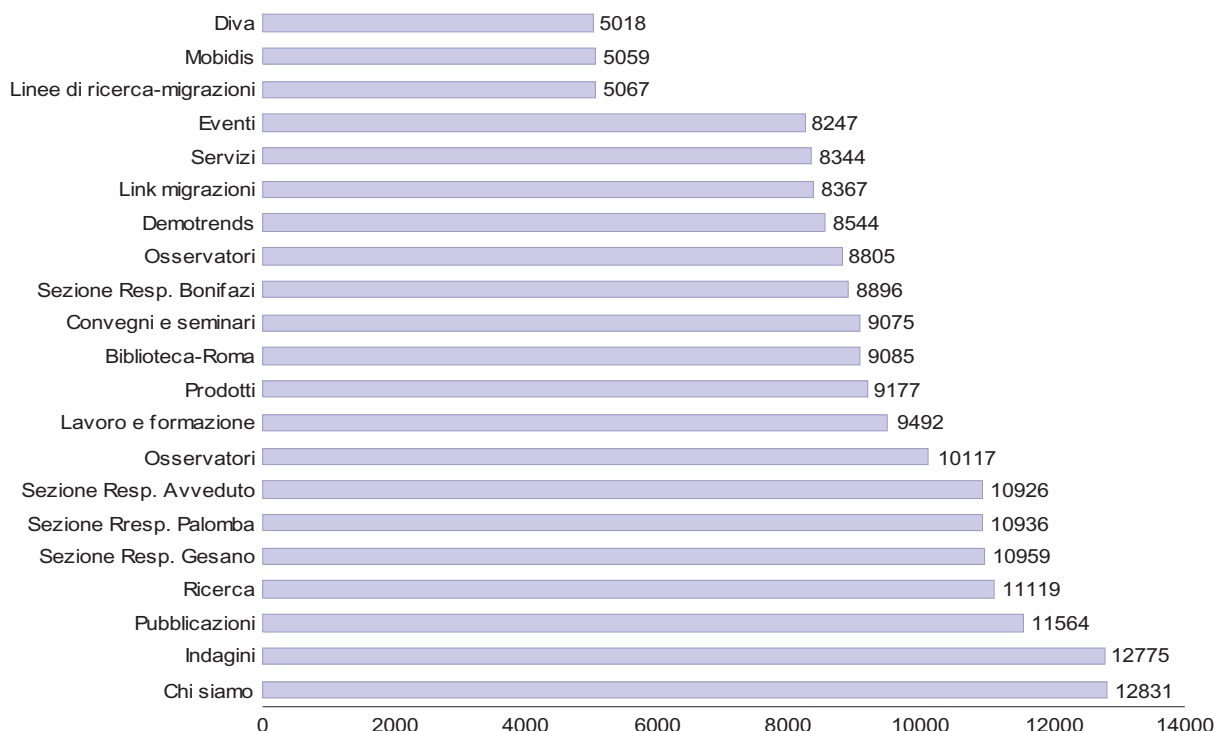


Grafico 7 – Download: i file più scaricati

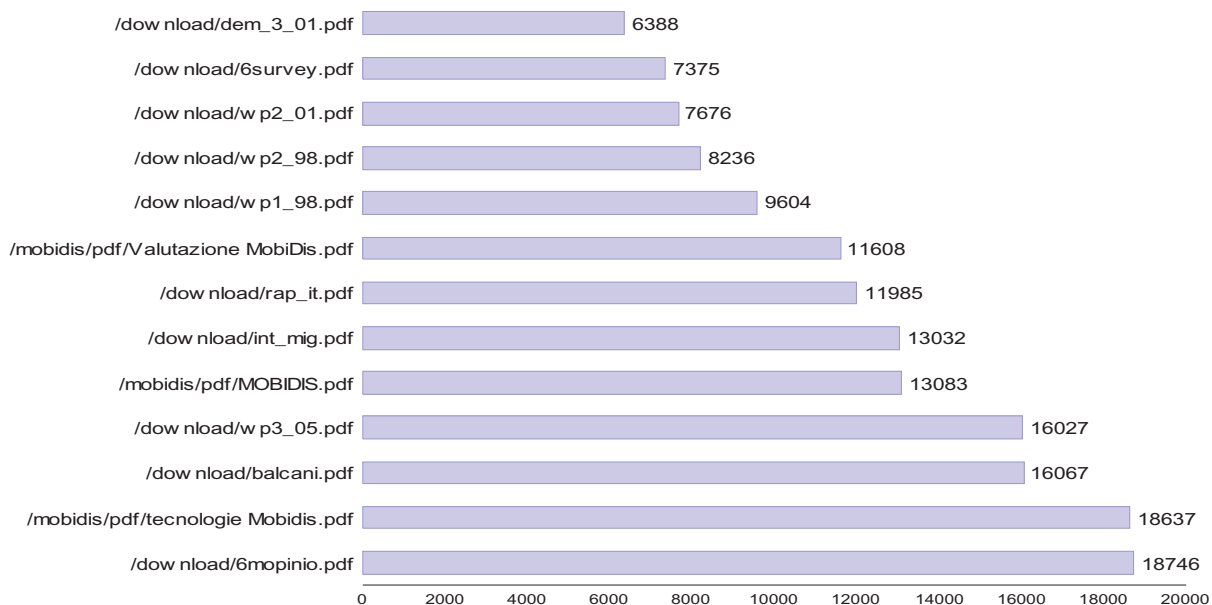
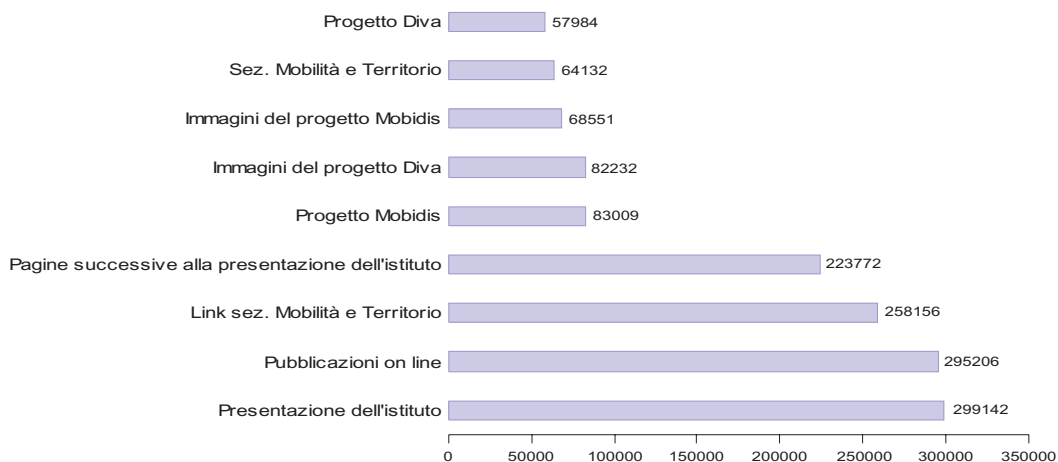


Grafico 8 – Aree più visitate



Le pagine di entrata rappresentano quelle aree del sito che con una buona probabilità i visitatori hanno memorizzato tra le loro pagine preferite o provenienti da una ricerca sul web. In ambedue i casi, si tratta di pagine mirate ad argomenti specifici. Nel grafico 9 abbiamo escluso le visite alla sola home page (valore assoluto di ben 80657) per fare un focus sulle altre pagine che, essendo tutte specifiche di ciascun argomento, sono dovute certamente all'interesse particolare su determinati temi che il sito mette a disposizione. Similmente, le pagine di uscita (Grafico 10) ci fanno stimare il numero di utenti che hanno navigato effettivamente nel sito, andando oltre la prima pagina e uscendo da una pagina secondaria.

Grafico 9 – Le pagine di entrata più frequenti

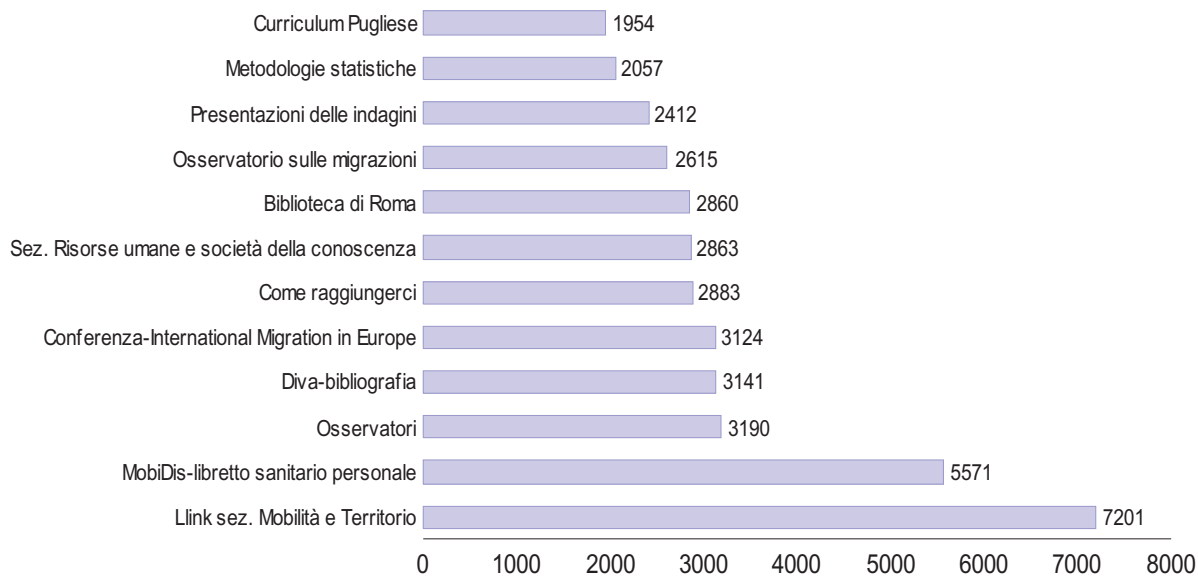
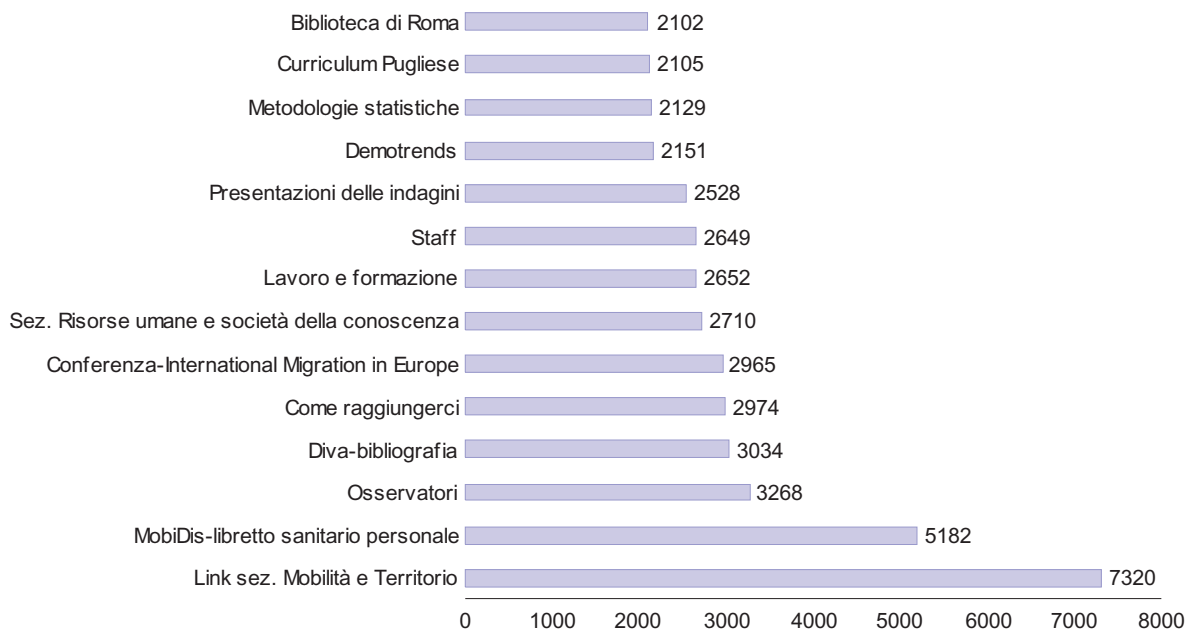


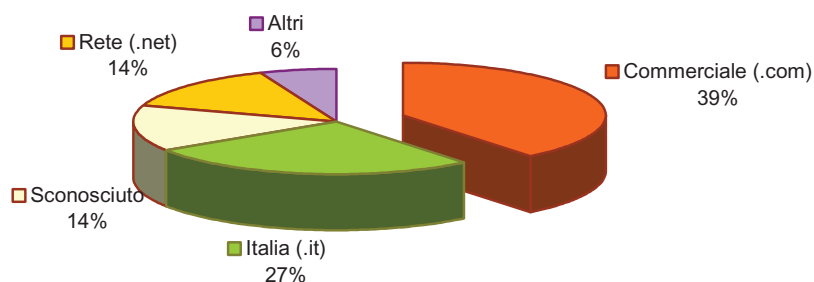
Grafico 10 – Le pagine di uscita più frequenti



4. Utenti visitatori

I domini maggiormente presenti (Grafico 11) sono quelli provenienti dal settore commerciale con il 39% dei visitatori (domini.com) mentre si riducono al 27% quelli italiani (domini.it). Seguono con il 14% quelli di cui non si conosce il dominio e i Network (domini.net) e infine con il 6% i restanti domini. Questa distribuzione percentuale si presenta diversa rispetto ai risultati 2002/03 (Working Paper 02/04). Mentre nei precedenti risultati avevamo al primo posto i domini .it, negli attuali risultati prevalgono i domini commerciali. Possiamo pensare ad un risultato inaspettato, ma allo stesso momento comprensibile. Probabilmente, lo stesso l'aumento del numero dei domini commerciali che si è verificato negli ultimi anni e, di conseguenza, la produzione di una nuova visione del web e di un suo utilizzo sempre più spesso inteso come strumento allargato per la comunicazione, ha indotto nuove tipologie di visitatori a cercare nel nostro sito delle cose più concrete come offerta di prodotti e di servizi che va al di là delle normali richieste di studenti, professionisti e università interessati particolarmente alla ricerca.

Grafico 11 – I domini più presenti



Il fatto che ben il 55% (Grafico 12) dei visitatori sia dovuto a traffico proveniente dagli Stati Uniti, dimostra come sia intensa l'attività dei programmi automatici dei motori di ricerca che fanno un monitoraggio periodico del materiale pubblicato, ma questi numeri tanto elevati dimostrano anche la nostra inevitabile presenza a livello internazionale. Anche questo dato, oltre ad una necessità insita nella natura dei dati pubblicati su un sito di ricerca, induce a ritenere quanto mai indispensabile la pubblicazione della versione bilingue (inglese e italiano) a tutto il sito nel più breve tempo possibile (attualmente la traduzione delle pagine in inglese è solo parziale). Anche se in percentuale molto bassa ci sembra interessante notare che l'1% dei visitatori proviene dalla Cina, dato che non è emerso nella precedente analisi (Working Paper 02/04), e che segnala l'importanza crescente di un Paese che è ormai arrivato ad uno sviluppo anche in termini tecnologici perché è presente in una discreta percentuale con i suoi provider nelle nostre liste di provenienza. Segue al grafico una lista delle Nazioni con le relative percentuali del totale dei visitatori.

Grafico 12 – I paesi di provenienza dei visitatori

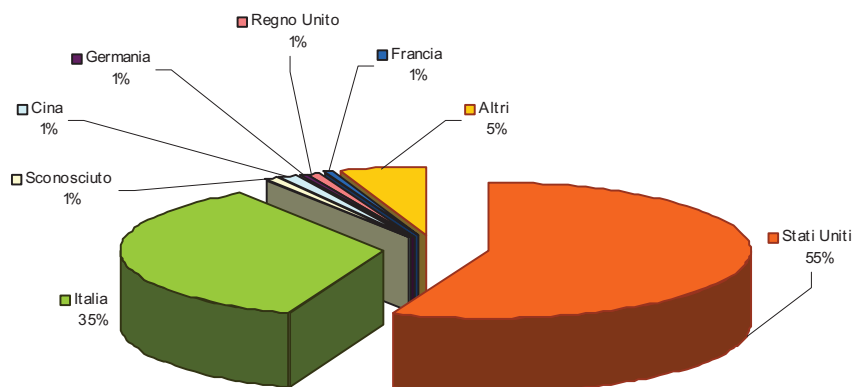


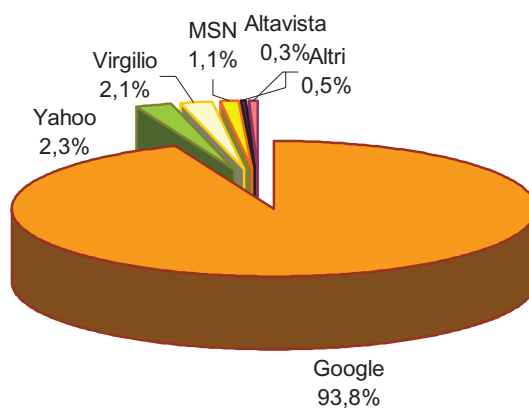
Tabella 2 – Dettaglio della provenienza dei visitatori

Nazione	Visitatori	% del Totale dei Visitatori	Nazione	Visitatori	% del Totale dei Visitatori
Stati Uniti	314,34	55.76%	Portogallo	516	0.09%
Italia	197,854	35.10%	Korea,	491	0.09%
Sconosciuto	5,066	0.90%	Grecia	470	0.08%
Cina	5,052	0.90%	Norvegia	447	0.08%
Germania	4,534	0.80%	Finlandia	431	0.08%
Regno Unito	3,816	0.68%	Irlanda	380	0.07%
Francia	2,918	0.52%	Marocco	371	0.07%
Etiopia	2,617	0.46%	Danimarca	343	0.06%
Olanda	1,706	0.30%	Federazione Russa	328	0.06%
Svizzera	1,694	0.30%	Singapore	285	0.05%
Polonia	1,522	0.27%	Arabia Saudita	271	0.05%
Canada	1,473	0.26%	Ungheria	265	0.05%
Belgio	1,263	0.22%	Iran	238	0.04%
Svezia	1,226	0.22%	Slovackia	220	0.04%
Australia	1,18	0.21%	Filippine	211	0.04%
India	979	0.17%	Tailandia	210	0.04%
Spagna	950	0.17%	Malesia	207	0.04%
Brasile	927	0.16%	Israele	205	0.04%
Giappone	839	0.15%	Taiwan	184	0.03%
Austria	666	0.12%	Indonesia	170	0.03%
Repubblica Ceca	599	0.11%	Pakistan	169	0.03%
Messico	584	0.10%	Croazia	158	0.03%
Romania	550	0.10%	Emirati Arabi	158	0.03%
Argentina	548	0.10%	Lituania	155	0.03%
Turchia	516	0.09%	Ucraina	148	0.03%

5. Motori e frasi di ricerca

Nel triennio 2005-2007, nel 93,8% dei casi il sito dell'IRPPS è stato trovato a fronte di una ricerca su Google, la restante percentuale si distribuisce tra Yahoo, Virgilio, MSN, Altavista ed Altri motori meno utilizzati in generale dagli utenti web. Questa alta percentuale trova riscontro anche nelle statistiche di descrittive che Wikipedia (l'enciclopedia libera) fa di Google. Infatti, sostiene che Google è riconosciuto come il più grande e affidabile tra i motori di ricerca, occupandosi attraverso il suo sito di oltre il 70% di tutte le ricerche effettuate su internet. Inoltre, Google non si limita a catalogare il World Wide Web, ma si occupa anche di immagini, foto, newsgroup, notizie, mappe, video, oltre a mantenere una copia cache di tutte le pagine che conosce (<http://it.wikipedia.org/wiki/Google>).

Grafico 13 - Migliori motori di ricerca



Come abbiamo detto precedentemente, una buona visibilità è stata registrata per il progetto "MobiDis", un progetto finalizzato in particolare all'analisi dell'uso della telefonia mobile all'interno di un processo di cura di pazienti non localizzati in ospedale. Buona parte di questa visibilità è dovuta alla frase "libretto sanitario", termine utilizzato all'interno del mini sito e che ha avuto un risultato sorprendente, per il nostro istituto che si occupa anche di molti altri temi: ben 3732 sono state le ricerche con questa frase chiave fatte dai nostri visitatori. Segue l'acronimo "irpps", oppure, frasi che sono attinenti ai temi delle nostre ricerche: immigrazione in Italia, popolazione, anziani, società della conoscenza, emigrazione italiana. Infine, non mancano ricerche su informazioni riguardo ad alcuni ricercatori dell'istituto. In totale nel triennio sono state eseguiti accessi dai motori di ricerca 145054 volte usando delle frasi di più di una parola e ben 520617 volte usando singole alcune parole chiave (popolazione, libretto, immigrazione, famiglie, donne, demografia).

Quindi è chiaro che il sito viene trovato a fronte di una ricerca che riguardi le aree di interesse in Istituto. Ciò ovviamente è quanto ci si auspicava ed è anche il risultato del

lavoro di indicizzazione di cui si è detto che ha condotto a questo risultato. Naturalmente, non si è mai dato particolare rilievo ad uno o a un altro settore del sito perché fosse più facilmente di altri trovato attraverso una ricerca sui motori più comuni; dunque non tutti i risultati erano prevedibili fin dall'inizio. Ad esempio il fatto che il sito fosse trovato in funzione della frase di ricerca 'libretto sanitario' che, se vogliamo, è un particolare dettaglio di un tipo di materiale pubblicato e non si tratta di un tema generale che riguarda molte aree del sito, è abbastanza sorprendente perché è un risultato che dipende essenzialmente dalla ricerca in sé e non dalla particolare indicizzazione.

Tabella 3 - Frasi di ricerca e parole chiave più frequenti

Frases	Numero di ricerche	Frases	Numero di ricerche	Parola Chiave	Numero di ricerche
libretto sanitario	3732	istituti di ricerca sociale	257	popolazione	6716
irpps	1988	gli anziani	236	libretto	4794
cluster analysis	1495	matrimonio e convivenza	231	sanitario	4777
biblioteca roma	1328	tecnica delphi	229	ethiopia	4768
immigrazione in italia	804	glossario informatica	223	immigrazione	4323
popolazione	676	sveva avveduto	208	roma	3900
cartella clinica	576	popolazione italiana	204	migration	3636
società della conoscenza	541	cartelle cliniche	202	italiana	3436
anziani	508	fernando ferri	196	anziani	3413
immagini anziani	507	delocalization	184	questionario	3010
ethiopia housing	496	microsoft photo editor	180	irpps	2962
popolazione mondiale	494	questionario anziani	178	cluster	2396
mainstreaming	442	brain drain	178	cartella	2363
enrico pugliese	432	demotrends	174	analysis	2325
photo editor	415	donne famose	173	famiglia	2291
diva	391	mobidis	172	clinica	2196
irpps cnr	374	urbanization in ethiopia	166	giovani	2136
medico on line	371	istituto di ricerche sulla popolazione	166	biblioteca	2119
housing in ethiopia	371	istituto di ricerche sulla popolazione e le politiche sociali	164	donne	2101
rossella palomba	300	corrado bonifazi	163	emigrazione	2082
nuove figure professionali	285	biblioteca salerno	161	istituto	2001
emigrazione italiana	281	migrazioni interne	159	ricerca	1999
metodologia delphi	274	ecg	159	dati	1913
tavola rotonda	274	figure professionali	157	sociali	1887
irp	272	ethiopia census	157	demografica	1849

6. Focus su alcune aree del sito.

Il sito principale è corredato da alcuni mini siti che offrono approfondimenti su tematiche particolari e di grande interesse per l'istituto. Alcuni di essi sono derivati da specifici progetti di ricerca finanziati a livello europeo, altri invece rappresentano linee di ricerca più complesse e composte da più progetti, alcuni europei, altri nazionali. Nonostante questa grande varietà di tipologie, essi sono tutti accomunati dal bisogno di utilizzare specifiche aree del sito per una comunicazione mirata e dedicata ad un certo tema.

L'analisi qui presentata è relativa ad un periodo di tempo compreso tra il 2006 e il 2008 e sono stati analizzati in particolare i file di log di quattro mini siti: "L'osservatorio sulle migrazioni italiane", il sito web della commessa del CNR "Comunicazione della Scienza ed Educazione", "DIVA" e "Light". Per ciascuno di essi era stata fatta una specifica richiesta di monitoraggio delle visite finalizzata all'analisi dell'utilizzo delle informazioni pubblicate per l'importanza strumentale delle pagine relative a queste aree.

Tabella 4 – Quanti sono i visitatori dei mini siti osservati?

Mini siti	Numero totale di pagine visitate	Media mensile
L'osservatorio sulle migrazioni italiane (18/01/2008 - 18/12/2008)	35000	3000
Sito web della commessa del CNR "Comunicazione della Scienza ed Educazione" (18/01/2008 - 18/12/2008)	18000	1500
DIVA (febbraio 2006-giugno 2007)	20000	1180
Light (giugno-ottobre 2008)	15390	3080

Riportiamo una sintesi delle informazioni ottenute dalle analisi di tali dati.

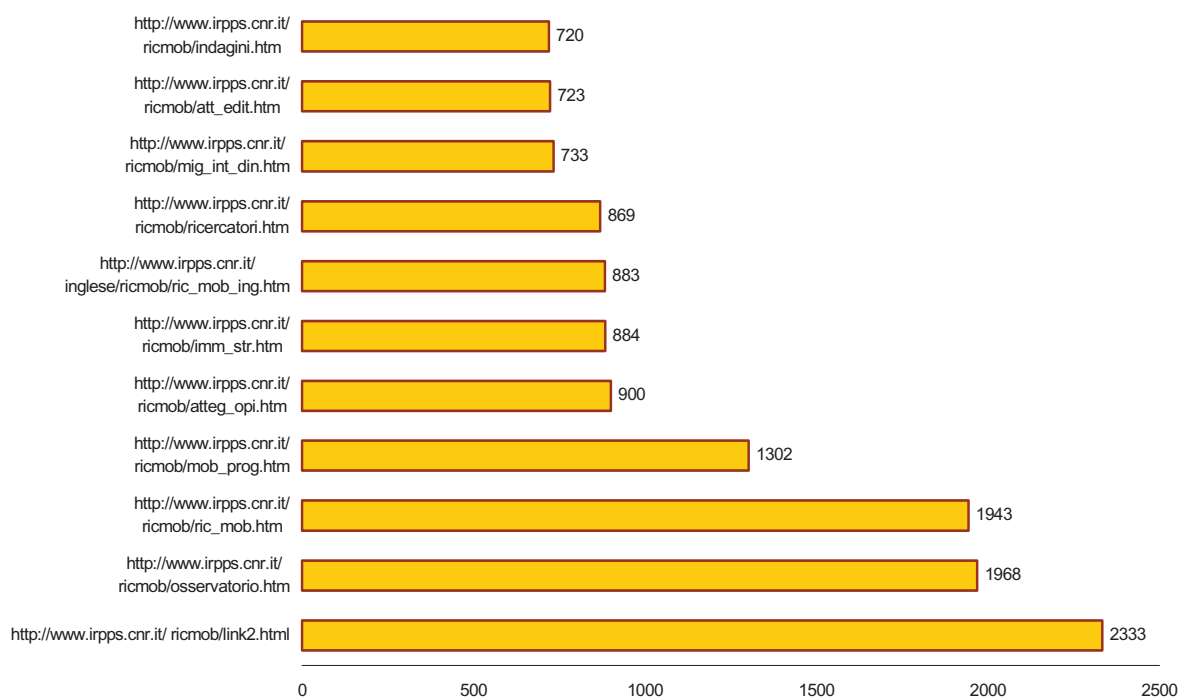
"L'osservatorio sulle migrazioni italiane" (Responsabile dottor Corrado Bonifazi) punta a seguire con continuità le diverse dinamiche migratorie che interessano il nostro Paese. In primo luogo, attraverso le ricerche che l'Istituto ha condotto e conduce sui diversi aspetti dei fenomeni migratori italiani. In secondo luogo, mettendo a disposizione degli utenti interessati le pubblicazioni dell'Istituto, il collegamento con alcuni siti che si occupano di migrazioni e raccogliendo dati statistici sul fenomeno. Anche in questo caso l'analisi delle visite al mini sito ha lo scopo di comprendere l'efficacia della comunicazione sul tema delle migrazioni attraverso il web. In particolare, dato che si tratta di un progetto sul lungo periodo, non interessa tanto la eventuale scoperta di picchi di visite in determinati periodi, cosa che poteva essere invece importante nell'analisi dei due mini siti: "DIVA" e "Light", quanto la presenza di un livello costante di visite, che dovrebbe sottendere una certa fidelizzazione degli utenti interessati al tema trattato.

Durante il periodo di osservazione delle viste a questa parte del sito (da gennaio a dicembre 2008) la pagina che ha avuto il maggior numero di contatti (2333 con una media mensile 194,41) (Grafico 14) è dunque dedicata ai *link* a siti contenenti informazioni sull'immigrazione italiana (istituzioni e produttori ufficiali di dati, centri di ricerca e riviste e organizzazioni non governative). Seguono le pagine di presentazione dell'osservatorio con 1968 visite e quella dedicata alle ricerche con 1943. Il totale delle visite al sito dell'osservatorio in tutto l'arco dell'anno supera di gran lunga le 35000 unità con una media

mensile di oltre 3000 visite. I dati dimostrano dunque quanto temi che trattano argomenti sulle migrazioni sono oggetto di interesse di molti utenti web.

Alcuni dettagli riscontrati in questa analisi fanno pensare che gli utenti siano alla ricerca di un certo tipo di servizi oggi non ancora del tutto implementati su queste pagine. Facciamo riferimento per esempio al fatto che sia proprio la pagina dei link ad essere tra le più visitate, e questo può costituire un indicatore del tipo di aspettativa che un utente generico potrebbe avere nei confronti del materiale pubblicato. Perciò, nonostante gli ottimi livelli delle visite al sito, si è pensato di ristrutturare le pagine ad esso dedicate attraverso l'inserimento di aree dedicate a servizi aggiuntivi che possono aumentare il livello di fedeltà degli utenti già acquisiti, ma anche spronare l'aumento delle visite complessive. Ovviamente, la pubblicazione di materiale in aree riservate, o la eventuale costituzione di una mailing list dedicata, saranno oggetto di una specifica progettazione che terrà conto della necessità di un aumento del livello di sicurezza delle pagine per la garanzia che dati, sia di natura personale, sia di altra natura, ricevano il trattamento adeguato e secondo quanto stabilito dalle leggi vigenti.

Grafico 14 - Visite alle pagine più popolari dell'osservatorio sulle migrazioni italiane nel periodo 18/01/2008 - 18/12/2008.



Sito web della commessa del CNR "Comunicazione della Scienza ed Educazione"
(Responsabile dottoressa Adriana Valente).

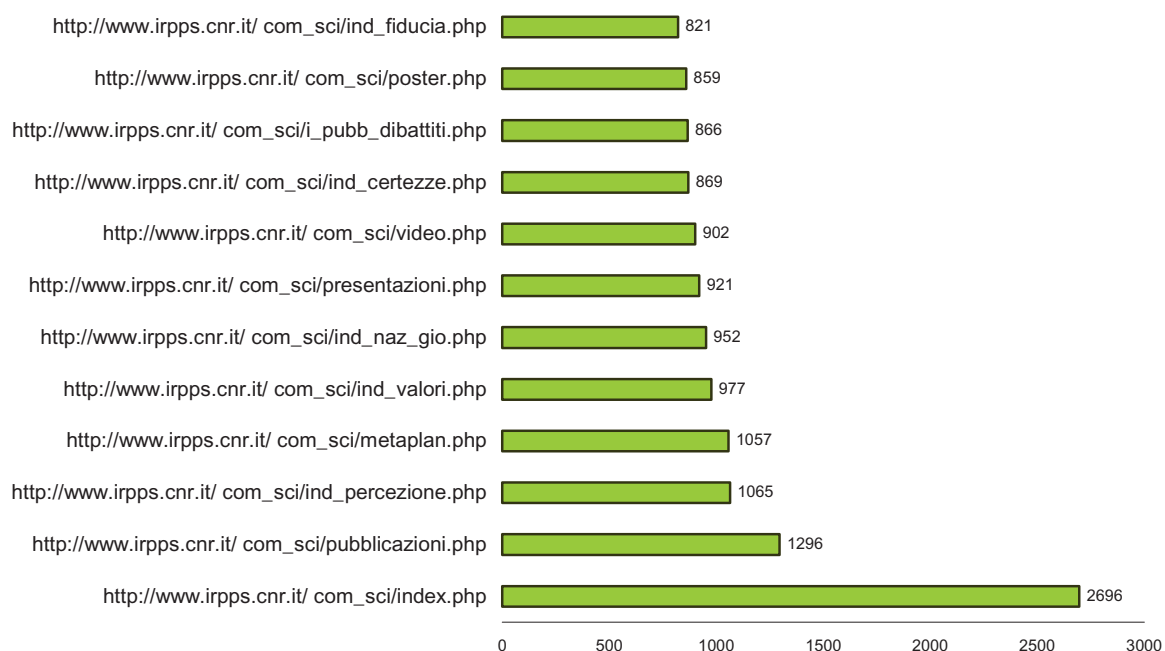
La commessa denominata 'Comunicazione della Scienza ed Educazione' svolge da anni attività di ricerca e di sperimentazione sui principali aspetti della comunicazione pubblica della scienza e della didattica delle scienze ai fini di una migliore comprensione e valorizzazione del rapporto scienza e società. In questa breve analisi vogliamo evidenziare la

visibilità che la commessa ha avuto attraverso il web, in quanto, anche, e soprattutto, in questo caso, ci interessa comprendere come agisce questo tipo di comunicazione. Inoltre nelle pagine dedicate a questa commessa sono presenti materiali scaricabili, video e foto che sono stati oggetto di divulgazione in occasione degli eventi organizzati dalla commessa stessa. Dunque corre l'obbligo, oltre che di soddisfare una legittima curiosità, di comprendere se e come sia avvenuto il trasferimento delle informazioni che sono state messe a disposizione sulle pagine di questo mini sito. Il totale delle visite in tutto l'arco dell'anno (dato riferito al 2008) arriva a quasi 18000 unità, con una media di circa 1500 visite al mese. La graduatoria delle sezioni più visitate del sito vede al primo posto (Grafico 15) la pagina principale con 2696 e una media mensile con 224,7, in cui vi è una sintesi sulla attività di ricerca e link a progetti, a indagini e ad altre informazioni. Segue al secondo posto la lista delle pubblicazioni con 1296, e infine le pagine che trattano presentazioni di progetti e di link a video.

Dunque, anche in questo caso, non c'è dubbio che le informazioni pubblicate siano state trasmesse agli utenti del web. Le modalità di ricezione a volte sono state testate attraverso gli incontri con i diretti interessati, specialmente studenti delle scuole superiori, ai quali la commessa rivolge in particolare la propria attenzione attraverso le attività di sperimentazione delle tecniche di partecipazione.

Le visite dell'osservatorio sulle migrazioni italiane risultano essere il doppio rispetto alle visite della commessa. Queste differenze sono spiegabili, perché, mentre l'osservatorio può vantare di una messa in rete che dura da circa 10 anni, il mini sito della commessa è nato solo da un anno, e probabilmente non ha ancora creato una sua rete di utenti abituali.

Grafico 15 - Visite alle pagine più popolari del Sito web della commessa del CNR "Comunicazione della Scienza ed Educazione" nel periodo 18/01/2008 - 18/12/2008.



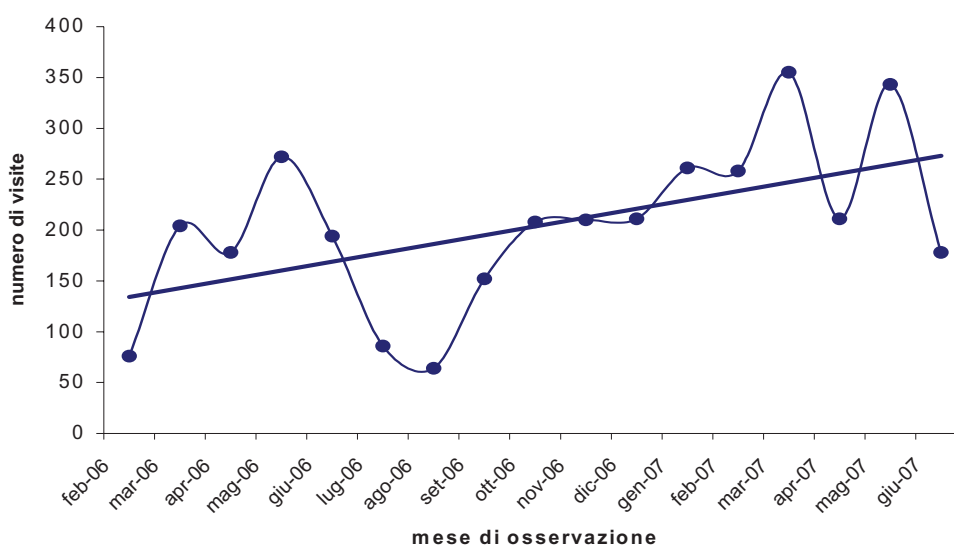
"DIVA" è un progetto finanziato dalla Commissione Europea nell'ambito dell'FP6- Programma "Strutturare lo spazio europeo della ricerca" - "Scienza e società: Donne Scienza" (Responsabile dottoressa Rossella Palomba).

Il progetto ha utilizzato il sito come principale mezzo di comunicazione tra l'istituto e gli utenti destinatari delle iniziative previste dal progetto, in primo luogo le scuole italiane, ma anche i centri di ricerca sparsi sul territorio. Risulta dunque del tutto superfluo spiegare l'importanza delle pagine web e, conseguentemente, dell'analisi delle visite alle stesse.

Dal momento della prima messa in rete, cioè dal febbraio del 2006, il sito di DIVA ha visto crescere l'attenzione dei visitatori con un trend sicuramente positivo pur risentendo di una certa stagionalità che sempre si riscontra quando si analizzano le visite alle pagine pubblicate sul web. Durante il periodo di osservazione, che copre l'arco di quasi un anno e mezzo, le pagine di DIVA sono state visitate oltre 20.000 volte con una media di 1180 accessi mensili. La graduatoria delle sezioni più visitate del sito - che vede al primo posto la sezione delle *biografie*, seguita al secondo posto da quella degli *stereotipi*, al terzo posto dalla pagina descrittiva del *progetto*, al quarto posto da quella che elenca gli *appuntamenti* dell'ambasciatore nelle scuole e nei centri di ricerca, al quinto posto dalla pagina sui *role models* e al sesto posto dalla pagina denominata *mainstreamings*, - dimostra che sono soprattutto le pagine a contenuto informativo ad interessare i visitatori, i quali, ovviamente si soffermano anche su tutte le altre aree del sito ricche di immagini (foto e grafici) di rapida consultazione.

La pagina più visitata in assoluto è dunque quella dedicata alle biografie delle scienziate famose che sono state visitate in media 204 volte al mese. Il grafico che segue (Grafico 16) mostra l'andamento delle visite nel periodo di osservazione e dimostra anche, con una linea di tendenza, quale sia l'andamento complessivo registrato durante tutto il periodo considerato.

Grafico 16 - Visite alla pagina www.irpps.cnr.it/diva/biografie.php nel periodo febbraio 2006-giugno 2007.

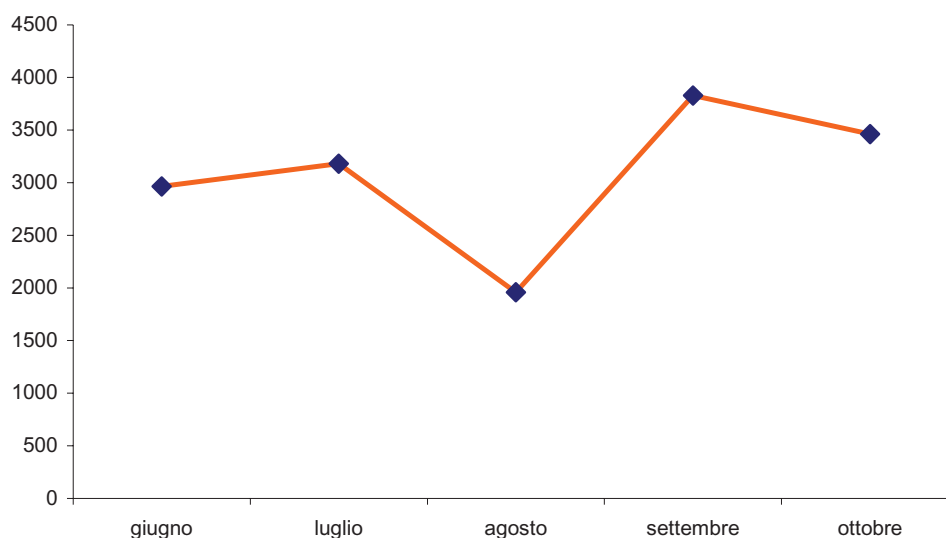


"Light" è una iniziativa della Commissione Europea finalizzata a promuovere l'interesse per la ricerca scientifica soprattutto tra i giovani (Responsabile dottoressa Adele Menniti).

Anche questa linea di ricerca ha utilizzato il sito come mezzo di scambio privilegiato delle informazioni tra il nucleo principale della ricerca e gli utenti finali. Dunque anche in questo caso l'analisi delle visite al sito è di primaria importanza per comprendere l'efficacia della comunicazione mediatica che si è messa in atto, in questo particolare caso, si è fatto uso dello strumento delle newsletter che sono apparse a intervalli di tempo regolari sul sito e che sono l'oggetto principale dello studio delle visite a questo mini sito.

Durante il periodo di osservazione (giugno-ottobre 2008), le pagine di Light sono state visitate 15390 volte con una media mensile di 3080. Come si evidenzia dal grafico (Grafico 17) il massimo delle visite si è concentrato a settembre, mese dell'evento finale del progetto che si è concretizzato in una iniziativa pubblica che è stata pubblicizzata principalmente attraverso il sito web. In totale, le newsletter, quattro in tutto, sono state consultate da circa 1070 visitatori e sempre con punte massime a settembre. Inoltre, in tutto il periodo di osservazione, la newsletter1 risulta essere maggiormente visitata rispetto alle altre, questa maggiore concentrazione è dovuta soprattutto a un periodo di messa in rete più prolungato rispetto alle altre. Però, nonostante le newsletter non siano state pubblicate contemporaneamente, non si è avuto l'effetto 'scadenza' delle newsletter precedenti perché i livelli delle visite sono tutti equiparabili mese per mese. Questo dimostra che le informazioni sul sito sono state ritenute sempre ugualmente interessanti al passare del tempo.

Grafico 17 - Visite al sito di Light nel periodo giugno-ottobre 2008.



6. Conclusione

Dai risultati dell'analisi dei file di log degli ultimi tre anni è emerso che le pagine visitate sono più che raddoppiate come allo stesso modo le medie delle visite giornaliere, anche se nell'ultimo anno hanno subito un decremento. Al contrario, l'aumento della numerosità degli utenti visitatori è stata costantemente crescente nel periodo di anni considerato, compreso l'ultimo. Lo stesso andamento si ritrova per la media giornaliera dei visitatori. A questo punto dobbiamo soprattutto soffermarci e cercare di dare una spiegazione al decremento delle pagine visitate in particolare nell'ultimo anno. Vale la pena di ripensare al senso che assume il sito dell'IRPPS in rete e di porsi ancora gli interrogativi che possono essere utili per stimolare l'utente a visitare le nostre pagine.

I nostri visitatori sono soprattutto di provenienza statunitense (ben il 55%), e questo valore, anche se non direttamente correlabile con la presenza di un pubblico oltre oceano, deve servire a sollecitare la pubblicazione delle pagine in lingua inglese, per ora troppo sporadiche. Oltre gli statunitensi molte sono le nazionalità di provenienza dei nostri utenti visitatori: tedeschi, francesi, inglesi e una buona percentuale di cinesi, ecc. Quest'ultimi, oggi, sono presenti al quarto posto nella lista delle nazionalità, mentre nella precedente lista non rientravano affatto tra i primi 50 (Working Paper 02/04), un fenomeno in linea con i recenti sviluppi delle competenze tecniche della Cina che si prospetta sempre più come un Paese ad alto potenziale tecnologico alla pari con molti altri Paesi del globo.

Infine sono in costante aumento gli utenti provenienti da domini commerciali (.com), un fenomeno atteso visto l'enorme sviluppo degli ultimi anni di questo tipo di domini, che è effettivamente più abbordabile perché meno soggetto a regole restrittive ed è tipico dei paesi di oltre oceano, che risultano molto presenti tra le nostre tipologie di visitatori.

Questo fatto deve però indurci a riflettere sulla prospettiva di una nuova progettazione del sito, quanto mai urgente, per poterlo adeguare alle esigenze odierne di una utenza internazionale, come quella che abbiamo nel nostro caso. Inoltre è necessario un continuo aggiornamento delle informazioni pubblicate per rendere il sito sempre più interessante sia per gli utenti affezionati che ci visitano periodicamente e che sono interessati in modo particolare ai prodotti della ricerca, sia per gli altri utenti della rete che possono essere avvicinati alla ricerca e ai suoi prodotti per pura curiosità e che stanno crescendo sempre di più. Infine anche l'archiviazione delle informazioni in un formato interfacciabile con altre applicazioni è auspicabile, per potenziare l'efficacia del lavoro e degli sforzi che si fanno per la manutenzione del sito e per rendere le stesse informazioni già disponibili per gli adempimenti periodici che i ricercatori sono chiamati ad assolvere e che fanno parte integrante delle attività previste.

Bibliografia

Baldi P., Fiasconi P., Smyth P., (2003) Modeling the Internet and the Web, Probabilistic Methods and Algorithms, Wiley.

Giudici P., (2001) Data mining – metodi statistici per le applicazioni aziendali, McGraw-Hill, Milano.

Lynch P.J., Horton S., (2001) Web – Guida di stile – Progettazione dei siti Web, Apogeo, Milano.

Pazzani, M. J. (2000), Knowledge discovery from data?, IEEE Intelligent Systems March/April 2000, 10–13.

Kohavi R., Provost F. (2001) Applications of Data Mining to Electronic Commerce, Data Mining and Knowledge Discovery, 5, 5–10, Kluwer Academic Publishers, Manufactured in The Netherlands

Rosenfeld L., Morville P., (1998) Information architecture for the World Wide Web, Sebastopol, O'Reilly.

Silani S., Tarantino M., (2001) Tecniche di Data Mining con SAS Enterprise Miner, SAS Institute.

http://it.wikipedia.org/wiki/Pagina_principale