

## 5 Considerazioni

Amministrare un *sito* Web articolato qual è quello dell'IVSLA richiede un impegno costante. Composto di tre settori distinti per funzioni e contenuti, ma nello stesso tempo integrati (insistono sulla stessa piattaforma e sfruttano gli stessi canali di connessione), il sistema deve essere capace di fornire a ciascuno un rendimento adeguato.

Nei capitoli precedenti ci si è occupati solamente delle due *Banche Dati Ambientali*, ma in effetti l'Istituto Veneto di Scienze, Lettere ed Arti ha realizzato, e gestisce autonomamente, anche il proprio *sito* Web istituzionale, dedicato alla presentazione dell'ente e delle attività svolte. Trattandosi di un Istituto che vanta una notevole tradizione culturale, è più che lecito attendersi non solo garanzie per quanto riguarda la qualità dei contenuti, ma anche un'elevata stabilità ed efficienza di connessione. Per fornire tali garanzie è necessario disporre di un apparato dimensionato in modo appropriato, tuttavia è impensabile pretendere di trovare l'equipaggiamento ottimale in un contesto di estrema dinamicità come quello caratterizzante i network informatici. La soluzione più razionale consiste piuttosto nel predisporre una struttura espandibile, capace di consentire un'agevole implementazione, e parallelamente esercitare un controllo continuo sulla resa del proprio prodotto, mantenendo sempre vigile l'attenzione alle proposte di ipotesi alternative.

Per quanto riguarda le funzioni di monitoraggio, un valido aiuto proviene da programmi appositamente concepiti i quali, sulla base dei rapporti registrati automaticamente in memoria dal Web *server*, operano i conteggi sugli accessi, identificandone provenienza e destinazione.

Di seguito si riportano alcune delle statistiche eseguite da «[AWStats](#)», un software “libero” distribuito sotto l'autorizzazione del GNU e progettato per monitorare *siti* Web, *canali* FTP, *mail server* e generare rapporti, restituiti in forma sia di grafici sia di tabelle numeriche.

Tra le informazioni interessanti che se ne possono ricavare citiamo il resoconto sulle pagine più visitate, indicativo delle preferenze del pubblico e/o del livello di indicizzazione nei motori di ricerca, e il volume di traffico che la consultazione comporta, utile a stabilire un adeguato dimensionamento dei canali di connessione. Il programma fornisce inoltre un valido aiuto per individuare quei *link* che, per una qualunque ragione, hanno perduto la corretta destinazione e non giungono a buon fine.

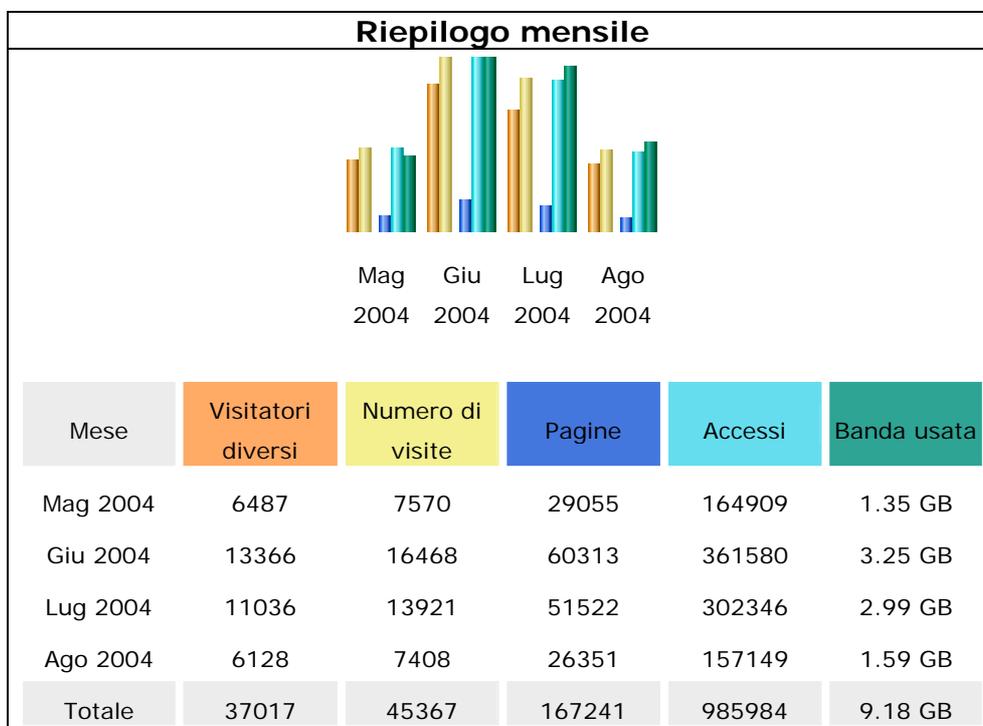
Le informazioni raccolte nelle tabelle sotto riportate si riferiscono al periodo: 19 maggio 2004 – 24 agosto 2004. L'intervallo considerato è limitato poiché questo specifico software è stato adottato solo di recente, in seguito ad aggiornamenti strutturali che hanno interessato le modalità di allacciamento alla rete telefonica.

Le diciture, che compaiono come intestazioni alle colonne nelle tabelle, ed i rispettivi colori presenti nei grafici, hanno i seguenti significati:

- **Visitatori diversi:** Numero di *client host* (indirizzi IP) utilizzati per accedere al *sito* e visualizzare almeno una pagina. Questa cifra riflette il numero di utenti differenti che hanno visitato il *sito* in un giorno.
- **Numero di visite:** Si considera una nuova visita ogni arrivo, per consultare una pagina, di un visitatore che non si sia connesso negli ultimi 60 minuti.
- **Accessi:** Numero di volte che una pagina, immagine o elemento del *sito* è visto o scaricato da un visitatore.

- **Pagine:** Numero di volte che una pagina del *sito* è stata vista. Questo valore tiene conto solamente delle pagine HTML complete e non dei singoli elementi costitutivi.
- **Banda usata:** Numero totale di kilobytes scaricati dal *sito* durante le visite.

**Tabella 2**



In tabella 2 viene esposto il resoconto della “connettività” complessiva al *sito*, espressa attraverso le varie categorie scelte dal programma come “indicatori”. È stata riportata, assieme al relativo grafico, solamente per fornire un esempio di ciò che il software è in grado di produrre, poiché la scarsa significatività statistica dei dati, riguardanti un periodo decisamente troppo breve, rende inopportuno estrapolare considerazioni di carattere generale.

**Tabella 3**

<b>Domini o nazioni dei visitatori (Prime 10)</b>					
	<b>Domini/Nazioni</b>		<b>Pagine</b>	<b>Accessi</b>	<b>Banda usata</b>
	Italy	it	46621	290617	2.67 GB
	Network	net	3978	24125	215.95 MB
	Commercial	com	1325	5939	93.12 MB
	France	fr	741	5272	38.11 MB
	Austria	at	422	1656	89.66 MB
	Canada	ca	421	1628	12.89 MB
	Germany	de	388	2210	24.36 MB
	United Kingdom	uk	340	1824	19.24 MB
	Spain	es	337	1982	26.50 MB
	Altri		3218	18283	222.92 MB

La tabella 3 riporta la provenienza nazionale degli accessi, in ordine decrescente. Buona parte proviene, com'è da aspettarsi in ragione della tipologia del *sito*, dall'Italia. Tuttavia non mancano le connessioni estere, che riteniamo di imputare alla stretta attinenza del *sito* con Venezia. A tale proposito si fa presente che la ricerca su «Google», relativa a tutto il Web (e non limitata alle «pagine in Italiano» o alle «pagine provenienti dall'Italia») attraverso le parole chiave «Venezia laguna», ha restituito (in data 24 agosto 2004) la «*Banca Dati Ambientale sulla Laguna di Venezia*» al primo posto!

**Tabella 4**

<b>Accessi di robot e spider (Prime 10)</b>			
<b>29 robot diversi*</b>	<b>Accessi</b>	<b>Banda usata</b>	<b>Ultima visita</b>
Unknown robot (identified by 'crawl')	54326	1.29 GB	22 Ago 2004 / 08:47
Googlebot	53522	2.31 GB	24 Ago 2004 / 06:30
MSNBOT	27677	624.57 MB	24 Ago 2004 / 03:08
Inktomi Slurp	5072	454.09 MB	24 Ago 2004 / 06:15
MSIECrawler	3198	39.75 MB	22 Ago 2004 / 16:36
Mirago	1830	475.32 MB	23 Ago 2004 / 11:54
AskJeeves	1066	16.18 MB	21 Ago 2004 / 14:59
Turn It In	947	23.57 MB	23 Giu 2004 / 14:39
Alexa (IA Archiver)	942	23.37 MB	21 Ago 2004 / 07:18
Openfind data gatherer	589	12.93 MB	22 Ago 2004 / 00:16
Altri	2264	33.41 MB	

\* I robots elencati hanno generato accessi o traffico “non visualizzato” dai visitatori, pertanto non vengono inclusi negli altri rapporti.

La tabella 4 indica che il *sito* è frequentemente visitato da *robot* e *spider*, ovvero quei dispositivi che ne curano l'indicizzazione all'interno dei motori di ricerca. Le notifiche di «Ultima visita» variano infatti da un minimo di poche ore prima ad un massimo di tre giorni prima rispetto all'istante di consultazione del rapporto (24 agosto 2004).

**Tabella 5**

<b>Tipi di file</b>				
<b>Tipi di file</b>	<b>Accessi</b>	<b>Percentuale</b>	<b>Banda usata</b>	<b>Percentuale</b>
 gif Image	474629	48.1 %	1.66 GB	18 %
 jpg Image	280107	28.4 %	2.69 GB	29.3 %
 htm HTML or XML static page	84255	8.5 %	1019.61 MB	10.8 %
 css Cascading Style Sheet file	47593	4.8 %	32.17 MB	0.3 %
 php Dynamic Html page or Script file	32128	3.2 %	1.19 GB	13 %
 pdf Document	17210	1.7 %	1.47 GB	15.9 %

	Js	Javascript file	15395	1.5 %	5.00 MB	0 %
	Sconosciuti		11948	1.2 %	240.52 MB	2.5 %
	html	HTML or XML static page	10778	1 %	79.72 MB	0.8 %
	Swf	Macromedia Flash Animation	8138	0.8 %	149.98 MB	1.5 %
	Doc	Document	1194	0.1 %	304.06 MB	3.2 %
	Png	Image	1019	0.1 %	10.01 MB	0.1 %
	xml	HTML or XML static page	327	0 %	168.64 KB	0 %
	PI	Dynamic Html page or Script file	276	0 %	9.42 MB	0.1 %
	Mp3	Audio	255	0 %	10.33 MB	0.1 %
	Ppt	Document	214	0 %	125.82 MB	1.3 %
	Txt	HTML or XML static page	193	0 %	67.10 MB	0.7 %
	Zip	Archive	150	0 %	106.22 MB	1.1 %
	Xls	Document	73	0 %	50.55 MB	0.5 %
-	Jar		59	0 %	2.05 MB	0 %
-	Ivr		25	0 %	18.16 KB	0 %
	Rm	Audio	7	0 %	7.17 MB	0 %
-	Tif		6	0 %	3.35 MB	0 %
-	wmf		2	0 %	0	0 %
-	com		1	0 %	4.46 KB	0 %
-	emz		1	0 %	5.37 KB	0 %
	Exe	HTML dynamic page or Binary runtime	1	0 %	2.96 MB	0 %

**Tabella 6**

<b>Pagine-URL (Prime 10)</b>				
<b>2696 pagine-url diverse</b>	<b>Accessi</b>	<b>Dimensione media</b>	<b>Pagine iniziali</b>	<b>Pagine d'uscita</b>
/	7263	4.71 KB	3394	1193
/venezia/immagini/satellite/immagini_da_satellite.htm	6251	11.09 KB	4106	2494
/iv/presentazione/soci/soci_trova.php	5623	29.09 KB	3581	3252
/venezia/divulgazione/valli/index.php	4221	17.63 KB	1309	1246
/venezia/immagini/foto_aeree/foto_aeree.htm	3998	7.62 KB	1446	899
/iv/presentazione/soci/soci_ricerca.php	3376	53.27 KB	118	189
<u>/iv/mappa/mappa.htm</u>	2598	21.63 KB	175	362
/venezia/home_bda.htm	2548	8.39 KB	564	178
/news_oriz.swf	2373	19.21 KB	8	363
/venezia/divulgazione/glossario/scheda.php	2355	4.87 KB	1365	1156
Altri	126635	34.24 KB	29301	34032

Le tabelle 5 e 6 sono indicatrici del fatto che in testa agli accessi si posiziona la pagina principale di ingresso a tutti e tre i settori del *sito* (<http://www.istitutoveneto.it/>). Seguono per importanza le pagine della «*Banca Dati Ambientale sulla Laguna di Venezia*» relative alla sezione *Immagini* (2° posto per le *Immagini da Stellite* e 5° posto per le *Foto Aeree*). Al 3°, 6° e 7° posto si collocano invece le pagine del *sito* istituzionale dell'IVSLA, cui seguono nelle preferenze le pagine dinamiche della sezione *Divulgazione* della Banca Dati «Ambiente Venezia» (4° posto per le *Valli da pesca* e 10° posto per il *Glossario*). Conoscere le pagine più richieste stimola a curarne con particolare attenzione contenuti e forma, inoltre aiuta a comprendere gli interessi dell'utenza, le cui inclinazioni vanno senz'altro seguite, ma anche indirizzate, qualora si intenda fornire uno strumento sia utile che formativo.

**Tabella 7**

Provenienza delle connessioni				
Provenienza	Pagine	Percentuale	Accessi	Percentuale
<b>Accessi diretti o via segnalibro</b>	42562	50.1 %	107097	68.6 %
<b>Accessi da un NewsGroup</b>				
<b>Accessi da motore di ricerca</b>	35868	42.2 %	40237	25.7 %
<b>Accessi da pagina esterna (altri siti eccetto i motori di ricerca)</b>	6227	7.3 %	8362	5.3 %
<b>Accessi di origine sconosciuta</b>	151	0.1 %	310	0.1 %

La tabella 7 fornisce il resoconto sul tipo di provenienza delle connessioni. Si osserva che per la metà (50,1%) si tratta di accessi diretti, ovvero di utenti che conoscono il *sito* e vi approdando intenzionalmente. Una buona percentuale (42,2 %) viene invece indirizzata dai motori di ricerca: questo fattore è sicuramente positivo in quanto è indice di un buon “grado di classifica” nelle liste di indicizzazione di questi dispositivi. Oltre il 7% deriva invece da rimandi *ad hoc* volutamente inseriti nelle proprie pagine da gestori di altri *siti*. Anche questo è un buon segnale in quanto notifica l'esistenza di diversi interlocutori che hanno ritenuto le pagine proposte dall'IVSLA meritevoli di considerazione. Le tabelle 8 e 9 forniscono ulteriori dettagli.

**Tabella 8**

Accessi da motore di ricerca	Pagine	Accessi
- Google	30356	34318
- Virgilio	3062	3079
- Google (Images)	782	1055
- Yahoo	759	845
- MSN	288	288
- AltaVista	207	223
- Unknown search engines	137	137
- Netscape	62	62

- Tiscali	54	54
- AOL	39	39
- Altri	122	137

**Tabella 9**

<b>Accessi da pagina esterna (altri <i>siti</i> eccetto i motori di ricerca)</b>	<b>Pagine</b>	<b>Accessi</b>
- <a href="http://www.tuttogratis.it/viaggi/immagini_satellitari.html">http://www.tuttogratis.it/viaggi/immagini_satellitari.html</a>	2257	2257
- <a href="http://arianna.libero.it/search/abin/search">http://arianna.libero.it/search/abin/search</a>	692	692
- <a href="http://arianna.libero.it/search/abin/google">http://arianna.libero.it/search/abin/google</a>	485	485
- <a href="http://www.devilfinder.com/find.php">http://www.devilfinder.com/find.php</a>	260	260
- <a href="http://www.lapnase.com/venise/vues_du_ciel/webcam_gd_canal.html">http://www.lapnase.com/venise/vues_du_ciel/webcam_gd_canal.html</a>	235	235
- <a href="http://www.calshop.biz/pianeta_terra/satellite.html">http://www.calshop.biz/pianeta_terra/satellite.html</a>	218	218
- <a href="http://www.gogle.it/search">http://www.gogle.it/search</a>	118	118
- <a href="http://www.thebestofnet.com/cgi-bin/topsecret.cgi">http://www.thebestofnet.com/cgi-bin/topsecret.cgi</a>	88	88
- <a href="http://ospiti.cilea.it/vertebrati/doc/riviste.htm">http://ospiti.cilea.it/vertebrati/doc/riviste.htm</a>	86	86
- <a href="http://www.tideproject.org">http://www.tideproject.org</a>	81	81
- Altri	1707	3842

**Tabella 10**

<b>Frase cercate (Prime 10)</b>		
<b>16636 frasi chiave diverse</b>	<b>Ricerche</b>	<b>Percentuale</b>
litosfera	1405	4.1 %
foto aeree	1237	3.6 %
immagini satellite	610	1.7 %
Igm	527	1.5 %
immagini da satellite	438	1.2 %
Mappa concettuale	278	0.8 %
Istituto veneto	270	0.7 %
technital	268	0.7 %
Arpav	265	0.7 %
laguna di venezia	255	0.7 %
Altre frasi	28584	83.7 %

**Tabella 11**

<b>Parole cercate (Prime 10)</b>		
<b>14317 parole chiave diverse</b>	<b>Ricerche</b>	<b>Percentuale</b>
di	3112	3.4 %
venezia	2985	3.2 %
litosfera	1777	1.9 %
foto	1736	1.9 %
satellite	1442	1.5 %
aeree	1395	1.5 %
immagini	1362	1.4 %
veneto	1228	1.3 %
laguna	1093	1.1 %
istituto	1067	1.1 %
Altre parole	73941	81.1 %

Le tabelle 10 e 11 riportano gli ingressi derivanti dall'inserimento, da parte degli utenti, rispettivamente di frasi o parole chiave all'interno dei motori di ricerca. Ne emerge che la

maggior parte degli accessi non è relegata ad un ristretto numero di alternative altamente preferite, ma è al contrario estremamente variegata. Questo è un dato sicuramente positivo poiché significa che le occasioni di giungere al *sito* sono molteplici e non limitate a pochi canali privilegiati. Ciò garantisce una maggior visibilità generale, requisito indispensabile per una struttura che si prefigge intenti divulgativi.

Tra le procedure utili a monitorare la visibilità di un *sito*, rientrano proprio le consultazioni periodiche dei motori di ricerca, mirate ad individuare qual è il livello di indicizzazione raggiunto attraverso frasi e/o parole chiave opportunamente scelte. Per quanto riguarda la «*Banca dati Ambientale sulla Laguna di Venezia*», si è riscontrato che i principali canali, attraverso i quali il *sito* viene raggiunto da chi utilizza questo genere di dispositivi, sono le parole chiave: i) *Atmosfera*; ii) *Litosfera*; iii) *Biosfera*. Poiché i risultati ottenuti, anche di recente, sono decisamente soddisfacenti, non si ritiene di dover adottare particolari strategie per aumentare il livello in classifica nelle liste di indicizzazione.