

## Come leggere le statistiche di traffico di un sito Web

Sono una preziosissima fonte di informazioni sui gusti dei propri visitatori, per poterli conoscere meglio. Le statistiche web sono sicuramente uno degli elementi principali per la verifica di ogni piano marketing legato ad un sito web, non solo ci sapranno indicare quali e quanti visitatori sono passati dal nostro sito web, ma ci possono dare importanti informazioni sulle preferenze di navigazione degli utenti del sito web indicandoci quali percorsi hanno fatto sul sito, quanto tempo restano su determinata pagina:

Di seguito trovate i dati fondamentali che possono emergere da una statistica

### Average page view duration

è il tempo medio speso da un utente unico su un'unica pagina del sito. Può essere calcolato in due modi:

o dividendo il tempo totale speso da un utente sul sito per il numero di page view registrate per quello stesso utente nel periodo considerato (ad esempio un giorno);  
o facendo la stessa operazione, ma con la differenza di considerare solo le page view la cui durata sia inferiore al timeout di sessione impostato. Nel presente caso, se il timeout è ad esempio di 40 minuti, una richiesta di pagina a cui non ne seguano altre da parte dello stesso utente per altri 40 minuti viene scartata: si considera cioè come una pagina che l'utente non sta più guardando.

Average page views per visit (o per utente session, numero medio di pagine viste per visita)

Il valore si ottiene dividendo il numero complessivo di pagine richieste da un utente unico per il numero di visite effettuate dallo stesso utente nel corso del tempo considerato. Incrociando i dati ottenuti per questo parametro con quelli relativi al tempo medio per visita, è possibile supporre il comportamento-tipo degli utenti del sito. Ad esempio, una media di poche pagine viste per utente, insieme ad una lunga durata media delle visite registrate, potrebbe dimostrare che i visitatori del sito trovano con relativa facilità ciò che cercano e leggono a fondo i contenuti reperiti. Diversamente, una media di molte pagine viste in veloce successione nel corso di poche e brevi visite potrebbe indicare che la

struttura del sito è caotica, che gli utenti non riescono a trovare ciò che stanno cercando e che pertanto non sono invogliati a ritornare. Chiaramente queste supposizioni devono essere avanzate a ragion veduta: ossia facendo la tara di tutti i possibili fattori di incertezza dei dati statistici rilevati.

Average time per visit (o average session duration)

è il tempo medio speso da un utente unico per una visita al sito. Il valore si ottiene dividendo il tempo totale speso dal utente sulle pagine del sito nel periodo considerato per il numero di visite che ha eseguito nello stesso periodo di tempo. Dal momento che il numero di visite effettuate da un utente unico in un certo corso di tempo dipende dal criterio arbitrario della durata di sessione (cfr. la voce session timeout), è chiaro che anche la durata media di una visita risulta influenzata dal valore assunto da questo criterio.

Browser used (Browser usati)

è la classifica espressa in valori percentuali dei browser utilizzati dagli utenti che si collegano ad un sito. Questa informazione è utile particolarmente ai responsabili tecnici, per regolare al meglio la struttura delle pagine e la presentazione dei contenuti, in maniera che siano navigabili per mezzo di ognuno dei vari tipi di browser che risultano presenti in questa classifica. Se, ad esempio, analizzando l'elenco dei browser utilizzati, si scopre che una discreta parte di visitatori utilizza un browser non compatibile con alcune soluzioni tecniche implementate sul sito, sarebbe appropriato ricalibrare le pagine in modo da renderle accessibili anche alla parte di utenza penalizzata dalle precedenti scelte tecniche. Va in ogni caso precisato che la verifica pratica di quali tipi di browser si colleghino alle pagine di un sito non dovrebbe aver nessuna importanza, se quelle pagine sono state codificate fin dall'inizio nei confronti di linguaggi standard per il Web definiti dal W3C.

Hit (richiesta di accesso)

Costituisce un hit qualunque richiesta di file pervenuta ad un server web. In tal modo, se una pagina web è costituita da un file HTML e da sei immagini, la visualizzazione completa all'interno della finestra di un browser sia della pagina sia delle immagini in essa contenute corrisponderà alla registrazione di sette hit nell'opportuno file di log del server. Un dubbio comune, consiste nel confondere le richieste di accesso con le pagine veramente caricate: un numero, quest'ultimo, che è in realtà quasi sempre di gran lunga inferiore al numero di hit registrato.

Hourly (daily, weekly, monthly, yearly) pages count

Un grafico o una tabella che mostra la distribuzione oraria delle page view registrate nel corso di una giornata (o la distribuzione giornaliera nel corso di una settimana oppure di un mese, o la distribuzione mensile nel corso di un anno solare).

Least requested, o popular, pages  
(pagine meno richieste)

è il supplemento del parametro most requested pages è cioè la classifica delle pagine meno richieste di un sito in un certo periodo di tempo. Lo studio di tale classifica è adatto per cercare di capire se i pochi accessi registrati per alcune pagine dipendono da contenuti non interessanti o dalla scarsa visibilità di quelle risorse nell'economia generale del sito.

Most common countries

è l'elenco, in ordine decrescente, delle nazioni da cui proviene il maggior numero di accessi ad un sito.

Most common operating systems

è la classifica espressa in valori percentuali dei sistemi operativi più utilizzati. Scorrendo i risultati forniti da questi resoconti si ha di consueto la riprova di quanto sia schiacciante il monopolio ormai raggiunto dalla Microsoft.

Most requested pages (pagine più richieste)

è la classifica, in ordine decrescente, delle pagine che hanno avuto più contatti in un determinato periodo di tempo. In base ai filtri impostati, possono essere considerati in questo elenco alcuni tipi di file - ad esempio HTML e ASP - e non altri. è utile in ogni caso, per i tecnici e per i responsabili editoriali

di un sito, considerare accuratamente la classifica delle pagine più richieste, sia per correggere eventuali problemi di natura tecnica - come un sovraccarico del server web dovuto ad errori di programmazione - sia per correggere problemi di struttura logica del sito: alcune pagine, ad esempio, potrebbero essere al apice della classifica delle più richieste non per i loro contenuti, ma perché sono delle strettoie obbligate da cui passare per raggiungere determinati altri contenuti; altre pagine, viceversa, potrebbero ricevere pochi contatti solamente perché non sono state rese abbastanza visibili agli utenti (vedi least requested pages).

## Page views (pagine viste)

Il numero di pagine viste su un sito è forse l'informazione più importante che le statistiche web possano fornire, ma è anche l'informazione di gran lunga più ambigua e complessa da determinare, sia per la difficoltà di definire univocamente cosa sia una pagina sia per l'impossibilità oggettiva di conoscere il rapporto esatto tra pagine servite e pagine caricate da un utente umano (a causa dell'interferenza di numerose variabili quali proxy, NAT, cache locali, spider, ecc.). L'oggetto principale registrato da un server web come una pagina è il file HTML, che contiene le chiamate a tutti gli altri oggetti - immagini, suoni, elementi multimediali, ecc. - che servono a completare quella cosa piuttosto astratta che è la "pagina". Ma non solo i file HTML hanno diritto ad essere considerati "pagine": anche i file chiamati "dinamici", cioè quelli che contengono elementi di programmazione in grado di creare contenuti differenti a seconda dei casi, sono a buon diritto da considerarsi "pagine". Rientrano in questa categoria i file con estensione ASP, PHP, PHP3, PL e simili. Ma possono rientrarvi - ed è una volontà dell'amministratore di rete impostare di conseguenza opportuni filtri sul server - anche i file TXT, i file RTF, i file DOC, i PDF ed altri ancora. In sostanza: non è per nulla semplice creare una categoria astratta chiamata "pagina", che comprenda alcuni tipi di file e ne escluda altri e che fornisca, allo stesso tempo, un parametro credibile per la valutazione del numero di pagine viste da utenti umani. Specialmente va tenuto presente che la comparazione delle page view registrate per due o più siti differenti può essere un'operazione dai risultati poco attendibili: infatti, pur ponendo come uguali gli strumenti di rilevazione del dato e i filtri impostati, la struttura dei siti - in termini di elaborazione delle pagine e di oggetti in esse presenti o da esse richiamati - può essere motivo sufficiente per creare, nel numero di page view rilevato, uno scarto decisamente superiore (o decisamente inferiore) alla reale differenza nella quantità di pagine viste da visitatori umani su ognuno di essi.

## Request By Organization Type

è la classifica in valori percentuali delle richieste di accesso ad un sito, ordinate in base al tipo di dominio da cui parte la richiesta (COM, NET, ORG, MIL, EDU, GOV, identificativi nazionali).

## Single access pages

è la classifica delle pagine uniche più richieste, visitate in un certo periodo di tempo. Si tratta cioè di quelle pagine che, per motivi che i responsabili di un sito dovrebbero studiare a fondo, suscitano l'interesse degli utenti, ma allo stesso tempo non li invogliano a portare avanti la visita appena iniziata. Potrebbe trattarsi di pagine con contenuti chiusi in se stessi (ad esempio una serie di collegamenti o una recensione), referenziate da altri siti. In tale caso andrebbe studiato il modo per invogliare il visitatore a continuare la navigazione all'interno del sito, ad esempio inserendo nelle pagine "incolpite" dei collegamenti ad altre sezioni con contenuti affini.

### Top directories

è l'elenco, in ordine decrescente, delle directory (di solito solo quelle di primo livello) che hanno ricevuto in tutto più richieste di accesso dagli utenti collegati. Questa classifica tende a dare un'idea dell'importanza reciproca delle sezioni in cui è suddiviso un sito. Poiché questo resoconto abbia un qualche valore conoscitivo, occorre che la costituzione logica del sito sia stata progettata in modo razionale, unendo i vari contenuti, in base alla loro omogeneità, sotto apposite directory.

### Top entry pages

è la classifica, in ordine decrescente, delle pagine iniziali più richieste per ciascuna visita al sito registrata in un certo ciclo di tempo. Di solito al vertice di questa classifica c'è la home page. Se non è così, diventa importante individuare i motivi per cui altre pagine funzionano meglio della home page come ingressi al sito. Questo può essere fatto, ad esempio, analizzando i referrer log, ossia i dati sulla derivazione delle visite, per capire se e da quali altri siti sono referenziate le pagine che si trovano al cima della classifica delle top entry.

### Top exit pages

è la classifica, in ordine decrescente, delle pagine più richieste in un certo periodo di tempo come pagine finali di una visita ad un sito. è, in altre parole, l'elenco delle pagine che sembrano più di tutte stimolare l'utente ad sospendere una visita in corso. Anche qui è importante uno studio dettagliato, alla fine di capire cosa c'è in quelle pagine che spinge i visitatori a lasciare il sito. Molto spesso il motivo è da ricercarsi in una serie di collegamenti diretti ad altri siti; altre volte può trattarsi di un cattivo sviluppo dell'albero di navigazione, che finisce con il portare gli utenti verso pagine-imbuto prive sia di informazioni utili sia di collegamenti

verso altre sezioni del sito.

### Top paths

è la classifica dei più comuni percorsi di navigazione seguiti dagli utenti nel corso delle loro visite ad un sito. Per ogni elemento della classifica vengono forniti normalmente:

la sequenza delle pagine visitate, che costituisce il path (percorso);

- la percentuale delle visite sviluppatesi seguendo quel path, rispetto alle visite totali registrate; il numero delle visite per quel path nel periodo quotato.

### Top referring pages (o URLs)

è la classifica, in ordine decrescente, delle singole pagine che hanno reindirizzato degli utenti verso un sito. è utile che il servizio di rilevazione statistica utilizzato permetta di associare i reindirizzamenti, in tal modo da differenziare quelli provenienti dall'interno del dominio di appartenenza del sito da quelli provenienti dall'esterno.

### Top referring sites

è la classifica, in ordine decrescente, dei siti che hanno reindirizzato il maggior numero di contatti a file presenti su un sito. Di solito un'elevata percentuale di contatti è accoppiata in questa classifica all'etichetta "no referrer": ciò significa che un utente si è collegato direttamente ad una risorsa su un sito, senza esserci arrivato per via di collegamenti. Questo caso si verifica, ad esempio, quando un visitatore, conoscendo l'indirizzo della pagina richiesta sul sito di destinazione, inserisce a mano la URL nella barra degli indirizzi del proprio browser.

### Top Search Engines (Motori di ricerca)

è l'elenco, in ordine decrescente, dei motori di ricerca che hanno generato più contatti al sito. Se il numero totale di contatti generato da motori di ricerca è basso rispetto al numero totale di contatti registrato per un sito, in quella occasione se ne può dedurre che le pagine e i contenuti di questo sito non sono sufficientemente indicizzati dai motori di ricerca. Servirebbe in questo caso effettuare le apposite procedure - molte volte a pagamento e ottenibili in blocco rivolgendosi a società specializzate - per migliorare l'indicizzazione dei contenuti messi in linea. Essere ai vertici delle classifiche generate dai principali motori di ricerca può essere, in realtà, un eccezionale strumento per incrementare il numero di visite ricevute.

#### Top Search Keywords (Parole chiave)

è l'elenco, in ordine decrescente, delle parole chiave con più frequenza utilizzate dai visitatori di un sito nell'interrogare i motori di ricerca; parole chiave che hanno creato, come risultato dell'interrogazione, dei collegamenti e delle conseguenti visite al sito. È fondamentale che un sistema di rilevazione del traffico sia in grado di fornire la classifica delle parole chiave più utilizzate dagli utenti. Studiare con accuratezza questa classifica è infatti molto utile al fine di comprendere che tipo di contenuti gli utenti riescono a trovare sul proprio sito grazie ai motori di ricerca. Per via di esclusione si può poi cercare di definire quali altri contenuti, pur presenti sul sito, non generano contatti mediante i motori di ricerca, e perché.

#### Unique visitor (visitatore unico)

è il visitatore, identificato per mezzo dell'indirizzo IP, che effettua una o più visite al sito nel periodo di tempo per cui si vogliono elaborare i report statistici. Se, nel corso di tempo considerato, il visitatore effettua più visite viene contato una sola volta. La rilevazione del numero di visitatori unici, nel caso di assegnazione di IP dinamici da parte di server DHCP, è soggetta a due limitazioni che vanno tenute presenti nella precisazione dei dati ottenuti.

Può infatti accadere che:

utenti diversi si colleghino in successione con lo stesso IP (e quindi il totale di visitatori unici viene sottostimato)  
lo stesso utente si colleghi più volte di seguito con un IP diverso (quindi viene contato più volte e il totale visitatori risulta sovrastimato).

I report di elaborazione statistica riguardanti i visitatori unici di un sito forniscono, oltre al totale, anche il numero di quelli che vi accedono una sola volta e di quelli che si collegano più volte. La percentuale di questi ultimi sul complessivo dei visitatori e il numero medio di visite per visitatore nel periodo di tempo considerato, rappresentano una valutazione del grado di fidelizzazione degli utenti.

### Visit (visita)

Tutte le richieste ricevute in progressione ininterrotta da un server web, provenienti da un medesimo indirizzo IP. Si considera terminata una visita da parte di un utente identificato se tra due successive richieste di pagina intercorre un tempo più in alto al timeout di sessione impostato: se questo, ad esempio, è impostato su 40 minuti, una richiesta di pagina da parte dello stesso utente che arrivi oltre 40 minuti dopo la precedente richiesta, sarà considerata come la prima pagina vista di una nuova visita al sito. Il totale delle visite che si ricava in questo modo dall'analisi dei file di log per un dato periodo di tempo rappresenta chiaramente un'approssimazione statistica, il cui indice di sicurezza rimane imprecisato, dal momento che non esistono mezzi per sapere - tramite la sola analisi dei log - né quanto siano durate in realtà le singole visite ricevute né se con un medesimo indirizzo IP si sia collegata costantemente la stessa persona oppure una serie di persone differenti.