

UNIVERSITA' DEGLI STUDI DI ROMA

TOR VERGATA



FACOLTA' DI LETTERE E FILOSOFIA

CORSO DI LAUREA IN

Scienze del Turismo Culturale

TESI IN

Informatica e Rappresentazione della Conoscenza

TITOLO

**Coltivare il turismo nei domini semantici: come pubblicizzare i siti
turistici acquistando parole chiave non ovvie su web**

Relatore:
Chiar.mo Prof.
Fabio Massimo Zanzotto

Laureando:
Lorena D'Angelo

Sommario

Introduzione.....	2
Capitolo 1	
Tecniche utilizzate per essere visibili su internet.....	8
1.1 Social Advertising.....	9
1.2 Tecniche SEO	14
1.3 AdWord:acquisto di spazi pubblicitari all'interno di Google	15
1.4 Utilizzare AdWord per essere competitivi sul web.....	20
Capitolo 2	
Aste elettroniche	21
2.1 Il panorama: Il mondo dei metodi di aste on-line	22
2.1.1 Aste <i>multi – item</i>	24
2.1.2 Aste combinatorie(CA).....	25
2.1.3 Il meccanismo Vickrey-Clarke-Groves (VCG)	27
2.2 Caso particolare: Le aste di query in Google.....	27
2.2.1 Sviluppo di un modello commerciale delle aste	30
2.2.2 L'asta Google degli spot	30
2.2.3 L'importanza della competizione.....	33
Riepilogo.....	34
Capitolo 3	
Utilizzo dei domini semantici nella ricerca sponsorizzata	35
3.1 Definizione di campi e domini semantici in linguistica.....	35
3.2 Migliorare la pertinenza dei risultati dei motori di ricerca attraverso i domini	36
Riepilogo.....	38
Capitolo 4	
Utilizzo dei domini semantici nelle aste di parola chiave.....	39
Capitolo 5	46
Conclusioni	46
Bibliografia	49

Introduzione

L'impatto che Internet ha avuto sul turismo sta cambiando le relazioni commerciali tra gli operatori turistici e gli utenti. La tecnologia di Internet si è rivelata, fin da subito, come portatrice di quella "globalizzazione" di cui oggi si parla tanto e che permette, ad esempio, a qualsiasi "navigatore", potenziale acquirente, indipendentemente dalla propria localizzazione, di accedere via rete a prodotti e servizi offerti da aziende situate in qualsiasi parte del mondo. Ciò ha fatto sì che operatori commerciali e utenti cominciassero ad avvalersi, sempre più spesso di mezzi informatici per la trasmissione e lo scambio di dati ed informazioni, all'interno di quello che viene genericamente denominato *Electronic Commerce* o meglio *Commercio Elettronico*. Uno dei settori economici che è stato in questi anni maggiormente stravolto dalle nuove tecnologie è quello turistico. È infatti ampiamente dimostrato che tale comparto economico abbia trovato in Internet uno strumento così importante da modificare le modalità tecniche e operative delle aziende turistiche e da consentire di migliorare il flusso di informazioni nei confronti del mercato.

Con la nascita del canale Internet sono nati numerosi siti di imprese del settore turistico (agenzie viaggi, tour operator, vettori, comunità locali), in risposta alla sempre maggiore attitudine dei consumatori ad utilizzare questo nuovo mezzo per ricercare informazioni di viaggio e acquistare prodotti turistici on-line. Ecco che a tal proposito è stato coniato il termine "Turismo online" ovvero le modalità con cui il Turismo si presenta e si realizza nella rete implicando una filosofia d'affari completamente nuova. Per affermarsi in questo scenario

competitivo infatti, ciascuna azienda turistica deve adottare un nuovo modello di business ed utilizzare efficienti *customer service* e nuovi canali promozionali. Internet non è semplicemente un nuovo mezzo di comunicazione su cui trasferire i contenuti che precedentemente erano veicolati al consumatore finale da altri media, ma richiede nuove ed appropriate strategie di comunicazione e di relazione con il cliente. I prodotti turistici hanno le caratteristiche ideali per il commercio elettronico: un prodotto complesso come quello turistico è adatto ad essere rappresentato in un sito web, dove possono essere sfruttate completamente le potenzialità della comunicazione multimediale ed ipertestuale (Passantino, 2010) . Il problema che si trovano ad affrontare le aziende turistiche è quello di riuscire ad essere competitivi nel mercato online per poter aumentare così il flusso delle vendite. Per “attirare i clienti” al proprio sito occorre essere visibili sui motori di ricerca; maggiore sarà la visibilità, maggiore sarà la possibilità di guadagno per un’azienda turistica.

I motori di ricerca compiono un’attività continua di scandaglio dell’intero web attraverso specifici software (I cosiddetti *spider*) che esplorano ed acquisiscono le informazioni relative e indicizzano tutte le pagine non ancora presenti nei loro archivi. Seguendo i link presenti nelle pagine web, gli spider passano da un sito all’altro. Una volta avuto accesso a una pagina, ne censiscono i contenuti in base alla rilevanza delle parole contenute in ogni pagina del sito, evidenziando quelle riportate più spesso, che si presume rappresentino l’argomento principale della pagina stessa. Ovviamente il funzionamento di questi software è regolato da algoritmi ben più complessi di quanto riportato qui, ma intuitivamente può essere paragonato al lavoro di un bibliotecario che continua ad acquisire nuovi volume, li legge, ne

cataloga i contenuti tramite parole chiave, inserendo infine riferimenti specifici in apposite schede di ricerca. Un importante elemento tenuto in considerazione dai motori di ricerca nel censimento delle diverse pagine è il numero e la natura dei link che da altre pagine puntano verso quella esaminata. Un ulteriore segno del fatto che sul web non conta solo il contenuto, ma anche le relazioni che il singolo elemento intrattiene con la rete. Il motore di ricerca, dunque, identifica i contenuti presenti nelle pagine e stabilisce la rilevanza di quella pagina rispetto a ciascun contenuto ivi trattato, compilando così un “ranking”, una classifica (Cederle, 2005) . I diversi siti concorrono per avere i primi posti nelle classifiche. L’obiettivo è quello di risultare nella prima pagina proposta dal motore di ricerca all’utente che fa una ricerca rispetto a uno specifico contenuto.

In questa tesi verranno quindi analizzate le tecniche che permettono sostanzialmente di aumentare il volume di affari delle *aziende turistiche*, aumentandone la visibilità e riducendo i costi. Per posizionarsi ai primi posti tra i risultati di un motore di ricerca esistono principalmente due tecniche: Il SEO (*search engine optimization*) e le tecniche a pagamento. Il SEO consiste in tecniche che permettono di costruire pagine web più facilmente leggibile e riconoscibile dai motori di ricerca, in modo che il contenuto della pagina sia più chiaramente compreso e il ranking risultante più alto; i siti verranno promossi dai motori di ricerca in base a diversi parametri, quali: *contenuti testuali, Keyword, popolarità*.

I limiti delle tecniche SEO sono: start Up e aggiornamenti lenti, nessuna garanzia sul posizionamento di una parola chiave, elevato budget senza la garanzia di una visibilità a lungo termine. L’utilizzo

delle tecniche SEO richiede dunque budget elevati senza la garanzia di ricavi dilazionati nel tempo.

In questa tesi si propone quindi di risolvere tali problematiche utilizzando un approccio semantico per pubblicizzare i siti turistici, ovvero sfruttare i domini semantici per acquistare parole chiave meno competitive, quindi meno costose, pur assicurandosi i posti migliori tra i link sponsorizzati con ricavi dilazionati nel tempo. Per esempio un inserzionista può decidere di acquistare la parola chiave albero da frutto per far comparire il suo sito che tratta di alimenti per cani tra i link sponsorizzati. Apparentemente la parola chiave non è conciliabile con l'oggetto in vendita, ma in realtà essi sono legati da uno stesso dominio semantico che in questo caso potrebbe essere quello del giardino o della campagna. Ciò comporta un evidente riduzione dei costi in quanto si utilizzano parole chiave non ovvie e quindi poco competitive, pur assicurando visibilità al sito web tra i link sponsorizzati.

Nel primo capitolo della tesi, quindi, verranno analizzate le principali modalità utilizzate per pubblicizzare un sito turistico sul web mediante l'utilizzo di differenti tecniche come: le tecniche SEO, la social advertising e infine l'utilizzo di spazi pubblicitari all'interno dei principali motori di ricerca attraverso l'acquisto di parole chiave nelle *aste di query*, molto importanti per determinare il costo e il vincitore di una determinata parola chiave.

Nel secondo capitolo verrà descritto il mondo delle aste online e in particolar modo verrà analizzata l'asta GSP, utilizzata dai principali motori di ricerca come Google, per l'acquisto di parole chiave da parte degli inserzionisti per la ricerca sponsorizzata, riportando uno specifico esempio per spiegare il meccanismo mediante il quale una

query viene messa all'asta ed evidenziare l'importanza dei domini semantici per la scelta delle parole chiave più strategiche da parte degli inserzionisti e maggiormente utilizzate dagli utenti.

Il terzo capitolo è dedicato all'utilizzo dei domini semantici per la ricerca sponsorizzata. Verranno definiti i domini semantici in linguistica e verrà mostrato come diversi studi hanno portato all'ipotesi di una possibile applicazione dei domini nell'ambito della ricerca sponsorizzata con evidenti vantaggi dal punto di vista computazionale, come la riduzione delle ambiguità lessicali, la discriminazione del senso di una parola e il miglioramento della pertinenza dei motori di ricerca.

Nel quarto capitolo si vuole proporre un approccio semantico per individuare nuove parole chiave per la ricerca sponsorizzata attraverso l'uso dei domini semantici, in quanto i metodi esistenti si basano su tecniche di abbinamento, ovvero forniscono solo suggerimenti che contengono la radice o l'intera parola chiave indicata, mentre i suggerimenti che non la contengono non sono mostrati. Ciò comporta suggerimenti molto competitivi e quindi costosi. Per esempio se si vogliono trovare suggerimenti per la parola chiave "isole Galapagos" il programma di Google AdWord propone "tour isole Galapagos" il cui costo per clic è di 5 euro, mentre utilizzando i domini semantici un possibile suggerimento potrebbe essere "Santa Cruz" il cui costo per clic scenderebbe a 0,35 centesimi (Alpa & Pennacchiotti, 2010) . Il suggerimento è attinente alla sfera semantica della parola chiave e non alla somiglianza sintattica permettendo una sostanziale riduzione dei costi.

Infine verrà mostrato come i suggerimenti ottenuti utilizzando i domini semantici potrebbero essere preziosi per gli inserzionisti nel

settore del turismo, in quanto gli permetterebbero di aumentare la visibilità dei loro siti turistici spendendo meno attraverso l'acquisto di parole chiave non banali e meno costose; maggiore sarà la visibilità dei siti turistici, maggiore sarà la possibilità di aumentare i guadagni.

Capitolo 1

Tecniche utilizzate per essere visibili su internet

In Italia la pubblicità su internet sta crescendo e per volume di investimenti ormai incalza quella suoi quotidiani, con gli spot tv che restano al primo posto. Come si vede nella fig.1.1 in Inghilterra invece la pubblicità online ha già scavalcato tutti gli altri media. Ormai sono ventisei milioni gli italiani che usano internet, con una crescita di circa il 25 per cento nell' ultimo anno. Sono il 54 per cento uomini, in gran parte in età lavorativa (il 56 per cento nella fascia 25-50 anni), con un reddito (31 mila euro) ed una istruzione più alta della media.

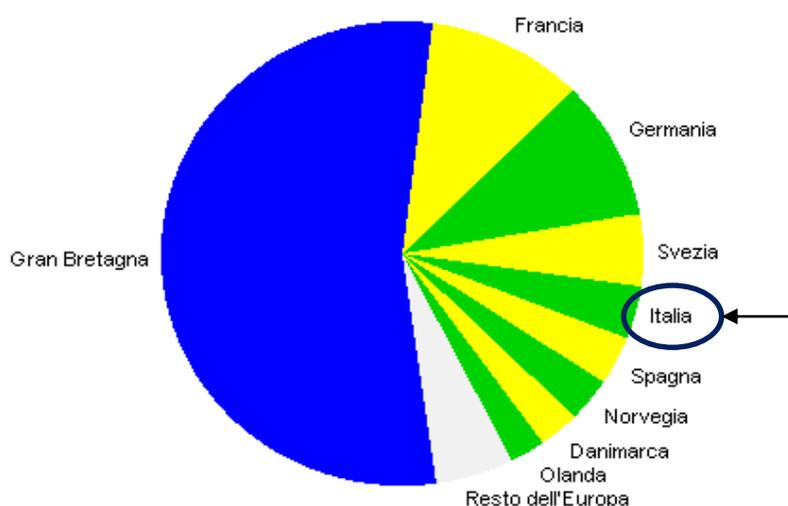


Figura 1.1 - *Percentuale di pubblicità on-line per i 9 paesi europei che ne detengono il 94% del totale.*

Secondo i dati resi noti da Nielsen relativi agli investimenti pubblicitari in Italia nel periodo Gennaio-Luglio 2010 gli investimenti totali su tutti i media sono stati oltre 5 miliardi di euro.

Per quanto riguarda Internet la crescita è stata del 17,8% rispetto ai primi sette mesi del 2009 ed è stato il mezzo con il numero più elevato di inserzionisti registrando un aumento del 24,0% (Nielsen, 2010) . Oltre l'87% delle persone ricerca informazioni con i motori di ricerca prima di un acquisto. Per raggiungere i potenziali clienti, quindi, anche nel settore del turismo occorre prima di tutto che un sito web ottenga un posizionamento nei motori di ricerca ai primi posti con le parole chiave più strategiche e utilizzate dagli utenti. Per fare ciò esistono differenti tecniche; nel primo paragrafo si parlerà dunque della social advertising, nel secondo paragrafo delle tecniche SEO e infine nel terzo paragrafo verrà descritto come avviene l'acquisto di spazi pubblicitari all'interno dei motori di ricerca.

1.1 Social Advertising

Il cosiddetto Social Advertising è un fenomeno recente che consiste nell'impiego di spazi pubblicitari all'interno delle reti sociali come Facebook, il social network ormai divenuto leader indiscusso del settore. Negli ultimi anni l'advertising sul web è stata in continua evoluzione partendo dalla pubblicità generica gestita indistintamente a quella contestuale ai contenuti presenti all'interno della pagina, sino ad arrivare ai giorni nostri allo sviluppo di una vera e propria *social-advertising* che si basa sulla conoscenza delle informazioni relative all'utente stesso. Le azioni sociali sono molto potenti perché agiscono e si basano sulle reti di fiducia che le persone instaurano tra di loro sul web; la social advertising non si basa su messaggi o pubblicità che vengono trasmesse ed erogate dalle aziende, ma si tratta di

informazioni che vengono condivise e co-costruite tra amici e conoscenti. Tali azioni sociali sono dunque la base per la costruzione, non solo di un nuovo modello di comunicazione e relazione, ma di un nuovo modello di marketing e di business.

Ormai il fenomeno dei social network sta cambiando drasticamente il comportamento delle persone che trascorrono il 10% del loro tempo navigando su internet offrendo così nuove sfide e opportunità sia all'industria dei media che a quella pubblicitaria. Nonostante ciò a parte alcune eccezioni la social advertising resta ancora un mezzo poco redditizio. Sono solo agli inizi i tentativi di modificare il modus operandi tradizionale del mercato pubblicitario per sfruttare questa rivoluzione nel comportamento dei consumatori.

I social network forniscono però anche molte opportunità: sfruttando infatti la volontà degli utenti di partecipare alla creazione dei contenuti gli editori possono infatti aumentarne il coinvolgimento e utilizzare i social media per diffondere i propri contenuti presso audience molto più vaste. Non è comunque ancora stata trovata la formula magica per rendere effettivamente redditizio il settore online che sta crescendo di più, ma uno degli ingredienti fondamentali sarà sicuramente che i messaggi pubblicitari dovranno essere diretti e autentici, costruiti sul principio della conversazione e non dell'imposizione, in modo da fornire al consumatore un valore aggiunto (Nielsen, 2009) .

Trattandosi di un contesto diverso, l'applicazione dei modelli di pubblicità standard non può funzionare. Bisognerà inventare e sperimentare modelli innovativi in cui la pubblicità non deve essere invasiva e interrompere l'esperienza del social media, ma partecipare ai discorsi. I messaggi pubblicitari sui social network devono venire da un punto di vista più autentico e sincero. Grazie ai social media, il

passaparola è infatti diventato la forma più autorevole di pubblicità, in un momento in cui l'advertising tradizionale vive una crisi di credibilità senza precedenti. Da una analisi realizzata a dicembre 2007 e a dicembre 2008 attraverso Nielsen BuzzMetrics, il servizio che monitora le conversazioni sui consumer-generated media, è emerso che il termine "falso" è quello più comunemente associato al concetto di "advertising". Alla base dell'esistenza dei social network c'è l'amicizia, il fatto che gli iscritti attraverso l'interazione danno qualcosa l'uno all'altro. Parlando di semplici numeri di visitatori tra gli utenti che si sono aggiunti a Facebook nell'ultimo anno, il maggiore incremento proviene dalla fascia di età 35-49 anni (+21.4 milioni). Inoltre tra i nuovi utenti quelli di età compresa tra cinquanta e sessantaquattro anni sono quasi il doppio di quelli minori di diciotto anni. La conseguenza è che nella composizione dell'audience di blog e social network diminuisce la percentuale rappresentata dagli under diciotto mentre aumenta quella degli over cinquanta. Il cambiamento del profilo degli utenti offre agli inserzionisti la possibilità di utilizzare i social network per raggiungere tutti i gruppi demografici. In Italia, ad esempio, hanno la propria pagina su Facebook brand come Maserati, tradizionalmente rivolti a target più maturi. I social network sono in realtà particolarmente adatti a creare coinvolgimento rispetto al brand attraverso le fan page, che in definitiva sono pubblicità. Facebook è il social network che più di tutti "colleziona" informazioni sul nostro conto: dai gusti alle amicizie, dal lavoro alla professione, dai contatti virtuali a quelli reali, dagli interessi alle attitudini. La social advertising si basa sulla capacità di tracciare ogni profilo presente all'interno della rete e di organizzare i contenuti promozionali a seconda della persona che impiega la piattaforma,

raccogliendo dati e informazioni sull'utenza. Possiamo distinguere due livelli di profilazione:

Profiling implicito: in questo caso la raccolta dei dati avviene senza che l'utente ne sia consapevole.

Profiling esplicito: questo tipo di profilazione si basa sulla richiesta di informazioni del sito all'utente che, inserendo dati, compilando moduli, scrivendo e-mail o contattando persone, rilascia al sito web in questione.

La profilazione non è fine a se stessa, ma serve a identificare, come abbiamo visto, tipologie di utenze specifiche che presentino determinate caratteristiche che siano poi "spendibili" in maniera utile da parte degli interessati. Una volta raccolte, le informazioni vengono analizzate e utilizzate. Ecco quindi che se un'azienda vorrà pubblicizzare un prodotto per i giovani, andrà a collocarsi in una fascia under venticinque; o se per esempio un'agenzia viaggi dovesse promuovere qualche offerta speciale per una location esotica, andrebbe a selezionare un tipo di utenza che ha espresso preferenze rispetto a quella località e/o indirizzerebbe i suoi sforzi sul tipo di utenza che riterrebbe statisticamente significativo per la propria promozione. Un discorso è fare pubblicità indistintamente verso chiunque; un altro caso è conoscere il panorama di utenze di cui si dispone, individuare il proprio target specifico di riferimento e indirizzare i propri sforzi in questo senso. Questo tipo di approccio per le aziende comporta un notevole vantaggio sia in termini di efficacia che in termini di efficienza: nel primo caso perché c'è maggiore probabilità che le utenze contattate possano essere interessate al prodotto o al servizio venduto/erogato; nel secondo perché il rapporto costi/benefici è senza dubbio minore rispetto a quello possibile per

una campagna di marketing tradizionale che porti gli stessi effetti. Accanto alla già citata piattaforma specifica di Facebook per l'advertising posta a lato della pagina e segnalata come tale, esistono tutta una serie di dinamiche pubblicitarie che si muovono all'interno della rete sociale senza che le utenze ne siano consapevoli. Pensiamo semplicemente alle pagine "Fan di...", ai gruppi o al fatto che qualunque sviluppatore possa creare un'applicazione personalizzata da diffondere all'interno di Facebook. Le pagine sono gratuite e potenzialmente realizzabili da chiunque, anche se, come si legge nel disclaimer tale creazione dovrebbe spettare solo ad un rappresentante autorizzato di tale marca, organizzazione o persona. Le pagine, rispetto ai gruppi, prevedono un'elevata personalizzazione, vengono indicizzate da Google e offrono un servizio di monitoraggio delle visite e delle iscrizioni davvero utile e ben fatto. La creazione di una pagina è davvero semplice e non richiede nessuna particolare competenza: è sufficiente recarsi all'indirizzo <http://www.facebook.com/pages> e seguire la procedura guidata. Per quanto riguarda i gruppi di Facebook, essi possono ruotare attorno ad un oggetto specifico o essere più generici, ma di sicuro sono dei luoghi privilegiati per il *Buzz marketing*, tant'è che moltissime aziende di comunicazione ne tengono sotto stretto controllo le dinamiche interne. Aprire un gruppo di promozione della propria identità non ha nessun costo, mantenerlo correttamente e pubblicizzarlo all'interno della rete è, però, un altro discorso. Certo è che un gruppo numeroso e ben nutrito è un'ottima chance per i suoi amministratori di contattare un buon numero di persone e di soggetti potenzialmente interessati. La forza di tali strumenti, come facilmente intuibile è l'estrema facilità che possiedono di raggiungere in brevissimo tempo un numero

enorme di utenze. Come abbiamo visto in passato è proprio su questo che si basano le nuove logiche di marketing. Non solo pagine e gruppi, che pur rappresentano i due veicoli principali, data anche la loro enorme semplicità di creazione e di fruizione (decisamente alla portata di qualunque utente), ma anche eventi, applicazioni, sondaggi. L'applicazione di Facebook per creare eventi, permette infatti di segnalare party di inaugurazione, congressi, fiere, meeting, insomma: tutto ciò che può essere direttamente o indirettamente collegato con il nostro *brand*. Di natura un po' differente sono le applicazioni già presenti in Facebook sia quelle che vengono sviluppate direttamente dalle aziende interessate sfruttando le API messe a disposizione dal social network. All'interno di questa categoria, possiamo segnalare sicuramente Make a Quiz, attraverso la quale è possibile creare quiz personalizzati da diffondere attraverso il proprio profilo; Quanto analizzato ci pone di fronte a questioni non indifferenti: per il consumatore lo "spazio pubblicitario" non è più qualcosa di definito e statico, ma rappresenta piuttosto un contorno (se non alle volte addirittura il contenuto stesso) sempre visibile e sempre presente all'interno delle differenti dinamiche comunicative (Besana, 2009) .

1.2 Tecniche SEO

Il SEO consiste in tecniche che permettono di costruire pagine web più facilmente leggibile e riconoscibile dai motori di ricerca, in modo che il contenuto della pagina sia più chiaramente compreso e il ranking risultante più alto; i siti verranno promossi dai motori di ricerca in base a diversi parametri, quali: *contenuti testuali*, più

l'argomento ricercato dall'utente viene trattato sul sito, e più il motore di ricerca spingerà il sito verso i primi posti della lista. È dunque consigliato produrre considerevoli quantità di testo e trattare qualunque argomento in maniera estesa e approfondita; *Keyword*, ovvero nel realizzare un sito web bisogna porre attenzione al fatto che nei testi delle pagine siano presenti anche quelle parole-chiave che presumibilmente gli utenti useranno come termini di ricerca sui motori; *popolarità*, calcolata in base al numero di link sparsi per il web che puntano ad esso. Più sono i link che puntano al sito (in un certo senso "consigliandolo") e più il sito è considerato popolare (Cascioli, 2010) .

I limiti delle tecniche SEO sono: start Up e aggiornamenti lenti, nessuna garanzia sul posizionamento di una parola chiave, elevato budget senza la garanzia di una visibilità a lungo termine. L'utilizzo delle tecniche SEO richiede dunque budget elevati senza la garanzia di ricavi dilazionati nel tempo.

1.3 AdWord:acquisto di spazi pubblicitari all'interno di Google

Migliaia di inserzionisti utilizzano il programma AdWords di Google per pubblicizzare i propri prodotti e servizi sul Web con annunci mirati per guadagnare nuovi clienti. Come si può vedere nella fig.2.1 Google vende il suo ospitare pubblicità, che è dichiarata e segnalata dalla scritta "collegamenti sponsorizzati", presente in riquadri accanto alla lista dei primi dieci siti. I "collegamenti sponsorizzati" occupano due zone della schermata di Google. La prima è quella su fondo giallo (a volte celeste) in alto, prima dei dieci indirizzi consigliati; qui

possono essere visualizzati uno o due siti. L'altra è sulla destra del monitor, tutte e due le zone sono precedute appunto dalla scritta "collegamenti sponsorizzati".

SEO (risultati "naturali")

Risultati sponsorizzati

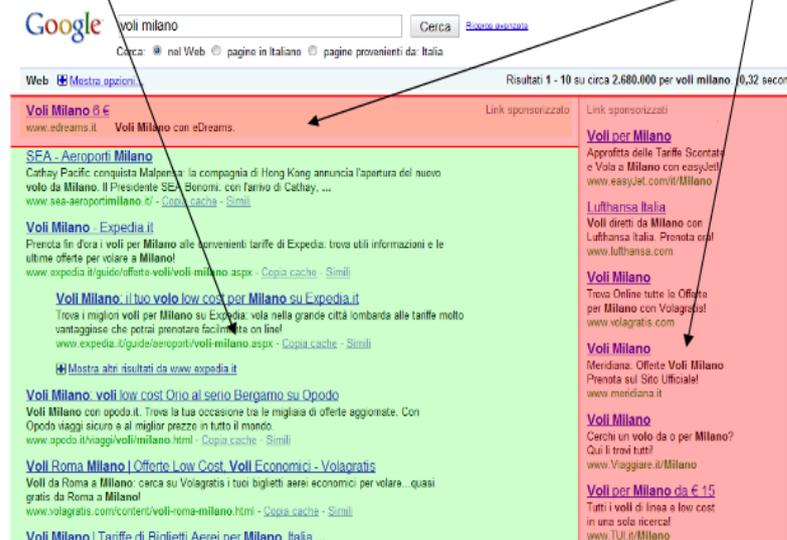


Figura 2.1: Spazi pubblicitari venduti da Google.

Attraverso AdWords, il sistema che gestisce le pubblicità sul motore, Google permette di inserire un annuncio all'interno dei collegamenti sponsorizzati. Il motore di ricerca garantisce il posizionamento in testa ai risultati pubblicitari per una certa parola chiave acquistata all'asta, dietro il pagamento per ogni clic. Molta gente pensa che il motore venda l'essere primi nei risultati, che si paga per ottenere un'alta indicizzazione dei siti internet sui motori di ricerca, mentre non è così. Non è possibile ottenere a pagamento l'inserimento nei "primi posti consigliati" da Google, ma solo acquistare gli spazi pubblicitari "collegamenti sponsorizzati" a destra dei risultati. AdWords di Google usa "parole chiave" selezionate e "comperate" dalle aziende per

distribuire annunci pubblicitari in maniera mirata agli utenti che sono alla ricerca di informazioni su un particolare servizio. Il programma è basato su due elementi: costo-per-clic (CPC) e performance dell'annuncio. Le offerte si basano su un sistema *pay per clic*, in base al quale ciascun partecipante all'asta specifica l'importo che è disposto a pagare ogni volta che un utente digiterà il suo annuncio, quindi sceglie un importo medio giornaliero, ad esempio 100 euro al giorno, per un totale di 3.000 mensili. Quando un utente compie una ricerca su Google o su uno dei motori collegati utilizzando una parola "acquistata" da un inserzionista, accanto ai risultati "naturali" compaiono degli annunci chiaramente definiti "collegamenti sponsorizzati". Ogni volta che un utente clicca sul collegamento, lo sponsor paga a Google una cifra variabile, da pochi centesimi a diversi euro, a seconda della richiesta e della concorrenza esistente per quella parola-chiave. Gli inserzionisti stabiliscono il prezzo che sono disposti a pagare per ogni clic fatto sui loro annunci, e pagano solo quando il loro annuncio viene cliccato.

Google visualizza gli annunci con una frequenza sufficiente per poter garantire che il *click-through* (ovvero il numero di persone che vedono l'annuncio e fanno clic su esso) non superi il budget giornaliero. L'annuncio viene mostrato finché non riceve un numero di cliccate degli utenti che esauriscono il suo budget giornaliero di 100 euro. In base a quanto è cercata la frase e a quanto è cliccata il budget si può esaurire in un'ora o in dieci. Gli inserzionisti si alternano durante le 24 ore, in modo che non ci siano mai più di sei o sette banner nella colonnina delle pubblicità del motore. All'inserzionista il costo della singola cliccata sopra il suo banner è determinata da un'asta fra tutti i gestori dei siti che aspirano alle prime posizioni per quella parola

chiave. Avere il proprio banner in alto, evidenziato da un fondo giallo o celeste, è un premio che Google dà agli inserzionisti migliori. Google sposta a rotazione gli annunci dalla posizione di destra a quella in alto su fondo colorato, secondo criteri suoi di profittabilità legati a: numero di click ricevuti; CPC impostato (costo per contatto; quanto è disposto a pagare lo sponsor per ogni click); budget massimo per la campagna. In sostanza, dato che Google sa che quelle in alto sono le posizioni che danno più click, le riserva a chi, secondo lui, gli farà guadagnare di più.

Come si vede in fig.2.2 Google quando classifica i collegamenti sponsorizzati, valuta anche i click-through degli annunci, mostrando per primi quelli con le valutazioni più elevate. Di conseguenza, non è sufficiente fare offerte più alte di altri per garantirsi la prima posizione infatti quando Google decide quale annuncio viene visualizzato per primo tra i collegamenti sponsorizzati, tiene conto delle offerte e del numero di clic fatti dagli utenti su ogni annuncio, dando la precedenza agli annunci più efficaci. Pertanto, non è possibile acquistare una posizione.

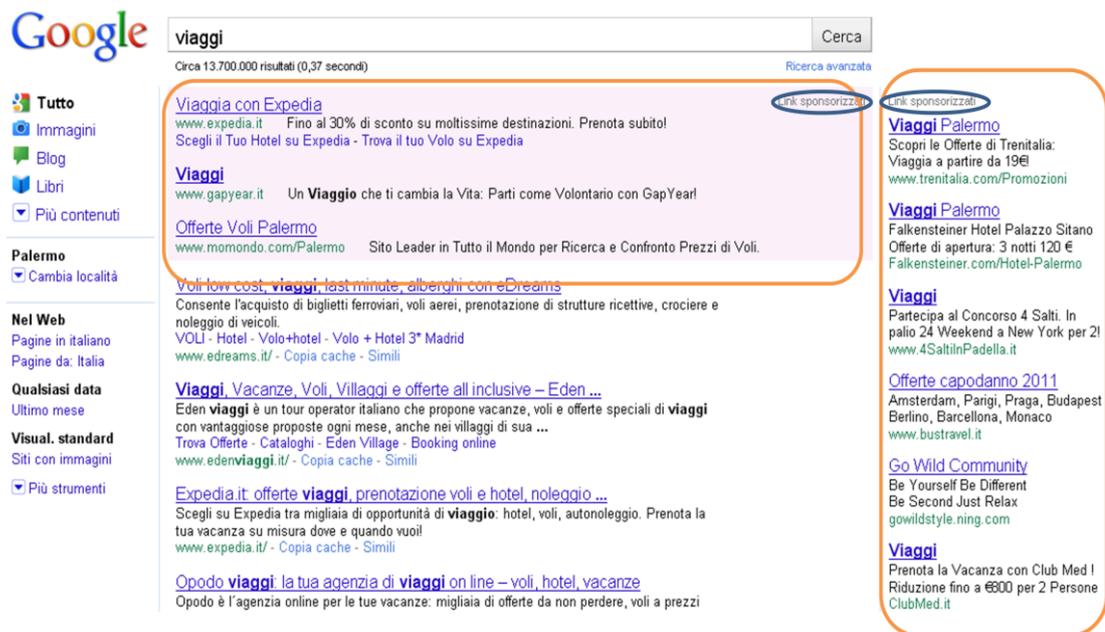


Fig. 2.2: *Link sponsorizzati da Google, tra i quali (in alto al centro), quelli che hanno ricevuto il maggior numero di clic.*

In effetti, se il proprio click-through è superiore a quello dei concorrenti, ci si potrebbe ritrovare in alto anche con un'offerta inferiore. Questa procedura potrebbe sembrare insolita ma è efficace. Dopotutto se il commerciante che ha pagato un euro genera cinque volte i click-through di quello che ha pagato due euro, dal punto di vista economico ha un senso dargli la prima posizione: sta facendo guadagnare più soldi a Google, che prende una percentuale per ogni click. I primissimi due posti, quelli su fondo colorato sono anch'essi "collegamenti sponsorizzati", ma Google abbastanza subdolamente li fa comparire sopra ai primi dieci. Quei due posti sono oggettivamente di grande suggestione e visibilità, e senz'altro ottengono più click dei altri "collegamenti sponsorizzati" nella colonna di destra (Cascioli, 2010).

I vantaggi di una campagna su AdWord di Google sono:

1. Ridotto tempo di attivazione;

2. Possibilità di controllo del proprio budget;
3. Possibilità di verifica delle performance per ogni keyword prescelta;

I principali svantaggi sono:

1. Per alcune parole chiave il costo del click è decisamente alto.
2. la campagna diviene molto onerosa se protratta nel tempo.
Il “Pay per click” è come una droga. Appena smetti di pagare Google i tuoi visitatori calano vertiginosamente.

1.4 Utilizzare AdWord per essere competitivi sul web

Nei paragrafi precedenti sono state analizzate le principali tecniche per pubblicizzare i siti turistici online, ovvero : la social advertising, le tecniche SEO e l'utilizzo dei link sponsorizzati all'interno dei motori di ricerca come Google. Sono state mostrate le caratteristiche di ciascuna modalità e alla luce degli svantaggi che comporta l'utilizzo della social advertising e delle tecniche SEO, nella presenti tesi si è scelto di sperimentare l'utilizzo di tecniche a pagamento, mediante l'acquisto di spazi pubblicitari offerti in vendita dai motori di ricerca per verificare se è possibile ottimizzare i costi acquistando parole non ovvie nelle aste online di query e quindi meno competitive. Nel prossimo capitolo verrà dunque presentata una panoramica delle aste elettroniche, concentrando l'attenzione sulle aste GSP, utilizzate dai principali motori di ricerca come Google per l'acquisto di parole chiave da parte degli inserzionisti per la ricerca sponsorizzata.

Capitolo 2

Aste elettroniche

Le aste stanno ottenendo uno strepitoso successo sia nel settore del commercio tra imprese, il cosiddetto *business-to-business*, sia nel commercio tra imprese e consumatori finali (*business-to-consumer*), esse hanno infatti assunto un ruolo predominante come strumento di *pricing* dinamico. Attraverso il Web fornitori e banditori d'aste possono rivolgersi ad un pubblico sempre più vasto, offrendo prodotti e servizi di qualsiasi tipo, anche di basso valore. Un'asta consiste fondamentalmente nel processo di vendere e acquistare beni in situazioni di indeterminazione dei prezzi. Questo consiste principalmente in tre momenti:

- mostrare i beni;
- ricevere le offerte;
- assegnare i beni al miglior offerente.

Nel primo paragrafo dunque verranno descritte le principali tipologie di aste online, esaminando nel primo sottoparagrafo una nuova categoria di asta chiamata asta multi-item; nel secondo sottoparagrafo verranno definite le aste combinatoriali e nel terzo sottoparagrafo verrà descritto il meccanismo Vickrey-Clarke-Groves, l'asta combinatoriale più famosa.

Nel secondo paragrafo verrà analizzato un caso particolare, ovvero come avvengono le aste di query in Google, descrivendo nel primo sottoparagrafo quando è stato sviluppato da Google un modello commerciale delle aste, nel secondo sottoparagrafo l'asta di Google

degli spot e infine nel terzo sottoparagrafo l'importanza della competizione all'interno di un'asta.

2.1 Il panorama: Il mondo dei metodi di aste on-line

La teoria delle aste è molto studiata nell'ambito della micro-economia poiché le aste sono uno strumento fondamentale per l'implementazione di transazioni in ambito concorrenziale.

Alcuni tipi più comuni di asta sono i seguenti:

- *Asta inglese*: in questo tipo di asta ciascun partecipante invia pubblicamente le proprie offerte una dopo l'altra in modo tale che il prezzo di ogni offerta è superiore a quello della precedente. L'asta termina quando nessun partecipante vuole più fare offerte o quando viene raggiunto un certo prezzo soglia. Il venditore, inoltre ha la facoltà di poter scegliere un certo prezzo di riserva che, nel caso non venga raggiunto, può comportare l'annullamento dell'intero processo di vendita e quindi dell'asta.

- *Asta olandese*: questa tipologia di asta inizia con un prezzo alto che viene poi ribassato dal banditore fino a quando un partecipante non lo accetta. Il partecipante paga l'ultimo prezzo annunciato.

- *Sealed-bid con first-price*: in questo tipo di asta tutti i partecipanti sottopongono contemporaneamente le proprie offerte in modo che nessuno conosca le offerte degli altri. L'offerente che ha presentato l'offerta più alta si aggiudica l'asta e paga il prezzo da lui sottoposto.

- *Sealed-bid con second-price* (conosciute anche come aste di Vickrey): questo tipo di asta è uguale alla precedente con la differenza

che il vincitore paga la seconda offerta più alta invece del prezzo dell'ultima offerta.

Internet è ormai diventato lo strumento principale per effettuare scambio di informazioni e transazioni commerciali tra le persone. Proprio questo crescente uso di internet nella nostra società ha accresciuto l'interesse verso le aste e verso le sue nuove applicazioni. Molti meccanismi ed applicazioni pratiche che vedremo in seguito fanno uso, infatti, delle aste per venire incontro a varie esigenze.

Alcuni vantaggi delle aste on-line sono i seguenti:

- *Disponibilità.* Le offerte possono pervenire in qualsiasi momento, 24 ore su 24. Solitamente gli oggetti vengono mostrati per un periodo che va da 1 a 10 giorni (a discrezione del venditore) in modo da dare agli acquirenti il tempo di visualizzare l'asta e di inviare l'offerta.

- *Nessun vincolo geografico.* I venditori e gli offerenti possono partecipare da qualsiasi parte del mondo. Questo fa sì che l'asta abbia un audience più ampio aumentando così il numero di potenziali offerenti.

- *Elevata interazione sociale.* L'intensità dell'interazione in questo tipo di asta è molto simile a quella che si ha con i giochi d'azzardo. Gli offerenti fremono nell'attesa di conoscere l'esito dell'asta e sono quindi invogliati a partecipare più attivamente al "gioco".

- *Mercato globale.* Il largo numero di offerenti incoraggia i venditori a creare un'asta; a sua volta i venditori incoraggiano più offerenti, creando così una sorta di circolo vizioso che coinvolge sempre più persone.

Ebay è il sito leader del settore dove poter effettuare aste online mettendo a disposizione degli utenti una piattaforma per creare aste attraverso un'interfaccia semplice, ma funzionale per creare aste e per

mostrare ai potenziali offerenti le caratteristiche dell'oggetto in vendita. Il venditore può configurare un prezzo di riserva e definire una scadenza per l'arrivo delle offerte (da 1 a 10 giorni). Una volta che il tempo limite per l'arrivo delle offerte è trascorso, l'offerente che ha inviato l'offerta migliore si aggiudica l'oggetto.

Le piattaforme per aste on-line attualmente più diffuse si prestano bene alla vendita di singoli oggetti; ma cosa succede se un venditore vuole mettere all'asta un insieme di oggetti contemporaneamente? Proprio per gestire situazioni in cui si vendono più oggetti, già a partire dagli anni '60, alcuni dei più noti economisti e studiosi di micro-economia come Vickrey, Clarke e Groves hanno studiato il meccanismo alla base di un nuovo tipo di asta, chiamata asta combinatoriale che appartiene alla categoria delle aste multi-item.

2.1.1 Aste *multi – item*

Una nuova tipologia di aste che si sta affermando negli ultimi anni è quella *multi-item*, dove si ricerca la migliore allocazione di un insieme di prodotti. Le aste multi-item possono essere classificate in aste sequenziali, aste parallele e aste combinatorie. Nella *sequential auction* vengono effettuate l'una dopo l'altra un numero di aste pari al numero di oggetti messi all'incanto. Nella *parallel auction* tali aste avvengono contemporaneamente. Le aste combinatorie permettono invece ai partecipanti di formulare le proprie offerte non sui singoli oggetti ma su combinazioni di essi. Determinare i vincitori sia per le aste sequenziali sia per le aste parallele è facile, in quanto occorre semplicemente scegliere, separatamente per ciascun prodotto, qual è

l'offerta migliore. Tuttavia sia le aste sequenziali che quelle parallele possono condurre ad allocazioni inefficienti delle risorse; al fine di porre rimedio a tale inconveniente sono stati proposti due diversi approcci. Un primo approccio è quello di istituire un mercato "a posteriori" in cui gli offerenti, una volta terminata l'asta, si scambiano i prodotti ottenuti. Un secondo metodo è quello che permette agli offerenti di "ritirare" la propria offerta se la combinazione ottenuta, al termine dell'asta, non corrisponde a quella desiderata. I prodotti, le cui offerte vincitrici sono state ritirate, diventano oggetto di una seconda asta. Se al termine di quest'ultima, il nuovo prezzo vincente è più basso del precedente, colui che ha ritirato la propria offerta dovrà comunque pagare la differenza. La stessa cosa non vale per le aste combinatorie, soprattutto quando si hanno delle preferenze circa una determinata combinazione di prodotti (T. Sandholm, 2000).

2.1.2 Aste combinatorie (CA)

Nel dizionario di *gametheory.net* l'asta combinatoria viene definita come *“Un’asta con più beni in vendita, in cui ogni offerente propone un prezzo per una collezione di beni (scelti dall’offerente) invece che proporre un’offerta per ogni bene separatamente. Il banditore seleziona un insieme di queste offerte in modo da ottenere il massimo guadagno ed evitando di assegnare lo stesso oggetto a più offerenti.”*

L'asta combinatoria è un particolare meccanismo d'asta che tiene conto di relazioni di complementarità e di sostituibilità fra gli oggetti messi in vendita; per questo motivo gli offerenti spesso non hanno interesse per tutti i prodotti oggetto d'asta bensì per un particolare

sottoinsieme di essi. In queste situazioni, le aste combinatoriali sono desiderabili poiché permettono agli offerenti di esprimere le loro reali valutazioni sulle possibili combinazioni di oggetti e di conseguenza portano ad un'allocazione più efficiente.

Tuttavia, questo tipo di aste, nonostante rappresentino l'ideale in molte situazioni, non vengono ancora utilizzate in larga scala (ad esempio il sito ebay non le supporta) perché soffrono di alcuni problemi legati alla difficoltà computazionale di trovare un'allocazione ottima e alla capacità di selezionare l'insieme delle offerte vincitrici una volta che tutte le offerte sono state inviate. Il problema di ottimizzazione delle allocazioni è noto anche come problema del *SET-Packing*. Purtroppo non è stato ancora scoperto un algoritmo polinomiale per la soluzione di un problema di packing e ci sono ragioni fondate per credere che non esista. Il SPP appartiene alla classe di problemi NP-hard. Molti studi sono stati indirizzati verso una restrizione del problema che ammetta un algoritmo efficiente per la risoluzione del CAP (combinatoria auction problem). Due sono le strade percorribili in questo senso. La prima risulta quella di restringere i sottoinsiemi su cui è possibile effettuare le offerte. Un secondo approccio è quello di effettuare restrizioni non tanto sulle combinazioni ammissibili quanto piuttosto sugli ordini di preferenza degli offerenti stessi (S. de Vries, 2000) .

2.1.3 Il meccanismo Vickrey-Clarke-Groves (VCG)

La più famosa asta combinatoriale è data dalla generalizzazione dell'asta Vickrey, noto come il meccanismo Vickrey-Clarke-Groves (VCG). Il meccanismo si basa sull'idea che ogni giocatore dell'asta paga il costo opportunità che la loro presenza introduce a tutti gli altri giocatori. Per capire meglio supponiamo che due mele vengano messe all'asta fra tre candidati. L'offerente A offre 5 per una mela, l'offerente B ne offre 2 e l'offerente C è disposto a pagare 6 a condizione che le ottenga entrambe. Per decidere a chi assegnare le mele viene considerato il costo opportunità imposto da ciascun offerente sul resto dei concorrenti. L'offerente B ha un programma di utilità di 2, l'offerente A ha un programma di utilità di 5, se A fosse stato assente C avrebbe vinto poiché aveva un programma di utilità di 6.

Una variante dell'asta Vickrey è chiamata generalizzazione dei prezzi seconda asta (GSP) differente dal meccanismo VCG e noto per essere usato da Google e Yahoo per programmi di pubblicità on-line (Wikipedia, 2010).

2.2 Caso particolare: Le aste di query in Google

Il GSP (*Generalized Second-Price*) è stato creato su misura per l'ambiente unico degli annunci on-line. L'interesse per lo studio di questo nuovo meccanismo è stato motivato dal suo grande successo. L'entrate di Google per esempio nel 2005 sono stati pari a 3,89 miliardi di dollari, il 98% delle entrate proviene dalle aste GSP. Quando un utente compie una ricerca su un motore di ricerca ottiene

una pagina contenente i risultati dove, dove vengono visualizzati sia i siti più rilevanti attinenti alla sua ricerca, sia i link sponsorizzati chiaramente distinguibili. Per esempio se un agente di viaggio partecipa ad un'asta per parole chiave e si aggiudica la parola Hawaii, ogni volta che qualche utente digiterà la parola Hawaii il suo annuncio comparirà nella pagina dei risultati tra i link sponsorizzati. Quindi l'inserzionista paga il motore di ricerca per inviare l'utente al suo sito web, da qui il nome "pay-per-click" dei prezzi. Il numero di annunci che il motore di ricerca può mostrare ad un utente è limitato, inoltre un annuncio nella parte superiore della pagina è più probabile essere cliccato rispetto ad un annuncio mostrato in fondo. Di conseguenza, i motori di ricerca hanno bisogno di un sistema per assegnare le posizioni agli inserzionisti, e le aste sono una scelta naturale. Attualmente, il meccanismo più utilizzato dai motori di ricerca si basa sulle aste GSP. Come viene illustrato nella fig.2.3 nella più semplice asta GSP, per aggiudicarsi una determinata parola chiave, gli inserzionisti devono presentare offerte attestante la massima disponibilità a pagare per un clic. Quando un utente inserisce una parola chiave, riceve i risultati della ricerca lungo i link sponsorizzati, questi vengono indicati in ordine decrescente rispetto alle offerte che hanno ricevuto dagli inserzionisti. In particolare, l'annuncio con la offerta più alta è visualizzata nella parte superiore, l'annuncio con l'offerta immediatamente superiore è visualizzato nella seconda posizione, e così via.

OFFERENTE	OFFERTA (euro per clik)
1	0,60 €
2	1,20 €
3	0,80 €
4	1,00 €

POSIZIONAMENTO	OFFERENTE	COSTO PER CLIK
Primo Risultato	2	1,01 €
Secondo Risultato	4	0,81 €
Terzo Risultato	3	0,61 €

Fig. 2.3: Esempio di asta GSP

Con più posizioni nell'asta GSP, un inserzionista in posizione I paga un prezzo per ogni clic uguale all'offerta di un inserzionista in posizione $(i + 1)$, più un incremento minimo (tipicamente 0,01 dollari). L'asta GSP sembra uguale al meccanismo Vickrey-Clarke-Groves, in quanto in entrambi meccanismi l'inserzionista paga in base alle offerte degli altri partecipanti e non in base alla propria di offerta. Ma nel GSP a differenza del meccanismo Vickrey-Clarke-Groves non si possono prevedere strategie dominanti; nel caso di un solo slot disponibile i due meccanismi sarebbero identici, ma con più slot disponibili differiscono in quanto con VCG ciascun agente paga un prezzo pari al valore totale che matura in funzione della presenza di altri agenti. Riconoscendo questi vantaggi anche *Yahoo* ha adottato questo sistema (Edelman, Ostrovsky, & Schwarz, 2005).

2.2.1 Sviluppo di un modello commerciale delle aste

Overture fu la prima società a sviluppare un modello commerciale delle aste, nel quale venivano messi all'asta degli spot pubblicitari. Due dipendenti di Google, Salar Kamangar e Eric Veach, osservarono quello che stava facendo Overture e decisero che potevano aggiungere dei miglioramenti. Nell'autunno del 2001 venne così presentata la *google ad auction* (l'Asta degli spot pubblicitari in Google). All'inizio, l'asta degli spot pubblicitari di Google veniva applicata solo agli spot che apparivano sul lato destro della pagina, in cui gli spot in cima alla pagina (l'area con il migliore rendimento) erano riservati ai prezzi negoziati da una forza vendite. Alla fine divenne chiaro che i prezzi generati dall'asta erano più congrui di quelli generati dalla negoziazione, così Google passò ad usare un'asta per tutti gli spot presentati (Varian, 2007).

2.2.2 L'asta Google degli spot

L'asta Google degli spot è probabilmente l'asta più vasta del mondo, con miliardi di aste eseguite ogni settimana. L'asta presenta una struttura teorica molto interessante: esistono diversi slot in cui possono inserirsi gli spot, ma alcuni ricevono più click di altri. In condizioni di equilibrio, ciascun offerente deve preferire lo slot in cui si trova a qualsiasi altro slot. Per vedere in che modo questo funziona, consideriamo un offerente che pensa di partecipare ad un'asta di parole chiave. I partecipanti offrono un determinato importo, quindi un nuovo offerente si trova di fronte ad una "curva dell'offerta dei

click". Più offre, e più dei precedenti offerenti verranno superati, arrivando così ad una posizione più alta e a più click di accesso. Nel determinare la sua offerta, l'inserzionista deve prendere in considerazione il costo incrementale per click: quanto denaro in più dovrà spendere per ottenere un maggior numero di click.

Varian, economista presso Google, afferma che per creare un asta occorre conciliare gli interessi di tre parti: inserzionisti, utenti e google. L'inserzionista desidera pubblicare annunci pertinenti in modo che gli utenti clicchino su di essi. Gli utenti desiderano visualizzare annunci pertinenti; non vogliono essere disturbati da spam o altro materiale per loro non rilevante. Google infine vuole offrire un buon servizio per essere sicuro che sia gli inserzionisti sia gli utenti tornino ad usare i suoi servizi.

Supponiamo vi siano tre spazi pubblicitari, ossia tre posizioni dove gli annunci possono essere visualizzati e quattro offerenti in competizione per questi tre spazi.

Ciò che accade è che ogni volta che viene effettuata una ricerca su google parte un'asta. L'annuncio che ottiene il punteggio più alto si aggiudica la prima posizione, gli inserzionisti non devono pagare l'offerta ma l'importo dell'inserzionista che si trova nella posizione immediatamente inferiore, quindi paga solamente il minimo indispensabile per battere la concorrenza.

Bisogna anche considerare il concetto di qualità dell'annuncio; l'obiettivo di google è quello di visualizzare annunci di elevata qualità. Il punteggio di qualità è costituito da tre elementi; di questi quello più importante è la percentuale di clic. Permettendo agli utenti di esprimere un voto mediante i loro clic, google consente a milioni di persone in tutto il mondo di decidere quali annunci sono più

appropriati per ciascuna ricerca effettuata. La filosofia di google si basa da sempre sul giudizio degli utenti usato come criterio fondamentale per prendere decisioni. La pertinenza è il secondo fattore più importante nel punteggio di qualità, ossia la pertinenza della parola chiave rispetto agli annunci quanto ai termini di ricerca utilizzati dagli utenti. Google utilizza la pertinenza per fare in modo che gli utenti visualizzino solo annunci utili, impedendo di far pagare agli inserzionisti per una ricerca che non è correlata al loro prodotto o servizio. La terza componente del punteggio di qualità è la qualità della pagina della destinazione. Un annuncio è utile solo se la pagina alla quale esso rimanda aiutano gli utenti a trovare informazioni che sta cercando. Una pagina di destinazione di elevata qualità deve presentare contenuti pertinenti originali, deve essere facilmente navigabile e deve avere tempi di caricamento rapidi. Il punteggio di qualità influisce sul funzionamento dell'asta poiché viene utilizzato per ottenere un altro dato, ovvero il "ranking" ottenuto dalla moltiplicazione dell'offerta dell'inserzionista per il livello di qualità dell'annuncio. Il ranking serve a definire quanto si paga effettivamente un clic; ovvero il ranking dell'annuncio dell'inserzionista che si trova sotto di lui diviso il punteggio di qualità dell'inserzione in questione. Se un inserzione aumenta il suo punteggio di qualità, ad esempio con un annuncio più pertinente, aumenta di conseguenza il suo ranking e diminuisce allo stesso tempo il prezzo pagato per ogni clic che riceve (Varian, 2007).

2.2.3 L'importanza della competizione

È ampiamente riconosciuto che il ricavo realizzato in un'asta dipende da quanta competizione è presente in quell'asta. Lo stesso principio vale per l'asta delle posizioni: i ricavi non decollano fino a che non c'è competizione. Nell'asta Google ci sono 8 slot per gli spot sul lato destro della pagina, e fino a 3 slot per quelli in cima alla pagina. Come detto in precedenza, l'ordine degli spot è determinato dalle offerte e dai tassi di accesso, ma gli spot che vengono “promossi” (spostati alla cima della pagina) devono soddisfare alcuni criteri addizionali, che riguardano anche la qualità degli spot. Per semplificare, se un'asta ha meno offerenti degli slot disponibili, o un numero di offerenti appena sufficiente a riempire gli slot disponibili, noi diciamo che è “svenduta”. Se ha più offerenti degli slot diciamo che è “venduta oltre la disponibilità”. Se un'asta è svenduta, il prezzo pagato dall'ultimo offerente sulla pagina è il prezzo di riserva, che supponiamo sia di 5 centesimi. Se la pagina è venduta oltre la disponibilità, il prezzo pagato dall'ultimo offerente sulla pagina è determinato dall'offerta del primo agente escluso, che può facilmente essere almeno 10 volte più alto del prezzo di riserva.

Il punto importante è che le pagine vendute oltre la disponibilità sono molto più redditizie delle pagine vendute parzialmente non solo perché ci sono più offerenti, ma anche perché le forze della competizione sono molto più forti.

Bisogna anche considerare l'importanza dell'algorithmo di abbinamento utilizzato per presentare gli annunci. L'utente immette una “interrogazione” e l'agente pubblicitario acquista delle “keyword”. L'agente può specificare “abbinamento esatto”, che

significa che l'annuncio è mostrato solo se l'interrogazione dell'utente combacia esattamente con le keyword dell'agente pubblicitario. Ma è più frequente che gli agenti specifichino «abbinamento approssimativo» che significa che l'interrogazione si abbinerà con varie espansioni della Keyword quali i sinonimi e le sottostringhe.

Gli annunci addizionali per via dell'abbinamento approssimativo beneficiano l'utente e l'agente pubblicitario, dato che rendono più probabile che l'utente effettuerà il click. Ma accrescono anche la competizione dell'asta, facendo aumentare i prezzi (Varian, 2007).

Riepilogo

Nei paragrafi precedenti si è presentata una panoramica delle aste elettroniche, concentrando l'attenzione sulle aste GSP, utilizzate dai principali motori di ricerca come Google per l'acquisto di parole chiave da parte degli inserzionisti per la ricerca sponsorizzata. È stato mostrato come il GSP (*Generalized Second-Price*) è stato creato su misura per l'ambiente unico degli annunci on-line, riscuotendo grande interesse grazie al suo successo. Quando un utente inserisce una parola chiave, riceve i risultati della ricerca lungo i link sponsorizzati, questi vengono indicati in ordine decrescente rispetto alle offerte che hanno ricevuto dagli inserzionisti. Il costo di una parola chiave aumenta in base a quanto essa viene contesa dagli inserzionisti. Nel prossimo capitolo si mostrerà quindi come è possibile individuare parole chiave non costose sfruttando i domini semantici, scegliendo parole strategiche e non ovvie.

Capitolo 3

Utilizzo dei domini semantici nella ricerca sponsorizzata

Diversi studi sui domini semantici in linguistica hanno portato all'ipotesi di una possibile applicazione dei domini nell'ambito della ricerca sponsorizzata con evidenti vantaggi dal punto di vista computazionale come la riduzione delle ambiguità lessicali e la discriminazione del senso di una parola per un miglioramento della pertinenza dei motori di ricerca.

Nel primo paragrafo verranno definiti i campi semantici e i domini semantici in linguistica, nel secondo paragrafo verrà mostrato come è possibile migliorare la pertinenza dei risultati dei motori di ricerca sfruttando i domini semantici.

3.1 Definizione di campi e domini semantici in linguistica

I campi semantici sono una questione di recente interesse della linguistica computazionale. La nozione linguistica di campo semantico è stata principalmente proposta da Jost Trier al quale si deve il merito di aver dato inizio ad una nuova fase della storia della semantica. Per campi semantici intendiamo delle regioni concettuali ripartite tra le varie parole. Ogni campo è visto come una regione parziale della distesa di idee contenuta dal vocabolario di una lingua. All'interno di ciascun campo il significato di una parola è determinato dalla rete di relazioni stabilite con le altre parole. La limitazione della teoria di Trier è che essa non fornisce alcun criterio oggettivo per

determinare i campi semantici. Per superare tale limite è stato introdotto il concetto di dominio semantico. I domini semantici sono aree comuni di discussioni umane, ad esempio economia , politica, scienze.... Il dominio semantico associato ad un particolare settore è l'insieme di specifici termini di dominio appartenenti ad esso. Il concetto di dominio semantico in linguistica, dal punto di vista computazionale, fornisce un utile strumento per ridurre le ambiguità lessicali e per migliorare la discriminazione del senso di una parola (Gliozzo, Strapparava, & Dagan, 2009) .

3.2 Migliorare la pertinenza dei risultati dei motori di ricerca attraverso i domini

Nella linguistica computazionale, la discriminazione del senso della parola è un problema di elaborazione del linguaggio naturale , che consiste nell'identificare il senso di una parola (cioè il significato) utilizzato in una frase , quando la parola ha molteplici significati (polisemia). La soluzione a questo problema comporterebbe il miglioramento della pertinenza dei motori di ricerca. I domini vengono utilizzati con un duplice ruolo di descrizione linguistica. Un ruolo è esprimere i sensi che caratterizzano la parola nei diversi domini(ad esempio la parola “gru” ha i sensi nei domini di “Zoologia” e “Costruzione”). Un secondo ruolo è quello di caratterizzare i testi, dando un generico livello di categorizzazione di testo (ad esempio per la classificazione delle notizie e degli articoli) , (Gliozzo, Strapparava, & Dagan, 2009).

Nel 2002 Magnini e Strapparava hanno studiato il ruolo dell'informazioni fornite dai domini semantici nell'ambito della disambiguazione del testo. Essi sono giunti all'ipotesi che l'etichette di dominio (come medicina, sport ecc), forniscano uno strumento per stabilire le relazioni semantiche tra i termini, le quali possono successivamente essere utilizzate per disambiguare il testo. In particolare essi affermano che i domini costituiscono una fondamentale proprietà semantica sulla quale si basa la coerenza del testo, così che i sensi delle parole che compaiono all'interno di una porzione di testo, tendono a massimizzare l'appartenenza ad uno stesso dominio. *WordNet Domains* rappresenta il primo tentativo di estendere l'insieme di etichette di dominio in un database lessicale, ovvero il WordNet, un database semantico-lessicale per la lingua inglese, elaborato dal linguista George Armitage Miller presso l'Università di Princeton, con l'obiettivo di organizzare, definire e descrivere i concetti espressi dai vocaboli. L'organizzazione del lessico si avvale di raggruppamenti di termini con significato affine, chiamati "*synset*" (dalla contrazione di *synonym set*), e del collegamento dei loro significati attraverso diversi tipi di relazioni chiaramente definite. All'interno dei *synset* le differenze di significato sono numerate e definite.

Nel *WordNet Domains* ogni *synset* viene annotato con una o più etichette di dominio; le informazioni fornite dalle etichette offrono un modo naturale per stabilire le associazioni tra i sensi delle parole all'interno di un testo, che possono essere proficuamente utilizzate durante il processo di disambiguazione per migliorare la pertinenza dei risultati proposti dai motori di ricerca (Gliozzo, Strapparava, & Dagan, 2009).

Riepilogo

Nei paragrafi precedenti sono stati analizzati diversi studi sui domini semantici in linguistica, i quali hanno portato all'ipotesi di una possibile applicazione dei domini nell'ambito della ricerca sponsorizzata all'interno dei motori di ricerca, con evidenti vantaggi dal punto di vista computazionale come la riduzione delle ambiguità lessicali e la discriminazione del senso di una parola digitata da un utente per un miglioramento della pertinenza dei motori di ricerca. Nel capitolo successivo si è quindi proposto l'utilizzo di un approccio semantico, con il quale individuare nuove parole chiave per la ricerca sponsorizzata attraverso l'uso dei domini semantici scegliendo parole chiave non ovvie e ottimizzando i costi.

Capitolo 4

Utilizzo dei domini semantici nelle aste di parole chiave

Con la presente tesi si vuole proporre un approccio semantico per suggerire nuove parole chiave agli inserzionisti che vogliono pubblicizzare il loro sito turistico, attraverso l'uso dei domini semantici, in quanto i metodi esistenti per la ricerca di parole chiave forniscono solo suggerimenti che contengono la radice o l'intera parola chiave indicata, mentre i suggerimenti che non la contengono non sono mostrati poiché si basano su tecniche di abbinamento. Ciò comporta suggerimenti molto competitivi e quindi costosi.

I motori di ricerca sono in continuo aggiornamento, al di là delle parole chiave tradizionali, essi ricercano nuovi metodi su come suggerire parole chiave agli inserzionisti nella ricerca sponsorizzata. Si tratta del superamento del classico sistema di abbinamento, migliorato da un sistema di relazioni tra termini. Gli inserzionisti che vogliono dare visibilità al proprio sito partecipano ad aste online per acquistare determinate parole chiave. Al vincitore dell'asta è permesso di mettere il proprio link e i propri annunci sulla pagina dei risultati che comparirà quando verrà digitata dall'utente quella specifica parola. Google e Overture stanno investendo sul miglioramento delle loro piattaforme di offerta per attrarre nuovi inserzionisti, utilizzando strumenti come *AdWord* e *YahooTool*. Essi aiutano l'inserzionista a selezionare la parola chiave per esprimere l'intenzione del suo annuncio e lo strumento mostra un elenco di suggerimenti che può utilizzare per l'offerta; Il problema principale dei metodi esistenti

(AdWords, 2009; YahooTool, 2009; wordTracker, 2009) è che producono solamente i suggerimenti che contengono le parole chiave indicate, mentre i suggerimenti che non la contengono non sono mostrati. Il compito di generare le parole chiave è automatico e negli ultimi anni ha ricevuto una crescente attenzione nella “search community” per il suo impatto sulle entrate delle aziende. Utilizzando un nuovo approccio semantico, è evidente come i suggerimenti proposti avrebbero delle alte performance, in quanto si basano su un *indice di non-ovvietà* tenendo conto dell’*user space*. Per non-ovvietà intendiamo quanti sono i suggerimenti per una determinata parola chiave che non contengano la radice stessa. Tale indice è calcolato automaticamente utilizzando algoritmi *string matching* e semplici *stemmer*. Per “user space” intendiamo invece le percezioni e gli interessi dell’utente. La non ovvietà ha come obiettivo di generare un gran numero di parole chiave che sono rilevanti per la parola digitata dall’utente, ma non banali, raggiungendo così un vantaggio economico e una minor concorrenza. Come possiamo vedere nella fig.4.1, ad esempio, per “isole Galapagos”, un tipico sistema suggerisce “tour isole Galapagos” il cui costo è di quasi 5€ per click, mentre il meno ovvio “Santa Cruz” costerebbe solo 0,35€.

	TERMINE SUGGERITO	COSTO PER CLIK
Google	tour isole Galapagos	5 €
Nuovo Metodo	Santa Cruz	0,35 €

Fig. 4.1 – Confronto tra i prezzi delle keywords suggerite

Altri suggerimenti potrebbero essere “caraibi, Messico” o ancora “perla dell’oceano” (Alpa & Pennacchiotti, 2010).

Tutti questi suggerimenti, pur essendo correlati semanticamente alla query, se ne differenziano per somiglianza sintattica, per esempio un utente cerca isole Galapagos allo stesso modo di come tenderebbe a cercare Santa Cruz o Messico. Come viene illustrato nella fig.4.2 è possibile fare tanti altri esempi: per pubblicizzare un sito sui “lettori mp3” si potrebbe acquistare la parola chiave “ipodnano”; o ancora per vendere un pacchetto turistico alle Maldive basterebbe acquistare la parola chiave “atollo”, così come per pubblicizzare un hotel converrebbe acquistare la parola chiave “albergo” per poter comparire tra i link sponsorizzati attraverso parole meno competitive che permettono di abbassare i costi.

PAROLE CHIAVE	POSSIBILI SUGGERIMENTI
Isole Galapagos	Caraibi, Messico, perla dell'oceano
Lettore mp3	Ipod-nano
Hotel	Albergo
Maldiva	Atollo

Fig. 4.2 – Esempio di altri suggerimenti che differiscono dalla radice della parola chiave

Un altro metodo per sfruttare i domini semantici è quello di acquistare parole chiave apparentemente non conciliabili con l'oggetto in vendita. Per esempio se un utente digita nel campo di ricerca la parola chiave "albero da frutta", come illustrato in fig. 4.3, tra i link sponsorizzati compare un annuncio per "alimenti per cani"; apparentemente la parola chiave non è correlata all'oggetto che l'inserzionista vuole vendere, ma in realtà essi sono legati da uno stesso dominio semantico, che in questo caso potrebbe essere quello della campagna o del giardino. Quindi potenzialmente l'utente interessato agli alberi da frutta, potrebbe possedere un cane e quindi essere interessato ad un sito che tratta di alimenti per cani. In questo

modo l'inserzionista riuscirebbe a dare visibilità al proprio sito tra i link sponsorizzati attraverso l'acquisto di parole chiave non costose e poco competitive.

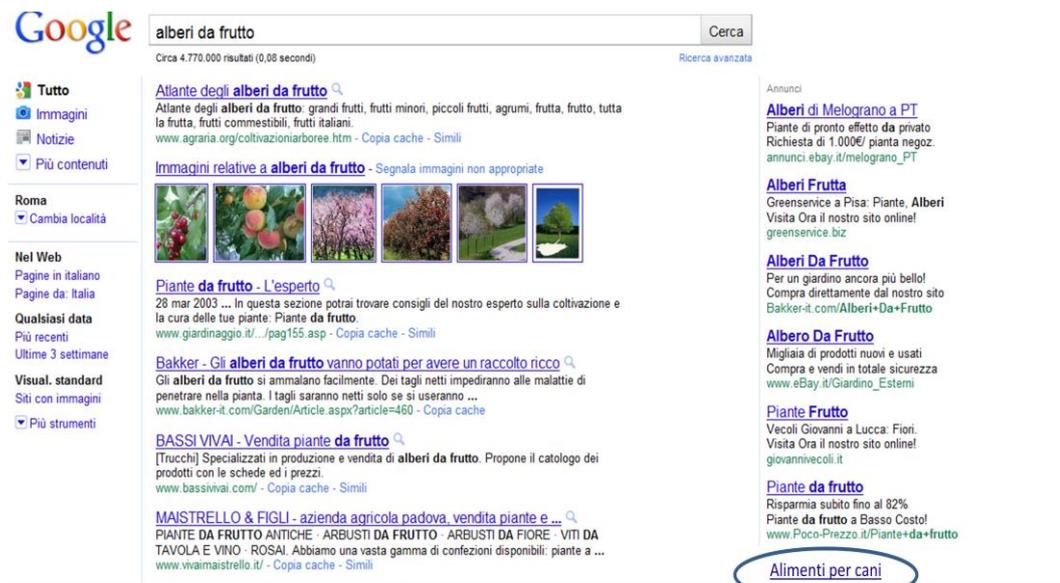


Fig. 4.3 : Esempio di oggetto in vendita non correlato alla parola chiave

Un esempio riferito al turismo potrebbe riguardare la parola chiave "coralli"; come è possibile vedere nella fig.4.4 se un utente digita questa query, tra i link sponsorizzati proposti da Google non compare nessun annuncio pubblicitario. È quindi evidente che si tratta di una parola il cui costo sarà molto basso in assenza di altri competitori; dunque l'acquisto della parola "coralli" da parte di operatori turistici sarebbe molto vantaggioso per far comparire tra i link sponsorizzati offerte, voli e pacchetti turistici ad esempio per il Mar Rosso o per l'Australia.

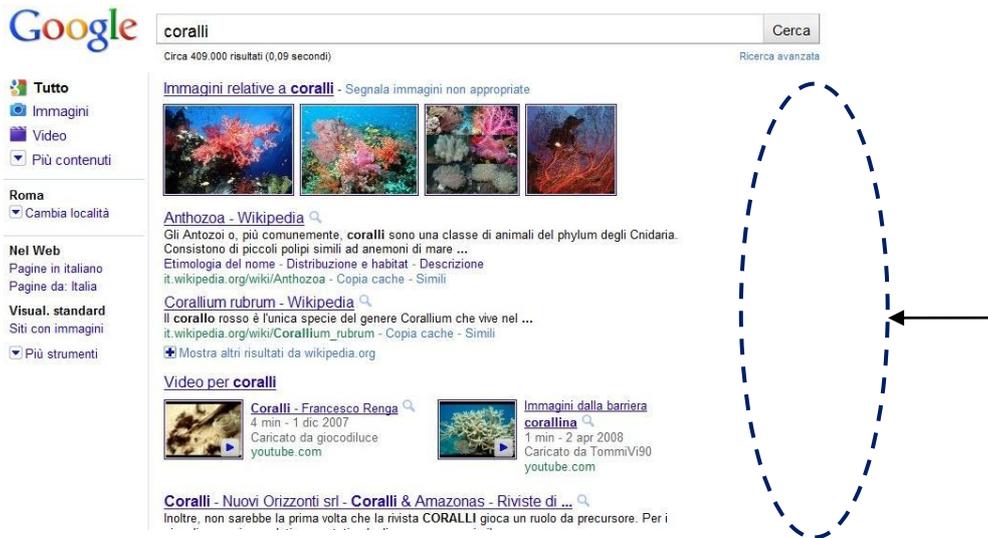


Fig. 4.4 - Esempio di parola chiave poco competitiva

Al contrario come possiamo vedere nella fig.4.5 se un utente digita per esempio la query “Mar Rosso” compariranno numerosi annunci pubblicitari tra i link sponsorizzati poiché si tratta di una parola chiave molto ambita dagli operatori turistici intenzionati a pubblicizzare le loro offerte su questa località e di conseguenza molto costosa.

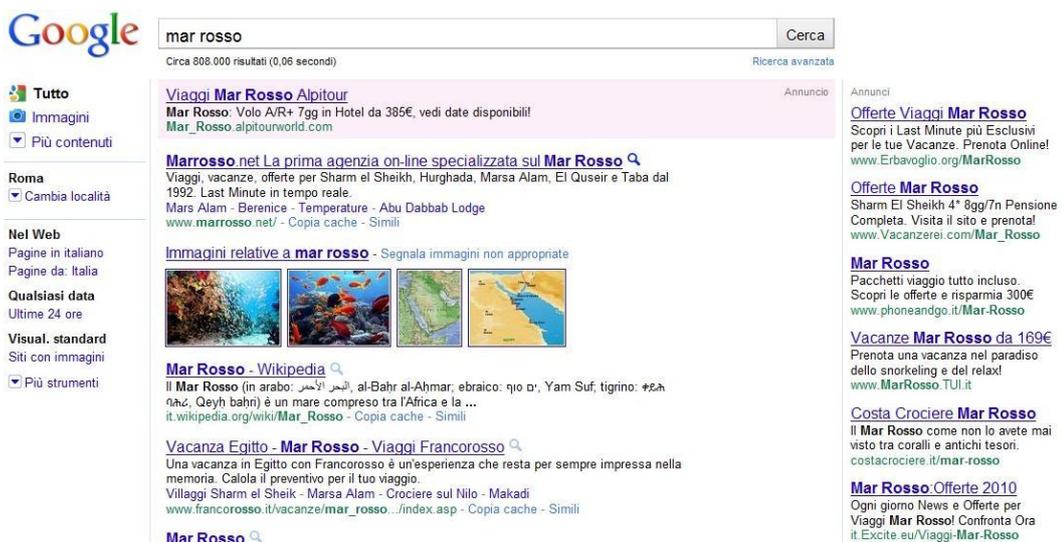


Fig. 4.5 - Esempio di parola chiave molto competitiva

L'acquisto della parola "coralli" quindi è vantaggiosa per gli operatori turistici sia in termini economici, in quanto il suo costo sarà molto basso, sia in termini di visibilità, in quanto un utente interessato a cercare informazioni e video sui coralli potrebbe essere potenzialmente interessato a fare un vacanza nel Mar Rosso per ammirarne la barriera corallina. Questi suggerimenti sarebbero molto preziosi per gli inserzionisti che nell'ambito del turismo vogliono migliorare la loro visibilità sul web attraverso parole non banali e meno costose.

Capitolo 5

Conclusioni

Nella presente tesi si è voluto studiare il turismo nei domini semantici, proponendo un metodo per pubblicizzare i siti turistici attraverso l'acquisto di parole chiave non ovvie da parte degli inserzionisti per la ricerca sponsorizzata.

Lo studio si è concentrato dapprima sulle caratteristiche della vendita di turismo online, sottolineando l'importanza e la necessità della visibilità di un sito internet. Per affermarsi in questo scenario competitivo infatti, ciascuna azienda turistica dovrà adottare un nuovo modello di business e nuovi canali promozionali.

E' stato mostrato che il principale problema delle aziende turistiche è quello di riuscire ad essere competitivi nel mercato online per poter così aumentare il flusso delle vendite. Per "attirare i clienti" al proprio sito occorre essere visibili sui motori di ricerca; maggiore sarà la visibilità, maggiore sarà la possibilità di guadagno per un'azienda turistica. In questa tesi verranno quindi analizzate le tecniche che permettono sostanzialmente di aumentare il volume di affari delle *aziende turistiche*, aumentandone la visibilità e riducendone i costi. A tale scopo la nostra indagine si è soffermata sulle principali tecniche per posizionarsi ai primi posti tra i risultati proposti da un motore di ricerca, ovvero le tecniche SEO (*search engine optimization*) e le tecniche a pagamento. E' stato mostrato come l'utilizzo delle tecniche SEO comporta diversi svantaggi come: start Up e aggiornamenti lenti, nessuna garanzia sul posizionamento di una parola chiave, elevato

budget senza la garanzia di una visibilità a lungo termine. L'utilizzo delle tecniche SEO richiede dunque budget elevati senza la garanzia di ricavi dilazionati nel tempo. E' stato quindi suggerito l'utilizzo delle tecniche a pagamento per rendere visibili gli annunci turistici tra i link sponsorizzati. Si è posta così l'attenzione sulle principali tecniche a pagamento per pubblicizzare un sito turistico sul web mediante l'utilizzo di spazi pubblicitari nei social network o all'interno dei principali motori di ricerca attraverso l'acquisto di parole chiave nelle *aste di query*.

Successivamente si è presentata una panoramica delle aste elettroniche, concentrando l'attenzione sulle aste GSP, utilizzate dai principali motori di ricerca come Google per l'acquisto di parole chiave da parte degli inserzionisti per la ricerca sponsorizzata. E' stato mostrato come il GSP (*Generalized Second-Price*) è stato creato su misura per l'ambiente unico degli annunci on-line, riscuotendo grande interesse grazie al suo successo. Si è descritto come nella più semplice asta GSP, per aggiudicarsi una determinata parola chiave, gli inserzionisti devono presentare offerte attestante la massima disponibilità a pagare per un clic; il vincitore dell'asta pagherà l'importo della seconda offerta più alta, più un incremento minimo per una determinata parola. Dunque con questa tipologia d'asta basta il minimo indispensabile per battere i concorrenti. Quando un utente inserisce una parola chiave, riceve i risultati della ricerca lungo i link sponsorizzati, questi vengono indicati in ordine decrescente rispetto alle offerte che hanno ricevuto dagli inserzionisti.

In seguito è stato introdotto il concetto di qualità dell'annuncio; l'obiettivo di google è quello di visualizzare annunci di elevata qualità, quindi la classifica dei link sponsorizzati viene realizzata da

Google non solo in base all'offerta di ciascun inserzionista, ma tenendo anche conto del punteggio di qualità, dato soprattutto dalla percentuale di clic che il link riceve. Permettendo così agli utenti di esprimere un voto mediante i loro clic, google consente a milioni di persone in tutto il mondo di decidere quali annunci sono più appropriati per ciascuna ricerca effettuata.

Successivamente sono stati analizzati diversi studi sui domini semantici in linguistica, i quali hanno portato all'ipotesi di una possibile applicazione dei domini nell'ambito della ricerca sponsorizzata con evidenti vantaggi dal punto di vista computazionale come la riduzione delle ambiguità lessicali e la discriminazione del senso di una parola per un miglioramento della pertinenza dei motori di ricerca.

Infine si è proposto l'utilizzo di un approccio semantico con il quale suggerire nuove parole chiave agli inserzionisti che vogliono pubblicizzare il loro sito turistico attraverso l'uso dei domini semantici, in quanto i metodi esistenti per la ricerca di parole chiave utilizzati dai principali motori di ricerca forniscono solo suggerimenti che contengono la radice o l'intera parola chiave indicata, poiché si basano su tecniche di abbinamento, mentre i suggerimenti che non la contengono non sono mostrati. Ciò comporta suggerimenti molto competitivi e quindi costosi.

I suggerimenti ottenuti utilizzando i domini semantici sarebbero molto preziosi per gli inserzionisti nel settore del turismo, in quanto aumenterebbero la loro visibilità attraverso parole chiave non banali e meno costose e di conseguenza accrescerebbero i loro guadagni.

Bibliografia

Alpa, & Pennacchiotti. (2010). Open entity extraction from web search query logs.

Besana, S. (2009). *Facebook e il social advertising*. Tratto da HTML:
<http://webmarketing.html.it/articoli/leggi/3067/facebook-e-il-social-advertising/>

Cascioli, F. (2010). Tratto dal sito: <http://www.ilpalo.com/>

Cederle. (2005). Appunti di database management e e-commerce turistico.

Edelman, B., Ostrovsky, M., & Schwarz, M. (2005). *Internet Advertising and the Generalized Second Price Auction: selling Billions of dollars Worth of Keywords*. Tratto da
<http://rwj.berkeley.edu/schwarz/publications/gsp051003.pdf>

Gliozzo, A., Strapparava, C., & Dagan, I. (2009). *Unsupervised and Supervised Exploitation of Semantic Domains in Lexical Disambiguation*. Tratto da
<http://wndomains.fbk.eu/publications/CSL-2004.pdf>

Nielsen. (2009). *Global Faces and networked Places*.

Nielsen. (2010). *Investimenti pubblicitari in Italia nel periodo Gennaio-Luglio 2010*.

Passantino. (2010). *Vendere turismo online*.

S. de Vries, R. V. (2000). *Combinatorial Auctions: A survey*.

T. Sandholm. (2000). *Approaches to winner determination in combinatorial auction*.

Varian, R. (2007). *L'economia politica dei motori di ricerca su internet*. Tratto da
http://www.rivistapoliticaeconomica.it/2006/nov_dic/pdf/Varian_ita.pdf

Wikipedia. (2010). *Vickrey asta*. Tratto da Wikipedia:
http://en.wikipedia.org/wiki/Vickrey_auction