

UNIVERSITÀ DI PISA

Corso di Laurea in Informatica Umanistica

Tesi di Laurea

**Agriturismo.com: restyling  
ottimizzato e analisi statistica  
sull'utilizzo del database**



**Relatori:**

prof. ANNA MONREALE

prof. BEATRICE RAPISARDA

**Candidato:**

RICCARDO GIOMI

matricola: 495996

ANNO ACCADEMICO 2015-2016

*per Elena*

*per i nonni*

# Indice

<b>1</b>	<b>La struttura di Agriturismo.com</b>	<b>5</b>
1.1	Le visite al portale . . . . .	5
1.2	L'operatore "site:" . . . . .	7
1.3	La scelta del dominio . . . . .	9
1.4	Il tag title . . . . .	9
1.5	Rich snippet: rendersi attraenti . . . . .	11
1.6	Il tag meta robots . . . . .	12
1.7	Il file robots.txt . . . . .	13
1.8	La sitemap . . . . .	14
1.9	Url rewriting: dove mi trovo? . . . . .	15
<b>2</b>	<b>Il corpo della pagina</b>	<b>17</b>
2.1	Le chiavi di ricerca . . . . .	17
2.2	Enfatizzare i contenuti . . . . .	19
2.3	L'effetto "matrioska" degli header tag . . . . .	20
2.4	Assenza del tag H1 . . . . .	21
2.5	Internet per tutti: gli attributi delle immagini . . . . .	22
2.6	Ottimizzare le immagini . . . . .	22
2.7	Attributi a confronto . . . . .	23
2.8	SEO Social: farsi conoscere oggi . . . . .	24
<b>3</b>	<b>Analisi statistica sull'utilizzo del database</b>	<b>27</b>
3.1	Estrazione delle tabelle . . . . .	28
3.2	Creazione dei grafici . . . . .	29
3.3	Grafici e risultati . . . . .	32
<b>4</b>	<b>Conclusioni</b>	<b>35</b>
4.1	Sviluppi futuri . . . . .	37
	<b>Bibliografia</b>	<b>38</b>

# Elenco delle figure

1	Esempio di SERP . . . . .	4
1.1	Agriturismo.com: visite totali degli ultimi 6 mesi . . . . .	6
1.2	Agriturismo.it: visite totali degli ultimi 6 mesi . . . . .	7
1.3	Agriturismi.it: visite totali degli ultimi 6 mesi . . . . .	7
1.4	Pagine indicizzate con dominio agriturismo.com . . . . .	8
1.5	Pagine indicizzate con dominio agriturismo.it . . . . .	8
1.6	Pagine indicizzate con dominio Agriturismi.it . . . . .	8
1.7	Esempio di rich snippet . . . . .	11
1.8	Homepage di Schema.org . . . . .	11
1.9	Frammento di robots.txt del portale Agriturismo.com . . . . .	14
2.1	Dettaglio SEMrush su analisi del 12 gennaio 2017 . . . . .	18
2.2	Numero delle interrogazioni mensile (colonna Volume) . . . . .	18
2.3	La terza posizione guadagnata dalla keyword “agriturismi” . . . . .	18
2.4	Eccesso di tag stilistici in agriturismo.com . . . . .	20
2.5	Dettaglio pagina Twitter di Agriturismo.com . . . . .	24
2.6	Dettaglio pagina Twitter di Agriturismi.it . . . . .	25
2.7	Dettaglio pagina Twitter di Agriturismo.it . . . . .	25
3.1	Lista delle tabelle con le relative occorrenze . . . . .	30
3.2	Grafico a torta con i tre diversi range . . . . .	32
3.3	Grafico a barre per le tabelle più utilizzate nel codice . . . . .	33
3.4	Altra visuale delle statistiche relative alle tabelle più usate . . . . .	33
3.5	Grafico a barre con tutte le tabelle utilizzate meno di 5 volte . . . . .	34
4.1	Dettaglio tabella relativa al tipo di stagione . . . . .	36
4.2	Dettaglio tabella su date di inizio e fine stagione . . . . .	36
4.3	Frammento tabella dei trattamenti . . . . .	37

# Introduzione

In rete al mondo d'oggi troviamo tutto. E' possibile interagire con i social network, leggere giornali, partecipare ad aste online ed effettuare acquisti su svariati e-commerce. Un'altra azione possibile, che viene effettuata circa 59.000 volte al secondo, è la ricerca di una risorsa o di un prodotto sui motori di ricerca. I prodotti costruiti da aziende o privati vengono pubblicati e modificati costantemente, ma chi gestisce i contenuti online spesso commette errori che ne compromettono la visibilità e il posizionamento all'interno di *Google*, *Yahoo* o *Bing*, che sono i motori di ricerca più famosi. Spesso vengono commessi errori di natura tecnico - strutturale che affondano un prodotto nella marea di risorse offerte dal web.

Il progetto prende in esame un portale per la prenotazione di agriturismi, chiamato "Agriturismo.com", che è stato acquistato dall'azienda "Uplink Web Agency SRL"<sup>1</sup>. Il portale presenta problemi di ottimizzazione SEO che ne compromettono il numero di visite e la potenziale clientela ed è importante un'analisi che riguarda i fattori strutturali della pagina. Verranno presi in esame altri *benchmark*<sup>2</sup>, come "Agriturismo.it" e "Agriturismi.it" e verranno effettuati alcuni confronti. Gli errori più evidenti vengono analizzati, con la proposta di modifiche ottimizzate. Una piccola descrizione verrà dedicata anche a quei fattori che risultano già essere ottimizzati al momento dell'acquisto del portale. Successivamente verrà effettuata un'analisi statistica delle tabelle più o meno utilizzate in Agriturismo.com, con la creazione di un programma in *Python*<sup>3</sup> per l'estrazione delle tabelle. I dati ricavati verranno visualizzati in tre diversi grafici, dove ci saranno alcuni esempi di casi interessanti. Il database risulta essere molto grande, portando poca fruibilità nelle ricerche e rende difficoltoso lo sviluppo di nuove *query*<sup>4</sup>.

---

<sup>1</sup>Azienda in cui ho svolto il tirocinio curriculare previsto dal piano di studi

<sup>2</sup>Insieme di metodi utili a comparare un prodotto con altri prodotti del solito ambito

<sup>3</sup>Linguaggio di programmazione pseudocompilato

<sup>4</sup>Le query sono delle interrogazioni che l'utente fa ad un database per estrarne contenuti

---

## Search Engine Optimization

La *Search Engine Optimization*, (ottimizzazione dei motori di ricerca), è un insieme di tecniche interne ed esterne al sito che ne migliorano la visibilità fra i motori di ricerca. La SEO fa parte del *Search Engine Marketing* (SEM), uno dei punti fondamentali del Web Marketing. Seguire gli standard SEO per ottimizzare al meglio una pagina web, garantisce una posizione di rilievo, tra i risultati non a pagamento, del proprio sito web.

## Come funzionano i motori di ricerca

I dati trattati dai motori di ricerca sono spaventosamente enormi e vengono elaborati e “digeriti” da tanti server, che utilizzano i *link* (collegamenti), per raggiungere le innumerevoli risorse in internet. I *crawler* (chiamati anche *spider*), sono robot automatici per la lettura dati che lavorano per il motore di ricerca, riuscendo a leggere più pagine interconnesse tra loro. Il report di queste analisi viene salvato in giganteschi hard drive, per essere ritrovati in altre ricerche. Questa modalità di salvataggio di milioni di dati, avviene nei “*Data center*”, strutture enormi situate in tutto il mondo, dotate di server che elaborano inimmaginabili quantità di dati.

I processi principali di un motore di ricerca sono divisibili in:

- **scansione** (o *crawling*),
- **indicizzazione**,
- **ranking** e creazione di pagine risultato (**SERP**).

## Scansione

Gli spider passano in rassegna i contenuti sul web, analizzando tantissime pagine, collegate fra loro dai link. Le risorse che vengono trovate sono visibili tramite un linguaggio di marcatura, chiamato HTML, acronimo di *HyperText Markup Language*. I dati ricercati possono essere visibili, oppure non visibili e di solito sono contenuti non testuali o animazioni che risultano invisibili ai motori di ricerca. Dare una forma ai nostri prodotti tramite l’uso di HTML, utilizzato in contemporanea ad altri metodi, garantisce impaginazioni molto semplici e statiche, ma anche grandi prodotti

---

all'avanguardia. In HTML vengono sviluppati alcuni tag come `<title>`, gli attributi "alt" delle immagini, i tag `<meta description>`; tutti importanti elementi che fanno parte dell'intestazione e del corpo della pagina. All'interno di questi fattori, approfonditi più avanti, lo spider cercherà le parole chiave ricorrenti e rilevanti, che servono per indicizzare i risultati<sup>5</sup>.

## L'importanza delle keyword per l'indicizzazione

I pilastri fondamentali presenti in ogni ricerca sono senza dubbio le parole chiave. Le *keyword* semplificano molto il lavoro dei motori di ricerca, perché inserendole in una ricerca, viene comunicato al crawler di accedere al database in cui è contenuta quella parola chiave, invece di indicizzare ed analizzare miliardi di contenuti che non si avvicinano nemmeno alla risorsa che ci interessa. Il recupero dei dati è quindi molto più veloce e preciso, fattore di importanza incredibile, poiché il tempo di risposta del motore di ricerca, per l'utente, è fondamentale.

## Search Engine Result Page

Cercare qualcosa sul web, per un comune utente, è un'operazione molto veloce e banale, ma in quella semplice ricerca c'è un'analisi di dati incredibile, oltre ad una grande selezione di pagine rilevanti o utili a ciò che viene cercato, aggiungendo anche una classificazione qualitativa delle risorse mostrate. Quando si ricerca qualcosa, si inseriscono delle parole nella *textbox* e in automatico viene creata una query, ovvero un'interrogazione di un utente ad un database. Il crawler sa a quale tipo di dato attingere per restituire un risultato e preleva i documenti più rilevanti alla ricerca, classificandoli. In risposta alla ricerca vengono restituite molte SERP, *Search Engine Result Page*, ovvero le pagine risultato, che saranno poi classificate in base alla loro rilevanza e alle visualizzazioni che hanno raggiunto<sup>6</sup>. In una ricerca su Google, sotto la barra di input, dove verranno inserite le parole relative all'argomento/risorsa ricercata, appariranno le SERP. Fra queste pagine di risultato sarà possibile affinare la ricerca selezionando alcuni filtri offerti da Google. Ad esempio, se viene cercata un'icona, verrà selezionato il filtro denominato "Immagini". Un altro filtro, chiamato "Strumenti di ricerca", sempre sull'esempio di ricerca di un'icona, servirà per inserire le misure precise che interessano al webmaster, la possibilità di selezionare solo immagini con sfondo trasparente e molto altro ancora. A piè di pagina,

---

<sup>5</sup>[www.seocenter.it/learn/seo/abc/come-lavorano-i-motori-di-ricerca.html](http://www.seocenter.it/learn/seo/abc/come-lavorano-i-motori-di-ricerca.html)

<sup>6</sup><http://digitalmarketing.h-farm.com/guida-seomoz/funzionamento-motori-di-ricerca.php>

---

finite le prime SERP indicizzate, vengono mostrate delle ricerche correlate, strumento importante per affinare o modificare la ricerca effettuata. Oltre ai correlati, sarà possibile selezionare altre pagine di risultati di Google e ne verrà fornita una lista, in base ai risultati maturati<sup>7</sup>.



The image shows a search result snippet for 'Informatica Umanistica'. The title is 'Informatica Umanistica | Le cose da dire, gli strumenti per dirle - Pisa'. Below the title is the URL 'www.fileli.unipi.it/infouma/' followed by a dropdown arrow. The meta description reads: '25 ott 2016 - \*La fortuna di chi pratica l'informatica umanistica è guardare il proprio territorio con gli occhi degli altri. Francesco Varanini. \*Il miglior modo di ...'. At the bottom of the snippet are several links: 'Laurea', 'Laurea Magistrale', 'Calendario accademico', and 'Orario delle lezioni 2016/17'.

Figura 1. Esempio di SERP

La Fig.1 è un esempio di SERP ottenuta digitando “Informatica Umanistica” all’interno della textbox di Google. Le SERP sono formate da tre parti fondamentali:

- il titolo, che in caso di click ci porterà alla pagina;
- l’URL, relativo all’indirizzo della pagina web
- una breve descrizione della pagina, identificata nel tag *meta description*

Nell’esempio sono presenti anche i *minilink*, ovvero alcuni link che si trovano sotto lo snippet relativi a pagine correlate alla nostra ricerca.

---

<sup>7</sup>[www.seocenter.it/learn/seo/abc/serp-google.html](http://www.seocenter.it/learn/seo/abc/serp-google.html)

# Capitolo 1

## La struttura di Agriturismo.com

Una pagina web per raggiungere buone posizioni, deve rispettare alcune regole, su cui però spesso non può intervenire il webmaster<sup>1</sup>. Il portale preso in esame per questo progetto è Agriturismo.com, una pagina dedicata alla prenotazione e visualizzazione dei più famosi agriturismi italiani. I fattori che determinano un buon posizionamento SEO sono divisi in due grandi insiemi: fattori *On-page* e i fattori *Off-page*. Il macro settore On-page contiene fattori interni, di natura strutturale. Nella composizione della pagina HTML di un sito verranno analizzati dal crawler i tag più interessanti rispetto ad altri, come i tag `<title>`, i tag `<meta>`, gli attributi `alt` delle immagini e molti altri ancora. Invece, nel settore Off-page, vengono trattati i fattori esterni di un sito web che non sono in nessun modo modificabili da un webmaster. I fattori che fanno parte di questo settore si occupano della popolarità della pagina, che viene determinata dal numero di link, dalla qualità dei link e da ciò che trattano i link. Se una pagina riceve molti link di buona qualità, sarà sicuramente apprezzata nel web. Lo spider, una volta scansionate queste informazioni, determina se la pagina è più appetibile ed affidabile di altre e la inserisce fra le SERP di risultato.

### 1.1 Le visite al portale

Prima di iniziare un'analisi SEO è importante capire chi sono i competitors del portale analizzato e stimare la quantità di traffico che riceve. Per effettuare un'analisi di questo tipo è possibile usare un *tool*<sup>2</sup> online chiamato *SimilarWeb*. Utilizzando la versione gratuita è possibile stimare quali sono i 5 paesi che visitano di più la pagina e confrontare il numero di visite per ogni mese. Questi dati possono essere confrontati fra loro e mostrano subito grandi differenze.

---

<sup>1</sup>responsabile della gestione della pagina web

<sup>2</sup>programma di ausilio per attività specifiche

Percentuali visite	
<b>Italia</b>	63,01%
<b>Francia</b>	14,39%
<b>Turchia</b>	6,01%
<b>Germania</b>	5,71%
<b>Messico</b>	2,82%

Tabella 1.1. Elenco dei primi 5 paesi che più visitano Agriturismo.com

La tabella mostra i cinque stati che più visitano il portale. Se questo dato é utile per valutare la potenziale clientela e chi porta più traffico alla pagina, un altro dato, che sicuramente sarà interessante, è il numero di visite mensili che il portale riceve. Come evidenziato in Fig.1.1, il periodo di Luglio - Agosto arriva a sfiorare le



Figura 1.1. Agriturismo.com: visite totali degli ultimi 6 mesi

20.000 visite, mentre il periodo con il minor numero di visite è Ottobre - Novembre. In questo periodo si raggiungono a malapena le 4.500 visite sul portale. I dati di Agriturismo.com, per avere un primo riscontro sulla competitività del *brand*, devono essere confrontati con i dati reattivi ad Agriturismo.it e Agriturismi.it. I numeri in Fig.1.2 e Fig.1.3 mostrano sin da subito una grande varietà di dati per i tre portali. I numeri di Agriturismo.it sono senz'altro elevati. Il picco massimo di visite, in Agosto, arriva a circa 480.000 visite, mentre il periodo con il minor numero di visite arriva a 110.000. La differenza con Agriturismo.com è enorme, basti pensare che il numero minimo di visite in Agriturismo.it è 5 volte più grande del numero di visite massime su Agriturismo.com. Agriturismi.it invece raggiunge, in media, più



Figura 1.2. Agriturismo.it: visite totali degli ultimi 6 mesi

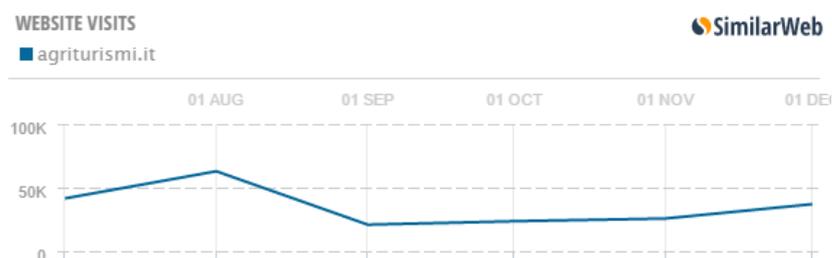


Figura 1.3. Agriturismi.it: visite totali degli ultimi 6 mesi

del doppio delle visite rispetto ad Agriturismo.com, con un picco di 62.568 visite totali in Agosto.

## 1.2 L'operatore "site:"

Quando si effettuano analisi SEO i numeri danno sempre indicazioni importanti sui prodotti e se nel paragrafo precedente è stato trattato il traffico mensile dei portali, è interessante fare anche un'analisi di tutte le pagine di Agriturismo.com, Agriturismo.it e Agriturismi.it che vengono indicizzate da Google. Questa operazione è possibile digitando il comando **site:** nella barra di ricerca di Google, seguito dal nome della pagina che vogliamo analizzare. Questo procedimento permette al webmaster di controllare e vedere il numero di pagine, nelle varie lingue disponibili, che il portale restituisce nella ricerca, in base ai filtri e alle ricerche di ogni utente. Avere un gran numero di pagine permette molta visibilità e varietà di scelta di prodotti o offerte, oltre ai target che possono essere interessati o possono usufruire del prodotto<sup>3</sup>. Di seguito è possibile vedere la grande differenza di pagine indicizzate dei tre portali analizzati.

---

<sup>3</sup><https://www.posizionamento-seo.com/search-engine-optimization/comandi-google/>

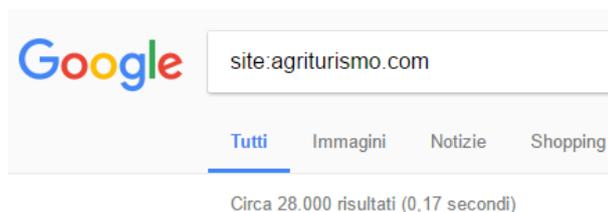


Figura 1.4. Pagine indicizzate con dominio agriturismo.com

Il motore di ricerca riesce ad indicizzare 28.000 pagine relative al dominio Agriturismo.com. Oltre a queste pagine è possibile fare una stima delle pagine del portale relative alle varie lingue in cui il sito è navigabile. Per Agriturismo.com vengono indicizzate circa 5.000 pagine per ogni lingua diversa dall'italiano e le lingue disponibili sono lo spagnolo, il tedesco, il francese e naturalmente l'inglese.



Figura 1.5. Pagine indicizzate con dominio agriturismo.it

La pagina Agriturismo.it invece è composto da 3 lingue, l'italiano, il tedesco e l'inglese. In Fig.1.5 è visibile il gran numero di pagine che viene indicizzato e la grande differenza tra Agriturismo.it e Agriturismo.com. Per quanto riguarda la lingua inglese vengono indicizzate 29.000 pagine per Agriturismo.it, mille pagine in più rispetto alle pagine italiane di Agriturismo.com, mentre le pagine in tedesco indicizzate sono 27.900. Ultimo portale analizzato, che rivela dati interessanti rispetto ai primi due benchmark citati, è Agriturismi.it.

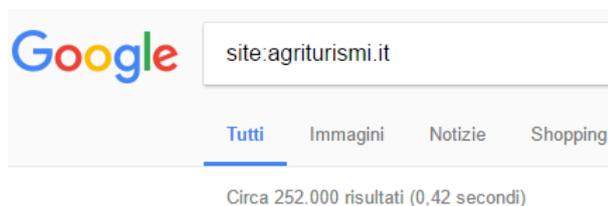


Figura 1.6. Pagine indicizzate con dominio Agriturismi.it

Le 252.000 pagine indicizzate da Google per Agriturismi.it rivelano una grandissima varietà di scelta e una grande visibilità delle offerte relative al portale. Questo dato è molto significativo rispetto agli altri due benchmark, ma la varietà di lingue disponibili e il numero di pagine indicizzate per ogni lingua è sorprendente. Agriturismi.it offre, oltre alle tradizionali spagnolo, inglese, francese e tedesco, altre tre lingue, che stanno diventando sempre più importanti al mondo d’oggi. Una di queste è il cinese, che offre circa 8.200 pagine indicizzate, seguito dal russo con 7.000 pagine disponibili e il giapponese che con 6.000 pagine indicizzate rende il portale navigabile da molti utenti diversi

### 1.3 La scelta del dominio

La scelta del dominio risulta fondamentale se si decide di pubblicare una pagina web. Sicuramente è importante inserire parole pertinenti all’argomento o al prodotto trattato dalla pagina e non più di tre parole chiave. In questo caso il dominio Agriturismo.com è senza dubbio ottimo, poiché tratta esclusivamente la visualizzazione e prenotazione di agriturismi. L’estensione “.com” conferisce al brand una copertura mondiale, a differenza di altri competitors, come Agriturismo.it o Agriturismi.it che utilizzano un dominio esclusivamente italiano. Considerando il dominio del portale e riprendendo in esame il paragrafo precedente relativo alle pagine indicizzate è doveroso creare e gestire più contenuti, in modo da avere più pagine indicizzate dal portale. L’anzianità del dominio rafforza ancora di più il livello SEO, facendo però attenzione a gestire al meglio aggiornamenti della pagina, poiché un sito vecchio che non viene aggiornato da anni viene comunque penalizzato<sup>4</sup>.

### 1.4 Il tag title

Il tag title è una parte fondamentale dell’ottimizzazione SEO. All’interno del portale sono riscontrabili alcuni errori che fanno diminuire l’ottimizzazione.

```
<title>Agriturismo.com - I migliori agriturismi per le tue vacanze in Italia</title>
```

Nell’analisi risulta essere troppo lungo, poiché non deve superare i 68 caratteri. La lunghezza consigliata per raggiungere buoni punteggi è fra i 57 e i 63 caratteri e la mancanza del tag H1 nella pagina, come visibile nei capitoli 2.3, 2.4, crea grandi problemi di ottimizzazione. Il primo fattore a cui il webmaster deve fare riferimento è se, all’interno del tag, è presente il simbolo di interpunzione dei due punti (:).

---

<sup>4</sup>Francesco De Nobili, 2015. SEO Google, guida pratica per farsi trovare con Google. Hoepli Editore pp. 23-24

Questo simbolo è molto importante poiché comunica al crawler una separazione del contenuto: una prima parte che forma un'introduzione dell'argomento e una seconda parte più descrittiva delle attività o funzioni svolte dalla pagina. In alternativa è utilizzabile il separatore "pipe" ( | )<sup>5</sup>, che ha la stessa funzione dei due punti. Il tag <title> preso in analisi comprende 69 caratteri, considerando anche gli spazi ed è ottimizzabile nel seguente modo:

```
<title>Agriturismo.com: I migliori agriturismi, le tue vacanze in Italia</title>
```

Il principale benchmark con cui possiamo fare un confronto è senza dubbio il portale Agriturismo.it. Questa pagina web risulta essere molto più ottimizzata rispetto ad Agriturismo.com, soprattutto per quanto riguarda il tag title:

```
<title>Agriturismo.it - Solo Agriturismi di qualita'!</title>
```

Il tag title è lungo solo 45 caratteri, ma utilizza in maniera ottima le keyword della pagina, garantendo una buona ottimizzazione. L'altro competitor, Agriturismi.it, presenta questo tag <title>:

```
<title>Agriturismi.it - Offerte in Agriturismo!</title>
```

Anche in questo caso Agriturismi.it ha scelto un tag title molto corto ma con una buona concentrazione di keyword: 40 caratteri sono molto pochi rispetto alla soglia consigliata per una buona ottimizzazione (almeno 55 caratteri). Entrambi i portali riescono a raggiungere una buona ottimizzazione grazie alla combinazione tag title – header tag, trattata più avanti nel capitolo 2.3.

```
<meta name="description" content="Agriturismo.com La migliore guida italiana all'agriturismo. Ricerca nel nostro portale turistico tutti gli agriturismi per la tua vacanza in Italia">
```

Il tag meta description visto sopra, è utile a dare un'anteprima e una descrizione all'utente che visita la pagina e risulta essere ottimizzato, restituendo una bella e precisa descrizione del portale. La lunghezza di 147 caratteri del tag meta description è perfetta per la creazione di snippet<sup>6</sup>, poiché i contenuti che superano i 155 caratteri verranno in parte tagliati dal motore di ricerca, facendo perdere informazioni alla pagina.

---

<sup>5</sup>Francesco De Nobili, 2015. SEO Google, guida pratica per farsi trovare con Google. Hoepli Editore pp. 27

<sup>6</sup>consiste nella descrizione della pagina che appare in ogni SERP

## 1.5 Rich snippet: rendersi attraenti

Se il tag meta description fornisce un assaggio del contenuto della pagina è importante inserire informazioni aggiuntive e che arricchiscono il contenuto delle SERP tramite i **rich snippet**. Nella Fig.1.7 è visibile un esempio di rich snippet, identificabile dalla presenza di cinque stelline relative a feedback o recensioni di esperienze su un determinato argomento, da parte dell'utente.



Figura 1.7. Esempio di rich snippet

I tre portali analizzati mostrano la totale mancanza di microdati o microformati e quindi la mancanza di rich snippet. Per inserire questi arricchimenti stilistici è possibile usare *Schema.org*. Schema.org è una pagina web costruita per i più grandi motori di ricerca, in cui vengono elencate le istruzioni e i marcatori utili ai motori di ricerca. Il webmaster deve fornire le informazioni aggiuntive del proprio sito secondo il formato voluto e contenuto in Schema.org. Nell'atto pratico si tratta di inserire i marcatori all'interno di tag HTML e questa operazione aiuta il motore di ricerca a meglio indicizzare un contenuto<sup>7</sup>.

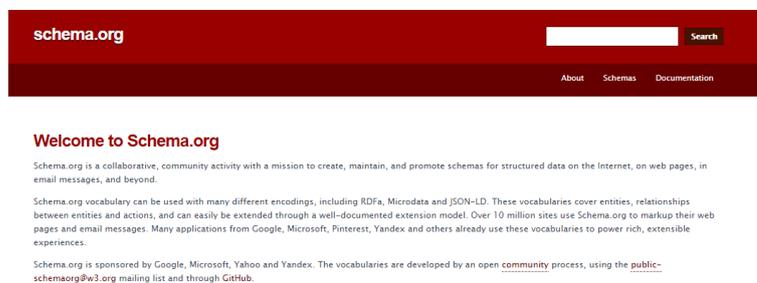


Figura 1.8. Homepage di Schema.org

La mancanza di microdata in tutti e tre i portali porta a problemi di ottimizzazione, ma rende anche possibile, per Agriturismo.com, di guadagnare posizioni sugli altri benchmark. I microdata vengono usati per etichettare dei contenuti e questo tipo di markup risulta utile per inserire i rich snippet nelle SERP. Per prima cosa è importante costruire un contenitore principale, che si identifica nel seguente modo:

---

<sup>7</sup><http://www.cmswiki.net/seo-posizionamento/microdata-e-schemaorg>

```
<div itemscope itemtype="http://schema.org/Struttura"></div>
```

Il contenitore principale contiene alcuni attributi, come l'attributo *itemprop* che risulta utile nella valorizzazione dello snippet <sup>8</sup>. Essendo un portale per prenotare e visualizzare gli agriturismi italiani sarebbe opportuno creare rich snippet per aziende, che viene formato da attributi come la geolocalizzare nel risultato, indicando la via del punto di riferimento, la città, il numero di telefono, per ogni struttura del portale. Un'altra tipologia utile ad Agriturismo.com risulta essere il rich snippet per prodotti in vendita, che gestiscono molti attributi, tra cui immagini, prezzi e disponibilità. Recensioni, giudizi e punteggi sono altri fondamentali dati aggiuntivi per attirare clienti ed essere ancora più visibili tra le SERP <sup>9</sup>. Un esempio di inserimento e visualizzazione di recensioni di una struttura è il seguente:

```
<div itemscope itemtype="http://data-vocabulary.org/Review">

<span itemprop="itemreviewed">Nome del cosa si sta recensendo</span>
<span itemprop="rating" itemscope
itemtype="http://data-vocabulary.org/Rating">Voto
<span itemprop="average">4</span> di <span itemprop="best">5</span>
</span> su un totale di <span itemprop="votes">1276</span> voti.
<span itemprop="count">767</span>
recensioni positive degli utenti.</div>
```

## 1.6 Il tag meta robots

Per garantire un buon posizionamento della pagina è importante gestire al meglio ciò che lo spider deve eseguire o meno durante la scansione del sito. Questa operazione è possibile tramite questa sintassi:

```
<meta name="robots">
```

Che, in aggiunta di alcuni attributi, contiene le informazioni necessarie per far muovere il motore di ricerca tra le pagine web da scansionare. Gli attributi che vengono inseriti, possono essere di tipo *Follow*, *NoFollow*, *Index* e *No Index*. Modificando la sintassi precedente in questo modo:

```
<meta name="robots" content="NoIndex">
```

---

<sup>8</sup><http://www.alessiopomaro.com/microdati-microformati-rich-snippet-cosa-sono/>

<sup>9</sup><http://www.internetpost.it/seo-rich-snippet/>

la pagina potrà o meno esser scansionata dal motore di ricerca. Cambiando invece l'attributo in Follow o NoFollow, informeremo il crawler che può visitare o non visitare i link contenuti nella pagina. Sia Index che Follow risultano di poca importanza all'interno del tag meta robots, poiché i motori di ricerca sono creati appunto per indicizzare e seguire i link delle pagine. È utile sottolineare che per evitare che venga seguito un link o che una pagina non venga scansionata, è compito del webmaster scegliere il giusto attributo per il tag<sup>10</sup>.

## 1.7 Il file robots.txt

A differenza del tag meta robots, che gestisce le istruzioni di una singola pagina, è possibile gestire le URL del sito nel suo complesso e scegliere se farle scansionare o meno dal motore di ricerca, tramite il file `robots.txt`. I risultati delle ricerche possono così essere gestiti dal webmaster, informando il crawler su quali sono le pagine più importanti da indicizzare. Per impartire un'istruzione allo spider si usano due semplici comandi, *User-agent* e *Disallow*. Il comando *User-agent* specifica a quale spider viene applicata una certa istruzione. Ad esempio, se vogliamo comunicare al crawler di Google una certa istruzione, verrà inserito nel file robots.txt questa riga:

```
User-agent: Googlebot
```

In aggiunta alla riga di esempio si può inserire l'attributo *Disallow*, che serve ad impedire l'accesso ad una cartella o file del sito che vogliamo indicizzare o far vedere. Nell'esempio seguente è possibile vedere un'esempio di restrizione

```
User-agent: Googlebot  
Disallow : /foto/
```

In questo modo non verrà indicizzata la cartella relativa alle foto del sito.

Nel frammento di Fig.1.9 è possibile vedere il comando User-agent utilizzato con il valore \*(asterisco), che viene utilizzato per dare le seguenti istruzioni a tutti i motori di ricerca. Successivamente il comando Disallow informa gli spider di non seguire le directory relative ai file *Javascript CSS* e le cartelle relative alle immagini. Gli standard da seguire per usare e gestire al meglio il file robots.txt sono:

- la corretta scrittura del nome del file robots.txt (scritto in minuscolo)
- il corretto posizionamento del file, che deve essere nel percorso principale del sito<sup>11</sup>.

---

<sup>10</sup><http://www.seocenter.it/learn/seo/abc/meta-tag-robots.html>

<sup>11</sup>[www.seocenter.it/learn/seo/abc/file-robots-txt.html](http://www.seocenter.it/learn/seo/abc/file-robots-txt.html)

```
User-Agent: *
Disallow: /adm/
Disallow: /agriinclude/
Disallow: /agriturismo-site-map.html
Disallow: /ajax/
Disallow: /APPOGGIO/
Disallow: /aspnet_client/
Disallow: /backup/
Disallow: /css/
Disallow: /css1/
Disallow: /curiosita/
Disallow: /download/
Disallow: /ele/
Disallow: /Emilia_Romagna/
Disallow: /englisch/
Disallow: /ext/
Disallow: /gif/
Disallow: /images/
Disallow: /imgs/
Disallow: /includes_az/
Disallow: /js/
Disallow: /js1/
Disallow: /mappe/
```

Figura 1.9. Frammento di robots.txt del portale Agriturismo.com

## 1.8 La sitemap

Realizzata per aiutare la navigazione dell'utente e del crawler, la *sitemap* è un semplice indice del sito web. È di fondamentale importanza per il motore di ricerca poiché cataloga tutti gli URL delle pagine del sito e aiuta il webmaster nella gestione degli aggiornamenti di ogni singola pagina, oltre a definire se una pagina è più importante di un'altra all'interno del sito<sup>12</sup>. Viene creata in XML<sup>13</sup> ed è utile a dare significato e fornire un controllo ai contenuti del documento trattato.

I parametri fondamentali per una sitemap sono gli indirizzi delle pagine, la frequenza degli aggiornamenti, catalogabile in *daily*, giornaliera, *weekly*, settimanale e *monthly*, mensile. Il tag relativo alla priorità va da un valore massimo di 1.0 ad un valore minimo di 0. Quest'ultimo fattore è molto importante per il crawler perché, come vediamo nella sitemap di Agriturismo.com, è stata inserita una grande priorità di indicizzazione per la homepage e le pagine relative ad offerte e last minute.

```
<url>
  <loc>http://www.agriturismo.com</loc>
  <changefreq>daily</changefreq>
  <priority>1.0</priority>
</url>
```

---

<sup>12</sup>[www.seocenter.it/learn/seo/tuts/creare-una-sitemap.html](http://www.seocenter.it/learn/seo/tuts/creare-una-sitemap.html)

<sup>13</sup>un linguaggio marcatore, chiamato eXtensible Markup Language

```
<url>
  <loc>
    http://www.agriturismo.com/offerte-agriturismo/index.asp
  </loc>
  <changefreq>daily</changefreq>
  <priority>1.0</priority>
</url>

<url>
  <loc>
    http://www.agriturismo.com/lastminute-agriturismo/index.asp
  </loc>
  <changefreq>daily</changefreq>
  <priority>1.0</priority>
</url>
```

Se la precedente analisi evidenziava le carenze strutturali di Agriturismo.com rispetto agli altri benchmark, la sitemap del portale ci restituisce qualche punto in più. Infatti la sitemap risulta essere reperibile e ben strutturata per Agriturismo.com, mentre per Agriturismo.it e Agriturismi.it non risulta essere inserita oppure non reperibile, penalizzandone il posizionamento finale.

## 1.9 Url rewriting: dove mi trovo?

Un'altra tecnica che permette una buona ottimizzazione della pagina web, oltre ad aumentare la *user experience* dell'utente, è la creazione di url "parlanti". Nell'immagine seguente possiamo subito vederne un esempio per chiarirne le idee.

```
http://www.fileli.unipi.it/infouma/laurea-triennale/tirocini/
```

Dopo un po' di navigazione nella pagina capita spesso di non capire in che punto o sezione della pagina un utente si trova. L'url "parlante", come possiamo notare nell'esempio, fa capire chiaramente che l'utente si trova all'interno della sezione dedicata ai tirocini, che si trova nella sezione relativa alle lauree triennali del corso di laurea in Informatica Umanistica che, a sua volta, è nella pagina relativa al dipartimento di Filologia Letteratura e Linguistica. Questi accorgimenti sono molto importanti sia per garantire una buona ottimizzazione, sia per quanto riguarda l'esperienza portale – utente, che risulterà più facile e quindi migliore. Agriturismi.it presenta per i propri indirizzi informazioni precise, con un *path*<sup>14</sup> preciso e chiaro.

```
http://www.agriturismi.it/it/toscana/signa/agriturismo_il_borghetto.html
```

Agriturismo.it presenta un buon path, seppur lungo, con tutte le categorie e le cartelle ben specificate

```
http://www.agriturismo.it/it/agriturismi/toscana/siena/
SanGiovanni-8620908/index.html
```

---

<sup>14</sup>percorso del file o di una risorsa all'interno della pagina

Per quanto riguarda Agriturismo.com gli indirizzi relativi a dove si trova o meno un utente nella pagina, le informazioni si fermano quasi subito. Come possiamo notare nel codice seguente, l'url afferma che ci troviamo in una pagina relativa ad un agriturismo, ma non sappiamo né il tipo di agriturismo, che in questo caso è rappresentato da un "id" con numero 4794, né sappiamo la regione o la provincia in cui si trova, diminuendo quindi le informazioni utili all'utente.

`http://www.agriturismo.com/dettaglioAgriturismo.asp?idLingua=1&id=4794`

Per rendere i link molto informativi per l'utente e per aumentarne l'ottimizzazione è importante impostare questi indirizzi in modo che contengano sempre le parole chiave contenute sia nel tag `title`, sia nel tag `meta description`. Oltre a questo sarà importante inserire il separatore "-" dopo ogni sezione, rendendo appetibile il link per il crawler.

# Capitolo 2

## Il corpo della pagina

Il contenuto che verrà visualizzato dall'utente in ogni pagina web è inserito in una sezione della pagina HTML, delineata da un tag aperto e chiuso, chiamato `<body>`. Immagini, testo e molto altro contenuto, saranno visualizzati in questa pagina e sarà possibile gestire il tipo di *font* (tipologia di carattere per il testo), grandezza dell'immagine e colori.

### 2.1 Le chiavi di ricerca

Nelle pagine del portale sarà di fondamentale importanza inserire le keyword (parole chiave) nei tag di intestazione come `<title>` e `<meta description>`, ma anche all'interno del tag `<body>`. Esiste un tool online molto famoso ed utilizzato da esperti SEO, chiamato *SEMrush* che fornisce molti dati relativi ai volumi di ricerca in base alle parole chiave ricercate. SEMrush permette a chiunque di analizzare il traffico e le keyword più ricercate per periodi di tempo diversi. La schermata che più interessa per controllare il numero di ricerche effettuate e le percentuali di ricerca per ogni parola chiave è la schermata della **ricerca organica**<sup>1</sup>. E' formata da 12 colonne ed offre molti spunti interessanti.

Le colonne offrono molte informazioni ma solo alcune sono utili alla nostra analisi. La prima colonna rappresenta la keyword e la seconda colonna, relativa alla posizione, ha come valore due numeri, di cui uno fra parentesi. Il primo numero ci indica la posizione che la keyword fa guadagnare al portale nei motori di ricerca mentre il secondo numero, fra parentesi, rappresenta la vecchia posizione che la keyword faceva guadagnare al portale prima dell'ultimo aggiornamento. E' chiaro che la keyword "agriturismo.com" ci fa essere in prima posizione per quella parola

---

<sup>1</sup> è la lista di risultati (SERP) che vengono forniti dai motori di ricerca

Keyword	Pos.	Volume	KD	CPC (USD)	URL	Traffico, %	Costi %	Con.	Risultati
agriturismo	3 (3)	90,500	63.32	0.65	www.agriturismo.com/	34.57	39.83	0.16	28,400,000
agriturismi	3 (3)	6,600	76.26	0.47	www.agriturismo.com/	2.52	2.09	0.48	8,840,000
agriturismo umbria	7 (7)	9,900	64.96	0.89	www.agriturismo...umbria	1.68	2.64	0.79	514,000
agriturismo.com	1 (1)	480	61.06	0.49	www.agriturismo.com/	1.63	1.41	0.12	314,000

Figura 2.1. Dettaglio SEMrush su analisi del 12 gennaio 2017

chiave. Altre keyword molto importanti che ci portano in terza posizione nei motori di ricerca sono “agriturismo” e “agriturismi”. E’ interessante notare che la paro-

Keyword	Pos.	Volume	KD	CPC (USD)	URL	Traffico, %
last minute	18 (18)	301,000	79.62	0.31	www.agriturismo...rismo/	1.04
agriturismo	3 (3)	90,500	62.81	0.65	www.agriturismo.com/	47.07
lastminute	31 (32)	74,000	82.81	0.34	www.agriturismo...rismo/	0.04

Figura 2.2. Numero delle interrogazioni mensile (colonna Volume)

la chiave “agriturismo” porta sul portale il 47,07% del traffico mensile, con 90,500 interrogazioni mensili. Altre colonne degne di nota e interessanti sono:

- la colonna relativa alle SERP, che ci restituisce un’anteprima della pagina con i risultati della ricerca per ogni keyword
- CPC, utile per effettuare analisi sui costi medi, in dollari, di ogni click

Infine è importante consultare la colonna “last update” gli aggiornamenti delle analisi su ogni keyword. Il report contiene dati che vengono costantemente aggiornati.

**Agriturismo.it - Solo Agriturismi di qualità!**  
[www.agriturismo.it/](http://www.agriturismo.it/)  
 Il portale sull'agriturismo più visitato in Italia: solo i migliori agriturismi scelti e verificati da noi, con commenti di migliaia di ospiti. Agriturismo Lastminute e ...  
 Agriturismo in Italy · Trentino Alto Adige · Veneto · Piemonte

**Agriturismo a Cecina**  
[www.agriturismo.it](http://www.agriturismo.it) > Toscana > Livorno  
 Agriturismo Cecina: scegli tra le migliori 3 offerte di Agriturismo a Cecina.

**Agriturismi.it - Offerte in Agriturismo!**  
[www.agriturismi.it/](http://www.agriturismi.it/)  
 Agriturismi.it è il portale degli agriturismi italiani. Prenota la tua vacanza in Agriturismo e scegli tra migliaia di Offerte e Lastminute in Agriturismo. Da 29 Euro a ...  
 Toscana · Pasqua In Agriturismo · Campania · Calabria

**Agriturismo.com - I migliori agriturismi per le tue vacanze in Italia**  
[www.agriturismo.com/](http://www.agriturismo.com/)  
 Agriturismo.com La migliore guida italiana all'agriturismo. Ricerca nel nostro portale turistico tutti gli agriturismi per la tua vacanza in Italia.

Figura 2.3. La terza posizione guadagnata dalla keyword “agriturismi”

## 2.2 Enfaticizzare i contenuti

Il contenuto che verrà inserito nella pagina web conterrà sicuramente informazioni più o meno rilevanti riguardo l'argomento trattato. È possibile enfaticizzare queste porzioni di testo utilizzando dei tag HTML, per arricchire e personalizzare le proprie pagine. Di seguito è possibile confrontare i tag e le funzioni che svolgono<sup>2</sup>:

- il tag

```
<strong>esempio</strong>
```

trasformerà la parola esempio in “**esempio**”;

- il tag

```
<em>esempio</em>
```

trasformerà la parola esempio in “*esempio*”;

- per quanto riguarda il tag

```
<u>esempio</u>
```

invece, la parola esempio verrà trasformata in “esempio”.

Gli spider sono molto interessati anche a questi arricchimenti stilistici aggiunti dal webmaster. È bene tenere presente però che un eccessivo uso di tag enfaticizzanti, come in Fig.2.4, porterà a problemi di sovra-ottimizzazione, oltre alla riduzione di facilità della lettura del testo da parte dell'utente<sup>3</sup>.

Gestire al meglio i fattori On-page porterà sicuramente risultati ottimi per l'indicizzazione e il ranking della pagina web. Riempire il testo con parole in grassetto, sottolineature e corsivi, indicherà al crawler che la nostra pagina sta superando i livelli di ranking assegnati dal motore di ricerca, penalizzandone il posizionamento<sup>4</sup>.

---

<sup>2</sup><http://www.seofaidate.com/tag-di-enfaticizzazione>

<sup>3</sup><http://www.seocenter.it/learn/seo/abc/ottimizzazione-seo.html>

<sup>4</sup>Fortis Daniele, 2013. Scrivere per il web. Maggioli Editore pp. 228

L'agriturismo si trova in un contesto aziendale di circa 70 ettari. Nell'Azienda si coltivano prevalentemente cereali (grano, orzo, avena) e si produce, con metodi tradizionali, un buon vino, frutto del bellissimo vigneto appena adiacente all'agriturismo. La Famiglia Marzoli vive nell'abitazione che si trova a 100 metri di distanza dall'Agriturismo.

La gestione familiare da parte della Famiglia Marzoli, che si dedica con particolare cura e attenzione ai propri ospiti, permette un contatto diretto e continuo con gli ospiti.

Possibilità di fare degustazioni di piatti della cucina tipica toscana preparati con i prodotti genuini dell'azienda, in particolare ortaggi e verdure dell'orto, e utilizzando integralmente prodotti toscani a km 0, rievocando la tradizione delle antiche ricette casarecce.

Produzione di vino e olio, con metodi della sapienza contadina tradizionale. Possibilità di visitare le attività dell'azienda e gli animali della fattoria: galline, conigli, caprette e maiali.

Figura 2.4. Eccesso di tag stilistici in agriturismo.com

## 2.3 L'effetto "matrioska" degli header tag

Gli Header Tag sono dei tag nei quali sono contenuti i titoli che vengono dati ai contenuti della pagina web. Gestire al meglio l'utilizzo e la densità di questi tag è importante, facendo molta attenzione soprattutto ad una precisa gerarchizzazione. Esistono sei tipi di header tag:

- H1, che è la misura più grande che è possibile dare ad un titolo,
- H2, H3, H4, H5 che misurano grandezze in ordine decrescente,
- H6 che è la misura più piccola per un titolo di paragrafo.

Un focus particolare è per il tag H1: Questo tag è di vitale importanza in ogni pagina web, poiché richiede l'inserimento di parole chiave al suo interno. Per strutturare al meglio una pagina con i tag header, dobbiamo dar loro un ordine di tipo decrescente. È quindi importante scrivere titoli di paragrafo a partire da H1 per arrivare, se necessario, ad H6.

Scrivere header tags nel modo seguente:

- `<h2>Titolo paragrafo 1</h2>`
- `<h1>Titolo paragrafo 2</h1>`
- `<h5>Titolo paragrafo 3</h5>`

porterà invece a problemi di posizionamento, oltre che di natura grafico - organizzativa dei contenuti. Il layout della pagina perderà sicuramente a livello estetico, ma sarà comunque possibile gestire la grafica della pagina, con l'inserimento di fogli di stile<sup>5</sup>.

---

<sup>5</sup>Da file pdf Michele De Capitani, Guida posizionamento dei motori di ricerca pp. 13/14

## 2.4 Assenza del tag H1

Il problema presente in Agriturismo.com è la totale assenza di un tag H1. Nella homepage analizzata, risulta essere presente un tag `<p>`, un tag HTML che viene utilizzato come contenitore di testi o altri tag della pagina web. Di seguito è visibile il frammento di codice del tag p.

```
<p class="titolo">ACCOGLIENTI PER NATURA</p>
```

Questo errore compromette molto il posizionamento e sarà utile sostituire, all'interno della pagina, la riga di codice, in questo modo:

```
<h1>ACCOGLIENTI PER NATURA</h1>
```

Parlando sempre di titoli e sottotitoli, nella pagina principale, è presente anche un sottotitolo, il cui codice è il seguente:

```
<p class="titolo2">
  <strong>FERMATI!</strong><br>...ALLA TUA VACANZA PENSIAMO NOI
</p>
```

Il frammento non rispecchia minimamente gli standard SEO e questo tipo di errori, seppur di piccola dimensione, abbassano notevolmente il rank della pagina. Anche in questo caso è stato possibile sopperire a questo inconveniente sostituendo il tag `<p>` con il tag H2, senza togliere l'enfaticizzazione del testo resa dal tag `<strong>`.

```
<h2>
  <strong>FERMATI!</strong><br>...ALLA TUA VACANZA PENSIAMO NOI
</h2>
```

Riprendiamo adesso in considerazione il portale Agriturismo.it per un ulteriore confronto tra i due portali. Il tag H1 garantisce alla pagina una buona ottimizzazione SEO:

```
<h1 class="main-sprite agriturismo-logo">Agriturismo.it</h1>
```

Di sopra vediamo il tag H1 di Agriturismo.it, che però non è ottimizzato al meglio, poiché racchiude soltanto una keyword al suo interno. Anche in Agriturismi.it il tag H1 è presente ma non è ottimizzato al meglio:

```
<h1>Agriturismo in Italia dal 1999 </h1>
```

In Agriturismo.com circa l'82% delle pagine del portale risulta sprovvisto di tag H1.

## 2.5 Internet per tutti: gli attributi delle immagini

Inserire un'immagine nella propria pagina web è un incremento di informazione che diamo alla pagina, curando sia l'aspetto grafico, sia quello informativo. Quando si inseriscono nuove immagini nella pagina web, è bene scegliere nomi distinti per ogni immagine e personalizzare l'attributo `alt`, conosciuto anche come “testo alternativo”.

La sintassi seguente chiarirà meglio questo importante aspetto:

```

```

Nell'attributo “`src`”, possiamo visualizzare il nome del file e la sua allocazione (se presente). Nell'attributo “`alt`” invece viene fornita una descrizione dell'immagine stessa. Questi accorgimenti sono importanti per il crawler, poiché se verrà inserita una descrizione nell'attributo “`alt`” più simile possibile al nome del file, l'ottimizzazione sarà ottima; senza però eccedere nell'uso di keyword, per evitare penalizzazioni di sovra-ottimizzazione. È importante anche per gli utenti perché in caso di problemi con la visualizzazione dell'immagine, le informazioni presenti nell'attributo `alt` dell'immagine verranno stampate sulla pagina al posto della stessa. È buona norma creare una cartella in cui sono contenute le immagini, per agevolare lo spider nella ricerca e salvarle con estensioni JPEG, PNG o GIF, che risultano essere compatibili con i principali motori di ricerca e browser<sup>6</sup>.

## 2.6 Ottimizzare le immagini

All'interno del progetto preso in esame sono presenti molte immagini che non risultano ottimizzate con l'inserimento dell'attributo “`alt`”. Le immagini che non hanno implementato l'attributo “`alt`”, penalizzano la *user experience* e il posizionamento della nostra pagina. Di seguito sono presenti alcuni frammenti di codice che rispecchiano la mancanza dell'attributo “`alt`”.

```
  
  

```

In entrambi i casi è presente l'attributo “`src`”, che indica dove è situata una certa immagine. Da questi screenshot è possibile estrapolare ciò che va inserito nell'attributo “`alt`”, inserendo quindi una descrizione dell'immagine. Di seguito i nuovi interventi utili per ottimizzare le immagini

---

<sup>6</sup><https://static.googleusercontent.com/media/www.google.com/it//intl/it/webmasters/docs/search-engine-optimization-starter-guide-it.pdf> pp. 18

```
  
  

```

Oltre all'aggiunta dell'attributo alt è importante inserire all'interno del tag relativo alle immagini anche l'attributo "title". Anch'esso risulta importante per l'ottimizzazione delle pagine. Di seguito è possibile vedere alcuni esempi che riportano errori

```
  
  
  
  

```

Sia l'attributo "alt" sia l'attributo "title" non sono presenti in questi tre esempi. Gli errori sono stati individuati e saranno modificati con i giusti attributi

```
  
  
  
  

```

## 2.7 Attributi a confronto

Una sezione che accomuna tutti e tre i portali e che porta maggiori visite, è la sezione delle "offerte", che possiamo trovare nel menù della pagina principale dei portali. Questa categoria è utile per poter fare una stima delle immagini presenti in questa sezione, estraendone i risultati utili alla nostra analisi. La mole di immagini nella sezione offerte di Agriturismo.it è molto grande, con 2003 immagini nella pagina, di cui 2002 ottimizzate con l'attributo "alt"! Per quanto riguarda Agriturismi.it le immagini della sezione esaminata sono veramente poche (52), e risultano tutte perfettamente ottimizzate. Agriturismo.com presenta 102 immagini nella sezione "offerte", di cui 62 ottimizzate con attributo "alt".

## 2.8 SEO Social: farsi conoscere oggi

Il web marketing utilizza una miriade di tecniche per promuovere la propria pagina. Una particolare caratteristica del Web 2.0 è senza dubbio la promozione dei propri prodotti tramite i social network. L'integrazione di *Fanpage*, la condivisione di articoli o notizie relative all'argomento trattato nelle pagine, porteranno alla creazione di fidelizzazioni e collaborazioni utili alla visibilità stessa della pagina. Le esperienze, le foto, le notizie vengono condivise dagli utenti, innescando una serie di "feedback" sia positivi che negativi sull'utilizzo o sull'esperienza del sito web.



Figura 2.5. Dettaglio pagina Twitter di Agriturismo.com

*Facebook* è senza dubbio il social network più famoso ed utilizzato al mondo e promuoversi utilizzandolo può portare grande visibilità alla nostra pagina. Le *Fanpage* sono state create su Facebook per svolgere attività di promozione pubblicitaria ed è possibile condividere tutto ciò che ha a che fare con l'argomento o con i prodotti offerti dal portale. Questa pagina è visibile anche ai non utenti Facebook e a chi non è presente sulla pagina, garantendo una grande visibilità, poiché il numero di fan che una pagina può avere è illimitato. Agriturismo.com presenta una pagina Facebook costantemente aggiornata ed è seguita da circa 170.000 utenti. Agriturismo.it ha circa 122.000 like alla pagina, anche se il nome della pagina, chiamata soltanto "Agriturismo", non è facile da trovare. I contenuti di questa pagina sono stati condivisi circa 5549 volte, creando una grande visibilità fra i profili sia interni, sia esterni alla pagina. Il portale Agriturismi.it invece supera di poco le altre due pagine, poiché registra circa 195.000 like alla pagina.

Un altro social network molto importante e famoso nel mondo è *Twitter*, un sito di micro-blogging in cui è possibile far seguire o farsi seguire da altri utenti, chiamati *followers*. Questo social network consente la creazione di un profilo, in cui è possibile inserire notizie o brevi messaggi, poiché la possibilità di scrittura per ogni



Figura 2.6. Dettaglio pagina Twitter di Agriturismi.it

messaggio è molto limitata (circa 150 caratteri a messaggio) . All'interno del profilo che viene creato sarà possibile inserire il link della nostra pagina web e la scelta del nome che useremo per il profilo sarà importantissima. Twitter viene quindi usato per aggiornarsi, per seguire e per essere seguiti da profili interessati a ciò che viene offerto. Emergere su Twitter è molto difficile, soprattutto se dobbiamo competere con altri benchmark, poiché il segreto di Twitter è l'utilizzo e l'aggiornamento quotidiano di notizie e offerte, poiché i *tweet* creati dovranno coinvolgere ed interessare i nostri follower. A differenza di Facebook, i tweet utilizzano un particolare simbolo che precede le parole chiave delle nostre ricerche: il cancelletto, o più conosciuto come *hashtag*.



Figura 2.7. Dettaglio pagina Twitter di Agriturismo.it

Dei tre portali sicuramente il meno attivo risulta essere Agriturismi.it, mentre il numero dei tweet di Agriturismo.com e Agriturismo.it è quasi equivalente. Nonostante questi numeri, è proprio Agriturismo.it ad avere il più alto numero di follower, mentre Agriturismo.com registra poco più di 1.000 follower. Il prezzo che però viene pagato da Agriturismo.com in fatto di ottimizzazione SEO è che non ha la *Fanpage*

di Facebook e il profilo Twitter collegate direttamente alla pagina. In questo modo quindi il prodotto sarà pubblicizzato e condiviso sulle pagine, ma per quanto riguarda il posizionamento risulterà non attivo né su Facebook, né su Twitter, interessando ancor meno gli spider dei motori di ricerca.

## Capitolo 3

# Analisi statistica sull'utilizzo del database

E' importante effettuare anche uno studio sull'utilizzo delle tabelle del database di Agriturismo.com. Il database contiene molte tabelle, alcune relative a dati aziendali o relative ad offerte. Su Agriturismo.com è possibile filtrare le proprie ricerche in base alla regione, consultare le offerte last minute, una sezione dedicata alla cultura, ai sapori e alle curiosità. Inoltre è possibile accedere al portale come cliente registrato o come gestore della struttura, in modo da poter aggiornare le foto, le informazioni o le date di inizio di alta, media e bassa stagione. Per riservatezza e in linea con le regole aziendali è possibile mostrare solo alcuni frammenti di tabelle che non contengono dati sensibili. Di seguito sono elencati i passaggi necessari all'estrazione e alla visualizzazione delle tabelle utilizzate in Agriturismo.com:

- creazione di un programma Python utile all'estrazione dei nomi delle tabelle
- i dati sono stati inseriti in un database locale ed estratti in formato Json<sup>1</sup>, poi sono stati inseriti in 3 diversi grafici utilizzando Highcharts<sup>2</sup>, uno a torta e due a barre
- il primo grafico divide in diversi range le tabelle trovate, nel nostro caso i range saranno tre, con tabelle usate più di 20 volte, usate tra le 20 e le 5 volte e quelle usate meno di 5 volte
- il grafico a barre raffigura le tabelle più utilizzate

---

<sup>1</sup>Acronimo di JavaScript Object Notation, tipo di formato dei dati utile allo scambio di informazioni client-server

<sup>2</sup>tool Javascript che, in aggiunta di JQuery, permette la creazione di grafici interattivi

- il secondo ed ultimo grafico prende in esame le tabelle meno usate nel codice.

Agriturismo.com è stato sviluppato utilizzando il linguaggio ASP<sup>3</sup> e presenta query di grandi dimensioni

```
strSQL = "SELECT agri_new.inserzioni.* FROM
agri_new.inserzioni INNER JOIN agri_new.tcomuni
ON agri_new.inserzioni.IdComune = agri_new.tcomuni.IDCOMUNE
INNER JOIN agri_new.tprovince ON
agri_new.tcomuni.IDPROVINCIA = agri_new.tprovince.IDPROVINCIA
INNER JOIN agri_new.tregioni ON agri_new.tprovince.IDREGIONE
= agri_new.tregioni.IDREGIONE WHERE agri_new.tregioni.IDREGIONE = "
```

I dati utili ed interessanti per la nostra analisi statistica sono i nomi delle tabelle

### 3.1 Estrazione delle tabelle

Il codice sviluppato estrae i nomi delle tabelle all'interno del codice e li stampa nel foglio di output, senza prendere in considerazione gli attributi della tabella.

```
file = open("codice/leggiAgriturismi.asp", "r")
output = open("risultati/output.csv", "w")

for line in file:
    if "SELECT" in line:
        x = line.split("FROM")
        z = x[1].split("WHERE")
        y = z[0].split("INNER JOIN")
        for i in y:
            rew = i.split(" ")
            dc = rew[1].split(" ")
            hh = dc[0].split("agri_new.")
            for linea in hh:
                print >> output, linea
```

Entrando nel dettaglio del codice precedente, possiamo certamente vedere la grande versatilità e usabilità del linguaggio Python. In pochissime righe abbiamo estratto i nomi delle tabelle dai fogli di codice.

---

<sup>3</sup>Acronimo di Active Server Pages, pagine formate da codice HTML e script utili alla creazione di pagine dinamiche

```
file = open("codice/leggiAgriturismi.asp", "r")  
  
output = open("risultati/output.csv", "w")
```

Nel *parser* Python è possibile notare due diversi file, uno di input, che sarà il file da analizzare e uno in output, che sarà il file in cui il programma scriverà i nomi delle tabelle

La prima iterazione del codice analizzerà riga per riga il file e per ogni riga controllerà se è presente la stringa “SELECT”. Le query inizieranno sicuramente con la parola SELECT, quindi per iniziare vengono ricercate tutte le righe che iniziano con questa parola, prendendo tutte le query di ogni pagina di codice. Successivamente viene utilizzata la funzione *split*, che serve per dividere una stringa in una lista di parole<sup>4</sup>. Il modulo *split* viene modificato, inserendo all'interno delle parentesi un delimitatore che nel nostro caso sarà “FROM”. Successivamente “splitteremo” ancora il risultato per ben due volte, delimitando la tabella e i relativi attributi.

La realizzazione di un nuovo ciclo *for* utile alla rimozione degli attributi delle tabelle prende in considerazione solo la tabella. Tutto ciò è possibile grazie allo *split* che, come abbiamo visto prima, divide un input in una lista. Il prodotto finito verrà quindi scritto tramite il comando *print* sul file chiamato *output*. L'analisi è stata effettuata su 268 fogli di codice scritti con linguaggio ASP. I risultati ricevuti dalle pagine analizzate mostrano circa 700 nomi di tabelle estratti, con molta differenza di utilizzi all'interno del portale, in quanto appartenenti a tabelle diverse.

## 3.2 Creazione dei grafici

I dati sono stati inseriti in un nuovo database locale e saranno utili alla creazione dei grafici per confrontare i risultati dell'analisi. I dati vengono estratti con un codice Php<sup>5</sup>, che trasforma in formato Json nel modo seguente:

```
<?php  
header('Content-Type: application/json');
```

---

<sup>4</sup><http://www.python.it/doc/Howtothink/Howtothink-html-it/chap08.htm>

<sup>5</sup>Acronimo di PHP Hypertext Preprocessor, free software utile alla creazione di pagine dinamiche (come ASP)

```
require("conn.php");

$res = select($mysqli, "SELECT tabella, COUNT(*) as totale FROM
tabelle GROUP BY tabella ORDER BY totale DESC");

echo json_encode($res);

?>
```

Il risultato del codice precedente restituirà una lista di elementi in formato Json, esempio visibile in Fig.3.1. Il progetto ha richiesto l'utilizzo di 3 differenti grafici. Per prima cosa era importante suddividere in diversi range i dati in output, quindi l'opzione del grafico a torta ha reso il lavoro e l'analisi più precisa. Le tabelle utilizzate sono state divise in:

- tabelle che superano le 20 occorrenze nel codice;
- tabelle che sono nel codice tra le 20 e le 5 volte;
- tabelle che non superano le 5 occorrenze nel codice

```
[
  {
    inserzioni: "75"
  },
  {
    tregioni: "71"
  },
  {
    tipoins: "71"
  },
  {
    tprovince: "57"
  },
  {
    tcomuni: "55"
  },
  {
    posgeo: "41"
  },
  {
    utentiweb: "34"
  },
  {
    prodacq: "32"
  },
  {
    lingue: "29"
  },
]
```

Figura 3.1. Lista delle tabelle con le relative occorrenze

Come visibile nel codice php precedente, è presente una semplice riga, in cui vengono contate tutte le tabelle incontrate nei fogli di codice in linguaggio ASP e per ognuna di queste verranno contate le volte che occorre una tabella, ordinando i risultati in ordine decrescente. Successivamente il risultato di questa pagina php dovrà essere collegato alla pagina del grafico, sviluppata in codice Javascript.

```
var contenitore = {
    max: 0,
    med: 0,
    min: 0,
};
var max_value = 0;
var med_value = 0;
var min_value = 0;
```

Per prima cosa, nel codice Javascript del grafico, vengono creati tre diversi contenitori che divideranno i risultati in tre diversi range.

```
function grafo1() {
    $.getJSON("./conn_tab.php", function (resp){
        $.each(resp, function(i, item){
            if(item.totale > 20){
                max_value++;
            }else if(item.totale <= 20 && item.totale > 5){
                med_value++;
            }else{
                min_value++;
            }
        });
        contenitore.max = max_value;
        contenitore.med = med_value;
        contenitore.min = min_value;
    });
}
```

La funzione che vediamo sopra prende i risultati Json del file “conn\_tab.php”. Ogni tabella e il relativo numero di volte che compare nei file di codice, viene analizzato dalle istruzioni condizionali (*if-else*<sup>6</sup>) e , in base al numero di occorrenze, incrementa la variabile relativa al range.

---

<sup>6</sup>Le istruzioni condizionali consentono, in modo selettivo, l'esecuzione o meno del codice

```

var series= [{
  type: 'pie',
  name: 'Percentuale: ',
  data: [
    ['Maggiori di 20', contenitore.max],
    ['Tra 20 e 5', contenitore.med],
    ['Minori di 5', contenitore.min],
  ]
}];

```

### 3.3 Grafici e risultati

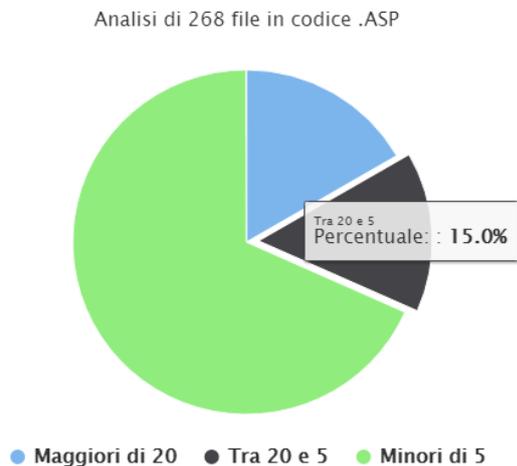


Figura 3.2. Grafico a torta con i tre diversi range

In Fig.3.2, è possibile notare molto equilibrio tra i due range che riguardano le tabelle con almeno 21 occorrenze e le tabelle fra le 5 e le 20 occorrenze. Il dato che sicuramente è rilevante è il raggiungimento del 69% delle tabelle che non superano le 5 apparizioni. Nei grafici in Fig.3.3 e Fig.3.4 è possibile vedere i nomi delle tabelle che superano le 20 occorrenze nel codice, dove la tabella più utilizzata è sicuramente quella delle “inserzioni”, ovvero la tabella che contiene tutti i dati di ogni singolo agriturismo del portale.

Nell'ultimo grafico in Fig.3.5, è stato studiato il range relativo alle occorrenze delle tabelle che non superano le 5 occorrenze nel codice del portale. Ancora più significativo è il numero delle tabelle che occorrono una volta sola nel codice. Un focus



Figura 3.3. Grafico a barre per le tabelle più utilizzate nel codice

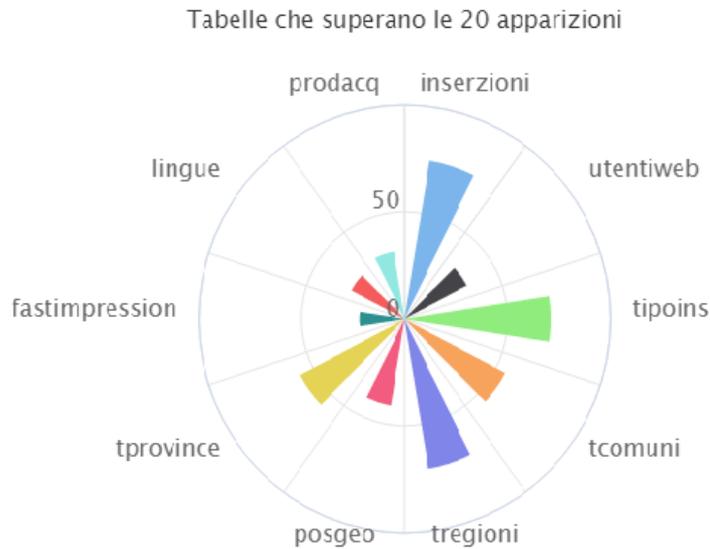


Figura 3.4. Altra visuale delle statistiche relative alle tabelle più usate

particolare è per le tabelle che compaiono meno di 5 volte nel codice. Queste tabelle, soprattutto quelle che occorrono solo una volta nel codice, sono formate da tre attributi, che sono tre diversi ID, utili al collegamento e alla creazione di nuove tabelle tramite il comando “JOIN”. Oltre a questi dati sono state trovate tabelle particolari e di poco senso, come la tabella “menuexperience”, che risulta non terminata o magari accorpabile ad altre tabelle, poiché contiene 3 attributi e 3 righe al suo interno. I tre attributi in tabella sono un id, un nome (nel nostro caso è presente “Sport” o “Enogastronomia”) e un attributo relativo ad un’immagine. Un’altra tabella, chiamata “metaprep\_province”, presenta molti record relativi a delle preposizioni, come “-a” o “-ad”, nelle 5 diverse lingue del portale.



# Capitolo 4

## Conclusioni

Questo progetto parla di SEO On-page e confronta il portale Agriturismo.com con altri competitors. La struttura della pagina mostra alcune carenze importanti per il posizionamento, come il tag H1, che deve essere presente in ogni pagina del portale. Questo errore porta problemi di ottimizzazione anche per il tag title, riducendone il posizionamento finale. Per quanto riguarda la creazione e la gestione dei link da seguire e la presenza di una sitemap, Agriturismo.com risulta avere una sitemap reperibile e ben impostata, oltre alla creazione del file robots.txt e del tag meta robots. I contenuti della pagina e le keyword analizzate tramite SEMRush mostrano come Agriturismo.com sia penalizzato sul posizionamento rispetto ad altri benchmark. Per quanto riguarda l'attributo "alt" delle immagini, in molte pagine del portale circa il 70% delle immagini è ottimizzato, mentre il restante 30% mostra carenze sia per quanto riguarda l'attributo "alt", relativo al testo alternativo in caso di problemi di visualizzazione, sia la mancanza dell'attributo "title". Anche i link interni alla pagina, seppur in minoranza, non sono molto ottimizzati, in quanto anche in questo caso l'attributo "title" non è inserito, penalizzandone l'indicizzazione. L'aspetto social di Agriturismo.com è ben curato e condiviso anche se non correttamente impostato, così come la Twitter Card relativa alla pagina Twitter del portale. Sia la pagina Twitter che la Fanpage di Facebook sono molto seguite e condivise, ma non sono collegate direttamente alla pagina web.

Nella seconda parte del progetto è stata studiato l'utilizzo delle tabelle del database, con l'estrazione delle occorrenze delle tabelle all'interno dei fogli di codice, restituendo un risultato molto alto di tabelle occorse meno di 5 volte nel portale, circa il 70%, in 268 pagine di codice. Un focus particolare è stato fatto proprio su questo risultato poiché ci sono molte tabelle che risultano essere veri e propri "connettori" fra tabelle, utile alla creazione di un output del risultato di diverse "JOIN" fra tabelle. La tabella "ins\_keywords" ha moltissimi record (circa 15.000) e occorre nel codice solo in una pagina. Se in questo caso si parla di spreco di dati o di poco utilizzo di

quest’ultimi, nella tabella “menuExperience” invece sono presenti 3 diversi attributi: un ID, un attributo contenente un nome e un attributo relativo all’immagine del record. Il caso interessante è che i record in questa tabella sono soltanto 3, quindi è ipotizzabile che questa tabella sia stata creata e poi non conclusa, oppure che possa essere servita in un’occasione soltanto. Così come per la tabella “menuExperience”, anche la tabella “metaprep\_province” presenta un curioso numero di record. Le preposizioni “a” e “ad” sono i record di questa tabella, e gli attributi si dividono in “id” e 5 attributi diversi per ogni lingua. Per concludere l’analisi relativa alle tabelle occorse meno di 5 volte, la tabella “tag\_cloud\_testi” riporta un solo record, risultando inutile o opportunamente costruita per un particolare caso nel codice. Le

idStagione	idInserzione	sNome_ita	sNome_eng	sNome_fra	sNome_deu
1	0	Stagione Unica	Unique season	Saison unique	Einzel Saison
2	0	Bassa Stagione	Low season	Saison basse	Nebensaison
3	0	Media Stagione	Middle season	Moyenne saison	Durchschnittliche Saison
4	0	Alta Stagione	High season	Haute saison	Hochsaison

Figura 4.1. Dettaglio tabella relativa al tipo di stagione

tabelle che invece appaiono frequentemente nel codice sono senza dubbio la tabella “inserzioni”, che presenta tutte le strutture nel portale, creata però con 70 attributi. Sarebbe utile gestire questi attributi in maniera diversa, magari con la creazione di una nuova tabella relativa a comunicazioni, eventi e altre informazioni su ogni struttura. Le tre tabelle “tregioni”, “tcomuni” e “tprovince” ricorrono molte nel codice e sono molto utilizzate per filtrare le ricerche delle strutture ricettive. Un attributo che può essere eliminato dalla tabella “tprovince” è il dato relativo agli abitanti per provincia, un dato che non utilizziamo all’interno del portale. Un altro caso particolare è l’utilizzo delle tabelle “stagioni” e “stagioni\_desc”. Queste due tabelle sono collegate alla tabella relativa alle inserzioni in questo modo:

idStagione	idStagDesc	idInserzione	dal	al	minnotti
13	2	2758	2013-01-07	2013-04-04	3
14	2	3139	2014-01-09	2014-04-01	1
21	3	3139	2014-04-02	2014-06-24	1
23	1	58	2008-09-05	2008-09-05	2
43	4	2645	2010-07-01	2010-08-31	1

Figura 4.2. Dettaglio tabella su date di inizio e fine stagione

Nel dettaglio di Fig.4.2 è possibile vedere la tabella relativa alle date di inizio di ogni stagione. Queste due tabelle possono essere riorganizzate creando una sola tabella contenente sia gli attributi della prima, sia gli attributi della seconda tabella, cercando di avere meno tabelle per estrarne i risultati. Un'altra tabella che può essere accorpata ad altre tabelle è la tabella “**trattamenti**”, che si presenta in questo modo:

IdTrattamento	tDescr_ita	tDescr_eng	tDescr_deu	tDescr_fra	tDescr_esp
1	Colazione	Breakfast	Frühstück	Petit-déjeuner	Desayuno
2	Mezza pensione	Half board	Halbpension	Demi-pension	Media pensión
3	Pensione Completa	Full board	Vollpension	Pension complète	Servicio de Pensión completa
4	Nessuno	None	Keine	Aucun	Ninguna

Figura 4.3. Frammento tabella dei trattamenti

Anche la tabella in Fig.4.3 può essere trasformata in un attributo di un'altra tabella (“**inserzioni**” ad esempio), limitando il numero di tabelle che servono alla configurazioni dei servizi offerti dalla struttura.

## 4.1 Sviluppi futuri

Per completare il quadro relativo all'analisi SEO è sicuramente opportuno effettuare un'analisi dei fattori Off-page, ovvero esterni alla pagina e non gestibili dal webmaster. Il progetto è in fase di costruzione.

# Bibliografia

- [1] Francesco De Nobili. 2015. *SEO Google: Guida pratica per farsi trovare con Google*. II° edizione, Hoepli
- [2] Daniele Fortis. 2013. *Scrivere per il web*. Maggioli Editore
- [3] Michele De Capitani. 2015. *Guida posizionamento dei motori di ricerca*. Edizioni Master
- [4] Bruce Clay. 2015. *Search Engine Optimization All-in-One For Dummies*. Third Edition, Published by John Wiley Sons, pp. 289-293
- [5] Paolo Maffei. 2012. *Web che funziona. Guida ai fondamenti del marketing digitale..* Lulu.com Editore. pp. 149
- [6] SEO Center. *Come lavorano i motori di ricerca*. [www.seocenter.it/learn/seo/abc/come-lavorano-i-motori-di-ricerca.html](http://www.seocenter.it/learn/seo/abc/come-lavorano-i-motori-di-ricerca.html)
- [7] Digital marketing. *Funzionamento motori di ricerca*. <http://digitalmarketing.h-farm.com/guida-seomoz/funzionamento-motori-di-ricerca.php>
- [8] SEO Center. *SERP Google*. [www.seocenter.it/learn/seo/abc/serp-google.html](http://www.seocenter.it/learn/seo/abc/serp-google.html)
- [9] SEO Center. *Meta tag robots*. [www.seocenter.it/learn/seo/abc/meta-tag-robots.html](http://www.seocenter.it/learn/seo/abc/meta-tag-robots.html)
- [10] SEO fai da te. *Tag di enfaticizzazione*. <http://www.seofaidate.com/tag-di-enfatizzazione>
- [11] Google User Content. *Search Engine Optimization Starter Guide*. <https://static.googleusercontent.com/media/www.google.com/it//intl/it/webmasters/docs/search-engine-optimization-starter-guide-it.pdf>
- [12] Moz. *Beginners guide to SEO*. <https://moz.com/beginners-guide-to-seo>
- [13] Posizionamento SEO. *Comandi Google* <https://www.posizionamento-seo.com/search-engine-optimization/comandi-google/>
- [14] CMS Wiki. *Microdata e Schema.org* <http://www.cmswiki.net/seo-posizionamento/microdata-e-schemaorg>
- [15] Alessio Pomaro. *Rich snippet* <http://www.alessiopomaro.com/microdati-microformati-rich-snippet-cosa-sono/>

- [16] Python. <http://www.python.it/doc/Howtothink/Howtothink-html-it/chap08.htm>
- [17] Internet Post. <http://www.internetpost.it/seo-rich-snippet/>