



Innocent Ignorance or Crime and Punishment: The Perils of Duplicate Web Content

| By Max Starkov

Duplicate Web Content Defined

Duplicate Web content is when two web pages with different URLs contain "largely identical content." These two pages may reside on the same website (internal duplication) or on two different sites with two completely different core URLs (external duplication).

Here is Google's definition: "Duplicate content generally refers to substantive blocks of content within or across domains that either completely match other content or are appreciably similar."

What are the reasons for duplicate Web content?

In some cases duplicate content is a result of hoteliers' "innocent ignorance". Example: a hotel signs up with an online travel agency (OTA) and completes the OTA questionnaire, providing the OTA with hotel descriptions which are an identical copy of the content descriptions from the hotel own website. When the OTA publishes these descriptions on its own website without editing or altering them in any significant manner, then we have a clear case of duplicate content under two different core URLs (external duplication).

An example of internal duplication is printer-only versions of web pages on the hotel website.

In other cases content is deliberately duplicated across the Web in an attempt to manipulate search engine rankings or win more traffic. These deceptive practices are usually performed by shady SEO vendors or self-taught in-house "experts".

In all cases, in the eyes of the search engines such duplicate content practices, if served within the search engine results, result in a poor user experience.

Duplicate Content and the Search Engines

Generally speaking, search engines hate duplicate content. Why? The search engines are working very hard to index and show pages with distinct information. For them duplicate information is not beneficial to the search engine users and inhibits the user experience. Duplicate content is "spam." This is the reason why when two or more Web pages are identified as "too similar", one or more of those Web pages usually disappear from the Search Engine Results Pages (SERPs). Google's own guidelines to webmasters are very clear: "Don't create multiple pages, subdomains, or domains with substantially duplicate content."

The search engines perform duplicate content filtering throughout the three main parts of the search engine process: spidering or crawling, indexing, and query processing. In this way

some duplicate content is filtered out before Web pages are even added to the search engine index, some during the process of cataloguing new content (indexing), and later, when responding to user searches and serving the SERPs. The end result is that typically duplicate content is not displayed in the search results.

What do the search engines do when they discover duplicate content?

Internal Duplication:

In the case of a simple internal duplication e.g. if your site has a "regular" and "printer" version of each page, and neither of these is blocked with a "noindex meta tag", the search engines will choose one of them to list. If your site contains multiple pages with largely identical content (e.g. large e-commerce sites or poorly structured hotel brand or OTA websites), there are a number of ways you can indicate your preferred page and its URL to Google and the other search engines. This process of identifying the preferred URL is called "canonicalization" and may involve the use of a 301 re-direct or noindex meta tag.

External Duplication:

As mentioned above, a typical example is a hotel providing hotel descriptions to an OTA that are identical to the ones from the hotel's own site. When the search engines discover this duplicate content

about the hotel on both sites, which one prevails? Typically a site that is better known, has a larger user audience, and a better link popularity (backlink structure in Google-speak), is likely to trump any other website. In other words, the search engines will include the OTA pages about the hotel in the search results and ignore the hotel's own website.

Deliberate Content Duplication:

In cases of content duplication in which the search engines perceive that duplicate content may be shown with intent to manipulate search rankings and deceive users, the search engines make "appropriate adjustments in the indexing and ranking of the sites involved". As a result, the ranking of the site may suffer, or the site might be removed entirely from the search engine index, in which case it will no longer appear in search results.

Hotel Websites and Duplicate Content

Providing content about the hotel to other sites is inevitable on the Web. This is not a new phenomenon. All hotels provide hotel and room descriptions to various third-party sites, branding and distribution partners, etc:

- Major hotel brands, soft brands and hotel rep companies
- OTAs, GDSs, booking engine vendors, etc.
- Hotel listings on CVB and Chamber of Commerce sites, hotel directories, destination portals, etc.

In addition to distributing the hotel inventory to a wider audience, some of these content listings provide added benefits to the hotel's own website. Hotel listings on directories, destination sites, CVB sites, etc which feature a URL link to your own hotel's website directly affect how your hotel is ranked on the search engines. Search engines love such incoming links to your site (so-called strategic linking; Google calls them backlinks) as each such link is considered a vote of confidence in your website's content.

In other words, hoteliers cannot avoid sharing content descriptions about their hotel with other sites. But hoteliers have to be very smart about it and avoid providing duplicate content to any external distribution or marketing partner site. Unfortunately, this has been a serious problem in the industry for many years: hoteliers provide the third-party sites with exactly the same content descriptions found on their websites. Why? It is much easier to "copy and paste" than write "significantly different" hotel and room descriptions.

The inevitable result of the existing practices is that many of these third-party sites like the OTAs know SEO (search engine optimization) techniques far better than the hotels do, and the result is that the OTA listings end up higher in the search engine rankings than the property's own content.

Unique Content to the Rescue

So what should hoteliers do to avoid their own sites being excluded from the search engine results or not indexed at all by the search engines, or being ranked lower than third-party sites?

To begin with, avoid duplicate content at all cost. Make sure that all content descriptions about the hotel are “significantly different” across the Web:

- Franchised Hotels: on the brand website vs. the hotel’s own vanity site
- Independent Hotels: on the hotel’s own site vs. the management company’s website vs. the hotel rep company’s site
- All Hotels: the hotel own site vs. hotel’s descriptions on the OTA sites.

Understand that the real threat is not in providing content in the form of hotel listings on a CVB site, or in the form of re-seller listings on Expedia.com, but to use the same content on your own website.

By providing unique hotel and hotel product descriptions on the hotel’s own website, your site will have a clear advantage in the eyes of Internet users and search engines alike, compared to all other sites that provide duplicate content descriptions about your hotel. Therefore, **creating the best, deepest, most unique and relevant content (textual and visual) about your hotel on your own website, naturally optimized for the search engines (SEO) as per best practices, should become a top priority for any hotel in 2010.**

Case Studies:

A typical Marriott Hotel:

- Expedia has 4-6 pages of content for each Marriott hotel on its website.
- Marriott.com has in average 30-50 pages of content for each Marriott hotel with good SEO
- The result? Property content pages from Marriott.com are usually ranked better on Google and the other search engines than Expedia’s content pages on the same properties.

Mandarin Oriental:

- Expedia has 4-6 pages of content for each Mandarin Oriental hotel on its website.
- MandarinOriental.com has in average 100-150 pages of content for each Mandarin Oriental hotel with good SEO
- The result? Similar to the Marriott.com case, property content pages from MandarinOriental.com are usually ranked better on Google and the other search engines than Expedia’s content pages on the same properties.

In other words, the content on your own website has to be able to “outshine” any other description of your hotel product and services on any other website. If the CVB site or a hotel directory or an OTA site has a single page of content/description of your property, your own website should have at least 25 pages of deep, unique, relevant and SEO-friendly content about the property.

Conclusion

Sharing textual and visual content about your hotel with third-party distribution and marketing sites is inevitable. In the same time hoteliers should avoid at all costs having content from their own website duplicated on other sites to avoid the hotel site from being marginalized in the search engine results or de-listed by the search engines altogether.

Hoteliers should strive to create the best, deepest, most SEO-friendly and unique content (textual and visual) about your hotel on your own website, complimented with the following action steps in 2010 aiming to position your hotel website as the most sought-after source of information about your hotel:

- Re-design the hotel website as per industry's best practices
- Create fresh and deep textual and visual content on the sites (the best content about your hotel should be on your website)
- Implement robust Web 2.0 functionality on the hotel website:
 - o Blogs
 - o Customer reviews/comment card on the site
 - o Interactive calendars about local events/happenings, special offers and packages
 - o Photo and experience sharing
 - o Interactive contests and sweepstakes
 - o Rich media and videos
- Implement solid SEO on the website
- Implement robust strategic linking from relevant, preferably non-paid sites and directories
- Implement social marketing initiatives (Twitter profile, fan pages on Facebook, etc), and maintain regular engaging postings and new content creation – Google now delivers real-time results from these social networks to its search result pages
- Implement comprehensive Internet marketing: email marketing, search marketing, banner advertising and sponsorships, mobile marketing, etc), each piece featuring unique content about the property
- Implement an online PR campaign and property news dissemination campaign, featuring unique content about the property
- Implement an enterprise level website analytics and campaign tracking (e.g. Omniture SiteCatalyst and Omniture SearchCenter) to measure the results and ROI from your efforts.
- Partner with hotel Internet marketing experts who understand the intricacies of how search engines work and provide in-house copywriting and SEO expertise as per industry's best practices, as well as guide your direct online channel strategy, website re-design and optimization, search and email marketing, social marketing and mobile Web initiatives.

About the Author and HeBS:

Max Starkov is Chief eBusiness Strategist at Hospitality eBusiness Strategies (HeBS), the industry's leading Internet marketing strategy consulting firm for the hospitality vertical, is based in New York City (www.hospitalityebusiness.com). HeBS has pioneered many of the "best practices" in hotel Internet marketing and direct online distribution. The firm specializes in helping hoteliers build their direct Internet marketing and distribution strategy, boost the hotel Internet marketing presence, establish interactive relationships with their customers, and significantly increase direct online bookings and ROIs.

A diverse client portfolio of over 500 top tier major hotel brands, luxury and boutique hotel brands, resorts and casinos, hotel management companies, franchisees and independents, and CVBs has sought and successfully taken advantage of the firm hospitality Internet marketing expertise. Contact HeBS consultants at (212)752-8186 or sales@hospitalityebusiness.com.