

Search engine optimization (SEO)

1. 何謂 SEO

搜尋引擎優化，是一種透過了解搜尋引擎的運作規則來調整網站，以期提高目的網站在有關搜尋引擎內排名的方式 (Wiki)

2. 為甚麼需要SEO

由於不少研究發現，搜尋引擎的用戶往往只會留意搜尋結果最前面的幾個條目，所以不少網站都希望透過各種形式來影響搜尋引擎的排序，讓自己的網站可以有優秀的搜尋排名。當中尤以各種依靠廣告維生的網站為甚。(Wiki)

3. SEO善與惡

SEO的技巧當中有善與惡之分，像是正規的方式提高排名，而當中極大部份的網站都是使用不正規的方式達到排名技巧，故搜尋引擎的探測技術以及自動懲罰系統也越來越聰明，雖然有變聰明的趨勢但不正規的方法也一同進化。

4. 搜尋引擎排行原理

搜尋引擎的排行分數計算方式一直是商業機密，若算式外洩可能造成針對算法的分數惡意攻擊，所以市面上的資訊都是借由測試摸索嘗試出的大方向，而搜尋引擎為了避免被破解，每隔一段時間便會更新算法，導致搜尋結果可能一段時間後會有大變動，但也不都是不可被得知的，搜尋引擎 Google 有提出他們的評分方向還有大致上的計算方式。

在公佈的資訊中可以列出大概以下幾點

1. HTML DOM 必須被正確的Parse出來
2. 吻合 HTML 標準
3. 依照標籤等級決定關鍵字的重要性 <h1>,<h2>...
4. 詞組頻率
5. 資料是可以被電腦給解析的 (Ex: Flash, Video, Image 都是不可被理解的)

5. 優化方法

5.1 正規的方法

1. 利用 <h1>, <h2> 等標題標籤確切的讓搜尋引擎得知權重
2. 使用<p> </p> 適當的讓文章分段
3. Image 適當的加入 alt 作為補充描述
4. 與權重高的網站進行交換鏈結
5. 少抄襲文章

6. 重複頁面資源少
7. 無效連接要處理掉
8. 利用 Span 控制文字切格方式
(Ex: 非常有趣, 可能會被解析成 [非常][有趣], 但可以利用span確保會被視為單一詞彙)
9. 合理的關鍵字密度 (Ex: 蘋果蘋果蘋果 這樣的字串會被扣分)
10. 連線速率快
11. DNS 穩定度高
12. 高更新頻率
13. 網站資源固定化 (Ex: news.php?id=3 搜尋引擎會視為單一網頁多變種, news/3 會被視為獨立頁面, 相較于前面的方法, 後者會獲得較高的分數)
14. 子網域積分法

以前有效, 但現在無效

1. meta 標籤
2. 註解
3. 屬性不可見的內容

5.2 不正當的方式

1. 建立大量垃圾網站指向自己
2. 使用 CSS 樣式 display:none 隱藏大量關鍵字文章
3. 大量關鍵字使用 size:0 隱藏在畫面當中
4. 堆疊在最下層的關鍵字
5. 關鍵字排列組合組字
6. 無意義關鍵字佈局文章
7. 將所有內文都用 <h1>, <h2> 等高權重標籤包住, 再利用CSS 使其正常顯示
8. iframe 灌流量
9. 用機器人自動點自己的搜尋結果來提升訪問量

5.3 個人發現的不正當方式

1. 借由免費Blog 張貼關鍵字文章來分享大網域的積分
2. 借由判定爬蟲使其看到故意準備的大量關鍵字網頁, 一般使用者不會看見
3. 利用無效 js code 來藏匿關鍵字

4. 靜態有關鍵字頁面, 動態生成時再利用 js 移除(避免影響使用者)
5. 利用 Facebook 的預覽功能訪問搜尋結果, 誘導FB的檢測程序掃描自己的搜尋結果達到提高訪問量
6. 利用子網域經營高流量內容, 將分數導回主網域, 再由主網域導回真正想衝分數的網站
7. 利用沙盒評分法來獲取不正當之高分數

6. 搜尋引擎額外的參考來源

搜尋引擎除了參考自家搜尋引擎的資訊外, 還會參考其他價值估算網站, 目前網站的價值評估公司還有 Aleax, 該公司藉由在使用者端安裝Tool來偵測使用者的訪問行為, 有些搜尋引擎會參考, 而Google會使用 Chrome 來搜集結果。

7. 不正當方法的應對方式

由於惡意的操作導致結果可能不如使用者預期, 為了避免這些狀況, 搜尋引擎本身會不斷變化算法外, 還有一些方式用來探測各種作弊方式。

1. 開放用戶檢舉
2. 藉由未公開的IP範圍派出爬蟲驗證內容
3. 搜尋引擎會參照文章抄襲狀況來對網站扣分, 扣分也會擴散出去
4. 獨門算法
5. Google V8引擎(PhantomJS) 實際執行JS後查看結果
6. 實際渲染網頁後對於被遮蔽或不可見得元素進行忽略

8. 排名對於新網站不公平狀況之修正

對於新網站來說, 網址的分數都會低得可憐, 這對新網站不是件好事, 使用者永遠不會知道自已的存在, 搜尋引擎對於這點幾乎都有設計沙盒(SandBox)系統, 概念上非常單純, 當新網站被索引到時, 搜尋引擎會給予一個非常非常高的獎勵分數讓新網站在關鍵字搜尋中獲得非常好的結果, 若網站沒有在短時間內獲取點擊便大量扣分, 這個方式是建立在假設使用者對於有用的網站會願意訪問, 對於不喜歡的網站會忽視的特性所設計, 因為這樣的特性所以又稱流沙(Quicksand)系統。