

**Responsible Practices for Search Engines in Reducing Online Infringement  
Proposal for a Code of Practice**

**Executive Summary**

Consumers rely on search engines to find and access entertainment content and they play a vital role in the UK digital economy. At present, consumers searching for digital copies of copyright entertainment content are directed overwhelmingly to illegal sites and services. This causes consumer confusion and significantly impedes the development of licensed digital entertainment markets in the UK. Search engines, as trusted intermediaries, should assist consumers in finding legal services and should not contribute to copyright infringement.

This paper proposes the introduction of a voluntary Code of Practice for search engines, overseen by Government, which would help to ensure that consumers are directed to safe and legal sources for entertainment content online and grow the UK digital economy.

In particular, it proposes that search engines should:

- assign lower rankings to sites that repeatedly make available unlicensed content in breach of copyright;
- prioritise websites that obtain certification as a licensed site under a recognised scheme;
- stop indexing websites that are subject to court orders while establishing suitable procedures to de-index substantially infringing sites;
- continue to improve the operation of the ‘notice and takedown’ system and ensure that search engines do not encourage consumers towards illegal sites via suggested searches; related searches and suggested sites; and
- ensure that they do not support illegal sites by advertising them or placing advertising on them, or profit from infringement by selling key words associated with piracy or selling mobile applications which facilitate infringement.

**Introduction**

The growth of the UK digital economy is presently held back by the pervasive nature of online infringement of copyright, particularly for digital entertainment content. Innovation in new digital content services is hampered by the fact that such sites have to compete against large numbers of unlicensed, free competitors, some of which have become well-known brands themselves while continuing to evade the law. Investment in original content is undermined because it is difficult to get a fair return on investment. And consumers are faced with a confusing array of legal and illegal services, and may not always know for certain which are legitimate and which are not.

## Private & Confidential

Much of the illegal activity in the digital economy is facilitated and encouraged by money making rogue sites. Intermediaries, unwittingly or by willfully turning a blind eye (or in some cases, by encouraging such activity), play a key role in enabling content theft and often even profit from it. Only a comprehensive approach can address this issue.

It is recognised that rights holders must be active in licensing new digital services, supporting innovation and offering the consumer flexibility and choice as to how they access digital entertainment. It must be as simple and as affordable as possible for consumers to access legal content online and legal services must be communicated to consumers.

However, it is also essential that other players in digital markets help create an environment that supports and promotes legality on the internet. The recent *Newzbin2* judgment confirms that service providers have a responsibility to cooperate in addressing sites that favour or encourage copyright infringement – so called structurally infringing sites. These are sites a substantial part of whose activities actively and knowingly encourage, induce, assist and/or are designed for infringement. The clear judgment rendered in this case should inform and give impetus to a codification of an effective and balanced approach to addressing such sites.

As recognised by DCMS in its recent announcement on *Next Steps for Implementation of the DEA*,<sup>1</sup> it is also appropriate for intermediaries, such as search engines, internet advertising networks, online payment services, domain registries etc, to play a greater role in promoting an environment of legality on the internet by adopting responsible practices and procedures.

In this paper, rightsholder organisations from the music, film, television, football and publishing sectors (BPI, Motion Pictures Association, PACT, The Premier League, Publishers Association) outline the main elements of a possible Code of Practice for search engines, which would have a significant impact in helping to grow the UK digital economy and in ensuring that consumers are not drawn into illegal behaviour online. These measures would:

- assist consumers in identifying and accessing legal entertainment content on the internet;
- encourage websites which host or facilitate access to illegal content towards improved online behaviour;
- ensure that consumers reduce their exposure to malware, viruses, insecure financial transactions; scams and other risks associated with illegal sites;
- ensure that the existing system of removing illegal content from search results works to optimum effect; and
- help ensure that search engines do not unwittingly profit from illegal content on the internet through the placement of advertising or the sale of mobile applications.

---

<sup>1</sup> <http://www.culture.gov.uk/images/publications/Next-steps-for-implementation-of-the-Digital-Economy-Act.pdf>

## Private & Confidential

We believe that parallel Codes of Practice should be developed with (for example) internet advertising networks and online payment providers, so that all are contributing properly to grow the UK digital economy.

### The role of search in online infringement

Major search engines such as Google, Yahoo and Bing are vital portals through which consumers access information and digital content. At present, it is trivially easy for UK consumers to find and access illegal entertainment content via search engines. Indeed, in many cases it is much more difficult for consumers to find and access *legal* services via search engines, since they are swamped in search results by illegal sites. As a result, a high percentage of the traffic that goes to illegal content sites is directed there from major search engines.

Research by Harris Interactive in September 2010 for BPI found that 23% of UK consumers regularly download music illegally using Google as their means to find the content - this is the same percentage of UK consumers that use P2P networks to download music illegally.

The scale of the problem is easily illustrated. For example, if one searches neutrally for digital music to download (i.e. without suggesting whether one wants it from a legal music store or an illegal free site), search results on the first few pages of major search engines direct consumers overwhelmingly to illegal websites in preference to legal websites. On 26 September 2011, BPI made test searches on Google for the name of each of the UK's top 20 singles and albums, followed in each case by the word "mp3" (the dominant legal and illegal file format for digital music). On average, 16 of the first 20 Google results for chart singles and 15 of the top 20 search results for chart albums linked to known illegal sites. These results were virtually the same as those obtained for the equivalent search undertaken a year earlier.<sup>2</sup>

Research undertaken by the Publishers Association, conducting free search on Google and Bing for the 50 bestselling books of the week 24.04.11 – 30.04.11 showed that:

- Google returned an average of 41% non-legal links in the top ten (first page) results;
- Bing returned an average of 21% non-legal links in the top ten (first page) results;
- Google's top ten free search results now contain 18% more non-legal links that were found in a comparable survey conducted by the PA in October 2010;
- the average position for the first non-legal link on Google was 3.48, whilst the same on Bing was 4.10; and
- the average position of the first legal link on Google was 1.32, with Bing coming in at 1.14.

Similar results can be demonstrated for films, TV programmes and highlights of Premier League football. For example, according to a June 2011 Envisional Briefing report, 77% of sites that commonly link to or host infringing film and television material get more traffic from Google than

---

<sup>2</sup> When the same test was conducted in November 2010, 17 of the first 20 Google results for singles and 14 of the top 20 search results for albums (on average) linked to known illegal sites.

## Private & Confidential

any other site online, according to a sample of thirty key sites. Notorious pirate sites such as Pirate Bay and Isohunt continue to appear as search results.

As time goes on, the situation is getting worse rather than better. Illegal sites are proliferating and they are becoming increasingly sophisticated and convincing, so that consumers are lured into using them. Illegal sites now commonly carry fake logos, carry advertising from known and respected brands, and use 'trusted' payment processors which all inspire confidence in the user.

Consumers do not wish to be left in this position. An IPSOS Omnibus survey conducted in March 2011<sup>3</sup> found a very high degree of consumer support for search engines directing users to legal sites over illegal ones. 84% of consumers surveyed said that they expect the sites their search engine identifies at the top of its results to be legal. The same percentage agreed with the statement that "search engines should direct consumers to legal sites rather than illegal ones". 85% also said that, when purchasing goods (including music) search engines should direct consumers towards legitimate websites.

These results are not surprising. Aside from the negative consequences for the consumer by potentially breaking the law - perhaps inadvertently - by acquiring content from or providing content to illegal websites, such websites also pose a number of other risks. Spyware, malware, and viruses are all common on sites featuring unlicensed entertainment content and the cost to consumers of repairing the damage that they unwittingly cause to their computer may be considerable or even irreparable in the case of identity theft or theft of personal information. Harris Research in September 2010 asked illegal file sharers to disclose unwanted problems that arose as a result of their use of unauthorised services. 41% of these respondents had downloaded spyware, 39% had downloaded a virus or Trojan. More worryingly, 17% said that their PC or laptop crashed and was unusable for a period of time, and nearly one in eight (12%) divulged that they had to have their PC or laptop repaired or replaced.

There are also not just risks to consumers, but to the integrity and security of networks. Computers infected by malware can be the source of Distributed Denial of Service Attacks – the method by which commercial services are attacked to ensure that legitimate services cannot be used by consumers and businesses. This is a source of financial loss to UK businesses and helps seed a lack of trust in online (particularly in financial services). Consumers that are guided towards illicit sites can unwittingly put networks at risk through downloading from those sites.

We believe that search engines have a role to play in protecting consumers and businesses by directing users to sites which comply with the law and do not propagate illegal content, host viruses and other damaging or inappropriate content.

While it will remain necessary for rightsholders to send takedown notices to illegal sites, and to continue to work with search engines to improve the procedures for de-listing sites and individual content items in search results, it is clear that the current regime will not be sufficient to ensure that search engines direct consumers first and foremost to legal places to acquire digital content. To achieve that, search engines will need to do more.

---

<sup>3</sup> BPI Music Online Omnibus 11th to 14th of March 2011, based 1009 adults 16-64.

## Private & Confidential

### Potential elements of a Code of Practice

#### A. Search rankings

##### i. De-ranking sites that persistently make available unlicensed content in breach of copyright

Proprietary search algorithms rank web pages according to “relevance” to the user; search companies have confirmed to content owners that they currently do not take into account the legality or illegality of content on a web page as a factor in determining its ranking.

At times search engines have sought to present the outcome of their algorithms, in the ranking of search listings, as something over which they does not exert control. In fact, search engines are well placed to influence the outcome of search results since they control the algorithm that produces them. They also directly control which sites they choose to crawl, index, and link. Google’s own “Web History” service learns a consumer’s preferences and re-lists sites according to individual preferences. Moreover, there are several documented instances of search engines intervening to modify their search rankings to achieve a commercial or policy goal.

For example, in February 2011, Google is reported to have taken action against JC Penny over alleged ‘gaming’ of search results.<sup>4</sup> Google has also reportedly reordered search on suicide to guide users towards sources of advice, such as the Samaritans.<sup>5</sup>

Google is taking action against content farms, as reported widely in January 2011 and confirmed by Google itself<sup>6</sup>. In this example Google makes a subjective judgement about the quality of sites it is linking to and “absolutely takes action on sites that violate our quality guidelines regardless of whether they have ads powered by Google”. Given Google makes such value judgements, it should be easy to do so in the face of objective evidence. The other element of this of note is that Google is making an assessment of the quality of the entertainment content. We believe that this is particularly relevant when considering content on sites on the basis of legality.

We propose that in order to further protect consumers and to encourage responsible behaviour among websites, the extent of illegal content on a website should become a factor influencing the ranking of that website in search results returned to consumers. In addition, where a site has been found by a court to be substantially infringing, it should no longer be crawled, indexed or linked at all.

---

<sup>4</sup> [http://www.nytimes.com/2011/02/13/business/13search.html?\\_r=1](http://www.nytimes.com/2011/02/13/business/13search.html?_r=1) These, and other aspects search ranking, were investigated by Congress at the House of Representatives Subcommittee on Intellectual Property, Competition and the Internet *Promoting Investment and Protecting Commerce Online: Legitimate Sites v. Parasites, Part II* 6 April 2011.

<sup>5</sup> [http://www.samaritans.org/media\\_centre/latest\\_press\\_releases/google\\_one\\_box.aspx](http://www.samaritans.org/media_centre/latest_press_releases/google_one_box.aspx)

<sup>6</sup> <http://googlepublicpolicy.blogspot.com/2011/09/making-copyright-work-better-online.html>

## Private & Confidential

We should explore simple and objective mechanisms to enable a search engine to take account of these factors. Search engines already respond to valid de-listing requests from copyright owners.<sup>7</sup> One proposal would be that the number of URLs that the search engine has de-listed as a result of takedown requests from right holders, in respect of a particular website, should be reflected in the ranking accorded to that website.

Another factor to consider might be the type of infringing content – i.e. if a site consistently gives access to particularly damaging content such as pre-release material - that might be a factor that would lead to increased de-ranking. There may also be a threshold of serious repeated or egregious infringement beyond which a site should be de-listed entirely or at least for a period of time, to allow the copyright infringement issues to be properly addressed.

The ability of websites to serve counter-notices for URL de-listing under US DMCA procedures, and the fact that URL de-listing is routinely publicised by Google on the Chilling Effects website, ensure both that sites are able to challenge de-listing requests that could have an impact on their ranking, and that there is transparency about the alleged infringements that may impact on a site's ranking.

Given that Google already de-ranks and de-lists sites that do not meet its own “quality guidelines” or otherwise violate its policies, we do not believe that search engines would face significant legal exposure if they were to de-rank or de-list sites using an objective measure, based on their actions in response to legal DMCA complaints, in pursuit of the legitimate objective of preventing their service being used to facilitate copyright infringement.

### ii. Prioritising legal sites in searches for content

The group of rights holders believes that, where content owners are able to objectively and reliably certify websites as providing legal access to digital content, those sites should be prioritised by search engines in their search rankings above sites which are not so certified, for certain specific types of search. Industry could take the lead on this by ensuring there are appropriate certification programmes in place. Rigorous care would have to be taken to ensure that any such schemes would be transparent, open and non-discriminatory.<sup>8</sup>

We recognise that the purpose of consumers' searches relating to digital content may be quite diverse and that it would not be appropriate for prioritisation to be implemented for every type of search. Rather, prioritisation should be focused on searches where the consumer is clearly trying to access digital content to download or stream, rather than simply looking for information.

---

<sup>7</sup> As reflected in Google's policy statement of November 2010, see <http://googlepublicpolicy.blogspot.com/2010/12/making-copyright-work-better-online.html>

<sup>8</sup> One example of such a certification programme is the *Music Matters* scheme <http://www.whymusicmatters.org>. This scheme, which is open to any website to join at zero or nominal cost, awards Music Matters certification to sites making available recorded music that satisfy objective criteria, in terms of having appropriate procedures in place to ensure that the content they make available is properly licensed.

## Private & Confidential

To use the example of music, we would propose that prioritisation be enabled for searches that contain any of the following key search terms: "mp3", "flac", "wma", "aac", "torrent", "download", "rip", "stream" or "listen", "free", when combined with an artist name, song or album title contained on a list to be regularly updated and provided to a search engine by a recognised and properly mandated agency representing rights holders for a particular sector, such as BPI.

Such a measure should not be complex to implement from a technology perspective and would go a long way towards ensuring that consumers are directed towards legal sites on the Internet rather than falling, sometimes initially unwittingly, into content piracy.

### iii. Making the existing process work better

Google has announced some recent welcome improvements to its processes for receiving and acting on DMCA notices submitted by rightsholders. In our experience, average takedown times have improved and the process of submitting notices has been made easier. Google has also removed from Autocomplete a number of terms that encouraged consumers towards infringing sites.<sup>9</sup> However, these steps do not go far enough. There is more that can and should be done to improve this process:

- Artificial limits on the ability of rightsholders to search for infringing content should be removed;<sup>10</sup>
- infringing links should be removed within a maximum of 4 hours of being reported by an authorised agent. Once removed, there should be effective measures to prevent re-appearance of infringing links and of pirate sites and blog spots;
- if a website is subject to a blocking order or has been adjudicated by a court to be a structurally infringing site, the sites in any case should be automatically de-listed and removed from the cache, without any time delay while appropriate procedures and criteria should be implemented to delist sites that are persistent offenders;
- to ensure an expedited process, rights holders should be able to take advantage of automated tools, which should be implemented by search engines to allow rapid removal and disabling of infringing links; and
- Autocomplete or "suggested searches" should not in any instance direct consumers towards illegal sources for content. Search engines should continuously review autocomplete/suggested search terms and should respond expeditiously to requests from rights holders to remove terms that direct consumers to illegal sites.

---

<sup>9</sup> <http://googlepublicpolicy.blogspot.com/2011/09/making-copyright-work-better-online.html>

<sup>10</sup> Google currently limits the number of de-listing submissions that a rights holder can make to 10k per day.

## **Private & Confidential**

### **B. Advertising**

Search itself is a form of advertising and the ranking and listing of sites is affected by search engines through the sale of words from search terms to advertisers to help serve adverts. Some search engines go further, and have operations that serve adverts such as Google's DoubleClick, AdSense and AdMob and Microsoft's AdCenter. Such adverts generate revenue for providers of websites that are based on illegal use of content. Meanwhile reputable firms find themselves as banner ads on sites in the UK and abroad that are set up to infringe copyright.

A Code of Practice for search engines should include measures to ensure that search engines do not support the business models of substantially infringing sites by supporting them with advertising and that they do not themselves profit from online infringement. To that end, we propose that search engines should:

- implement active, ongoing, effective screening procedures for ad partners (both those that buy ads on the search engine's website or run ads from the search engine's ad network on their third party sites);
- eliminate all adverts on substantially infringing sites and on search results pages that contain links to substantially infringing sites;
- establish an effective and simple complaints procedure so third parties can notify the search engine/ad network of ad placement on sites that facilitate infringement, or search result pages that do so;
- continuously review key search words;
- ensure that it does not place advertisements for Google products and services on substantially infringing sites;
- end the practice of selling key words that are closely associated with piracy; and
- prevent re-registration of terminated ad partners through active, effective internal procedures.

As indicated above, we believe that a separate track of the roundtable process should be established to develop and agree a Code of Practice for internet advertising networks generally.

### **C. Mobile Applications**

In addition to its search and advertising activities, Google also operates an applications platform for Android devices. Given Google's direct financial interest in applications sold through its Android Market, Google (or any search engine that operates an apps platform) should:

- effectively screen applications to see if they are likely to substantially facilitate or encourage infringement or otherwise designed to facilitate infringement;

## **Private & Confidential**

- take down mobile apps where it is aware that they are designed or known to be used to illegally download entertainment content, either via p2p applications or unauthorized lockers;
- prevent apps from being reposted, or “copycat apps” being posted, once an app has been removed as above; and
- terminate the accounts of developers that repeatedly post apps that facilitate infringement and prevent them registering new accounts.

### **D. Governance**

We believe that a Code of Practice could be administered and overseen on a voluntary basis, with DCMS in attendance, so that Government remains well informed about the functioning of the Code. It will also give strength to the process if Government maintains its role.

Moreover to ensure transparency, accountability and democratic scrutiny, we would suggest that DCMS prepare an annual report to the parliamentary Select Committee on Culture, Media & Sport on the functioning of the Code of Practice. This would help to ensure that, while meeting its objective of ensuring that search engines play a more responsible role in reducing online infringement of copyright, the Code of Practice would act proportionally within a wider public policy objective of freedom of general access to information online and healthy competition and access to online markets.