



Digital Advertising on Suspected Infringing Websites



Digital Advertising on Suspected Infringing Websites

Report commissioned by OHIM through the
European Observatory on Infringements of Intellectual
Property Rights to WhiteBullet Solutions Ltd.

WWW

CONTENTS

Background to this Report	05
Scope	08
Executive Summary	10
Trends in Digital Advertising	13
Digital Advertising Ecosystem	15
Advertising Analysis and Key Findings	21
Conclusions	31
Part II: Appendices	34

BACKGROUND TO THIS REPORT

Intellectual property (“IP”) crime significantly impacts the European and global economies, affecting businesses, jobs and tax revenue. An OECD study concluded that international trade in counterfeit and pirated goods could account for up to US\$250 billion with the share of counterfeit and pirated goods in world trade estimated at 1.95%.¹ Another study focusing on the economic value of IP industries to the European Union, calculated that 39% of total economic activity in the EU is generated by IP intensive industries and approximately 26% of all employment in the EU is provided directly by those industries.² A significant contribution to the problem of IP crime is websites suspected of infringing third party IP, for example by offering consumers unauthorised media content. These websites are located throughout the world and operate in a dynamic environment. Enforcement against these websites is critical, but only addresses the worst offenders due to volume, jurisdictional limitations and a dynamic digital environment where websites regularly change domains, or adapt to action against them.

Therefore an important complementary strategy adopted by rights holders across the world is ‘Follow the Money’: to tackle the revenue sources providing most suspected IP infringing websites with their financial lifeline. This strategy includes drying up support from advertising and making these websites less commercially viable.³

Research has documented that up to 86% of IP infringing websites allow web users to download or stream infringing media content for free with the websites earning money via online advertising.⁴ These popular websites often have high web user traffic as free content -- such as music, films, TV shows, games, software and eBooks -- has a broad appeal to a wide audience. Websites therefore capitalise on this high traffic by selling the advertising space of their webpages to advertisers keen to reach that audience. As digital advertising⁵ is dependent in part on the number of users allegedly seeing the ads displayed on the webpages, these websites can receive substantial digital advertising revenue, earning as much as €5.3 million each annually.⁶

In many cases major brands inadvertently advertise on suspected IP infringing websites, lending these websites credibility, possibly funding infringement and risking brand damage. Often this is due to a lack of understanding as to which websites pose an IP infringement risk.

1 - See The Economic Impact of Counterfeiting and Piracy, Organisation for Economic Cooperation and Development (“OECD”) study, November 2009 update. Figures are for 2007 and exclude domestically produced and consumed products and non-tangible pirated digital products. See also estimated value projections for 2015 of between US\$80 billion and US\$240 billion globally for digitally pirated music, movies and software, BASCAP, Global Impacts Study, 2011. See also Observatory quantification studies in specific IP sectors, e.g. The Economic Cost of IPR Infringement in Sports Goods, 2015. (<https://oami.europa.eu/ohimportal/en/web/observatory/quantification-of-ipr-infringement>)

2 - See Intellectual Property Rights intensive industries: contribution to economic performance and employment in the European Union, a study carried out by the European Patent Office and the Office for Harmonisation in the Internal Market acting through the EU Observatory, September 2013. <https://oami.europa.eu/ohimportal/en/web/observatory/ip-contribution>. ‘IP rights intensive industries’ are defined as those having an above-average use of IP rights per employee.

3 - See e.g. Follow the Money: Financial Options To Assist In The Battle Against Online IP Piracy, Mike Weatherley MP and IP Adviser to the Prime Minister, 2014.

4 - 86% of peer-to-peer infringing websites exist due to advertising revenue, see Six Business Models of Copyright Infringement, PRS for Music/Google report, 2012.

5 - Globally, Programmatic spend grew to 42% of total display-related ad spend in 2014, compared to 33% in 2013, and growth is expected to remain strong with an average annual growth rate of 27%, reaching an annual total of US\$53 billion in 2018, see Global Media Suppliers Advertising Revenue Forecasts, Magna Global, 2014.

6 - 76% of advertisers tend to view Programmatic as being less transparent than traditional trading and 85% of brands have concerns about ad misplacement as a consequence of automated trading, see World Federation of Advertisers’ surveys within WFA Guide to Programmatic Media, 2014.

This problem is exacerbated by the increased use of “Programmatic” transactions: targeted ad campaigns deployed according to software rules and enriched by data. Programmatic advertising is facilitated by advertising exchanges, where website advertising space is bought and sold via electronic transactions in real-time and otherwise. Programmatic transacting brings efficiency and increased automation to online advertising and is the future of digital media trading, with a strong annual growth rate of 27%. However, these advances bring potential lack of visibility into the trading process and little transparency around costs of trading, which, without proper checks and balances, can result in inadvertent brand advertising on inappropriate websites, such as adult websites, hate websites or suspected IP infringing websites. Therefore, preventing ad support for suspected IP infringing websites should be a significant issue for many global brands, themselves IP rights owners.

As part of a broader initiative to enhance cross-border brand and IP protection within the EU, the European Observatory on Infringements of Intellectual Property Rights (the “Observatory”) commissioned whiteBULLET Solutions Limited (“whiteBULLET”) to undertake a ‘snapshot’ study of the digital advertising landscape detailing in a report the scale of the problem of ad-based funding of suspected IP infringing websites affecting the EU market. This report aims to enhance understanding of how digital advertising supports suspected IP infringing websites in the EU and to provide strategic information to assist the development of innovative and effective strategies to tackle the problem.

Specifically, this report covers:

1. A brief overview of how digital advertising works and how suspected IP infringing websites make money from advertising;
2. Identifying the brands and ad sectors supporting suspected IP infringing websites affecting the EU market and how they operate across the EU; and
3. Identifying the advertising companies placing ads on those suspected IP infringing websites.

SCOPE

This study focuses on a selection of significant websites suspected of IP infringement, which were to meet the following criteria:

1. websites of significant **popularity** in one or more of the 28 EU Member States;
2. websites offering a variety of **suspected infringing copyrighted content**⁷ (individually or collectively) representing a range of IP industries (including film, TV, sports, music, games, software, eBooks);
3. websites to include a range of **website types** (i.e. offering content via a range of means, such as streaming or direct download); and
4. websites with some form of digital **advertising revenue**.

A combination of automated tools and human review was used to gather and assess the data for the study. whiteBULLET's AdWiseable™ ad monitoring system collected advertising data over six weeks on 1,400 webpages from 280 websites that met the above criteria set by the Observatory (the "Selected Websites"). The Selected Websites were scanned from each of the 28 EU Member States in order to generate locally targeted advertising. The advertising profiles of each website were also analysed. Details of methodology, including website selection, advertising monitoring and ad analysis processes are provided in Appendix B. See Appendix A for a description of AdWiseable™.

This version of the report does not call out or name the advertisers whose ads appeared on the Selected Websites nor provide details of the ad companies that placed those ads. This version also does not identify the Selected Websites. These details were recorded, however, for the purposes of analysis and are included in an unpublished version of this report.

Website Types

Suspected IP infringing websites distribute content files through different technical means, some more popular with certain demographics of the population depending on ease of technical use and speed of data transfer. A range of website types were included in the candidate website database. To do this, the candidate websites were categorised by the following website types according to the type of technology used and business model:

"BitTorrent Portals". These websites use the peer-to-peer ("P2P") file distribution technology to permit users to share content. The websites act as aggregators of P2P links, which users can search for and access via the website. When a user clicks on a link, the P2P technology allows the user to download media files stored on other P2P users computers across the P2P network. This is undertaken at no charge to any of the P2P users. Note that a user in a P2P network will not only download files from other users, but will make his or her media files available for upload to the P2P network so that other users can download files from them as well.

⁷ - This study was to monitor only websites suspected of offering infringing copyrighted content, not those offering suspected counterfeit goods. Websites offering counterfeit goods tend not to have advertising revenue, but rather make money through direct sales of products.

"Hosting Websites". Hosting websites are storage servers that permit users to upload and store media files in the cloud. Once a media file is uploaded, a user can generate a link to access that file again, either themselves, or for others to use the link to obtain the content. Hosting Websites permit content to be downloaded directly (often referred to as "Direct Download Hosting Websites") and may be ad-supported, offering the content for download at no charge, or they may charge for premium access, whereby users pay a subscription fee to avoid advertising and for faster download speeds. Other types of Hosting Websites permit the content files to be streamed via embedded media players ("Video Streaming Hosting Websites") and are usually ad-supported but may also offer premium subscriptions⁸. In addition, there are hybrid websites that allow immediate streaming while content also is downloaded.

"Linking Websites". These websites aggregate links to media content that is stored on Hosting Websites. Linking Websites index the links, usually categorising links by content type (e.g. movie files) and then offer search facilities and suggestions so users can quickly find content they want or have content suggested to them. These websites do not host any content themselves but link the user through to the Hosting Websites at no charge. These websites are usually entirely funded by advertising.

⁸ - Over 70% of Video Streaming Hosting Websites (also known as streaming "Cyberlockers") rely on ad revenue according to Digital Citizens' Alliance and NetNames, report Behind the Cyberlocker Door, Sept 2014.

EXECUTIVE SUMMARY

The Observatory commissioned this 'snapshot' study of the digital advertising landscape to detail the scale of the ad-based funding of suspected IP infringing websites affecting the EU market. In this study 280 suspected IP infringing websites were monitored from each of the 28 EU Member States over a six week period between May to July 2015. The study collected over 180,000 ads from 1,400 webpages. Ads were found in 131 different sectors, some websites having up to 57 unique sectors.

The Selected Websites included 64% Linking Websites, 23% Hosting Websites and 13% BitTorrent Portals. 62.5% of the Selected Websites offered a mix of different types of IP content (e.g. included music, film, games and software) and 37.5% of the Selected Websites offered only one type of IP content exclusively (e.g. eBooks only). BitTorrent Portals had the highest percentage of Mainstream ads (52%) and 92% of BitTorrent Portals used additional ad revenue generation techniques, such as Pop-up Ads or Pop-under Ads. Linking Websites had the highest percentage of High Risk ads (57%).

This study looks at how digital advertising supports suspected IP infringing websites and quantifies the prevalence and profile of such advertising. It analyses in detail the brands and sectors supporting the websites with their advertising and the ad companies placing those ads.

1. Brands are the key: Brands are able to direct ad placement and control how ad companies manage their campaigns. Despite this, suspected IP infringing websites are a brand-rich environment: this study identified over 1,500 unique brands. Mainstream advertising alone made up 46% of all ads collected in this study (for definitions see Glossary in Appendix C). Brands may inadvertently advertise either because they don't know which websites pose an IP risk, or they cede full placement responsibility to Intermediaries. Whilst a very large number of brands was identified, analysis shows there are two small sub-sets that have the potential to significantly impact the issue. First, premium brands belonging to 46% of the top 100 companies by global ad spend were identified. These companies potentially lend credibility to websites. Second, 70% of ads collected for named brands were for just 97 brands, all of which appeared in 20 or more EU Member States. Ad misplacement may impact brands negatively as it can affect return on investment for their ad spend. In addition, their own brand may be tarnished by being placed next to inappropriate content or malware.

2. High Risk ads: a complementary focus. In this study 54% of ads collected were in High Risk sectors: malware, fraud and adult (for definitions see Glossary in Appendix C). These ads pose a risk to consumers and generate income for websites. This type of advertising does not lend itself to outreach to the brand. Therefore the intermediary (the ad company placing the ad) is the focal point for effective action. This study found just 10 out of 232 intermediaries placed 91% of all High Risk ads collected.

3. Intermediaries and data management: Intermediaries will usually take steps to avoid ad misplacement if a brand requests it. However, the security of the ad supply chain is breaking down in this fast evolving era of data-driven advertising and real-time ad decisions. Intermediaries have a responsibility to effectively pass along brands' instructions and to implement systems and tools to ensure compliance with contractual obligations. In this study, a core group of 25 out of 232 intermediaries are involved in placing 73% of the Mainstream ads. This statistic, together with the fact stated above that just 10 Intermediaries placed 91% of all High Risk ads, indicates that these 35 Intermediaries may have a significant role to play in dealing with the problem.

4. Ad fraud boosts revenue for websites. Potential ad fraud was identified on 41% of the websites. These websites could potentially generate multiple ads on each webpage that are never visible to a consumer, thereby defrauding the brands that paid for those ads to appear. This is another reason why brands may find these types of websites pose a risk.

There have been a number of studies over the last few years looking at ad-funded online IP infringement. The findings of this study, which took a wider view across more territories than any previous studies on this subject, confirm that the problem has not yet been solved. The entire digital ad ecosystem has a role to play and can choose to make IP protection a critical component of the online ad compliance toolkit.

TRENDS IN DIGITAL ADVERTISING

Digital advertising is a fast growing market: global advertising revenue from digital media grew 17% to €127.9 billion (US\$142 billion) in 2014 and is expected to increase another 15% in 2015⁹. The Internet Advertising Bureau (“IAB”) Europe stated in its latest report that online digital ad spend in Europe doubled in the last five years¹⁰. In 2014 it grew **11.6% to a market value of €30.7 billion**, consolidating online ad spend as the second largest media category after TV¹¹. According to the same report, display advertising (as compared to online search advertising or online classifieds) outperformed other categories with a growth rate of 15.2%, and the total value of the European online display ad market being €10.9 billion.

Growth in European digital advertising continues to be driven by the larger, mature online advertising markets, such as the UK, Germany and France. This is a direct result of investment in targeting, data and format capabilities. Digital media is already the number one media category in 14 of the 73 markets analysed by Magna Global, including the UK, Germany, Sweden and the Netherlands¹². The Central and Eastern European regions are also growing strongly as online advertising in these markets benefits from improvements in broadband infrastructure and penetration, giving access to a larger online audience.

<i>IAB Europe Top 10 online ad markets by market size:</i>		<i>IAB Europe Top 10 online ad markets by individual growth:</i>	
1. UK	€8.9bn	1. Slovenia	43.1%
2. Germany	€5.4bn	2. Ireland	33.3%
3. France	€3.7bn	3. Belarus	32.8%
4. Italy	€1.9bn	4. Turkey	20.8%
5. Russia	€1.8bn	5. Hungary	19.5%
6. Netherlands	€1.5bn	6. Bulgaria	18.7%
7. Sweden	€1.0bn	7. Russia	17.3%
8. Spain	€0.9bn	8. Greece	16.3%
9. Denmark	€0.7bn	9. Austria	16.0%
10. Norway	€0.7bn	10. Belgium	16.0%

Figure 1: Data from IAB Europe AdEx 2014 Report showing top ten European online ad markets by size and growth.

9 - See Global Media Suppliers Advertising Revenue Forecasts, Magna Global, December 2014.

10 - See IAB Europe AdEx Benchmark 2014/ HIS Technology Study. Twenty seven markets participated in the study and twenty markets showed double digit growth with all markets showing positive growth. See also www.iabeurope.eu/files/9614/3228/6945/IAB_Europe_AdEx_Benchmark_2014_preliminary_results_released_at_Interact_2015.pdf.

11 - See IAB Europe AdEx Benchmark 2014/ HIS Technology Study.

12 - See Global Media Suppliers Advertising Revenue Forecasts, Magna Global, 2014.

In Europe, Programmatic advertising has increased in importance to a market value of more than €2 billion or 20% of the total European display market, developing beyond a tactical tool to being a key capability for every online brand¹³. This raises challenges as automated, data-driven advertising risks placement of ads outside of brand requirements (including on suspected IP infringing websites). The World Federation of Advertisers (“WFA”) and brand associations are increasingly calling out for enriched data to help brands better avoid such ad misplacement, which also wastes marketing budgets¹⁴. eMarketer figures indicate global digital ad spend is expected to be €191 billion (US\$213.89 billion) by 2018¹⁵, with Europe to have **over 24%** ad spending share of the worldwide total¹⁶. European digital ad spend will surpass TV ad spending by 2018 as the largest media category.¹⁷

13 - See IAB Europe Programmatic Market Sizing Study, September 2014.

14 - The WFA 2014 Guide to Programmatic Media specifically calls out the dangers of non-transparent Programmatic buying and references the desire for trading desks to become ‘better equipped’ than their competitors to distinguish their services.

15 - See eMarketer Digital Ad Spending Worldwide 2012-2018, June 2014, www.emarketer.com/Article/Global-Ad-Spending-Growth-Double-This-Year/1010997.

16 - See eMarketer Digital Ad Spending Share Worldwide by Region 2012-2018, March 2014, www.emarketer.com/Article/Digital-Ad-Spending-Worldwide-Hit-3613753-Billion-2014/1010736.

17 - See IAB Europe Programmatic Market Sizing Study, September 2014.

DIGITAL ADVERTISING ECOSYSTEM

Understanding how digital advertising works is key to finding effective solutions to the problem of ad-funded IP infringement. The process by which ads end up on suspected IP infringing websites, often due to poor data management and lack of transparency, is therefore considered below¹⁸.

Like other marketplaces, the online or ‘digital display’ advertising market is fundamentally characterised by buyers and sellers. Websites (or “Publishers”) sell the space on their webpages (“Ad Space”) to parties wishing to advertise (“Brands”) to users of the website. Therefore Brands buy multiple Ad Spaces (collectively known as “Inventory”) either directly from a website or via brokered supply chains. Digital display advertising continues to evolve towards increased automation and reliance on data to pair advertising with audiences, with computer servers now making decisions about ad placement in milliseconds. This evolution has made the digital ad supply chain much more complex, with many “Intermediaries” (or “Ad Companies”) involved in the ad placement process.

A. Intermediaries Involved in Ad Placement

Historically, Brands or their advertising agency representative (“Agency”) would buy Ad Space directly from a Publisher website. Each time an ad is viewed online is termed an “Impression”, so buyers would buy thousands of Impressions from Publishers and then fill the Publishers’ Ad Space with their own creative digital ads. Publishers would sell millions of potential Impressions on their websites, resulting in advertising brokers (“Networks”) developing technology to buy unsold Ad Space from websites and selling it on as packaged Inventory to Brands and Agencies within targeted markets (e.g. the travel industry). Networks thus became part of the selling market but are also buyers.

Ultimately, the relationship in the digital marketplace is between the Brand as buyer of Ad Space and the Publisher as seller of Ad Space. However, the growth in sophistication of the marketplace has created a number of Intermediaries, as shown in Figure 2.

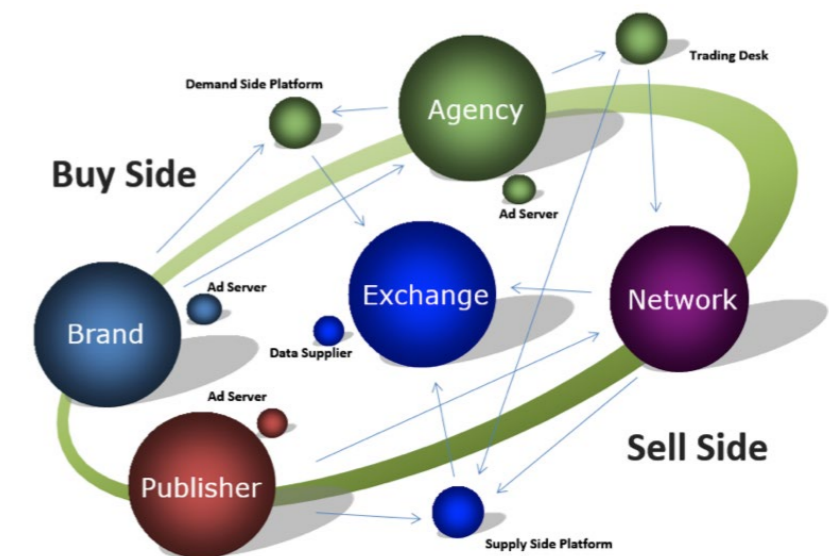


Figure 2: Digital Display Intermediary universe.

18 - The price paid by brands to a website for displaying ads is influenced by a range of factors, such as: (i) which different types of ad companies are involved in ad targeting and placement; (ii) size and display format choices (e.g. banner or pop-up and image or video); and (iii) positioning on a webpage. Website revenue is not within the scope requested for this study.

Competition for pricing and increased automation eventually created “Exchanges”, through which Publishers make available their Ad Space packaged for a particular audience and buyers bid on it with the winner having their ads served to that audience on that website. Exchanges also use “Real-Time Bidding” (or “RTB”), which is an advertising process during which Inventory is bought and sold via almost instantaneous programme-managed auctions (see further detail below under section Automated and Data Driven Advertising). These developments brought other efficiencies, such as the creation of “Trading Desks” (either independent or subsidiaries of Agencies) and “Demand Side Platforms” (“DSPs”), used to better integrate purchasing of Ad Space on Exchanges through audience data often purchased from “Data Suppliers” (companies that collect and organise information about websites and audiences). Publishers and Networks can also use “Sell Side Platforms” (“SSPs”) to facilitate trading through Exchanges by matching ads with desired audience. Each Intermediary takes a cut of a Brand’s budget and it is estimated that as much as 60% of a Brand’s digital ad spend is taken to pay for the role of all Intermediaries in placing an ad¹⁹.

The **increasing complexity** of the ad supply chain and the **need for specialisation** has led to vast **growth** in the number of players in the digital ad ecosystem, with the number of marketing companies alone doubling from 2014 to 2015²⁰. This in turn risks **breaking the security** of that ad supply chain, with Brands not always knowing which Intermediaries have been involved in placing their ads.

B. Automated and Data Driven Advertising

Exchanges use data driven processes to match audiences and advertising based on information identifying suitability, desirability and demand either from their own data sources or via Data Suppliers. “Behavioural Advertising” relies on a range of methods to improve the effectiveness of advertising by matching users’ Internet browsing information with ads that may appeal to those users’ interests. This is done via tools like cookies and similar technologies. “Cookies” are small pieces of web code placed by websites and stored on a user’s web browser. Cookies send information about a user’s web habits on that website back to the website to remind the website of the user’s previous activities and to remember their interests. The information collected by Cookies, and similar technologies, then can be used to target advertising specifically to that user.

Real-Time Bidding has raised particular concerns about misplacement of advertising on inappropriate websites, including suspected IP infringing websites, due to the fact it can be controlled by inaccurate or incomplete data and decisions are made in fractions of seconds. When a web user visits a website, their browser connects to the Publisher’s content server to retrieve the content of the webpage. Within the HTML web code received by the user’s browser, there is specific code representing opportunities for advertising to be displayed on the webpage sought and data about where to retrieve those ads. The Publisher also has its own “Ad Server” which manages queries about the advertising to be placed, including whether the Impression is already reserved to be filled by a particular buyer (‘premium’ Inventory) and whether the user is recognised (e.g. via Cookies) as suitable audience for a particular ad.

19 - See WFA Guide to Programmatic Media, 2014, chart demonstrating ‘working media’.

20 - See e.g. Chiefmartec.com’s Marketing Technology Landscape, January 2015, showing the number of marketing technology companies in January 2014 to be 947 and jumping to 1876 companies by January 2015.

If the Impression is reserved for a premium buyer, then that Brand’s ad will appear. If the Publisher’s Ad Server recognises the user, then specific advertising may be targeted to that user’s interests (e.g. by a Network). If neither applies, then the Publisher’s Ad Server may offer the Impression as an opportunity on the open market for other buyers to bid via an Exchange. In this case, the Publisher’s Ad Server connects to an SSP, which runs additional queries and connects with the Exchange. The Exchange then communicates with the open buyers’ market (e.g. DSPs and other Exchanges) providing information about the type of Impression represented (e.g. user demographic, webpage type or content available). Buyers then have an opportunity to inform the Exchange how much they are willing to pay for that Impression, all of which takes around 10 milliseconds. The Exchange selects a winning bid and the winner informs the Exchange how to retrieve the ad, which instructions are in turn passed to the Publisher’s SSP and Ad Server and eventually to the user’s browser to retrieve the ad from the appropriate buyer’s Ad Server. The user’s browser then renders and displays the ad to the user.

Increasing levels of automation with almost **instantaneous ad placement decisions**, combined with **poor data management** and **lack of domain transparency**, risks ads being placed on suspected IP infringing and other inappropriate websites.

C. Ad Formats

In order to understand the analysis of the ad data collected for this study, it is important to have a basic knowledge of the different digital ad display formats.

Display ads may be text, audio, static images, video (including video ads served before, during or after video content that is streamed to a user within an online media player) or rich media content (e.g. Flash files that animate, rotate or expand when the user interacts with the ad or hovers the mouse cursor over it). The formats are categorised as banners, skyscrapers or mid-page units (“MPUs”) that appear at the top, bottom or at the size of content on the webpage. Examples of these format categories are shown in Figure 3. Efforts have been made to standardise the size of these formats to assist Publishers and Brands²¹.

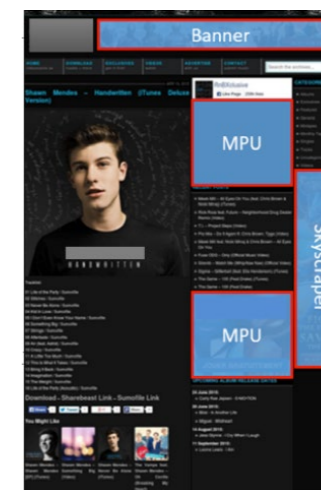


Figure 3: Screenshot demonstrating ad formats on a single webpage with suspected infringing content.

21 - See e.g. IAB UK Display Ad Standards 2010, a set of six ad size standards offered by the majority of Publishers in the UK. See also www.iabuk.net/resources/standards-and-guidelines/iab-display-ad-standards-2010.

Other formats include ads that appear in front or behind the webpage (“Overlays” or “Interstitials”). These include “Pop-up Ads” display in a standalone window superimposed over the main browser window, which itself is not a new browser window and may be closed.

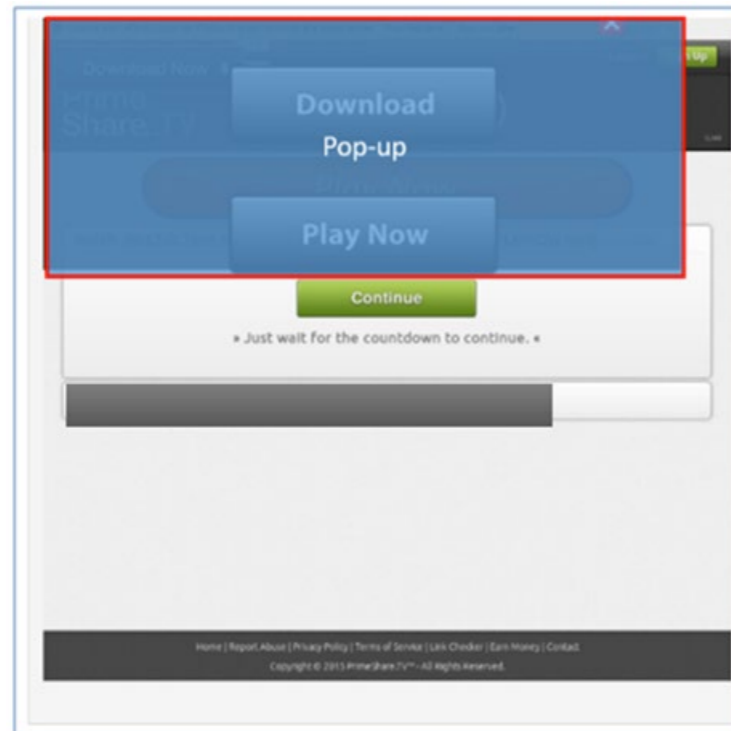


Figure 4: Screenshot showing Pop-up Ad appearing as a secondary viewing pane over the main browser window.

“Pop-under Ads” display in a new browser window that opens behind the main or originating browser window (rather than superimposed in front of it) and may not be noticed by the user until the main browser window is closed

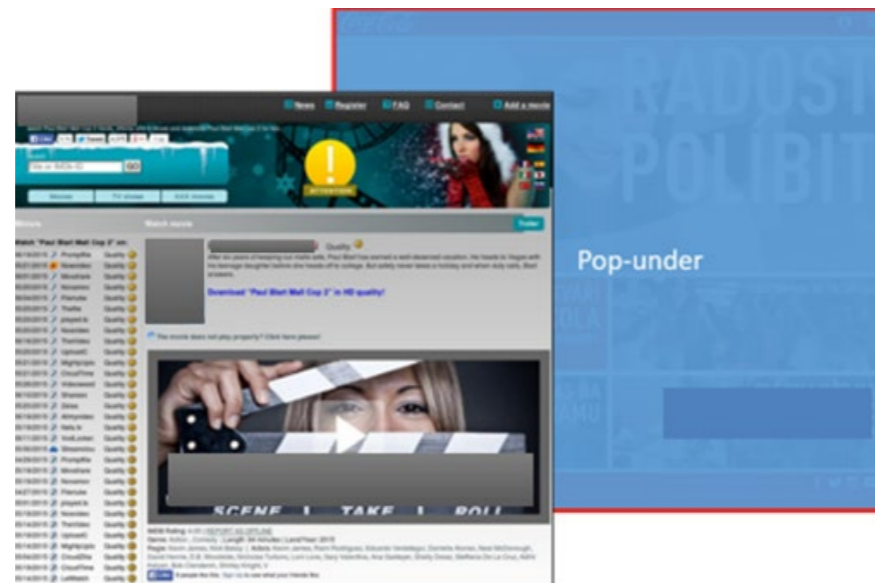


Figure 5: Screenshot showing how a Pop-under Ad appears in a new web browser window following a mouse click on the

originating webpage with suspected infringing content.

Pop-up Ads and Pop-under Ads may be generated upon interaction with the originating webpage, either by moving the cursor and resting it on the hot spot of an ad for at least one second (“Mouse-over”), which may trigger an event such as expanding the ad or initiating an animated sequence within the ad, or by mouse clicking on any part of the webpage (“On-click”).

Some webpages also have “Embedded Ads”, also known as “Sponsorship Ads”. This static embedded advertising is hard-coded into the webpage and will appear on the webpage until the code is removed or changed by the website administrator. As a result these ads are not engaged with Intermediaries or calls from Ad Servers. They take many formats and can include adding a logo or branded image to the design of a third party webpage.

Figure 6 shows an example of an Embedded Ad outlined in a red box. The ad appears as a banner ad and can only be identified as an Embedded Ad through analysis of the underlying web code, which shows no Intermediary engagement.

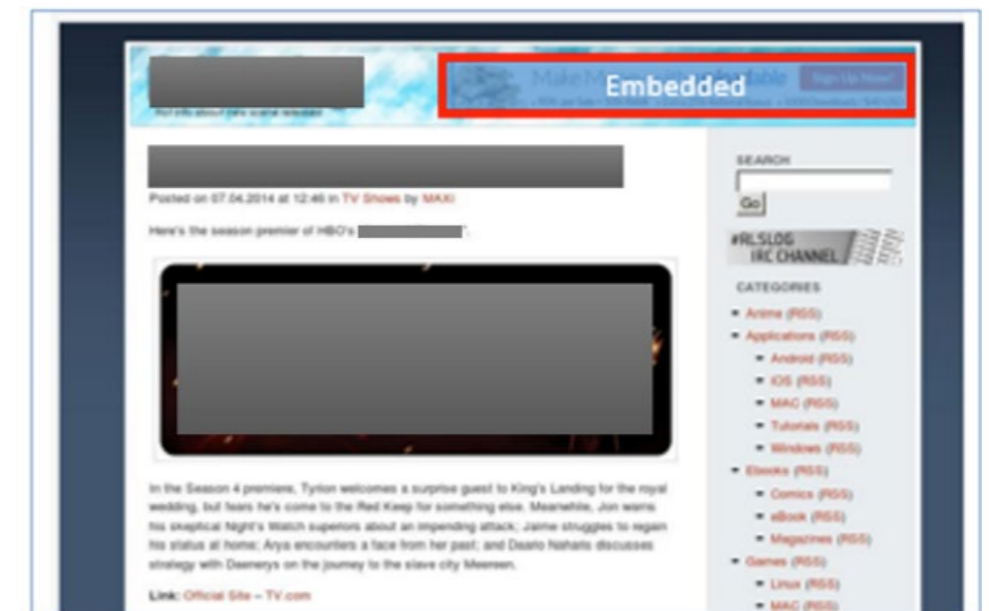


Figure 6: Screenshot showing Embedded Ad appearing on a webpage with suspected infringing content.

Another format is via “Content Marketing” of images, articles or news allegedly related to the webpage. This content is regularly updated or may change for each visit to the website and is displayed in dedicated boxes within the webpage by specialist content marketing platforms.



Figure 7: Screenshot showing Content Marketing appearing on a webpage with suspected infringing content.

ADVERTISING ANALYSIS AND KEY FINDINGS

A. Prevalence of Advertising and Ad Profiles

Over 180,000 ads were collected across the 1,400 webpages monitored from each of the 28 EU Member States.

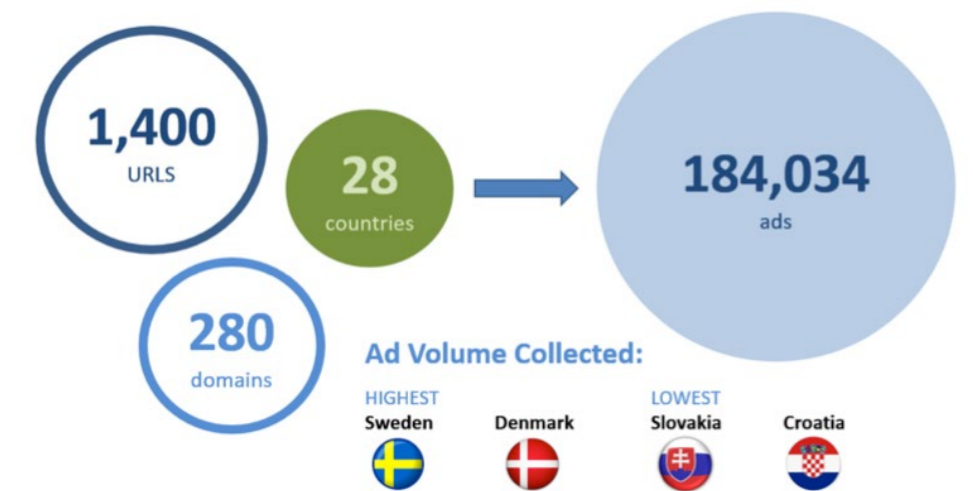


Figure 8: Summary of scope of cover including domains, webpages and ads collected.

The variation in the ad volume collected by country as shown in Figure 9 is similar to the profile of ad revenue per country as reported by IAB Europe, but the differences are interesting to note. The top digital ad revenue territories are the UK²², Germany, France, Italy, Netherlands, Sweden, Spain, Denmark and Norway; the highest volume of ads collected from the Selected Websites were from Sweden, Denmark, Belgium, Ireland, Spain, Germany, Netherlands, Czech Republic, Austria and Latvia. While the UK has the top digital ad revenue it was 15th out of 28 territories for ads collected on the Selected Websites²³.

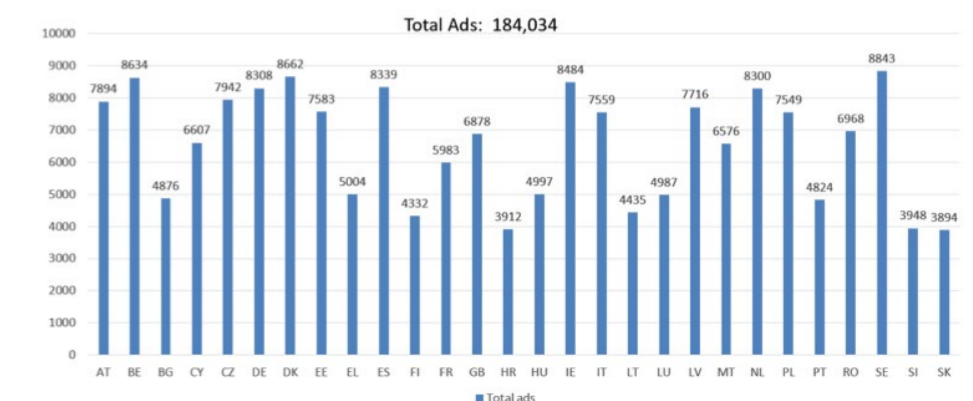


Figure 9: Volume of ads collected by country.

22 - The country analysis charts in this study all use the country code “GB” rather than UK.

23 - Some of the Selected Websites are blocked by certain ISPs in the UK (and other countries) pursuant to court orders. However, locally based proxies not subject to the court orders were used to access these websites for the purposes of this study.

The Selected Websites used a **range of advertising techniques to augment the number of ads** associated with their webpages and so likely would generate higher advertising revenues. These included display ads and Interstitials as well as forms of potential ad fraud.

Figure 10 shows the advertising profile across the Selected Websites. The average number of display ads (banners, skyscrapers and MPUs) per webpage was over two (2.6), with several of the Selected Websites having up to 14 display ads per webpage. **Pop-up Ads were recorded on 24% of the Selected Websites** and 70% had Pop-Under Ads (all generated through On-click interaction). Potential “Pixel Stuffing” (see definition below) was identified on **41%** of the Selected Websites.

“Pixel Stuffing” is an ad fraud technique involving tiny Ad Spaces sized at just 1x1 pixel (or sometimes 5x5 pixels) hidden within the top or bottom of webpages. These Ad Spaces are ‘stuffed’ with hidden ads, which are served to them just as they would be to regular sized Ad Spaces. However, a visitor to the website would not see these tiny ads, although they would be logged by the ad ecosystem as an Impression. Websites implement Pixel Stuffing in order to increase Impression counts and get paid for these fraudulent ads even though the Brand’s ad is never actually seen by the consumer. Such ads are often ‘stacked’ in multiple layers on top of each other in a single ad placement, so generating even more money for the website and of course the Intermediaries serving those ads also generate revenue.

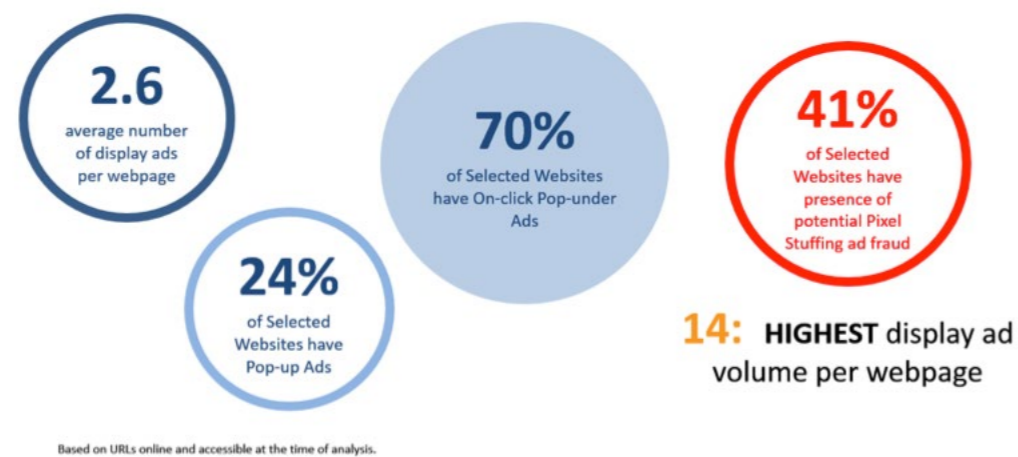


Figure 10: Advertising profile analysis for the Selected Websites.

Deeper analysis into the ad profile by website type revealed that the **average number of display ads on Linking Websites was higher** (at 2.8 ads) than on BitTorrent Portals (2.16 ads) and Hosting Websites (2.14 ads). However, nearly **92% of BitTorrent Portals used additional ad revenue generation techniques**, such as Pop-up Ads or Pop-under Ads, compared to the other website types (75% each for Hosting Websites and Linking Websites).

B. Diversity of Ad Sectors

The Selected Websites were found to have a **broad diversity of advertising**, with many different advertising sectors represented. In total **131 different ad sectors/ sub-sectors²⁴** were identified

²⁴ - There are 131 top level sector/ sub-sector combinations, which provide more granularity for deeper analysis.

across the EU as a whole, with the median number per website being seven (7). The highest number of ad sectors on any single website was 57 sectors. The full list of ad sectors collected is included in Appendix D.

Figure 11 shows the top level ad sectors by volume of ads across the EU. The top 15 top level sectors (out of 25 top level sectors) represent 99% of the total ad volume collected during the study. The top sector overall was **Click Generators/ Malware²⁵** at 51% of all ads. **Gambling** and **Arts & Entertainment** were in second and third place. A full list of the top level ad sectors are in Appendix E.

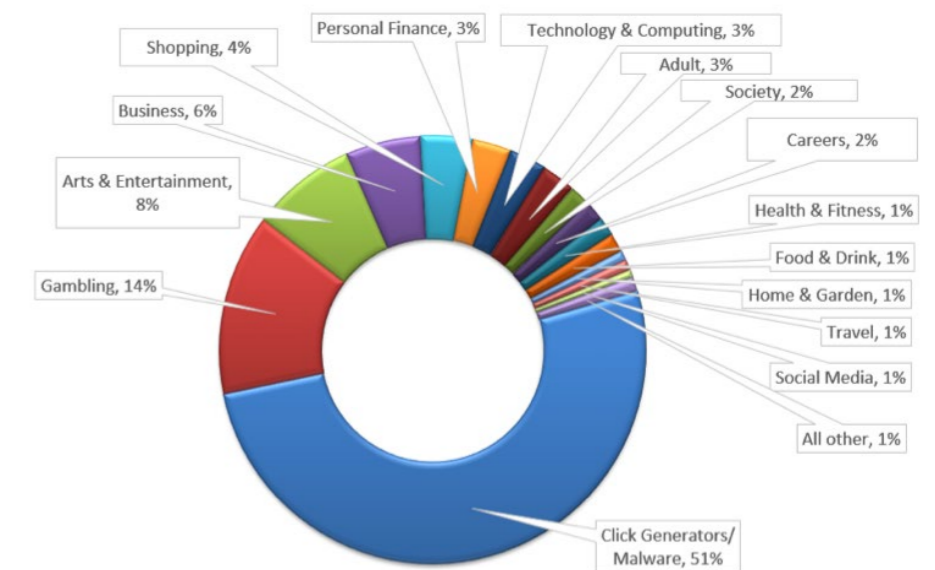


Figure 11: Top level ad sectors overall by volume of ads across the EU.

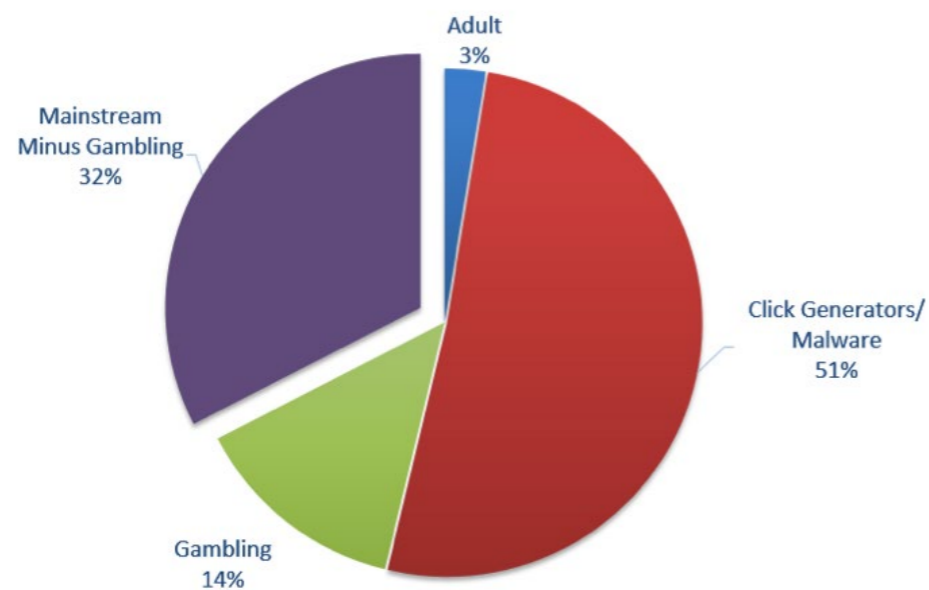
Mainstream Versus High Risk Ad Sectors:

Analysis involved dividing sectors into those that incorporated “High Risk” ads and all others, known as “Mainstream” sectors. Advertising falling into “**High Risk**” sectors could potentially have a **harmful impact on consumers** or can be designed to confuse consumers and defraud Brands. These High Risk ad sectors include ads with adult content, malware and those making use of fraudulent techniques, such as system fraud, click generator fraud or prize fraud (see below for descriptions). Click generator fraud and malware ads were grouped together as these ads often led to both activities or could not be separated.

Gambling sector ads were called out separately, as in some territories gambling is regarded as a high risk activity in terms of consumer safety and advertising may be regulated or restricted. For the purposes of this study, gambling sector ads are included with Mainstream sector ads.

Figure 12 shows the breakdown of Mainstream versus High Risk sector ads. Overall, **High Risk sector ads made up 54%** of all the ads collected, with click generators and malware found in 51% of the ads and adult content in 3% of ads. **Mainstream sector ads** (including gambling) made up 46% of all ads.

²⁵ - See definition in the Mainstream Versus High Risk Ad Sectors section.



■ Adult ■ Click Generators/ Malware ■ Gambling ■ Mainstream Minus Gambling

Figure 12: Mainstream versus High Risk ad sectors.

Malware and Fraud Ads:

Malware ads often encourage the download of software allegedly required to repair application bugs, upgrade systems or to continue to view content (forms of system fraud). These ads then initiate the installation of malware on the consumer's computer either through download of dangerous or unwanted software or redirection to a malicious website.

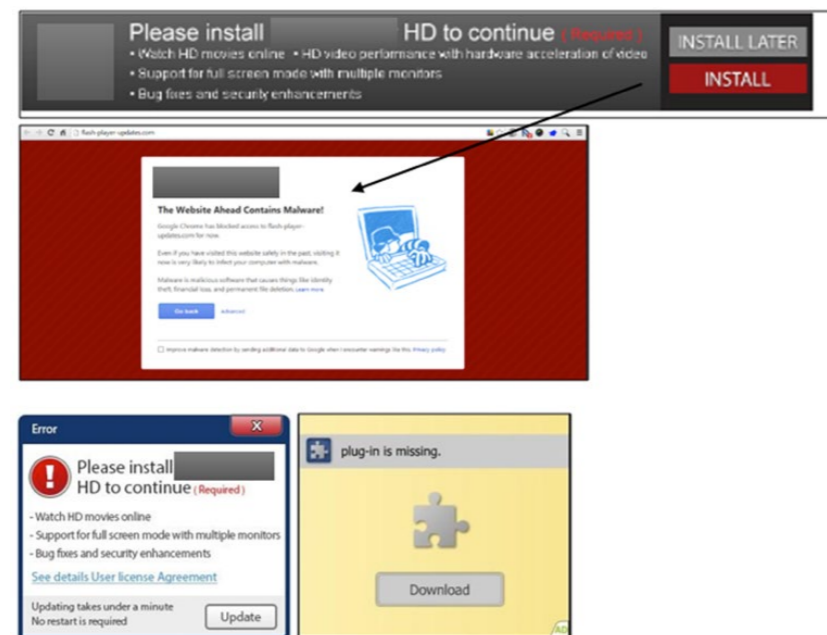


Figure 13: Examples of malware ads fraudulently directing installation of system upgrades. Note clicking on 'Install' shows a malware flag raised against the website to which the user is redirected.

Other malware ads take the form of innocent looking action buttons suggesting 'Download' or 'Play' of content but instead lead to the installation of malware. Many appear to be branded as software tools, which may be legitimate offerings when found on certain websites, however, when found via advertising on suspected IP infringing websites, they are usually associated with various forms of malware, including "PUPs" ("Potentially Unwanted Programmes"), browser hijackers, adware and even more serious forms of malware.

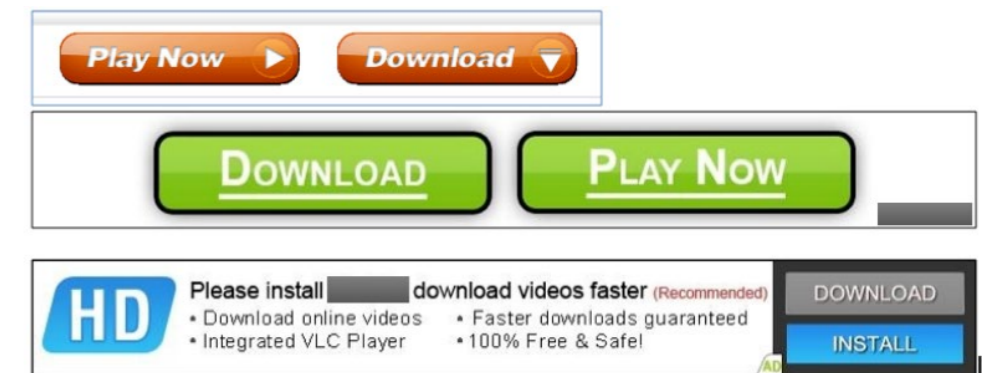


Figure 14: Examples of malware ads designed as action buttons.

Click generator ads are designed to accumulate 'clicks' mimicking a user clicking on an ad. Often the website displaying the ad will collect payment from a third party website or Brand. The third party website, Brand and consumer may have no knowledge of the fraud. The ads may appear harmless and will often suggest actions such as 'Download', 'Play Now' or fraudulently encourage the download of software (system fraud). These ads may redirect the consumer to third party websites, generating more clicks, even though the consumer had no intention of ever accessing the third party website. Others may generate many alleged clicks for phantom ads hidden behind the image without requiring any action or redirecting consumers.

Some ads do not seemingly offer any discernible product or Brand. These may include fraudulent ads involving prize fraud, where the ad is unrelated to any Brand although it may depict a known product as the prize. Such ads again usually lead to malware being downloaded to a consumer's computer or click generation.



Figure 15: Example of unbranded and prize fraud ads.

Mainstream Versus High Risk Ad Sectors by Country:

Figure 16 shows the Mainstream and High Risk sector ads broken down by country. High Risk ads on suspected IP infringing websites are more prevalent in certain countries, such as Bulgaria, Hungary, the Netherlands and Portugal. Adult ads in particular are more prevalent in Germany, France, Italy and the Netherlands. Mainstream ads on suspected IP infringing websites are significantly higher in Romania as percentage of ads found in that country, with France and Italy following. Still within Mainstream, gambling ads are prevalent in Austria, Germany, Finland, Sweden and Slovenia.

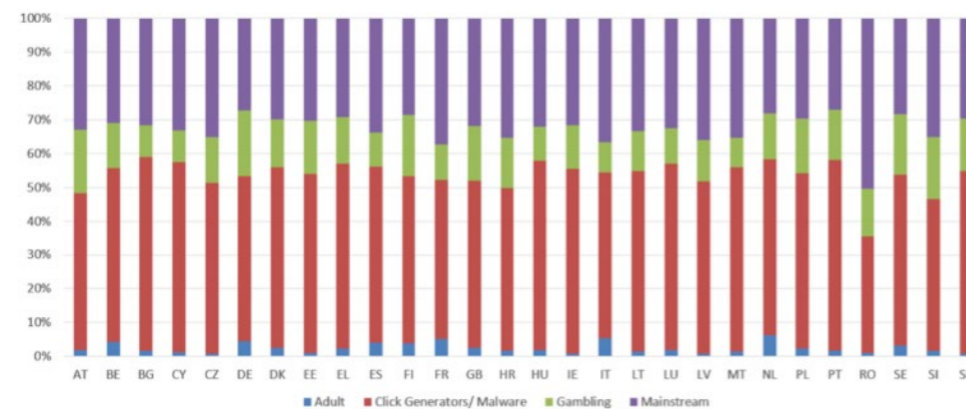


Figure 16: Mainstream versus High Risk ad sectors by country.

Mainstream Versus High Risk Ad Sectors by Website Type:

Figure 17 shows that BitTorrent Portals have the highest percentage of Mainstream ads (52%), whilst Linking Websites had the highest percentage of High Risk ads (57%). Gambling sector ads are 13-14% across all website types.

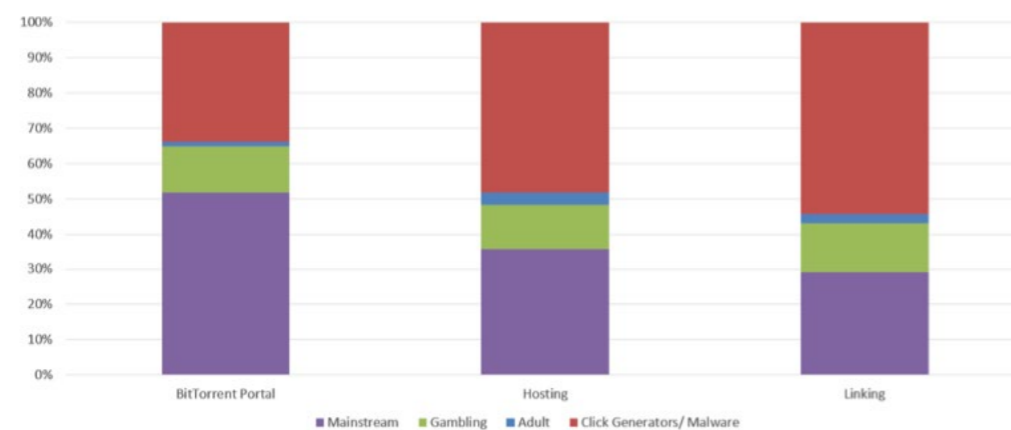


Figure 17: Comparing Mainstream to High Risk sector ads by volume by website type.

Top Mainstream Ad Sectors:

The Mainstream sectors across the EU and by country were analysed. This excluded High Risk ads and gives an indication of the advertising that supports the Selected Websites from

Mainstream and premium Brands. For this analysis a combination of top level sectors and sub-sectors was included to provide more granularity, therefore a top level sector may feature more than once. Figure 18 shows that **Gambling-Betting** sector ads rank first at an EU level at 20% of Mainstream ads and **Arts & Entertainment-Gaming** sector ads rank second at 13%. **Business-Advertising** sector ads are also popular at 11% and tend to include ads for Ad Companies.

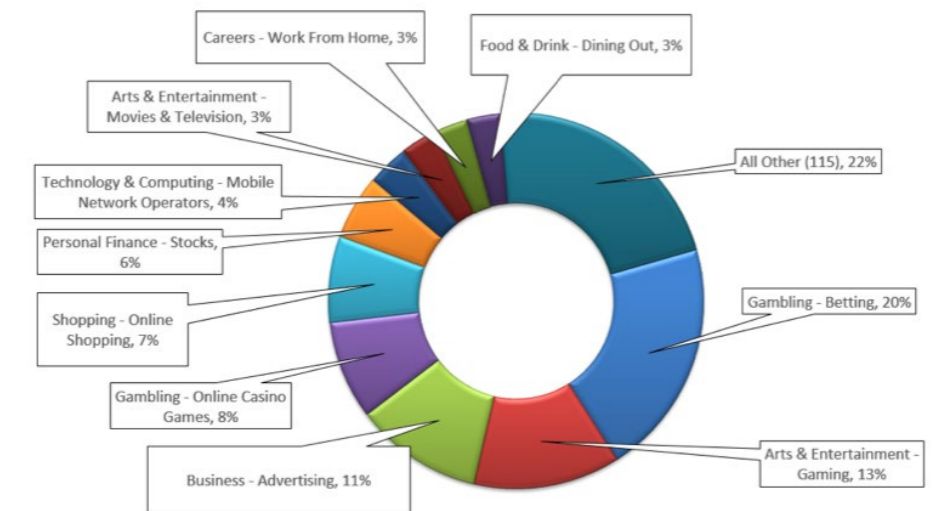


Figure 18: Top 10 Mainstream sectors overall by volume of ads across the EU (showing primary level sector and sub-sector).

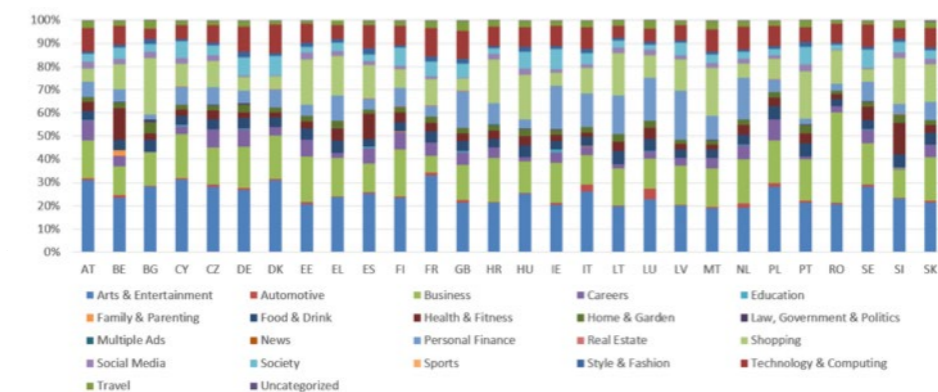


Figure 19: Variation in Mainstream sectors by ad volume by country (excluding gambling).

C. Diversity of Brands

The Selected Websites are a Brand rich environment with **1,581 unique named Brands** identified with ads placed during the period of the study. Some ads do not include sufficient Brand information (such as Brand or product names or identifying Brand features) that permit the ad to be attributed to a Brand. In such cases ads are attributed to the brand category “Generic”. The study also included 38 Generic Brands, taking potential unique Brands identified to over 1600²⁶. The highest number of unique Brands found on a single website was 204 Brands. The **top 25 named Brands** by ad volume across the EU accounted for 52% of all named Brand ads.

Of the top 25 named Brands, 18 (72%) appear to be based or headquartered in Europe. The remaining seven are based in China, the US, Russia and Curaçao.

26 - Generic Brands represented 54% of all ads identified.

Mainstream Brands:

Excluding High Risk Brand ads, there were **1,514 unique Mainstream Brands**, which includes premium Brands and household names across the EU. Figure 20 shows analysis by country and the significant differences in volume of Mainstream Brands. The UK, Germany, France and Ireland had more Mainstream Brands advertising than other countries. These markets are in the IAB Europe top 10 size or growth market lists. Greece and Lithuania had the fewest Mainstream Brands, but also had generally low volumes of ads overall.

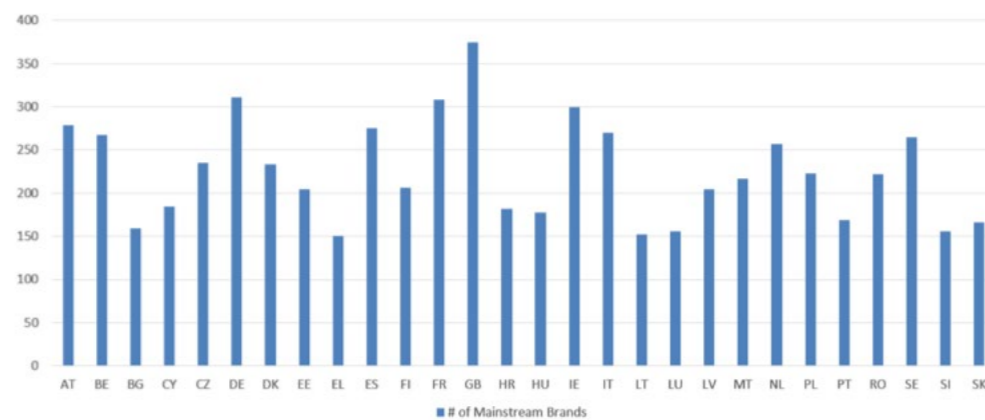


Figure 20: Variation in number of Mainstream Brands by country.

By looking at the number of websites on which each Mainstream Brand appeared, it is possible to gauge those Brands that dominate in terms of breadth of advertising cover. Analysis shows that **four (4)** of the top 10 Mainstream Brands by number of websites were **gambling Brands** and **three (3)** of the top ten were **gaming Brands**. The Brand with the widest breadth of ad coverage appeared on 27% of the Selected Websites (75 of the 280).

With guidance from advertising associations, advertising spend figures were used to determine the most significant Mainstream Brands on the Internet (or 'premium' Brands). This study used the 2014 Top 100 Global Marketers List published by AdAge²⁷. Forty six percent (46%) of companies on that Top 100 list had **at least one** of their Brands advertised on the Selected Websites.

EU-wide and Cross-border Advertising:

A number of Brands advertised across every country in the EU, whilst others (usually local national Brands) advertised only in specific local markets. Thirty seven (37) named Brands advertised in all 28 of the EU Member States and 97 named Brands advertised in 20 or more countries. Ads for these **97 Brands made up 70% of all named Brand ads** identified, meaning the volume of EU-wide advertising is high.

In some cases local Brands with local language advertising appeared in countries seemingly unrelated to the target market. Of the 97 Brands advertising in more than 20 markets, 8% were national Brands with little apparent relevance to other markets within the EU.

27 - See <http://adage.com/datacenter/globalmarketers2014>.

D. Intermediaries Involved in Ad Placement

Many different Intermediaries may be involved in placing an ad. As well as identifying the chain of different entities involved, the "Primary Intermediary" was also identified, which is defined as the Intermediary responsible for engaging with the website itself and often the party that engages in the financial transaction with the website (as opposed to the Ad Server that may simply deliver the ad image e.g. from a content delivery network). Throughout this analysis, references to Intermediaries are references to Primary Intermediaries.

Overall **232 unique Intermediaries** placed ads across the Selected Websites²⁸. The highest number of Intermediaries engaging with any single website was 29 and the median per website was two (2).

Figure 21 shows the variation in number of Intermediaries by country. The highest number of Intermediaries for any country was 143 in the UK. By way of comparison the median number of Intermediaries per country was just over 121.

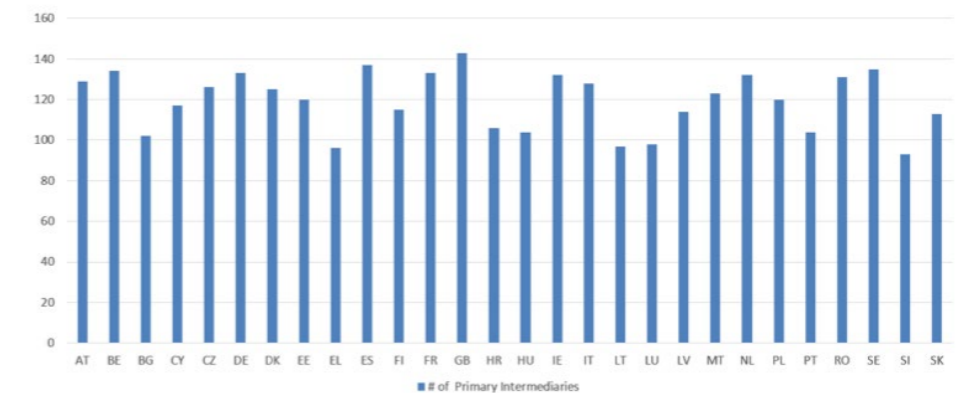


Figure 21: Variation in number of Intermediaries by country.

The **top 15 Intermediaries placed 64%** of all the ads and 60% of those (9 Intermediaries) have operations in or are located in Europe.

An assessment was undertaken as to which Intermediaries placed the majority of Mainstream ads and those that placed the majority of High Risk ads. Of the Intermediaries placing named Mainstream ads (which includes premium Brands) the **top 25 Intermediaries by ad volume placed 73% of those Mainstream ads**. When analysing the Intermediaries placing **High Risk ads**, the **top ten (10) Intermediaries by volume placed 91%** of those ads. This indicates that certain Intermediaries are more likely to place High Risk ads and may require a different approach when tackling the issue of suspected ad-funded IP infringement.

Intermediary Group Breakdown:

Some Intermediaries are members of national or regional advertising trade associations, such as IAB Europe. Members of such trade associations are dedicated to the promotion and growth of the interactive advertising marketplace and as a requirement of membership may be obliged to adhere to guidelines supporting responsible ad placement.

28 - In this study, Intermediary data for On-click ads was excluded from Intermediary statistics due to the varied methods for generating On-click ads, which may result in a different identification process for Intermediaries.

Intermediaries which are members of IAB Europe were identified and non-IAB Europe members²⁹ were further divided into groups:

1. **IAB Europe members:** Intermediaries with membership of IAB Europe;
2. **Verifiable contact details:** Ad Companies that are non-IAB Europe members with contact details readily available via public sources, including office addresses, telephone numbers or email contacts;
3. **Non-verifiable contact details:** Ad Companies that are non-IAB Europe members with no contact details readily available through public sources or only offering contact forms available on their website; and
4. **Other ad technology:** includes Publisher Ad Servers, in some cases belonging to suspected IP infringing websites using self-managed open source ad serving (also known as micro-networks); these do not have verifiable contact details.

Figure 22 shows the breakdown of Intermediaries by group. Of the 232 Intermediaries, **54% may be amenable to outreach** via IAB Europe or directly.

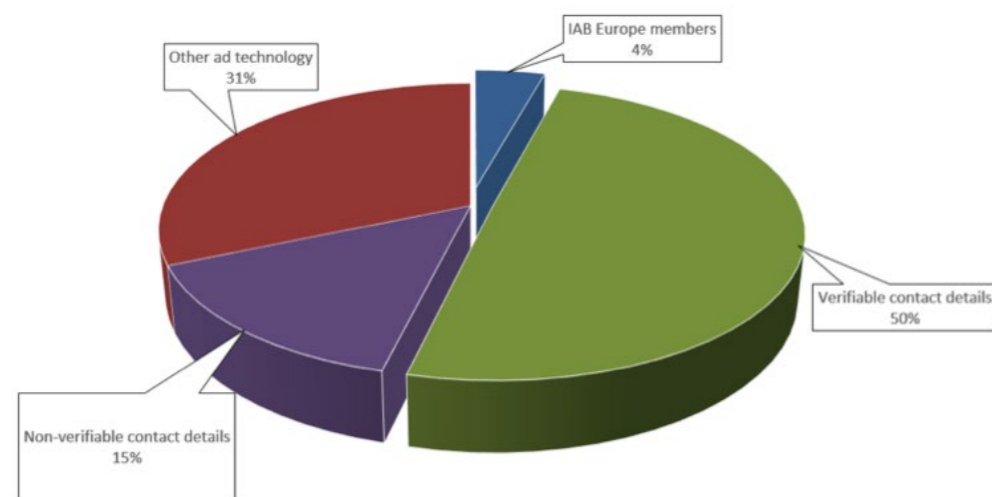


Figure 22: Group breakdown of Intermediaries and those amenable to outreach

CONCLUSIONS

This study represents data captured during a specific six week time period across 1,400 separate webpages on 280 websites suspected of infringing IP. The study collected and analysed over 180,000 ads. BitTorrent Portals had the highest percentage of Mainstream ads (52%) and 92% of BitTorrent Portals used additional ad revenue generation techniques, such as Pop-up Ads or Pop-under Ads. Linking Websites had the highest percentage of High Risk ads (57%).

A. Diversity of Advertising

The broad diversity of advertising on suspected IP infringing websites is evidenced by the 131 unique sectors identified advertising in this study. It is worth noting that 46% of advertising found on the Selected Websites was Mainstream, although other studies have shown this figure to be as high as 71%³⁰. High diversity of advertising and Mainstream sector ads on suspected IP infringing websites might increase the credibility of those websites in the eyes of consumers. Higher ad diversity may indicate a greater level of success in attracting advertising. In turn that may result in greater competition for Ad Space on those websites, potentially increasing the prices they can command.

B. Brand-rich Environment

Suspected IP infringing websites are a brand-rich ad environment with 1,581 unique named Brands identified. Well-known and premium European and global Brands are advertising on these websites. In this study, 1,514 Mainstream Brands were identified and 46 of the top 100 global companies by ad spend had at least one Brand advertised. These Brands may be funding infringement, risking their own reputations through ad placement associations and lending the websites credibility. In some cases Brands are inadvertently paying for their ads to appear next to malware or may be unknowing victims of ad fraud, duped into buying Ad Space on innocent sounding websites (note 41% of the Selected Websites appeared to engage in Pixel Stuffing ad fraud). Brands may be unaware that their ads appear on suspected IP infringing websites, particularly if the digital ad supply chain is not fully transparent meaning Brands are not familiar with the websites on which their Intermediaries are choosing to place their ads.

C. High Risk Advertising

Of the advertising identified on the Selected Websites 54% was High Risk (ads with malware, fraudulent and adult content). These ads may pose a risk to consumers and premium Brands that are advertising on suspected IP infringing websites risk further tarnishment by being placed next to High Risk advertising. Intermediaries involved in placing High Risk ads range from those that predominantly deal with High Risk content and those that conduct little due diligence regarding the Brands with which they engage. This study found just 10 out of 232 Primary Intermediaries identified placed 91% of all High Risk ads collected.

29 - Intermediaries may be members of national IABs or similar organisations but not members of IAB Europe. National level organisation analysis has not been undertaken for this report.

30 - Unpublished whiteBULLET study for UK IP Office on the scale and scope of ad supported IP infringement in the UK.

D. Intermediaries Placing Ads

A diverse range of 232 unique Primary Intermediaries has been identified in this study. This number comprises well-known European and global Intermediaries. Primary Intermediaries can be divided into (i) those Ad Companies that are members of IAB Europe with commitments to avoid placing ads on inappropriate websites; (ii) Ad Companies that are contactable but not necessarily members of IAB Europe; (iii) Ad Companies with no verifiable contact details; and (iv) other ad technologies, often non-transparent, that appear to have no contact details and no identifiable public presence. Primary Intermediaries with verifiable contact details made up 54% of all Primary Intermediaries identified in this study, inclusive of IAB Europe members. A core group of 25 out of 232 Intermediaries are involved in placing 73% of the Mainstream ads.

E. Putting the Study In Context

Dynamic Ecosystem:

A substantial number of suspected IP infringing websites appear to be relevant to each of the EU Member States and the universe of websites changes frequently. Websites regularly change domains to attempt to stay ahead of potential enforcement action and even websites that are blocked by certain ISPs use mirrors and proxies to get around such restrictions. Even during this study several websites changed domain, requiring use of monitoring techniques to ensure consistent ad capture. Therefore static lists of suspected IP infringing websites may have limitations, while dynamic tools may keep up with changes in this evolving ecosystem.

Current Efforts:

It should be noted that several efforts are underway across the world to tackle the issue of ad funded support of IP infringing websites.

In the UK the Digital Trading Standards Group published a set of Good Practice Principles for the Trading of Digital Display Advertising in 2013 suggesting that ad trading agreements should include the buyers' and sellers' intention as to where the advertising should (or should not) appear.³¹ These principles suggest use of independently-certified tools or 'appropriate/inappropriate schedules', which might comprise lists of websites usually maintained by Intermediaries internally. The UK City of London Police IP Crime Unit has created an Infringing Website List (or "IWL") populated with names of websites originally provided by IP rights holders and vetted by the police³². The IWL can be accessed voluntarily by Intermediaries and used to avoid ad placement to those offending websites. The list is updated periodically by right holders.

In March 2015 in France, a Digital Advertising Charter was drafted by the French Ministry of Culture and Communication, creating recommendations in the area of ad safety and IP and a committee to monitor progress on limiting funding to IP infringing websites³³. Further information and developments are expected.

³¹ - See www.jicwebs.org/agreed-principles/digital-trading-standards-group-good-practice-principles.

³² - See www.cityoflondon.police.uk/advice-and-support/fraud-and-economic-crime/pipcu/Pages/Operation-creative.aspx.

³³ - See www.culturecommunication.gouv.fr/Presse/Dossiers-de-presse/Signature-de-la-charte-des-bonnes-pratiques-dans-la-publicite-en-ligne-pour-le-respect-du-droit-d-auteur-et-des-droits-voisins.

In other countries, like Italy, Germany, Spain and Denmark, various announcements have been made about tackling suspected ad-funded IP infringement. These include engagement with major tech companies and brands committed to avoiding ads going to suspected IP infringing websites, talks with advertising stakeholders to determine best practices and potential regulation via judicial bodies.

Outside Europe, the US Trustworthy Accountability Group (or "TAG"), formed in September 2014, is an IAB led cross-industry group including Brands and Agencies³⁴. TAG seeks to deliver a safer digital advertising environment by eliminating fraudulent traffic, combatting malware, fighting Internet piracy and promoting brand safety through greater transparency. Efforts are underway to encourage Brands and Intermediaries to use tools to minimise the risk of ad placement on inappropriate websites (including suspected IP infringing websites) and to certify entities that can help Brands and Intermediaries do this.

³⁴ - See www.tagtoday.net.

PART II: APPENDICES

CONTENTS

Appendix A:	About whiteBULLET	33
	• IPI Index®	
	• AdWiseable™	
Appendix B:	Methodology	35
Appendix C:	Glossary	40
Appendix D:	Ad Sectors/ Sub-sectors Identified	43
Appendix E:	Top Level Ad Sectors Identified	47

APPENDIX A: ABOUT WHITEBULLET

whiteBULLET is an IP solutions company and a leading global provider of innovative Internet safety and compliance solutions. Our solutions appeal to a broad range of Internet companies, including advertising companies and brands.

whiteBULLET has created two leading products used by IP rights holders, law enforcement and a range of digital display advertising companies:

- **IPI Index®: the Standard for Online IP Risk Assessment.** The IPI Index® (IP Infringement Index) is the first comprehensive, global website IP risk assessment solution; and
- **AdWiseable™: Advancing Integrity in Digital Advertising.** AdWiseable™ is a high volume advertising monitoring tool that provides insight and transparency on digital ad placement.

whiteBULLET won an IP innovation award from the UK Government IP Office for the IPI Index® in 2013 and a UK Government Innovate Award in 2015.

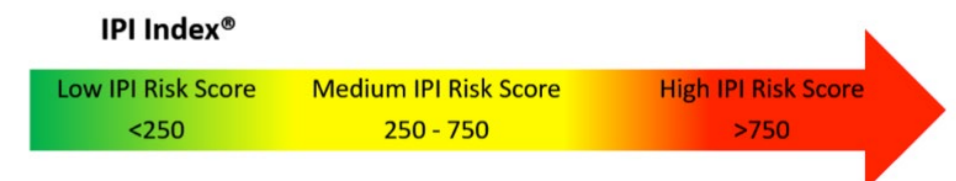
For further information visit www.white-bullet.com.

A. About the IPI Index®

whiteBULLET's proprietary IP Infringement Index ("IPI Index®") is an index of IP risk scores for websites developed through a unique standard for online IP risk assessment. It is a standardised means of rating websites by reference to IP infringement and assesses websites based on a broad range of over 400 data points and consistent core criteria.

Amongst others, these data points include (i) identification of suspected IP infringements on websites (using keyword and metadata scanning to find over 450 regularly updated content titles of a range of media types), and verification of infringements using a range of tools, including metadata, keyword and file matching; (ii) data as to the infrastructure of each website, including hosting and registration data; (iii) information about the means by which websites market suspected infringing files; (iv) measures websites may take to evade enforcement or outreach efforts; and (v) any measures websites may take to mitigate the likelihood of infringements being uploaded or distributed via their platform.

Information from these data points contribute to core criteria by which each website is assessed, with those criteria carrying a weighted score. Using a scoring algorithm, each website is assigned an "IPI Risk Score", which scores rank from low (<250) to medium (250-750) to high (>750). Any website can be assigned an IPI Risk Score. These scores guide online users as to the IP infringement risk posed by websites, allowing these users to make an informed choice on how to interact with a given website (e.g. advertising companies can know which websites may pose a risk to brands) or to inform IP protection and compliance strategies.



B. About AdWiseable™

AdWiseable™ is whiteBULLET's proprietary ad monitoring system. It is a high-volume tool that can be used to aid safe and transparent placement of online advertising.

Unique software identifies advertising on websites (e.g. on suspected IP infringing websites that pose a risk to brand equity) by visiting the webpages multiple times daily and from different IP addresses. AdWiseable™ detects not only the brands advertised but also the chain of Intermediaries placing advertising on the webpages.

The system can be used as an auditing tool for the advertising industry to avoid lost revenues and brand tarnishment due to such misplacement of ads on inappropriate websites. Brands can identify when ads are placed on inappropriate websites and not pay the websites or the broker fees. It can also assist IP rights holders and law enforcement to identify companies funding suspected IP infringement through advertising.

Summary of features:

- Monitors default websites or customised website lists;
- Crawls from various territories to capture locally targeted ads;
- Defined crawling schedule visits each website multiple times daily;
- Identifies advertising served in association with suspected infringing or inappropriate content and captured in context of the webpage;
- Proprietary system combines automated identification and 'learning' technology trained by experts;
- Identifies brands and advertising sectors (e.g. malware, fraud, adult, financial, fashion, travel, technology);
- Deep ad industry expertise guides identification of complex ad placement paths;
- Analyses ad delivery codes to identify chains of intermediaries involved in ad placement and the primary intermediary responsible for engaging with the website;
- Optional automated notice programme sends notices or alerts to intermediaries, brands and/or interested parties with appropriate evidence.

Data available for each webpage and ad includes brands, ad sectors, intermediaries, as well as evidence including timestamps, content suspected of being infringed or content risking brand damage, screenshots and ad or webpage software code.

APPENDIX B: METHODOLOGY

A. Project Scope

The Observatory requested this study focus on a selection of significant websites suspected of IP infringement (the "Selected Websites"), which were to meet the following criteria:

1. websites of significant **popularity** in one or more of the 28 EU Member States;
2. websites offering a variety of **suspected infringing copyrighted content**¹ (individually or collectively) representing a range of IP industries (including film, TV, sports, music, games, software, eBooks);
3. websites to include a range of **website types** (i.e. offering content via a range of means, such as streaming or direct download); and
4. websites with some form of **advertising revenue**.

The advertising profile for the Selected Websites would be assessed and the Selected Websites monitored for advertising from each of the 28 EU Member States in order to generate locally targeted advertising, using whiteBULLET's AdWiseable™ ad monitoring system supplemented by direct observation by analysts.

B. Questions

The following questions were outlined to be answered as part of the study:

1. **How significant is advertising on the Selected Websites?**
 - a. How prevalent is advertising on the Selected Websites?
 - b. What is the diversity of advertising sectors across the Selected Websites?
 1. Which advertising sectors appear on the Selected Websites (e.g. automotive, financial, adult, malware)?
 2. Which advertising sectors dominate by volume?
 3. What proportion of 'high risk' advertising sectors appear on the Selected Websites (defined as adult, malware and fraudulent advertising)?
 4. What proportion of mainstream advertising sectors appear on the Selected Websites?
 - c. What is the diversity of Brands across the Selected Websites?
 1. Which Brands appear on the Selected Websites?
 2. Which Brands dominate by volume?
 3. Are well-known household names among the advertised Brands and in what proportion relative to generic or un-named products?
2. **How is advertising being placed on the Selected Websites?**
 - a. Which Intermediaries (including digital advertising networks, exchanges or servers) place Branded advertising on the Selected Websites?
 1. Which Intermediaries dominate by volume?
 2. Which Intermediaries place the majority of the mainstream advertising?
 3. Which Intermediaries place the majority of 'high risk' advertising (see 1 above in relation to advertising sectors)?
 - b. Can these Intermediaries form part of the solution to advertising funded IP infringement?
 1. Are these Intermediaries based in the EU?
 2. Are these Intermediaries members of EU advertising trade associations?

¹ - This study was to monitor only websites suspected of offering infringing copyrighted content, not those offering suspected counterfeit goods. Websites offering counterfeit goods tend not to have advertising revenue, but rather make money through direct sales of products.

with commitments to protect Brands from placement on inappropriate websites?

3. Are these Intermediaries contactable or amenable to outreach?

3. **How significant is cross-border advertising across the Selected Websites?**

a. Is advertising from one territory being placed on the Selected Websites when viewed from a different territory?

b. Which territories have the highest level of mainstream advertising?

C. Website Selection Process

Thousands of websites impact the IP industries and consumers in the EU. A range of sources and processes were used in order to identify appropriate candidate websites to be included in the study meeting the criteria for the scope requested by the Observatory. The selection process was undertaken during the two week period 23 March to 3 April 2015 and involved the steps described below.

Website Popularity:

The Observatory requested websites popular in one or more of the 28 EU Member States be identified. The following sources were used to identify significant suspected IP infringing websites popular in one or more of the 28 EU Member States:

1. whiteBULLET's IPI Index®;
2. Alexa Top 500 website popularity ranks for each of the 28 EU Member States;
3. Alexa global Top 500 website popularity rank;
4. Searching IP infringement related terms within search engines popular in the EU Member States; and
5. Google's Transparency Report.

Alexa Ranks. Alexa is a web metrics company that provides data about the measure of a website's popularity compared with all of the other websites on the Internet. This data considers both the number of visitors and the number of pages viewed on each visit. Alexa collects traffic data daily from millions of users who have installed the Alexa toolbar, and from direct measurements from websites that have incorporated Alexa code, and then uses a proprietary formula to create a popularity ranking for each website. A website's Alexa Rank can be interpreted as the website's position in a league table, with the most popular website given a rank of 1, the next 2 and so on through millions of websites. Alexa provides information about the ranking of websites by country and creates top 500 most popular website lists by country (including for all EU member States). This data can be used to identify the most popular websites in a given country.² Alexa also provides a global top 500 ranking representing the most popular websites in the world according to Alexa³.

Google Transparency Report. Google regularly receives requests from copyright owners and their agents and organisations that represent them to remove search results that link to content allegedly infringing copyright. Google makes available online a report that specifies the number of requests it receives to remove search results, and indexes the results by domains, copyright owners, reporting organisations and requests⁴. The Google Transparency Report

2 - See www.alexa.com/topsites/countries.

3 - See www.alexa.com/topsites/global.

4 - See www.google.com/transparencyreport/removals/copyright.

indicates the volume of infringement takedown requests sent by parties to Google for search takedowns in relation to websites that may infringe copyright. Volume of takedown requests will depend on those parties' priorities, budgets and capacity to send volume notices to Google. Therefore, used alone, the Google Transparency Report may not provide an accurate indication of which websites are actually most infringing.

Initially a candidate website database of significant popular websites was created by selecting all websites within the IPI Index® (see description and details of the IPI Index® in Appendix A) with a medium or high risk IPI Risk Score (i.e. above 250 in the IPI Index®). These were then cross-referenced against Alexa top 500 most popular websites for each of the 28 EU Member States. Any websites in the top 500 lists suspected of having IP infringing content but not in the IPI Index®, were checked manually. Indicators of suspected IP infringing content included websites with keywords in the domain name associated with IP content (e.g. movie, book, sport, TV) or copyright infringement (e.g. pirate, download, stream, rip, mp3). The same exercise was undertaken with the websites on the Alexa top 500 most popular global website list.

Popular content titles were searched together with keywords related to copyright infringement (such as pirate, download, stream, free copy) within search engines popular in the EU Member States, including local website variations, e.g. www.google.fr or www.bing.de. Links to websites in the top ten search results were then cross-referenced with the candidate website database and any new websites then added as these were deemed to be popular in search engine results. Finally, websites in the candidate list were cross-checked with the Google Transparency Report for specific organisations to identify websites with the highest number of infringing link notices sent to Google by key IP rights holders and IP content protection associations for content types not currently included in the IPI Index®, such as games or eBooks.

To ensure a breadth of websites relevant to all EU Member States, each website was assigned to either an EU-wide status, or to a specific EU Member State using Alexa rank data. Around one third of the candidate websites were popular across the EU generally and so were not assigned to a single primary EU Member State⁵ but instead assigned to the EU as a whole. At least one popular website was identified for each EU Member State even though websites may not actually be hosted in that territory⁶.

Suspected Infringing Copyrighted Content:

All websites in the candidate website database were then reviewed and assessed for suspected copyright infringing content including the latest film, TV, music, sports, eBook, software and game titles. Titles included in the IPI Index® to search for suspected infringing content were used, as well as popular content titles selected from databases such as iTunes and Amazon or from 'new features' sections on websites offering IP content. The candidate websites were cross-referenced against websites known to offer licensed content or those published on legitimate content indices (e.g. thecontentmap.com).

5 - Assigning a website to one specific country does not mean that website is not also popular in at least one other EU Members States

6 - This is due to the widespread practice of locating servers in countries different from the primary audience of a website.

Website Types:

Suspected IP infringing websites distribute content files through different technical means, some more popular with certain demographics of the population depending on ease of technical use and speed of data transfer. A range of website types were included in the candidate website database. To do this, the candidate websites were categorised by the following website types according to the type of technology used and business model:

1. BitTorrent Portals;
2. Hosting Websites; and
3. Linking Websites.

Website type descriptions are in the Scope section of the report.

Advertising Presence:

Finally, to meet the criteria of advertising revenue, each candidate website was checked manually for the presence of advertising.⁷ Numerous webpages of each website were checked, including the homepage, and webpages deeper within the website structure such as those webpages offering suspected infringing links and files. The advertising profile of each website was analysed to identify the type of advertising present, including the use of display ads, Pop-up Ads, Pop-under Ads and to ensure the advertising appeared adjacent to webpages offering suspected infringing IP content. In addition, the HTML code of each webpage was checked for the presence of suspected Pixel Stuffing (ad fraud via 1x1 pixel advertising), which was later cross-referenced against the HTML code collected during automated ad monitoring. See Analysis and Key Findings section of the report for details on this form of ad fraud.

From the candidate website database, 280 websites were identified that met the criteria set by the Observatory. Of these, 66% had Alexa scores (either global or EU country scores) of less than 10,000, and 24% less than 500.

The list of websites was approved by the Observatory for the study and became the Selected Websites. The Selected Websites included 64% Linking Websites, 23% Hosting Websites and 13% BitTorrent Portals (definitions included below). 62.5% of the Selected Websites offered a mix of different types of IP content (e.g. included music, film, games and software) and 37.5% of the Selected Websites offered only one type of IP content exclusively (e.g. eBooks only).

D. Advertising Monitoring

In this study, 1,400 webpages were monitored across the Selected Websites using the AdWiseable™ ad monitoring system. This captures high volume data about advertising placed on websites (see Appendix A for a description of AdWiseable™) by visiting several webpages of each of the Selected Websites multiple times daily.⁸ Five webpages for each of the Selected Websites were selected for monitoring, each webpage confirmed as having suspected IP infringing content

⁷ - For a variety of reasons, not all websites are suitable for ad monitoring, including but not limited to: (i) the website having no advertising, (ii) the website having no suspected IP infringing content identified, (iii) the website having coding or loading errors, or using coding techniques that negatively impact the ability to display or monitor advertising, (iv) the website coming offline or changing domain name, (v) the website being only accessible to private members via password. These factors were considered when making the selection of websites for the study.

⁸ - AdWiseable™ operates a cookie-less ad collection system, ensuring no cookies are saved when websites are visited. In this way the system avoids specific advertising being induced through Behavioural Targeting.

available on it directly or via links. Each of the webpages of the Selected Websites was visited from all 28 EU Member States using locally based proxy servers to capture locally targeted advertising from each territory. The system avoids specific advertising being induced through behavioural targeting by deleting Cookies after each webpage visit and scan. Some websites may be blocked by certain ISPs in some countries pursuant to website blocking court orders. Locally based proxies not subject to the court orders were used to access these websites for the purposes of this study.

AdWiseable™ identifies ads served on or associated with suspected IP infringing content on webpages and captures images of ads in context of the suspected infringing webpage. It also uses a unique proprietary system to identify Brands and advertising sectors (e.g. malware, adult, financial, fashion, travel) and the Intermediaries involved in the placement and delivery of ads.

The following was identified for each webpage visit:

1. name of the Brand whose products/services were promoted (data included Brand, product and sub-product);
2. ad sector into which the Brand is categorised, e.g. financial, automotive, clothing or malware sectors (data included sector, sub-sector and sub-sub-sector);
3. names of the Intermediaries involved in placement of the ad;
4. name of the Intermediary identified to most likely be responsible for engaging with the Selected Website ("Primary Intermediary");
5. image of each ad;
6. time and date of each ad (captured in Universal Coordinated Time);
7. URL on which each ad appeared;
8. title of the content suspected of infringement on that webpage where the ad appeared;
9. country from which the ad request was made;
10. screenshots of the entire webpage on which the ad appeared; and
11. full HTML code of both the ad and the entire webpage on which the ad appeared.

The advertising on the Selected Websites was monitored for six weeks from 29 May to 10 July 2015. Over 180,000 ads were collected across the 1,400 webpages monitored from each of the 28 EU Member States during the six week study. The data was then analysed to identify trends across the countries, specific information per territory and amongst other:

1. top Brands and products advertised;
2. top ad sectors and sub-sectors into which Brands are categorised; and
3. Primary Intermediaries responsible for engaging with the Selected Websites.

APPENDIX C: GLOSSARY

Ad Company: entity involved in placing or serving digital ads to a website.

Ad Server: computer server or network of servers that store and deliver digital ads to websites or make technology driven decisions about the placement of advertising. Publishers, Brands and Agencies may each have their own Ad Servers.

Ad Space: dedicated space available within a website in which a digital ad can be placed. These spaces are often of a standard size.

Agency: advertising entity used by Brands in the planning, creating and buying of digital media campaigns.

Behavioural Advertising: advertising targeting technique that relies on a range of methods to improve the effectiveness of advertising by matching users' Internet browsing information with ads that may appeal to those users' interests. Often managed through use of tools like cookies and similar technologies.

BitTorrent Portal: websites using P2P file distribution technology to permit users to share content. These websites act as aggregators of P2P links, which users can search for and access via the website.

Brand: advertiser having its brand and creative ad placed on a website.

Cookie: small pieces of web code placed by websites and stored on a user's web browser. These send information about a user's web habits on that website back to the website to remind the website of the user's previous activities and can be used to target advertising specifically to that user.

Content Marketing: images, articles or news displayed in dedicated boxes within a webpage by content marketing companies. This format of advertising seeks to be related to the webpage and is regularly updated or may change on each visit to the website.

Cyberlocker: a form of Hosting Website that may offer streamed or downloadable media content, requiring payment of a monthly subscription fee or offering ad-supported content for no charge.

Direct Download Hosting Website: Hosting Website that permits content to be downloaded directly. These websites may be ad-supported, offering the content to download at no charge, or they may charge for premium access, whereby users pay a subscription fee to avoid advertising and for faster download speeds.

Data Supplier: entities that collect and organise information about websites and audience for use in advertising targeting.

Demand Side Platform or DSP: technology platform that enables buyers of digital ad space to manage multiple ad exchange accounts through a single technology interface.

Embedded Ad: static form of advertising that is hard-coded into a webpage and will appear on

the webpage until the code is removed or changed by the website administrator. These ads are not engaged with Intermediaries or calls from Ad Servers and often feature a brand or logo.

Exchange: unified technology platform that facilitates bid-based buying and selling of digital ad space for a range of online parties.

Generic: used by whiteBULLET to describe ads that have no discernible product or brand name associated with them and cannot be attributed to a Brand.

Hosting Website: websites acting as storage servers that permit users to upload and store media files in the cloud. Once a media file is uploaded, a user can generate a link to access that file again, either themselves, or for others to use the link to obtain the content.

Impression: terms used when an ad is received from a source and can be counted for the purposes of delivery to an Ad Space.

Inventory: collective term used to describe Ad Spaces made available for sale via digital advertising processes.

Intermediary: Ad Company or other entity involved in placing or serving digital ads to a website.

Interstitial: ad that appears in front of or behind a webpage, e.g. Pop-up Ads or Pop-under Ads

Linking: websites that aggregate links to media content that is stored on Hosting Websites. They often categorise links by content type and then offer search facilities and suggestions so users can quickly find content.

Mainstream: advertising that is not High Risk. Mainstream will include premium brands, including well-known household names.

Mouse-over: the act of moving a computer screen cursor and resting it on the hot spot of an ad for at least one second, which may trigger an event such as expanding the ad or initiating an animated sequence within the ad.

Network: ad broker that aggregates digital ad space from a range of websites, packages it for a target audience and matches it with advertiser demand. Networks may be directly engaged with buyers or act as secondary affiliates offering supplementary ad space to other Networks.

On-click: action describing advertising that is generated when a user clicks the mouse cursor on an originating website. The generated ad is often displayed in a new browser window.

Overlays: see Interstitial.

P2P: peer-to-peer technology for file distribution making use of links aggregated by Linkings. When a user clicks on a P2P link, the P2P technology allows the user to download media files stored on other P2P users computers across the P2P network. This

is undertaken at no charge to any of the P2P users. A user in a P2P network will not only download files from other users, but will make his or her media files available for upload to the P2P network so that other users can download files from them as well.

Pixel Stuffing: ad fraud technique involving tiny Ad Spaces sized at just 1x1 pixel (or sometimes larger) hidden within the top or bottom of webpages. These Ad Spaces are 'stuffed' with hidden ads, which are served to them just as they would be to regular sized Ad Spaces. However, a visitor to the website would not see these tiny ads, although they would be logged by the ad ecosystem as an Impression. Often used by websites to generate additional fraudulent ad revenue.

Pop-under Ad: ad displayed in a new browser window that opens behind the main or originating browser window (rather than superimposed in front of it) and may not be noticed by the user until the main browser window is closed.

Pop-up Ad: ad displayed in a standalone window superimposed over the main browser window, which itself is not a new browser window and may be closed.

Potentially Unwanted Programme or PUP: likely unwanted software that is downloaded by web users usually unknowingly through webpage advertising. It may compromise the user's privacy or contain malicious content such as viruses or other forms of malware.

Primary Intermediary: Intermediary identified by whiteBULLET to be the Intermediary responsible for engaging with the website, and often the party engaged in financial transactions with the website.

Programmatic: targeted ad campaigns deployed according to software rules and enriched by data.

Publisher: website which is selling its Ad Space to potential buyers.

Real-Time Bidding or RTB: advertising process during which Inventory is bought and sold via almost instantaneous programme-managed auctions

Sell Side Platform or SSP: technology platform that enables websites to manage ad space on multiple ad exchange accounts.

Sponsorship Ad: see Embedded Ad.

Trading Desk: either independent or a subsidiary of an Agency (Agency Trading Desk) that comprises a team and technology used to optimise trading on Exchanges through use of data and other Intermediaries.

Video Streaming Hosting Website: Hosting Website that permit the content files to be streamed via embedded media players. These websites may be ad-supported, offering the content for streaming at no charge, or they may charge for premium access, whereby users pay a subscription fee to avoid advertising. Hybrid versions of these websites exist that allow immediate streaming while content also is downloaded.

APPENDIX D: AD SECTORS/SUB-SECTORS IDENTIFIED

The following is a list of the 131 ad sectors (combined sector/sub-subsectors) identified from the 184,034 ads collected for the study. These are listed alphabetically.

	Ad Sector
1	Adult
2	Arts & Entertainment – Blog
3	Arts & Entertainment – Books & Literature
4	Arts & Entertainment – Celebrity Fan/Gossip
5	Arts & Entertainment – Gaming
6	Arts & Entertainment – Movies & Television
7	Arts & Entertainment – Music
8	Arts & Entertainment – Stock Media
9	Arts & Entertainment – Theatre
10	Automotive – Auto Parts
11	Automotive – Buying/Selling Cars
12	Automotive – Buying/Selling Motorcycles
13	Automotive – Car Expertise
14	Automotive – Car Repair
15	Automotive – Petroleum
16	Automotive – Road-Side Assistance
17	Business
18	Business – Advertising
19	Business – Consumer Services
20	Business – Logistics
21	Business – Manufacturing
22	Careers – Career Planning/Advice
23	Careers – Job Search
24	Careers – Web Careers
25	Careers – Work From Home
26	Click Generators/ Malware
27	Click Generators/ Malware – Adware
28	Click Generators/ Malware – PUP
29	Click Generators/ Malware – PUP/Browser Hijacker
30	Click Generators/ Malware – PUP/Browser Hijacker u2013 Adware
31	Education – College Life
32	Education – Distance Learning

33	Education – Primary Education	70	Law, Government & Politics – Legal
34	Family & Parenting – Babies & Toddlers	71	Law, Government & Politics – Public Records
35	Food & Drink – Cocktails/Beer	72	Multiple Ads
	Ad Sector	73	News – International News
36	Food & Drink – Coffee/Tea	74	Personal Finance
37	Food & Drink – Dining Out	75	Personal Finance – Credit/Debt & Loans
38	Food & Drink – Food Recipes		Ad Sector
39	Food & Drink – Groceries	76	Personal Finance – Financial Planning
40	Gambling – Betting	77	Personal Finance – General Banking
41	Gambling – Bingo	78	Personal Finance – Insurance
42	Gambling – Lottery	79	Personal Finance – Payment Services
43	Gambling – Online Casino Games	80	Personal Finance – Price Comparison
44	Gambling – Poker	81	Personal Finance – Stocks
45	Health & Fitness	82	Real Estate
46	Health & Fitness – Alternative Medicine	83	Real Estate – Buying/Selling Property
47	Health & Fitness – Dental Care	84	Shopping – Consumer Reviews
48	Health & Fitness – Exercise	85	Shopping – Contests & Freebies
49	Health & Fitness – Female Health	86	Shopping – Couponing
50	Health & Fitness – Foot Care	87	Shopping – Market Research
51	Health & Fitness – Hearing Health	88	Shopping – Online Auctions
52	Health & Fitness – Hygiene	89	Shopping – Online Shopping
53	Health & Fitness – Male Health	90	Shopping – Store
54	Health & Fitness – Medical Care	91	Social Media
55	Health & Fitness – Medical Equipment	92	Social Media – Virtual Reality Community
56	Health & Fitness – Medication	93	Society – Charity/Non-Profit
57	Health & Fitness – Optical Health	94	Society – Classifieds
58	Health & Fitness – Smoking	95	Society – Dating
59	Health & Fitness – Weight Loss	96	Society – Religion
60	Home & Garden	97	Sports – Auto Racing
61	Home & Garden – Appliances	98	Style & Fashion – Accessories
62	Home & Garden – Furniture	99	Style & Fashion – Beauty
63	Home & Garden – Household Cleaning	100	Style & Fashion – Clothing
64	Home & Garden – Interior Decorating	101	Style & Fashion – Jewellery
65	Home & Garden – Pets	102	Technology & Computing – Antivirus
66	Home & Garden – Remodelling & Construction	103	Technology & Computing – Broadband
67	Home & Garden – Stationery	104	Technology & Computing – Cameras
68	Home & Garden – Utilities	105	Technology & Computing – Cloud Services
69	Law, Government & Politics – Immigration	106	Technology & Computing – Computer Accessories
		107	Technology & Computing – Computers
		108	Technology & Computing – Electronics

109	Technology & Computing - Gaming Console
110	Technology & Computing - General Telecoms
111	Technology & Computing - Mobile Network Operators
112	Technology & Computing - Mobile Phones
113	Technology & Computing - Search Engine
114	Technology & Computing - Shareware/Freeware
115	Technology & Computing - Software
	Ad Sector
116	Technology & Computing - Tablets
117	Technology & Computing - Televisions
118	Technology & Computing - Web Analytics
119	Technology & Computing - Web Hosting
120	Travel
121	Travel - Air Travel
122	Travel - Cruises
123	Travel - Ferries
124	Travel - Holiday
125	Travel - Hotels
126	Travel - Online Travel Reservations
127	Travel - Public Transport
128	Travel - Tourist Attractions
129	Travel - Tourist Board
130	Travel - Travel Agent
131	Uncategorized - Images

APPENDIX E: TOP LEVEL AD SECTORS IDENTIFIED

The following is a list of the 25 top level ad sectors identified from the 184,034 ads collected for the study. These are listed alphabetically.

	Ad Sector
1	Adult
2	Arts & Entertainment
3	Automotive
4	Business
5	Careers
6	Click Generators/ Malware
7	Education
8	Family & Parenting
9	Food & Drink
10	Gambling
11	Health & Fitness
12	Home & Garden
13	Law, Government & Politics
14	Multiple Ads
15	News
16	Personal Finance
17	Real Estate
18	Shopping
19	Social Media
20	Society
21	Sports
22	Style & Fashion
23	Technology & Computing
24	Travel
25	Uncategorized



WWW