





Click Fraud and Legal Issues

paying for illegal clicks
linking to pornography
linking to Nazi memorabilia
linking to websites with false, libelous information
placing ads on bogus sites







Elements of Pay-per-Click

- Search engines rely upon pay-per-click advertising for the bulk of their revenue
- Last year Google earned approximately \$110 billion, mostly from clicks
 - if an average click brings in \$2, that works out to almost 150 million clicks/day on ads;
- When a viewer clicks on an ad, it is called a "click thru"; CTR is click thru rate —
- Ad position is critical
- CPM stands for Cost Per Thousand
- The percentage of invalid clicks is a hotly debated topic

Average Pay-per-Click Costs, 2017-2020

Average PPC Costs 2017-2020

Metric	2020	2019	2018	2017
Cost per click (CPC)		\$1.03	\$0.99	\$1.18
Click through rate (CTR)		1.8%	2.1%	2.5%
Cost per mille (CPM)		\$18.71	\$20.90	\$29.46
Conversion rate		5.2%	3.1%	2.7%
Cost per conversion		\$19.64	\$31.79	\$43.52
Invalid click rate		5.7%	6.6%	7.5%

Smartphone and tablets now account for 32% of paid search clicks and 25% of all PPC ad spend

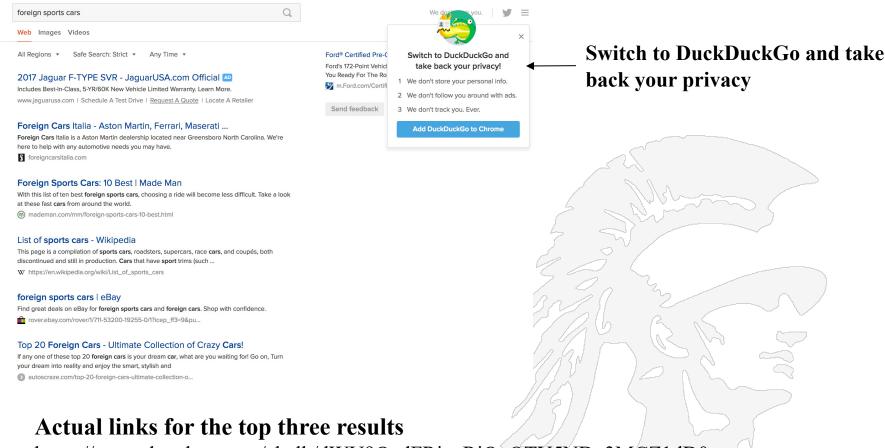


https://www.hochmanconsultants.com/cost-of-ppc-advertising/





DuckDuckGo Does Not Record Clicks



https://r.search.yahoo.com/cbclk/dWU9QzdFRjcxRjQyOTY5NDg3MCZ1dD0x

http://www.foreigncarsitalia.com

http://www.mademan.com/mm/foreign-sports-cars-10-best.html

Copyright Ellis Horowitz 2011-2022





Two Basic Problems with Pay-Per-Click Model

- 1. Although correlated, good click-through rates (CTRs) are still not indicative of good conversion rates, since it is still not clear if a visitor would buy an advertised product once he or she clicked on the ad
 - In this respect, the Cost Per Action (CPA)-based models (where the advertiser only pays when the product is purchased) provide better solutions for the advertisers (but not necessarily for the search engines), since CPA are more indicative that their ads are "working."
- 2. It does not offer any "built-in" fundamental protection mechanisms against click fraud since it is a challenge to determine which clicks are valid vs. invalid in general, (it can be done relatively easily in some special cases, but not in general).
 - For this reason, major search engines launched extensive invalid click detection programs and still face problems combating click fraud.





Definitions of Invalid Clicks

- **Definition:** "Click Fraud occurs in pay-per-click online advertising when a person, automated script or computer program imitates a legitimate user of a web browser clicking on an ad, for the purpose of generating an improper charge per click"
 - see Wikipedia, http://en.wikipedia.org/wiki/Click_fraud
- Google prefers to use the term "invalid clicks" rather than fraudulent clicks
- Alternate Definition: "Clicks or page impressions generated through prohibited means, and intended to artificially increase click or impression counts on a publisher account." see
 - https://www.techulator.com/experts/649-Explain-me-terminologiess-Google-Adsense
- A click is clearly valid if it leads to a purchase (or subscription), but this information is not always available





Some Examples To Ponder

- 1. A person clicks on an ad, leaves the landing page, but returns a while later; (is that two clicks or one?)
- 2. A person double clicks an ad in a short period of time, say p
 - If p < 5 seconds then it may be a mistake
 - If p > 10 seconds, then maybe he clicked once, hit the back button, and clicked again
 - Without correct tagging, one click and five page re-loads can easily be mislabeled as six clicks from the same visitor
- 3. An advertiser seeing many clicks from the same IP address surmises they must be fraudulent.
 - However many of these types of clicks are indeed valid, as so many people using corporate computers share the same IP, or ISPs assigning the same IP to more than one customer
- Conclusion: it is not possible to discern intent in all cases





7 Ways Click Fraud May Be Done

- 1. Individuals deploying automated clicking programs or software applications (called bots) specifically designed to click on ads
- 2. An individual employing low-cost workers or incentivizing others to click on the advertising links
 - Click2freemoney.com used to offer three options for people to earn money by clicking on website links via e-mails that they send, by clicking on banners and text ads in their paid-to-click section, and by referring others to the website
 - This website no longer is active
- 3. Website publishers manually clicking on the ads on their pages
 - attempting to generate revenue for themselves
- 4. Website publishers manipulating web pages in such a way that user interactions with the web site result in inadvertent clicks, e.g iframes
- 5. Website publishers subscribing to paid traffic websites that artificially bring extra traffic to the site, including extra clicking on the ads
- 6. Advertisers manually clicking on the ads of their competitors
- 7. Website publishers being sabotaged by their competitors or other ill-wishers







Google Detection of Invalid Clicks

Proactive

1. Filters

- •Automated algorithms which filter out invalid clicks in real time, before they are charged to advertisers' accounts
- •Analyze all clicks on AdWords ads
- •Automatically discard clicks and impressions from known sources of invalid activity
- •The above accounts for vast majority of invalid click detection

2. Offline analysis

- •Automated algorithms and manual analysis
- •Focuses on the AdSense network, but also includes Google sites and those in other networks
- •Accounts for a much smaller proportion of invalid clicks than that detected by our automated filters

Reactive

3. Investigations

- •Relatively rare
- •All advertiser inquiries about invalid clicks are investigated by Google and taken very seriously
- •The number of invalid clicks detected by investigations is negligible relative to the number detected by filters and offline analysis

Copyright Ellis Horowitz 2011-2022





What To Look For -Anomalous Behavior

- 1. Keyword performance
 - Words that normally don't do well suddenly become very popular
- 2. Abnormal number of clicks from the same IP address
 - Easy to identify, but still popular form of click fraud
- 3. Decline in the number of conversions (requires website cooperation)
 - Since fraudulent clicks don't lead to purchases
- 4. Large numbers of visitors who leave your site quickly (requires website cooperation)
 - Many search engines require that the user who clicks on an ad spend at least 30 seconds on the landing page
- 5. Large number of impressions, without accompanying clicks on your ad
 - If ad impressions are inflated, but not clicked then Google can cancel your ad because the CTR is so small
- 6. Abnormally high clicks and impressions on affiliate websites
 - If an affiliate competitor is using click fraud, the search engine will notice high payments to the affiliate and may cancel the affiliate's account
- 7. A large number of clicks coming from countries outside of your normal market area
 - Use sites such as dnsstuff.com to identify which country an IP address is coming from
- 8. Accidental click fraud Link checking software or search engine robots may click on ads as they parse a page

 Copyright Ellis Horowitz, 2011-2022

 Copyright Ellis Horowitz, 2011-2022





Google Reports To Advertisers

- When advertisers are billed by Google, they receive reports describing the clicking and billing activities.
- One problem with these reports, however, is that these statistics are aggregated by Google over some time period. The smallest unit of analysis is one day.
- For example, the number of invalid clicks on an ad detected by Google (or any other related statistic) can *only be reported on a daily basis*
 - In other words, advertisers cannot know if a particular click on a particular ad was marked as valid or invalid by Google, and Google refuses to provide this information to advertisers.
- The Fundamental Problem
 - The advertiser has the right to know why a particular click was marked as valid by Google (when the advertiser thinks that it is invalid)
 - However, if Google discloses this information it opens itself to click fraud on a massive scale because by doing so it provides certain hints about how its invalid click detection methods work





Tuzhilin's Fundamental Problem

- As part of a legal case (The Lane's Gifts v. Google), Prof. Alexander Tuzhilin (of NYU) wrote a report on Google's approach to detecting invalid clicks
 - http://googleblog.blogspot.com/pdf/Tuzhilin_Report.pdf
- Google settled the case on May 19, 2006
 - See the ruling at http://googleblog.blogspot.com/pdf/lanes_google_final_order.pdf
 - As part of the settlement Google agreed to \$60,000,000 of advertising credits and \$30,000,000 amount in attorney fees
 - Lane's Gifts and Collectibles LLC v. Yahoo! Inc., Case No. CV-2005-52-1 (Ark. Cir. Ct. complaint filed Feb. 17, 2005).
- Tuzhilin's main conclusion:
- There is no conceptual definition of invalid clicks that can be operationalized
- An operational definition cannot be fully disclosed to the general public because of the concerns that unethical users will take advantage of it, which may lead to a massive click fraud. However, if not disclosed, advertisers cannot verify or even dispute why they have been charged for certain clicks





Tuzhlin's Attempt to Define Performance of Online Filters

- A *true click is either valid or invalid*, assuming that we know the "absolute truth" about the validity of all the clicks (which is not the case)
- Google filters can *label a click as either valid or invalid*. These two dimensions (the actual click vs. click labeling by filters), give rise to the following *confusion matrix*

		Click classified by filters as	
		Invalid	Valid
Actual Click	Invalid	True Positive (TP)	False Negative (FN)
	Valid	False Positive (FP)	True Negative (TN)

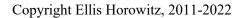
True Positive (TP) is an invalid click that is correctly identified as invalid True Negative (TN) is a valid click that is correctly identified as valid False Positive (FP) is a valid click that is incorrectly identified as invalid False Negative (FN) is an invalid click that is incorrectly identified as valid





Tuzhlin's Attempt to Define Performance of Online Filters

- Given the total number of clicks *N*, we can identify the number of *TP*, *TN*, *FP* and *FN* clicks.
- Note that TP + TN + FP + FN = N.
- and the accuracy rate of a filter is equal to
- (TP + TN)/N
- and the error rate to
- (FP + FN)/N.
- In addition to these measures, there are several other measures that can be used for determining performance of the filters
- But Google refuses to publish these numbers







Features of Pay-per-Click on Mobile

- 1. Google offers both pay-per-click and pay-per-call ads for mobile devices
 - an ad can have a link and a phone number
- 2. Google suggests that mobile ad campaigns should be distinct from desktop campaigns
- 3. Mobile searchers are generally more interested in making a decision rather than doing research
- 4. As a result choose keywords that emphasize an action
- 5. Bid for positions one and two, anything higher wont matter
- 6. Target based on location, device and action
- 7. Day-parting: off hours for desktop PPC are on-hours for mobile PPC
- 8. Include special offers for mobile users who click or call







Legal Issues and Search Engines







What is Intellectual Property

- Content of the human intellect deemed to be unique and original and to have marketplace value—and thus to warrant protection under the law.
 - Intellectual property includes: ideas, inventions, literary works, chemical, business, or computer processes;
- Intellectual property protections fall into four categories:
 - 1. copyright (for literary works, art, and music),
 - 2. patents (for inventions and processes),
 - 3. trademarks (for company and product names and logos), and
 - 4. trade secrets (for recipes, code, and processes).





What is Copyright

- The U.S. Copyright Act, 17 U.S.C. § 102(a), extends copyright protection to "original works of authorship fixed in any tangible medium of expression . .
- A work is protected by copyright law *automatically* when it is created and fixed in a tangible medium of expression. That is, the author owns the work and obtains rights to it by "default."
- The "©" symbol and the words "All Rights Reserved" are no longer required in America,
- A copyright owner has the following rights:
 - Reproduction Right
 - Adaptation Right
 - Distribution Right
 - Public Performance Right
 - Public Display Right
- The term of copyright protection is the length of the author's life plus fifty years.





What Items are Copyrighted

- Books
- Maps
- Charts
- Engravings
- Prints
- Musical compositions
- Dramatic works
- Photographs
- Web pages

- Paintings
- Drawings
- Sculptures
- Motion pictures
- Computer programs
- Sound recordings
- Choreography
- Architectural works







What are Patents

- In the United States, a person may patent any invention that is new, useful and, non-obvious.
- Unlike copyrights, the inventor must register the patent with the United States Patent Office.
- If the Patent Office determines that the patentability requirements have been satisfied, then a patent will be granted.
- The owner of a patent has the right to exclude all others from making, using, selling, or importing the invention.
- This exclusive right is granted for 20 years for most patents.





Patents and Prior Art

- In order to pass the non-obvious, or novel, requirement, the invention must not have been revealed in some form of "prior art."
 - Prior art is defined as the body of public knowledge that predates the date of invention.
 - That is, the invention would not have been obvious to a person skilled in the art with which the invention is concerned at the time the invention was made.
 - Prior art includes prior patents and publications and common knowledge in a field.
- In the United States, an application must be filed within one year of the date of first sale or offer of sale or the sale of anything incorporating the invention, within one year of a description of the invention, or within one year of a public or commercial use of the invention.





What are Trademarks

- A trademark is a word, words, phrase, symbol or design that is legally registered as representing a company or product.
- Applications for registering a trademark in the US are done through the US Patent and Trademark Office USPTO, www.uspto.gov
- The USPTO attempts to ensure that no one else receives a registration for an identical or similar mark
 - But they do not police the misuse of trademarks; that is up to the owner
- Some famous trademarks









What are Trade Secrets

- A trade secret is a formula, practice, process, design, instrument, pattern, or compilation of information which is not generally known or reasonably ascertainable, by which a business can obtain an economic advantage over competitors or customers. (Wikipedia.org)
- Trade secrets are often referred to as "confidential information" or "classified information"
- Companies with trade secrets must protect them from the public, e.g. by having people sign non-disclosure contracts and non-compete clauses
- Trade secrets can last indefinitely
- Computer programs whose only distribution is in executable form allows the source code to remain a trade secret
- Reverse engineering a program is often legal, and once exposed there is no longer a trade secret and no penalty ensues
 - There are laws that prevent reverse engineering, e.g. Digital Millenium Copyright Act (DMCA)





Is Google Responsible for the Content it Points To?

- Should Google not be allowed to provide links to
 - pornographic websites
 - child pornographic websites
 - Nazi, white supremacist, racist websites
 - Revenge porn websites
 - websites containing damaging or libelous information about an individual







Section 230 Communications Decency Act

- Section 230 says that
- "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider" (47 U.S.C. § 230)
- In other words, online intermediaries that host or republish speech are protected against a range of laws that might otherwise be used to hold them legally responsible for what others say and do.
- The protected intermediaries include not only regular Internet Service Providers (ISPs), but also a range of "interactive computer service providers," including basically any online service that publishes third-party content.
- CDA 230 creates a broad protection that has allowed innovation and free speech online to flourish, but has permitted sights like Google to continue to point to false and (in some cases) libelous information





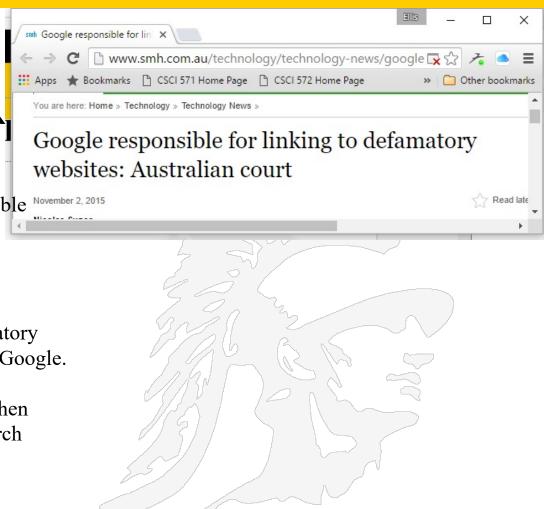


Some Recent Court Cases

Google IS Responsible

The South Australian Supreme Court this week found that Google is legally responsible when its search results link to defamatory content on the web.

In this long-running case, Dr Janice Duffy has been trying for more than six years to clear her name and remove links to defamatory material when people search for her using Google. The court found that once Google was alerted to the defamatory material, it was then under an obligation to act to censor its search results and prevent further harm to Dr Duffy's reputation









Some Recent Court Cases

Google IS NOT Responsible

Google has no obligation to "police (YouTube) content before it is put online as long as it informs users that posting television shows, music videos, concerts or advertisements without prior consent of the owner is not allowed."

So says a judgment by a court in France in the case of TF1, France's largest television network, against YouTube and parent company Google.









Some Recent Court Cases

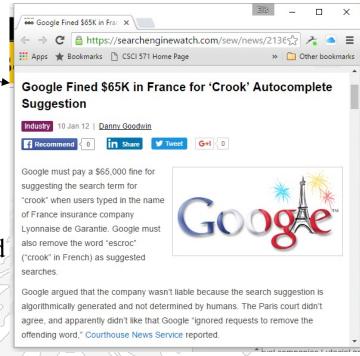
Google IS Responsible

remove the offending word,"

Google must pay a \$65,000 fine for suggesting the search term for "crook" when users typed in the name of French insurance company Lyonnaise de Garantie. Google must also remove the word "escroc" ("crook" in French) as suggested searches.

Google argued that the company wasn't liable because the search suggestion is algorithmically generated and not determined by humans.

The Paris court didn't agree, and apparently didn't like that Google "ignored requests to





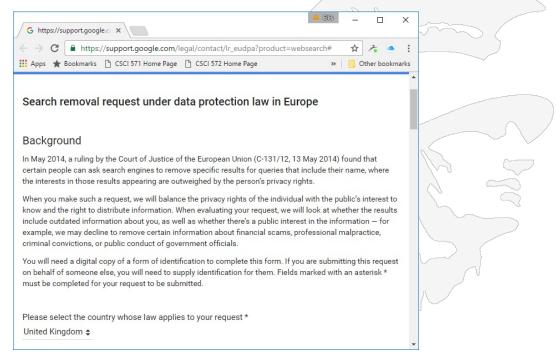




Google Required to Comply with EU's Right to be Forgotton

- In May 2014, the European Union found that certain people can ask search engines to remove specific results for queries that include their name.
- When you make such a request, Google will be looking at whether the results include outdated information about you, as well as whether there's a public interest in the information
- this website contains a form for content removal
- https://support.google.com/legal/contact/lr_eudpa?product=websearch#

this decision is limited to the European Union and does NOT currently apply to the United States







Google's Removal Strategy for US Customers

- It is not impossible to persuade Google to stop listing certain types of information in its search results,
- Google's Removal Policies page,
 https://support.google.com/websearch/answer/2744324
- Given a legal request Google will remove
 - material that violates copyright laws
 - Material that contains child sexual abuse imagery
 - Images that show bank or credit card numbers
 - personal signatures
 - Social Security numbers
 - revenge porn photos posted without the subject's consent.





European Union Proposes a Link Tax

- Article 11 of the proposed Directive for the Digital Single Market
 - Grants press publishers exclusive rights to control the reproduction and making available of online journalistic content by information society providers
- Google and others are making money by providing members of the public with free access to publisher's news, through links and snippets, without compensating the publishers
- When Spain passed such a law Google immediately shut down its Spanish Google News service (2014)
 - As a result, estimated traffic to Spanish news sites declined by 6-30%
 - http://csci572.com/papers/samuelson.pdf
- Spain in 2021 passed a new law allowing Google and publishers to arrange licensing fees for using their content
- https://www.engadget.com/google-news-spain-european-copyright-directive-eu-publishers-135511123.html







Google – Ads – Legal Suits



Google Responsible





Google Takes Ads From Companies Doing Illegal Things





Google's \$500 million forfeiture, which represents its revenue from displaying such ads

through AdWords, is one of the biggest in U.S. history.

According to the U.S. Attorney's Office for the District of Rhode Island which investigated Google's advertising practices in this case, the company was aware since 2003 that foreign pharmacies are advertising drugs to US consumers through its ad network.





pC Google Settles AdWords Laws C A Opaidcontent.org/article/419-google-settles-adwords-lawsuit-for-20-million/ CSCI 572 Home Page CSCI 571 Home Page CSCI 351 Home Page " Other bookmarks Google Settles AdWords Lawsuit For \$20 Million Comments (1) Text Size: A A ♣ Share 🖾 Email 📇 Print (PDF) The list of companies and By Tameka Kee individuals suing Google for Calcalle (a) Chic □ (a) CeekyChic the way it runs its search ad business is a long one: Apr 1, 2009 1:47 PM ET including everyone from biggies like American Airlines to one-off €0 advertiser lawsuits And settling its latest AdWords **≫** Tweet related lawsuit could ultimately cost Google (NSDQ: GOOG) up to \$20 million according to Mediapost. SEE ALSO: Google Sued Over Ad Rates, But Wins in Share +1 Separate, Street Views Court Challenge The class-action lawsuit was filed by two small business advertisers in 2005: Minnesota-based printing company CLRB Hanson Industries and N.J.-based Howard Stern (not the radio personality); they argued that Google had charged them more for ads on certain days than they'd agreed to pay for, Google countered that it had only overcharged them to make up for days when it under-delivered ads, but still hose to end the litigation with a settlement. Both plaintiffs vill receive \$20,000; Google agreed to pay their lawyers pre than \$5 million. Other affected advertisers will get Nords credits.

Google IS Responsible

Google's AdSense Displays Ads on Parked Sites and Charges Advertisers



The ads in question were those which appeared on "parked domains" which are registered but undeveloped websites, and on placeholder pages that appeared instead of error messages.

The plaintiffs claimed that these sites left a negative impression and that Google had failed to inform them their ads would appear there.





Can a Search Engine Sell Ads on Trademarked Terms?

- Some thorny examples:
- 1. In France, "Egg" is the name of a bank and "Orange" is a cell phone company. Do you ban anyone from bidding on those words?
- 2. What should be done when two or more people have a trademark on the same word, such as Apple Computers and Apple Corp, the Beatles' record label. Who trumps whom?
- 3. What should be done when someone wants to sell a product where the words used to describe it are a trademark. For example,
 - let's say you sell used cell phones.
 Are you prevented from using the word "Nokia?" Perhaps you can't bid on "Nokia," but can you bid on "used Nokia cell phones?"
- For the complete answer see Google's policy at
- https://support.google.com/adwordspolicy/answer/6118?hl=en

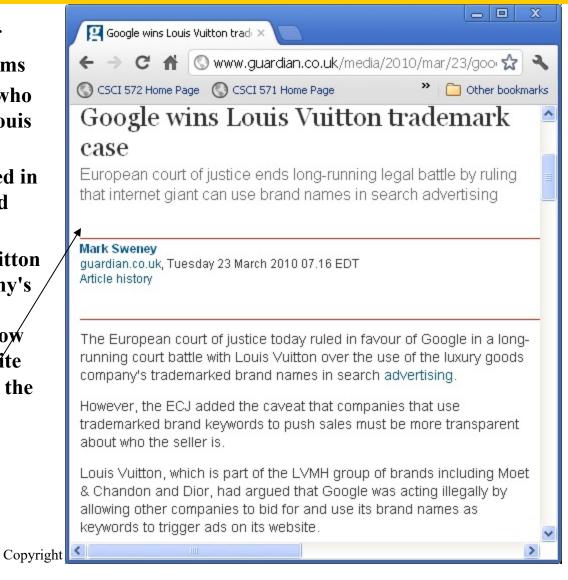




A Search Engine Can Sell Ads on Trademarked Terms

- Louis Vuitton is a famous maker of handbags, shoes, and other such items
- Google was sued by Louis Vuitton who said they could not sell ads using Louis Vuitton as the keywords.
- The European Court of Justice ruled in favor of Google selling ads on brand names
- In the Vuitton case, a search for vuitton at Google France shows the company's web site as the very first listing. It's hard to argue that Google is somehow preventing users from finding the site because it also carries ads linked to the Vuitton name.

Google IS NOT Responsible







Google Sued by European Union

- Three outstanding antitrust charges
- 1. April 2015 Google accused of favoring its own comparison-shopping services over others
 - Google's response is here https://blog.google/topics/public-policy/the-search-for-harm/
 - Google fined 2.4Billion Euros for promoting its own shopping sites
- 2. April 2016 Google is forcing smartphone manufacturers (that use Android) to pre-install Google apps
 - In 2018 Google fined \$5 billion by EU for Android antitrust
- 3. July 2016 Google is accused of preventing publishers using AdSense from displaying ads from other advertisers
- http://qz.com/798791/these-are-the-eus-reported-plans-to-break-up-googles-goog-monopoly-powers/
- 19 July 2018, EU fines Google €4.3 billion (about US\$5 billion) for illegal practices regarding Android
 - In 2021 Google lost its appeal and will have to pay
 - https://ec.europa.eu/commission/presscorner/detail/en/IP_18_4581





Oracle Sues Google Over Java

- At issue in Oracle v. Google is whether Oracle can claim a copyright on Java APIs and, if so, whether Google infringes these copyrights.
 - When it implemented the Android OS, Google wrote its own version of Java. But in order to allow developers to write their own programs for Android, Google's implementation used the same names, organization, and functionality as the Java APIs
- In May 2012, Judge William Alsup of the Northern District of California ruled that APIs are not subject to copyright.
- Oracle appealed Judge Alsup's ruling. On May 9, 2014, the Federal Circuit issued a decision reversing Judge Alsup and finding that the Java APIs are copyrightable, but leaving open the possibility that Google might have a fair use defense
- On October 6, 2014, Google asked the U.S. Supreme Court to review the Federal Circuit's decision. In June 2015 the Supreme Court denied Google's petition. The case will now return to the district court for a trial on Google's fair use defense
- On May 26, 2016 the jury decided in favor of Google declaring their use of the APIs to be "fair use".
- On March, 2018 Oracle appealed the decision and the judge decided in Oracle's favor, saying that Google's use of the APIs had NOT been fair use; The case was sent back to lower court to determine the damages
- On April 5, 2021 the Supreme Court of the US decided in Google's favor, see https://www.gibsondunn.com/supreme-court-holds-that-googles-use-of-oracles-java-software-interface-is-fair-use/

 Copyright Ellis Horowitz, 2011-2022

 37





US Dept of Justice Sues Google for Anti-Trust Violations

- Oct, 21, 2020. Justice files a 57 page complaint saying "Google has used anticompetitive tactics to maintain and extend its monopolies in the markets for general search services, search advertising, and general search text advertising the cornerstone of its empire."
- Joining the suit are attorneys general from Arkansas, Florida, Georgia, Indiana, Kentucky, Louisiana, Mississippi, Missouri, Montana, South Carolina, and Texas
 - The Governors of all of these states are Republicans
- You can download the actual complaint from here
 - http://csci572.com/papers/google_case.pdf







The Accusations

- Note: Simply having a monopoly in a given area is NOT a crime; what is a crime is when you use illegal means to create and/or maintain that monopoly
- 1. Google is dominant in search controlling over 80% of all search queries
 - Google has 88%, Bing 7%, Yahoo 4%, DuckDuckGo less that 2%, and
- 2. Google has suppressed competition by
 - Paying Apple Computer over \$8 billion/year to be the *default* search engine on iPhone's Safari browser as well as on iPads and iMacs
 - Requiring manufacturers of Android phones (LG, Motorola, Samsung) to include Google products (search engine, Gmail, Google Maps, etc) on their phones and to have Google search as the *default*; called *exclusionary* agreements
 - Paying browsers Firefox (Mozilla), Opera, UC Web (subsidiary of Alibaba) to make Google the *default* search engine; of course Google search is the default on Chrome
- 3. Google is dominant in digital advertising in part because of the data it amasses about users and their preferences; thus advertisers use this information to target their ads and other search engines cannot offer the same level of targeting

 Copyright Ellis Horowitz, 2011-2022





Google Defenses

- Google is not dominant in general search because many searches are done on Amazon, Facebook, eBay and others
- The deals Google makes with Apple are perfectly legal because any other company can make similar deals
- Users are free to switch to any other search engine, which are only "one-click" away
- As of now, the case is on-going