

Search engines and search neutrality: a primer



Dr. Jorge Padilla
Brussels, 16 March 2012

Agenda

- Search engines and algorithmic search
- Search bias
- Search neutrality and its limits
- Enforcing search neutrality

Search engines

- Search engines offer services for both Internet users and advertisers.
 - To Internet users, they offer free search services.
 - To advertisers, they sell advertising space (ad inventory) alongside their results pages through an advertising platform (which can be operated by the search engine itself or by a third-party vendor, referred to as an intermediary).
- A search engine is therefore a two-sided platform: it is more attractive to users if the search results are relevant, fast, and if the ads are well targeted; it is more attractive to advertisers the more users they gain access to and if their ads are clicked on by a large number of users.

Search algorithms

- The search algorithm is a key input into any search engine because it defines the results shown to users and thus establishes the usefulness of the engine in responding to queries. In this way, it determines the extent to which the engine can attract users. It also improves the value of the platform to advertisers, who value eyeballs.
- A search engine conducts three basic operations: (i) crawling the web, (ii) indexing the crawled content and (iii) matching indexed content to queries and then serving the results to a user.
- Search engines use proprietary technologies to crawl and index sites and proprietary algorithms to determine the relevance of the information on each crawled site for user's search query.

General purpose search platforms

- The main providers of general purpose search services are Google, Yahoo! and Microsoft's Bing, although Yahoo!'s core search business is being subsumed into Bing as the Microsoft/Yahoo! transaction is being implemented.
- Search engines attract users to their algorithmic search services by striving to provide highly relevant search results, but also by employing two types of exclusive agreements with third parties: syndication agreements and distribution agreements.
- Google is by far the leading general purpose search engine in terms of use.
 - Reach is defined as the proportion of Internet searchers that visit a particular web property at least once within a given month.

Vertical search engines

- Verticals are smaller or niche search engines that focus on a particular slice of content available in the web but which nonetheless compete with general search engines for users and search queries and, therefore, for a bigger share of the advertising pie.
- Vertical search sites offer users and advertisers more specialized and focused search services. They aim at addressing homogeneous users' needs and create value-added services from their knowledge of restricted users scope of expectations. Vertical search engines may focus on specific topics (e.g., travel, finance, healthcare, shopping comparison) and/or on specific media (text, images, news, videos, maps).

Search bias

[S]ince it is very difficult even for experts to evaluate search engines, search engine bias is particularly insidious ... For example, a search engine could add a small factor to search results from 'friendly' companies, and subtract a factor from results from competitors. This type of bias is very difficult to detect but could still have a significant effect on the market.

Sergey Brin and Larry Page, *The Anatomy of a Large-Scale Hyper-textual Web Search Engine* (1998)

Search bias (cont.)

- Several authors – e.g. Introna and Nissenbaum (2000) – have shown that the popularity of a website crucially depends on how it is treated by search engines.
- The literature –e.g. Thelwall (2008) – has also shown that most common search algorithms, like Google’s *PageRank*, are vulnerable to manipulation. For example, the *PageRank* algorithm induces a “rich-get-richer” phenomenon by creating a bias in favour of well-established web pages. See, e.g., Cho and Roy (2004).
- The implication of this finding is that even short-term manipulations of search results are likely to have substantial long-term implications on the allocation of internet traffic among competing websites.

Exploitative manipulation of organic search results

- Some recent papers have also shown that search engines may find it optimal to manipulate their organic search results in order to increase the profitability of their paid-search businesses. This is because “[O]rganic lists provide potential advertisers with prominence for free, which reduces their bidding incentive for sponsored exposure and thus hurts search engine revenue.” See Xu, Chen, and Whinston (2009).
- They can do so by (a) degrading the quality of their organic search results, (b) giving less prominence in the organic results part of the search results page to those advertisers who fail to bid high for sponsored links, and/or (c) giving more prominence in the organic results part of the search results page to those advertisers who pay more to be listed in the sponsored links.
- Algorithmic rankings appear to have a causal effect on both short-term and long-term traffic to websites even after controlling for their relevance. In other words, *caeteris paribus* users are more likely to click on a result at the top of the SERP than one that is ranked below.

Exploitative manipulation (cont.)

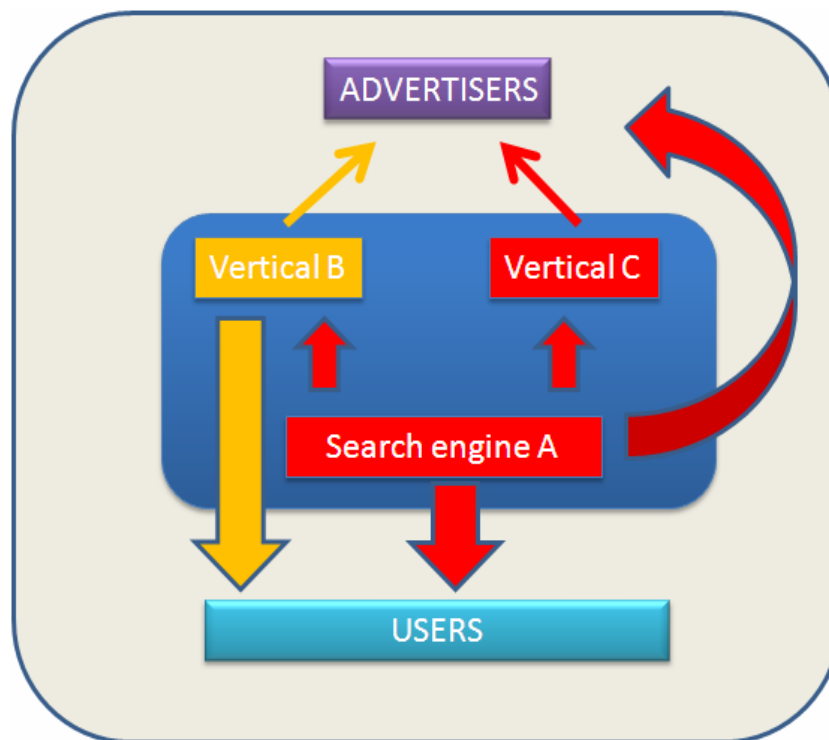
- Yang and Ghose (2010) have tested whether advertisers get more clicks when they appear on both sides of the page. Their empirical analysis employs a unique panel dataset of aggregate keyword-level data on clicks and conversions collected from a Fortune 500 firm that advertises on Google. They find a “positive and asymmetric interdependence between paid and organic listings” which “implies that the top ranking websites in organic search are likely to get a higher number of clicks in paid search as well”.
- Yang and Ghose conclude:
Because advertisers pay search engines on a per click basis, this has implications for search engines’ revenues. Indeed, there may be a moral hazard problem here as search engines may have an incentive to manipulate rankings in organic search and selectively present those firms on the top in organic search that experience higher click through rates in paid search.

Exploitative manipulation (cont.)

- White (2008), using a model of a monopoly search engine that provides both paid links and unpaid links, concludes that :

[the monopolist search engine has] an incentive, on the one hand, to offer results which are not as relevant as they possibly could be, even if improving them would come at no technical cost.

Naked exclusion



Consider the following market scenario:

- There is a general search engine A, which faces almost no competition from other general search engines.
- There is also company B which has a vertical search engine specialised in content of interest for competition policy practitioners and competition policy agencies.
- Search engine A has developed a vertical search engine C, which deals with the same content slice as search engine B.
- Verticals B and C fund their engines by selling ads. The absolute (and relative) appeal of verticals B and C for advertisers is a function of their user base.
- General search engine A is an indispensable “partner” for verticals B and C, at least at the early stages of their business life when few potential users are aware of their existence.
- In the absence of any discrimination by general search A in favour of its vertical C, verticals B and C should be expected to compete fiercely to appear at the top of the organic results on A’s SERP page.

Naked exclusion (cont.)

Input foreclosure

- A may have the incentive to foreclose vertical B in order to capture for itself all of the ad revenue that is generated by B. We denote this as the *monopolisation* motive for manipulation.
- By biasing its organic results against vertical B, A will ensure additional traffic to its own vertical C. Advertisers will respond by shifting their budgets from B to C. Through this strategy A denies both user and advertiser scale to the competitor of its own verticals. Since developing and maintaining a vertical search engine involves significant fixed costs, A's manipulation of its organic search results is likely to cause B's exit or, at very least, due to reduced advertising revenues necessary to improve its offering, its marginalisation.

Naked exclusion (cont.)

Input foreclosure

- This anticompetitive strategy involves no sacrifice for A, since it only entails a negligible implementation cost and A does not charge for the organic portion of its SERP.
- This is the very same reason why the “one-monopoly theorem” does not hold in this case: search engine A’s policy of not charging for organic placement implies that it cannot extract any of the rents that derive from its monopoly position in the supply of organic traffic to verticals.
- It could be argued that A’s actions do involve a sacrifice because they degrade the quality of its general search engine, which may have an impact on its usage. This possibility, though theoretically valid, is unlikely to make a difference in practice. This is because it only affects selected queries, allowing a search engine with a quality advantage to maintain its customer base and because users will not be able to detect any quality degradation in the short run.

Naked exclusion (cont.)

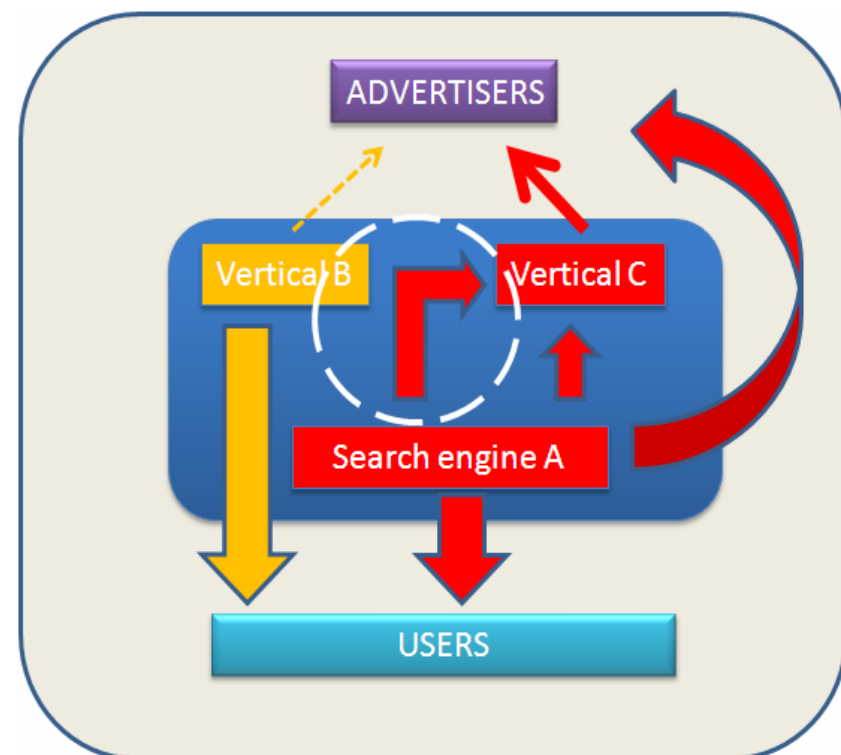
Defensive leveraging

- General reach engine A may also have the incentive to foreclose vertical B in order to protect its dominant position by eliminating the competitive threat posed by B, which could in collaboration with other verticals or general search engines develop a credible competing general search engine. This is the *defensive leveraging* motive for manipulation.
- Vertical B may find it very difficult to transform itself into a general search engine capable of competing with general search engine A, given the scale advantage of A. However, in combination with other vertical search engines and/or with an actual or potential general search engine competitor, it may be able to achieve the scale necessary to attract advertisers away from general search engine A or, at the very least, to steal some of the business that those advertisers patronise with A.

Naked exclusion (cont.)

Ability

- Search engine A can distort competition among verticals B and C by promoting one at the expense of the other in its SERP.
- It can do so by listing in the SERP on A the results from vertical C – its own vertical property – ahead of those of B as a matter of policy – i.e., without consideration to or in spite of their relative relevance for users.
- It may also interfere with its basic search algorithm, either algorithmically or manually, to “blacklist” vertical B from its SERP or, less radically, by lowering B’s ranking on A’s SERP below the ranking its relevance vis-à-vis a given query would justify.
- Furthermore, A may also make it more expensive for vertical B to bypass A’s algorithmic SERP by biasing its paid search algorithm against B (e.g., by algorithmically or manually reducing its relevance or by adjusting its quality score)



Naked exclusion (cont.)

Harm

- Consumers will be adversely affected when A acts on incentives to disadvantage vertical B by manipulating its search results to “starve” B of traffic.
 - The quality of general search engine A will suffer, as it will no longer render the most relevant results. While users may be unable to detect such degradation in the short term (indeed, they may never be fully cognizant of this), they still suffer the consequences (for instance, increased search costs, waste of time and effort, etc.) of degraded results that poorly match their needs.
 - In addition, Internet users will be restricted in their choice of verticals. They will be forced to rely on vertical C, which, in turn, will have fewer incentives to improve its offering by enriching its content, improving the presentation of that content or by providing new functionality.
- Advertisers will also be negatively affected by the exclusion or marginalisation of vertical B

Search neutrality

Strong search neutrality

- Search engines ought to employ “neutral” search algorithms that determine search result rankings based on some “objective” metric of relevance.

Weak search neutrality

- Search engines ought not to prefer their content on adjacent websites in search results but should instead employ “neutral” search algorithms that determine search result rankings based on some “objective” metric of relevance.
- WSN addresses the so-called own-content bias.

Why search neutrality?

Protecting users against search bias and manipulation

Protecting advertisers

- Search engines may have an incentive to bias their “algo” results in order to increase search advertising profits.

Protecting web publishers

- Search engines may have an incentive to bias their “algo” results in order to promote publishers that enter into profitable syndication agreements .

Protecting the competitive process

- Search engines may have an incentive to bias their “algo” results in order to discriminate against actual or potential competitors.

Limits to search neutrality

Search neutrality protects websites not users (Grimmelmann, 2011)

- “Scratch beneath the surface of search neutrality and you will find that it would protect not search users, but websites. In the search space, however, websites are as often users’ enemies as not; the whole point of search is to help users avoid the sites they don’t want to see.”

Governing bad behaviour (Evans, 2011)

- Search engines need to be able to punish bad behaviour in order to preserve the integrity of their platforms.
- Search engines, as all other multi-sided platforms, adopt governance rules to eliminate negative externalities.
- Exclusion is an effective punishment against misuse of the search platform.

Limits to search neutrality

The right to innovate (Crane, 2011)

- Search engines must have the freedom to make strategic choices about the design of their services ...
- ... Including the decision to blend proprietary functions traditionally performed by third-party websites in the engine's search properties.

Search bias is a normal feature of competition (Manne and Wright, 2011)

- Google's rivals also exercise editorial and algorithmic control over the information that is provided to consumers and in what manner.
- Vertical integration promotes competition and innovation

Enforcing search neutrality

Type I errors

- Evans – search bias may be justified in order to protect the integrity of the search platform.
- Crane – search bias is a form of product differentiation.
- Search bias is ubiquitous.
- Users can protect themselves as competition is one click away.

Type II errors

- There is no justification for bias aimed at excluding actual and potential competitors.
- The competitive process cannot discipline search platforms with a significant scale advantage.
- Distortions of competition in organic search are extremely costly due to spillover effects on paid search and on many other markets, including mobile telephony, tablets, PC software, etc.

Enforcing search neutrality (cont.)

Balancing type I and type II errors

- Weak search neutrality
 - Type II errors are likely and costly
 - Evans's right to protect your own platform's integrity is not endangered
 - Crane's freedom of innovation defence contains no limiting principle
 - Choice of remedies:
 - Unbundling remedies
 - Must carry remedies
- Strong search neutrality
 - Type I errors may become much more important
 - Some discrimination may be necessary to protect quality
 - Enforcement is bound to be complex

www.compasslexecon.com

