MANDATORY SHARING OF DATA AMONG ONLINE SEARCH ENGINES: NEED FOR PRO-COMPETITIVE INTERVENTIONS

INTRODUCTION

In today's digital World, the internet has become indispensable for information, communication, and commerce. The internet is the repository of a vast amount of information, which can easily be accessed using online search engines. This makes online search engines a central part of our daily lives. However, even after playing an integral role in all spheres of life, online search engines do not charge a penny for most of the services that they offer to the end users. This is made possible by network effects and the multi-sided nature of the market.[1] Online Search Engines typically collect data regarding search queries entered by the end users and behavioural data as to how they have interacted with the search results. This data is then used to train the algorithms to provide relevant search results and targeted advertisements.[2] The Online Search Engines earn revenue from the online advertisements on their platform. This provides dominant entities with a huge competitive advantage over other firms. A huge end-user base translates into huge amounts of user data and, consequently, better search results. It also translates into more demand among advertisers and better-targeted advertisements due to a better-trained algorithm.[3] To reduce the competitive advantages of the incumbents, the European Union (EU) passed the Digital Markets Act (DMA).[4] The Digital Markets Act provides for additional obligations and restrictions for enterprises that are designated as a Gatekeeper in any of the Core Platform Services ("CPS").[5] Online search engines have been included in the definition of CPS.[6] One of the obligations that only applies to Online Search Engines is the obligation to provide ranking, query, click and view data to any third-party undertaking under article 6(11) of the DMA.[7] The draft Digital Competition Bill ("DCB") in India is heavily influenced by the DMA and contains similar provisions on many of the issues.[8] However, the DCB does not contain the mandatory data sharing provision mentioned in article 6(11) of the DMA. Against this backdrop, it becomes necessary to investigate the need for a similar provision in India.

[1] Gunnar Niels, Helen Jenkins & James Kavanagh, *Economics for competition lawyers* (3rd Edition, OUP) [57-59]
[2] Online Platforms and Digital Advertising: Market Study Final report (Competition & Markets Authority, July, 2020) [43-46]
[3] ibid
[4] The Digital Markets Act, 2022
[5] The Digital Markets Act, 2022, arts 3, 5-7
[6] The Digital Markets Act, 2022, Art. 2(2)
[7] The Digital Markets Act, 2022, Art. 6(11)
[8] Pankhudi Khandenwal, "Tying, Self-Preferencing and the Digital Competition Bill: A Changing Landscape for Competition Intervention?" 19 IJLT 1

In this academic sphere, a lot of literature exists on the importance of data in the digital markets.[9] Some scholars have also argued that data is an essential facility and a refusal to provide the same can be considered as an abuse of dominance.[10] However, no literature exists on the necessity of a mandatory data sharing regime akin to Article 6(11) of the DMA in India. It is this gap that this paper seeks to fill. This paper deals with two research questions. The first question is whether a mandatory data sharing regime is necessary in the Online Search Engines Market in India. The second is whether mandatory data sharing is feasible considering possible conflicts with Privacy and Intellectual Property (IP) rights. In this paper, I shall firstly, argue that a mandatory sharing of Search queries, clicks, and view data ("Click and query") is necessary to protect competition in the market for Online Search Engines. Secondly, look at a possible way of enforcing such a provision and argue that it can be balanced with any privacy or IP rights concerns. I shall majorly rely upon Legal and Economic concepts as well as surveys conducted by different agencies across the globe to support my claims.

I. Online Search Engines: Need for data Sharing to break perpetual Monopoly

In this section, I shall firstly, provide a brief overview of the relevant markets, the players involved and the lack of competition in the markets; Secondly, discuss the importance of data for providing Online Search services; Thirdly, I shall highlight the loss of consumer welfare that the lack of competition is causing.

A. Market for Online Search Engines and Online Search Advertisements: More than a decade-long monopoly

As discussed in the introduction, Article 6 (11) only applies to Online Search Engines. Online Search Engines is further defined as: *"a digital service that allows users to input queries in order to perform searches of, in principle, all websites, or all websites in a particular language, on the basis of a query on any subject in the form of a keyword, voice request, phrase or other input, and returns results in any format in which information related to the requested content can be found"*.[11]

[9] Anush Ganesh, "Effective Remedies in Digital Market Abuse of Dominance cases" (2024) European Competition Journal 1; Alissa Johnson & Ors, "A Strategy for Challenging Anticompetitive Refusals to Deal in the Digital Economy (Department of Justice Conference, Yale University); Erik Hovenkamp, "The Antitrust Duty to Deal in the Age of Big Tech" (2022) 131(5) The Yale Law Journal 1385; Michael L. Katz, "Multisided Platforms, Big Data, and a Little Antitrust Policy" (2019) 54 Review of Industrial Organization 695

[10] Garima Gupta, "Antitrust Concerns in the Age of Data-Driven Economies: The Need to Revive the 'Essential Facilities Doctrine'" (2024) 45 Liverpool Law Review 209; Edouard Bruc, "Data as an essential facility under EU law: How to define the "Target" market and divert the data's pipeline" (2024) 15(2) European Competition Journal 177; Inge Graef, "Rethinking the Essential Facilities Doctrine for the EU Digital Economy" (TILEC Discussion Paper, 2019)

[11] Regulation (EU) 2019/1150 of the European Parliament and of the Council, Art. 2

The requirement of *"all websites"* in the above definition effectively means that the market only includes players that offer the services of a general web search rather than the services of specialized search services that are restricted to specific domains like hotels, restaurants that are offered by many other companies.[12] In effect, the market for Online Search Engines is similar to the relevant market of *"General Web Search"* defined by the CCI in the Google Ecosystem Case.[13] Therefore, for the purposes of this paper, reference to Online Search Engines shall mean the market for General Web Search. In the Google Ecosystem case, the CCI had traced the market share of different players in the market from 2009 to 2019. Google consistently had more than 95% of the market share while other players like Bing, Yahoo, etc, never had more than 3.5% of the market share in any of the years.[14] The trend has continued, and Google still has more than 95% market share in Online Search Engines in India.[15] In the European Union[16] and globally[17], the picture remains the same, with Google consistently having more than a 90% share for the past 10 years. I shall discuss the factors contributing to Google's high market share in the next sub-section.

Online Search Services and Online search advertising are two sides of the market for Online Search Engines. The revenue from the Online Search advertising is used to maintain and upgrade the Online Search Services.[18] Online Search advertising is, however, not the only source of revenue for Google. Google also generates revenue from Online display advertisements.[19] The CCI and other regulators have already differentiated between the two markets in their past decisions and reports.[20] For the purposes of this paper, I shall focus solely on the market for Online Search Advertisements. This is because the Online Search advertising market is open only to Online Search Engines. The players operating in the market for Online Search advertisements are the same as players operating in the Online Search Services market.[21] The Online Display advertisement market has other platforms like Meta, Amazon, etc.[22] Further, Google is present in the market for Online Display advertisement through YouTube, Google Maps, and other products

[12] In Re: Umar Javed & Ors. v. Google LLC & Anr (CCI, Case no. 39/2018) [205- 220]; Google Search (Shopping) [EC, case AT.39740] [5.2]
[13] Umar Javed (n 12) [205-206]
[14] Umar Javed (n 12) [222-224] [244]
[15] Market share of leading search engines in India from January 2018 to June 2024 (StatCounter, Statista)
[16] Market Share of Leading Search Engines in Europe from January 2015 to March 2024 (Statista)
[17] Market Share of Leading Search Engines worldwide from January 2015 to March 2025 (Statista)
[18] Michal S. Gal & Nival Elkin Koren, "Algorithmic Consumers" (2017) 30(2) Harvard Journal of Law & Technology 309
[19] CMA (n 2) [211-290] Prachi Kohli v. Whatsapp LLC (CCI, Case No. 05/2021) [224.3]
[20] CMA (n 2) [211 – 260], Prachi Kohli v. WhatsApp LLC (CCI, Case no. 05/2021) [109.2-110]
[21] CMA (n 2) [213]
[22] WhatsApp (n 19) [216]

that are not connected with the Online Search services.[23] Further, from the demand side, online search and display advertisements are not substitutable as they serve different purposes. Search advertisement is intent-based, that is, it is targeted towards customers who are already looking for the product. Online Display, on the other hand, is more suitable for raising brand awareness and venturing towards newer audiences.[24] Therefore, comparing the revenue generated by Google through online display advertisement with that of other Online Search Engines will not provide a correct assessment.

In the Google Ecosystem order, the CCI found that the revenue generated by Bing and Yahoo in the relevant market during the period of 2011 to 2018 was minuscule in comparison to Google.[25] This was because Google had more than 95% of the market share in the Market for General Web Search, which left no incentive for the advertisers to opt for Bing or Yahoo.[26] The condition of the market is expected to be similar since Google continues to hold more than 95% share in the market for general web search in India.[27] Therefore, Google's market share in the market for Online Display advertisement market should remain the same. The condition is not unique to India, as Google held more than 90% of search advertising revenue from 2010 to 2019 in the United Kingdom.[28] Therefore, Google enjoys a High Market Share not just in India but in developed markets like the United Kingdom as well. In the next sub-section, I shall look at the factors that have contributed to the continued high market share of Google and argue that access to some of the data collected by Google is necessary to build a product that can compete with the Google Online Search Engine.

B. Impossible to Compete: Access to Data a prerequisite for building a Competitive Online Search Engine

Online Search Engines collect aggregated click and query datasets. Click and query dataset is essentially information about the queries that the consumers have entered into the search box and the links that they have clicked from the results that were shown.[29] Click and query data is then used to train algorithms, which makes changes to the order of results shown (ranking).[30] The click

[23] Annual revenue of Alphabet from 2017 to 2024, by Segment (Statista) <https://www.statista.com/statistics/633651/alphabet-annual-global-revenue-by-segment/>
[24] CMA (n 2) [217 – 218]
[25] Umar Javed (n 12) [244]
[26] ibid [252]
[27] Statista (n 15)
[28] CMA (n 2) [224]
[29] 'How Does Google Determine Ranking Results - Google Search' (*How Search Works*) <https://www.google.com/intl/en_us/search/howsearchworks/how-search-works/ranking-results/> accessed 9 May 2025.
[30] ibid

and query data not only help in improving the quality of results shown but also help advertisers reach consumers who are more likely to avail their products or services. This click and query data can also be used to develop newer algorithm models that can suggest corrections to spellings.[31] This makes the Online Search Engine that has access to huge amounts of data highly superior in comparison to the ones that do not have access to the data. In a survey conducted by Duck-Duck-Go, consumers mentioned that relevance and quality of search results are the most important criteria for choosing a Search Engine.[32] Google's competitors like Duck-Duck Go and Bing have tried other ways of gaining consumers, like non-collection of personal data[33], providing rewards and incentives to consumers[34], respectively. However, all these strategies have failed, and DuckDuckGo and Bing could not effectively gain market share. This implies that a competitor will have to at least provide a similar standard of relevance in search results in order to compete with Google. The value of relevant search results seems to be so high that consumers are not responding to any other incentives provided by the competitors. However, as discussed earlier, the massive difference in market share and consequently the difference in the amount of data collected make it impossible to build a product of similar quality. In the next sub-section, I shall discuss some of the consumer harms that result from a lack of competition in this market.

C. The Cost of Online Search: Loss of Innovation and Consumer Welfare

The use of an Online Search Engine may come at zero cost to the consumers. However, the weak competition in the Online Search Advertising market is expected to significantly raise the cost of advertising for producers.[35] The higher cost of advertisement ultimately adds to the total cost of the producers and gets passed down to the consumers in the form of increased prices of goods and services. Consumers also face a lack of choice and are forced to accept whatever terms and conditions are put forward by the incumbent. In a more competitive market, we would expect competition for Consumer-friendly terms and conditions.

The highly concentrated nature of the market and the high entry barrier also result in a lack of innovation. As other players do not have the capability to make a product of similar quality, they lose the ability to innovate further. A lack of sufficient revenue also affects their ability to invest in R&D. Google also does not have any strong incentive to substantially upgrade or improve the

[31] CMA (n 2) [92]
[32] 'Search Preference Menu Immediately Increases Google Competitors' Market Share by 300-800%' (*Spread Privacy*, 30 October 2019) <https://spreadprivacy.com/search-engine-preference-menu/> accessed 14 May 2025.
[33] 'DuckDuckGo - Protection. Privacy. Peace of Mind.' (*DuckDuckGo*) <https://duckduckgo.com> accessed 14 May 2025.
[34] 'Bing Rewards | Microsoft Bing' <https://www.microsoft.com/en-gb/bing/features/rewards-features-tips/> accessed 14 May 2025.
[35] Robert S. Pindyck, Daniel L. Rubinfeld et. al, *Microeconomics* (Pearson, 9th Edition) [369-370]

quality of its product. As a result, the perpetual dominant position of Google has also translated into a lack of innovation.

II. Mandatory Data Sharing: Procedures and Conflicts

From the discussion in the previous section, it is clear that the unique nature of the Online Search Engines market is such that click and query data not only provides a competitive advantage but is a necessary raw material for the development of the product. Without access to this data, a competitor will not be able to develop a product of similar quality. This feature of the product is unique from most other markets. Digital markets are usually prone to high switching costs and entry barriers. However, the development of the product is usually not dependent on the sharing of a particular resource by a competitor.[36] This makes the nature of the product unique and ensures that a refusal to supply the same to the competitor is enough for Google to maintain dominance for years and decades. The only solution, therefore, is to mandate the sharing of click and query data that Google collects so that the development of similar or better products is not stopped. In this section, I shall firstly discuss ways of enforcing a mandatory data sharing regime; Secondly, address possible conflicts with privacy and Intellectual Property Rights.

A. Non-Personal Data Authority: Operationalizing the Non-Personal Data Governance Framework

After the DMA came into force, Google has complied with Article 6(11) by starting the European Search Dataset Licensing Program.[37] Under the program, Google provides access to a dataset of more than 1 billion queries in all countries across the European Economic Area.[38] This dataset includes click, view, query, and ranking data.[39] All the data shared is anonymized following a frequency threshold method as already provided in article 6(11). The licensing program, however, has faced criticism from competitors for the data anonymization policy followed by Google. DuckDuckGo has alleged that Google's anonymization technique reduces the benefits that can be derived from that data.[40] Therefore, a contractual licensing system may not be ideal for enforcing the data sharing obligation.

---

[36] For example, platforms for Hotel booking or shopping do not need a large amount of data, and they can gain at least make a product that is competitive. It is possible for a new entrant to make a similar product as MakeMyTrip, but the usefulness of the platform will not be substantially lower.

[37] Alphabet, "EU Digital Markets Act (EU DMA) Compliance Report" <www.transparencyreport.google.com/report-downloads?hl=en> accessed 9 May 2025

[38] ibid [185-187]

[39] 'Google European Search Dataset Licensing Program | Google Search Central | Support' (*Google for Developers*) <https://developers.google.com/search/help/about-search-data-program> accessed 14 May 2025.

[40] 'Roadblocks to Competition: Investigate Google's Non-Compliance with the EU's Digital Markets Act' (*Spread Privacy*, 20 November 2024) <https://spreadprivacy.com/investigate-google-dma/> accessed 14 May 2025.

A mandatory data sharing regime is not entirely novel in the Indian policy space. The Gopalakrishnan Committee had already recommended the same under the non-personal data governance framework.[41] The framework envisaged for the creation of a Non-Personal Data Authority, which will look at the policy and procedure for data sharing among entities.[42] This may provide for a better mechanism that will ensure that the data provided is useful, compliant, and free from risks like privacy. The Authority can also look at whether the price of the licence is reasonable. An excessively high price may defeat the purpose of the provision.

B. Data Sharing vs Privacy and IP: In Search of a Balance

A prominent argument against the adoption of a mandatory data sharing regime is a possible conflict with data privacy[43] and IP rights. In this sub-section, I shall deal with both arguments.

One of the possible reasons why the DCB has an obligation for Data Portability but not data sharing with third parties is compliance with the Digital Personal Data Protection Act, 2023 ("DPDPA").[44] The DPDPA requires consent from Data Principals before their personal data is processed or shared with any third party.[45] Data Portability does not require data to be shared with any third party, which is why compliance with DPDPA is not an issue.[46] For a mandatory data sharing provision akin to article 6(11) of DMA, data has to be shared with third parties. However, the search and query data is generally non-personal in nature and is not capable of identifying any person.[47] Further, personal data can be anonymised and identification of the data principal is not possible. So, DPDPA will not be a bar for the sharing of data, and privacy concerns do not exist as no personal data is being shared.

Another possible argument against adopting a provision similar to Article 6(11) could be the protection of IP rights and an adverse effect on the incentive to innovate. The Indian copyright law does not cover raw data or ideas.[48] It does not have a sui generis protection over databases, which is present in the EU.[49] Facts or raw data cannot be copyrightable. However, a compilation or database may enjoy copyright protection if the author has exercised their own skill and judgment

[41] Report by the Committee of Experts on Non-Personal Data Governance Framework (MEITY, 2020)
[42] ibid [20]
[43] Sophie Stalla-Bourdillon and Bárbara da Rosa Lazarotto, 'Search Queries and Anonymisation: How to Read Article 6(11) of the DMA and the GDPR Together?' [2024] European Law Blog <https://www.europeanlawblog.eu/pub/2uxr4anu/release/1> accessed 14 May 2025.
[44] Digital Personal Data Protection Act, 2023; Draft Digital Competition Bill, 2023
[45] Digital Personal Data Protection Act, 2023, s4(1)
[46] Draft Digital Competition Bill, 2023, s 12(3)
[47] CMA (n 2) [366]
[48] *Tech Plus Media Private Ltd. V. Jyoti Janda & Ors.* 2014 SCC OnLine Del 1819
[49] Directive 96/9/EC (EC, 1996)

and the database has sufficient originality.[50] A mere mechanical database containing the number of times a query has been searched or the number of times a link has been clicked will not fall under copyright protection. However, even if some form of copyright protection is established, the Online Search Engine is not being forced to give up the data for free. They can still charge a reasonable amount for licensing the search and query data. The benefits to consumers, innovation, and competition far outweigh the grant of absolute copyright protection that would result in a perpetual monopoly. Therefore, a reasonable restriction on copyright protection may be allowed by which a possible Systemically Significant Digital Enterprise (SSDE) in the Online Search Engines market may be forced to license its search and query data in exchange for a reasonable fee. This would not adversely affect the incentive to innovate, as the underlying algorithm or infrastructure has to be built by the competitor. It is only the raw data that the competitors will get.

III. Conclusion

The persistent dominance of Google in the Online Search Engine and Online Search Advertising market in India and abroad points towards the existence of structural barriers to entry. In this paper, I have shown how access to click and query data is so central to the development of Online Search Engines that without access to it, a similar product cannot be developed. As a result, the market suffers from stifled innovation, inflated advertising costs that are passed on to consumers, and a lack of meaningful consumer choice. This paper has argued that a mandatory data sharing regime akin to Article 6(11) of the EU's Digital Markets Act is necessary for India. I have also dealt with possible concerns related to privacy and IP rights. The privacy concerns are valid, and the proposed data sharing regime can adequately address them. Intellectual Property Rights violations should not arise as raw data is not copyrightable. However, even if some part of the database is copyrightable, the benefits of sharing the same far outweigh any possible negative effects of diluting the protection. Therefore, a data sharing regime similar to Article 6(11) needs to be introduced in the DCB to ensure that the Online Search Market can be made competitive.

[50] The Copyright Act, 1957 s2(o); *Emergent Genetics India Pvt. Ltd v. Shailendra Sharma* (47) PTC 494; *EBC v. D.B. Modak* (2008) 1 SCC 1