tripleC

On the Political Economy of Link-based Web Search

Deepak P*, James Steinhoff** and Stanley Simoes***

*Queen's University Belfast, Belfast, UK, deepaksp@acm.org

**University College Dublin, Dublin, Ireland, james.steinhoff@ucd.ie

***Queen's University Belfast, Belfast, UK, S.Simoes@qub.ac.uk

Abstract: Web search engines arguably form the most popular data-driven systems in contemporary society. They wield a considerable power by functioning as gatekeepers of the Web. Since the late 1990s, search engines have been dominated by the paradigm of link-based web search. In this paper, we critically analyse the Political Economy of the paradigm of link-based web search, drawing upon insights and methodologies from Critical Political Economy. We illustrate how link-based web search has led to phenomena that favour capital through long-term structural changes on the Web, and how it has led to accentuating unpaid digital labour and ecologically unsustainable practices, among several others. We show how contemporary observations on the degrading quality of link-based web search can be traced back to the internal contradictions with the paradigm, and how such socio-technical phenomena may lead to an eventual disutility of the link-based model. Our contribution is on enhancing the understanding of the Political Economy of link-based web search, and laying bare the phenomena at work, towards catalysing the search for alternative models of content organisation and search on the Web.

Keywords: web search, link-based web search, Political Economy, social relations

1. Introduction

The World Wide Web, conceived as a "universal linked information system" and "a place to be found for any information or reference which one felt was important" (Berners-Lee 1989), was heralded as an enabler for democratisation of information. On the eve of its thirtieth anniversary in 2019, its founder, Tim Berners-Lee, released an open letter (Berners-Lee 2019) outlining the "sources of dysfunction" affecting the Web, critiquing its evolution over the past three decades. Among those critical references is that it encompasses "system design that creates perverse incentives where user value is sacrificed". In this context, it is interesting to see that companies with a significant interest in the web have occupied four spots among eight in the list of trillion-dollar companies[1]. What led to the evolution of the Web from being an information seeking platform, to being what increasingly looks like a business platform? Towards understanding this, we start from the 'gatekeepers'[2] of the Web, the search engines (Germano and Sobbrio 2020), where most people start their Web journeys; search engines probably form the most popular information retrieval systems that have ever existed. In particular, we consider the Political Economy of the dominant paradigm of web search – link-based web search. We take a Critical Political Economy approach

---

[1] https://en.wikipedia.org/wiki/List_of_public_corporations_by_market_capitalization (accessed 21st October 2024)

[2] https://digital-markets-act.ec.europa.eu/index_en

towards understanding link-based web search, locating our work within Critical Algorithm Studies (Kitchin 2019).

We start with outlining the task of web search, illustrating the practical and technical challenges therein. From this vantage point, we briefly consider the first paradigm of modern web search, content-based search, and its various aspects and vulnerabilities. Against this historical backdrop, we look at its successor, the paradigm of link-based search, the dominant paradigm of web search since the early 2000s, the mainstream model in today's web. We look at the technology of link-based search and analyse the choice architecture embedded within it from technological, philosophical and political perspectives. We outline the variety of critical observations made about link-based search from a wide spectrum of scholarship and show that there is little analysis of its political economy. We motivate the need for an understanding of the Political Economy of link-based search that focuses on the totality of the socio-technical ecosystem and social relations embedded within it. As a step towards such a goal, we outline a theoretical framework that focuses on the variety of consequences of link-based search, of which the commodification of links is a key factor. We describe the various first order and second order consequences of the commodification of links within our framework, tracing how these relate to the variety of critical observations outlined earlier. We also consider how these effects synergize to reduce the utility of link-based search, and locate link-based search within the broader context of automation within Critical Political Economy. We further discuss directions towards conceptualising alternative models of organisation of web content and links, as well as search paradigms.

Our analyses of the Political Economy of link-based search lead us to conclude that there is an intrinsic alignment between libertarian political ethos and the socio-technical ecosystem of link-based search. Consequently, web search engines may be functioning as a substantive pathway for capital (as understood within Marxist Studies (Marx 1867)) to operate on the Web. We observe that the link-based model could be seen as offering a fundamental building block underpinning capital's advances on the Web. The intent of our work is to contribute to the scholarship in Critical Algorithm Studies and aid advance awareness of the politics of web search to levels where imagining responsible and fairer alternatives becomes possible.

Roadmap: Following this introductory section, Section 2 introduces Web search from a technical perspective while also discussing the historical context that involves content-based search. Section 3 discusses the model of link-based search detailing its technological aspects first followed by a detailed discussion critically analysing the link-based search model. In Section 4, we outline our conceptual framework characterizing the political economy of link-based search. Focusing on the centrality of the link commodity, we outline several direct and indirect consequences of the social relations embedded within the link-based search model. Section 5 discusses pathways forward against the backdrop of the changing technological landscapes of web search, followed by conclusions in Section 6.

2. Web Search: Conceptualisation, Socio-technical Challenges and Practice

We start with possibilities of conceptualising and technologically realising web search, and the various challenges and possibilities that exist therein. From this vantage point, we will also consider content-based search, the precursor of the link-based model.

2.1. Conceptualising Web Search

If we position web search as a gatekeeper of the Web, where the Web is envisioned to be 'a place to be found for any information or reference which one felt was important'

(Berners-Lee 1989), how should web search function? Towards conceptualising a possibility, we consider the notion of use-value from Critical Political Studies. For context, we may recollect that Marx distinguishes between two types of value in his discussion of the commodity. To quote from *Capital* (Marx 1867, 27), "the utility of a thing makes it a use-value". Elsewhere, Marx (1859, 6) says: "A use-value has value only in use, and is realized only in the process of consumption." Use-value, in Marxist literature, is contrasted with exchange-value, the worth of the thing expressed in monetary terms on the market. All commodities have use-value, but not all use-values are commodities.

Web search may be seen as a marketplace that brokers the matching between information needs expressed as web search queries and use-values that manifest as web pages. Use-values are obviously not absolute but relative to the need of a particular user; thus, each page may be regarded as having a use-value distribution over the space of needs. For example, a government page with immigration information may have a high use-value for a user seeking to apply for a visa, whereas it may not have any use-value for one who is searching for football scores of a recent match. When some web pages are returned by a search engine in response to a search query and the user clicks on one, the user can be regarded to have consumed the use-value represented by the web page that has been clicked on. Thus, the visibility of a web page through a search engine directly relates to realisation of use-values embedded within it.

The design of a search engine could reasonably be conceptualised as one that focuses on satisfying needs, which could be stated as "provide the best use-values for the needs expressed in the query". This would lead to a conceptually simple operation for need-based web search if needs were clearly specified, and the mapping from needs to use-values were straightforward. However, there are significant socio-technical impediments to realising such an operation. We outline some main considerations herein vis-a-vis practices in contemporary search technologies.

Query to Need Mapping: Going back to the early history of search engines, it has been observed that users tend to prefer short queries. Among the first quantitative studies on query logs (Silverstein et al. 1999) (for the AltaVista search engine) in 1999 indicates that user queries contain 2.35 words on an average. Similar statistics were observed for the Excite search engine in 2001 (Spink et al. 2001) where the average query was found to be 2.4 words. More recent studies have recorded an increase in average query length, with a 2009 study (Zhang, Jansen, and Spink 2009) reporting 2.9 and a 2012 study (Taghavi et al. 2012) recording 3.08. Yet, it is apparent that these remain significantly shorter than the size of verbal communications between humans while conveying information needs. The short query lengths pose a significant challenge to map queries to information needs that they relate to. While this has been a subject of research (e.g., Figueroa (2015)), there are potential intrinsic ambiguities. As authors imply in a recent paper (Haider, Rödl, and Joosse 2022), a query containing the names of two cities could correspond to a variety of information needs viz. identifying travel options, comparing cost-of-living or liveability, or other kinds of needs. The social preference for short search queries – which may arguably be accentuated with the prevalence of simple search interfaces promoted by search engines – predicates a deep technical challenge of mapping queries to needs. Note that it may be impossible to accomplish this mapping by any amount of sophisticated technology since the translation of the information need in the mind of the user to a representation in language may be intrinsically understood as a lossy process. As the noted philosopher Wittgenstein argues (Wittgenstein, Wright, and Anscombe 1964, 108), the fact that a language

needs to have certain properties makes it impossible to convey *everything* that is intended.

Estimating Use-value: The estimation of the use-value of a web page to a particular information need is another facet of socio-technical complexity. As pointed out earlier, the use-value of a web page needs to be assessed in relation to the information need. Use-value may be interpreted as correlating with the notion of *relevance*, one that the Cambridge dictionary defines[3] as: "the degree to which something is related or useful to what is happening or being talked about". The assessment of relevance is confounded by the fact that the query-to-need mapping is unlikely to be done unambiguously. This potentially means that relevance may need to be assessed over a distribution of needs (the distribution that the query corresponds to) rather than a sole need. Further, there exist many candidate notions of relevance, as outlined recently (Sundin, Lewandowski, and Haider 2022). While systemic relevance focuses on the match between the query and the web page, user relevance focuses on how well the individual users may subjectively evaluate the web page (a notion that is significantly more complex and nuanced). A third notion of societal relevance may require that web page relevance be also conditioned on beneficence to the society, thus excluding potentially pernicious web pages (e.g., sexist and racist ones (Noble 2018)). Of the several alternatives available, contemporary search engines arguably use only a narrow definition of relevance. Most search engines use an interpretation that stretches the notion of user relevance towards maximizing engagement; notably, engagement could be interpreted as a query-agnostic notion. For example, prior research from Microsoft (Song, Shi, and Fu 2013) has explored user engagement retention in the face of relevance drop (such as in data voids (Mager, Norocel, and Rogers 2023)).

Presenting Search Results: Search engines have, since early days, uniformly adopted the one-dimensional presentation model of search results as a list. This is typically accomplished as an ordering of search results in the descending order of relevance, arguably optimized for the user propensity to scan results from top to bottom (Chierichetti, Kumar, and Raghavan 2011). This contrasts with presentation of e-commerce and image search results in a two-dimensional matrix form. User perusal of results have been found to follow an F-shape pattern (Nielsen 2006), which may, in simple terms, be understood as the attention being more dense within a triangle with its corners in the top-left, top-right and bottom-right of the display (the golden triangle (Chierichetti, Kumar, and Raghavan 2011)). The rapid drop in the quantum of attention as one moves down the list, called position bias, has led to a practical difficulty in ensuring that user attention is apportioned to web pages based on their relevance, resulting in a stream of work on addressing the disparity between user attention and relevance (Zehlike, Yang, and Stoyanovich 2022).

In sum, the realisation of a relatively simple conceptualisation of web search poses several socio-technical challenges. We also saw how need-based web search can be operationalized as a set of key technical tasks, (i) a mapping from queries to needs, and (ii) an estimation of need-based use-values. The socio-technical impediments, engendered by a mix of human cognitive preferences and technical complexity, make it challenging to realize such an operation.

2.2. Content-based Web Search

While the focus of this paper is link-based web search, we glance quickly at content-based web search, its precursor, for a historical context. Early web search engines

---

[3] https://dictionary.cambridge.org/us/dictionary/english/relevance

such as AltaVista[4], the major pre-Google search engine launched in 1995, prided primarily on their comprehensiveness of coverage and expressivity of search[5]. While search algorithms tend to be proprietary, their workings can be inferred based on the nature of popular search engine optimization (SEO) tactics targeting them. The predominance of keyword stuffing and hidden text as major SEO strategies of the 1990s[6] indicate that pre-Google search engines such as AltaVista and Yahoo! were largely based on text/content-based search. The content-based paradigm would determine the inclusion of web pages in search results based on the extent of the match between keywords in the query and those in the web page. While the precise technical details may vary – some keywords may be considered as more important than other, or other variants – a web page would have a higher chance of figuring in results associated with search queries built out of words that (are similar to those that) appear in its content. This paradigm notably short-circuits the detailed identification of the need and yields a technologically simple operation.

While content-based search may have been a reasonable approximation of need based search, the emergence of SEO as a dominant force alters the landscape. With the pervasiveness of the Web and the pre-eminence of search engines as an arbiter of user attention on the Web, the determination of which products, services and messages get noticed is increasingly taken up by search engines (Lewandowski, Haider, and Sundin 2024). The correlation between visibility and ad revenue is also a factor at play; these have become more institutionalised in recent years (Davis and Iwanow 2009). This creates incentives for webmasters[7] to operate in ways that align with search engine algorithms. In attempting to enhance the visibility of their website over content-based search engines, they would need to use abundance of keywords aligned with user queries for which the web page would be relevant, to ensure visibility. Such practices orient the thinking towards considering web pages as exchange values (for they fetch visibility that can translate to ad revenue) as opposed to being viewed as use values.

2.3. Vulnerability to Goodhart Effects

It is well understood in literature that "any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes" (Goodhart 1975). This is often referred to as the Goodhart's law (Goodhart 1984). In the case of content-based search, the usage of content match as a proxy for relevance functions as a perverse incentive for web page creators to optimise for content match leading to a divergence between content match and actual relevance. Such effects are also variously referred to as Campbell's law (Campbell 1979) and Cobra effect (Siebert 2001). The latter refers to anecdotal experiences in colonial India where bounties were offered for dead cobras in order to stem the menace caused by the snake. As the bounty got entrenched in society, people began breeding cobras for the income, leading to an increase in the menace rather than a decrease. Content-based search, due to having a simple structure, could be gamed by individual webmasters using methods such as keyword stuffing as seen earlier. The emergence of SEO services at the time also indicates the entrenchment of such perverse incentives. In particular, this is an

---

[4] https://en.wikipedia.org/wiki/AltaVista (accessed 21st October 2024).
[5] https://medium.com/syssoc-fms/altavista-the-tale-of-the-first-mover-who-failed-3e008727440f
[6] https://www.techradar.com/news/black-hat-seo-7-tactics-to-leave-back-in-the-90s
[7] Webmasters are the creators/managers of web pages.

instance of adversarial Goodhart within the four categories of Goodhart effects (Manheim and Garrabrant 2018). While content-based search may be regarded as particularly vulnerable to Goodhart effects due to the effectiveness of simple adversarial actions by individual webmasters, Goodhart effects would remain an aspect of concern for any technical approximation of the intended model of web search through a proxy criterion. It has been pointed out that adversarial Goodhart effects are very difficult to avoid (Garrabrant 2017). The obvious way of keeping the proxy secret, an impractical option within the context of web search, would not scale to an intelligent plurality of adversaries. The alternative, as suggested in Garrabrant (2017), is to choose a proxy that is *conceptually simple and hard to optimise for*. For example, it could be to choose a proxy such that adversaries have minimal control over the world over which optimization for it would need to function.

3. Link-based Web Search

We now look at the model of link-based web search, first considering the technological aspects, followed by a critical analysis of the model.

3.1. The Technology of Link-based Search

The paradigm shift in web search pioneered by Google, in the late 1990s, was to move from keyword search to predominantly link-based relevance estimation (Brin and Page 1998). Thus, it is useful to use Google's famed PageRank algorithm as typifying the paradigm. The philosophy has been spelt out by Google[8] as: "*PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links ...*". PageRank associates every page with an *importance score*, one that is directly related to both the *number* and *importance* of links *towards* itself, often called backlinks. PageRank embeds a circularity within its formulation (Franceschet 2011). For example, two pages linking to one another would have their importance scores as dependent on each other. Such circular dependencies may also appear through a third page, or a longer chain of intermediate web pages. While this engenders technical difficulties in estimation, extensive research into PageRank over the years has led to very efficient estimation algorithms (Chung 2014).

The design of the link-based web page importance score is fully determined by number and importance of inward links to it. This is, notably, a query-agnostic estimation, independent of what query is being addressed by the web search process. Under link-based search, web pages are retrieved based on both: (i) their link-based importance score, and (ii) content-based match to the query. Thus, the link-based search paradigm, while dominated by the link-based estimation of importance scores, still has a content-based matching process within it.

This primacy of links in web search is pervasive in contemporary search services, and now pervades virtually all top search engines. For example, Bing has been observed to rely more on content relative to Google, but is still substantively dependent on leveraging links to assess page importance[9]. The shift of focus from only content to predominantly links in the late 1990s has not just endured but has virtually become the only existing paradigm of web search.

---

[8] https://bit.ly/how-google-search-works-archive
[9] https://bit.ly/rank-website-on-bing

3.1.1. Changing Contours of Link-based Search

While virtually all contemporary web search services are still underpinned by the link-based search model, the technical operational model of search engines have hardly remained static. We briefly glance over some highlights of the changing contours of link-based search.

Consider a user who, upon encountering the results for the search query they entered, goes ahead and clicks on the third result without any reluctance. This could be interpreted as an indication that the third result is more relevant than the first and the second within the context of their search. Such data, called click-through data, are often exploited in contemporary search engines to create virtual links between users or their search queries on the one side, and web pages that get clicked preferentially on the other. Such models extend the remit of link-based importance computation to beyond just links between web pages. Such assetisation of user-engagement data (Birch, Cochrane, and Ward 2021) including click-rates and length of stay has been noted to be dominant within Big Tech; similar phenomena have been referred to as behavioural commodification (Kang and McAllister 2011). It has been observed that Google has updated PageRank to involve click-through rate in the web page ranking process[10]. User behavioural data have been integrated, not just within the search operation, but also at the query-entry stage, through mechanisms such as auto-complete which have become pervasive, further fomenting platform power (Mager, Norocel, and Rogers 2023). As noted in Rieder (2022), the collection of user data has been observed to be a central element to feed expansionary goals of Big Tech players. In addition to employing link-based logics to other forms of links, the content around the actual links[11] on the web page have also been exploited within search and ranking.

Consequently, extending the remit of the link-based idea to link text, click data and other forms of user behaviour data can be seen as a fallout of the restlessness and expansionary nature of capital underpinning the operation of such search technologies. While we will restrict the remit of our analyses to the usage of links between web pages for the purposes of this paper, the changing contours of web search mean that such analyses could be extended.

3.2. Critically Analysing Link-based Search

We now consider link-based search from multiple critical perspectives.

3.2.1. Design Choices

If we look back at link-based search from the conceptualisation in Section 2.1, we may make some observations. First, much like the content-based search case, there is a short-circuiting of the *need* identification/matching step embedded within the content-based matching process that is implemented within link-based search engines. Second, the presence of the *importance* score in link-based search operationalises a splitting of the overall task of query/user relevance to two parts viz., a query-agnostic link-based importance score, and a query-based content-matching score. Such splitting is technologically attractive, since they enable the application of the general-purpose filter-and-refine strategy (Wood 2008) to speed up the search process. Under the filter-refine strategy, one of the criteria could be used to filter away many web pages upfront, so the second criterion needs to operate on a small set of web pages; restricting the

---

[10] https://marketbrew.ai/the-evolution-of-the-pagerank-algorithm
[11] https://www.seroundtable.com/google-context-around-links-30173.html

second phase to work on a small subset enables fast computational processing. While link-based search uses filtering based on link-based importance, it is notable there have been tremendous advances in indexing of web pages based on content (e.g., Guo et al. (2022)), to enable fast content-based filtering; thus, link-based filtering is a conscious design choice and not the only available technological choice.

In particular, there has been no suggestion that the dichotomous conceptualisation of relevance is in accordance with or motivated by any theory of human assessment of relevance, such as used in structuring automated decision making elsewhere (e.g., Gestalt theory for web page segmentation (Xu and Miller 2016), the psychologically motivated recency heuristic for trend prediction (Gigerenzer 2023)). Thus, the two-part conceptualisation is probably motivated by technological convenience than by anything else.

3.2.2. Historical Context, Philosophy and Ideology

This link-based paradigm has been widely noted to have been motivated by the information sciences discipline of bibliometric citation analysis (Lewandowski, Haider, and Sundin 2024). Considering the number of citations received by a paper as a measure of its impact can be traced back to 1927 (Gross and Gross 1927); it rests on a simple assumption that high quality scientific work will attract more citations. It has been argued that citations tend to be *unobtrusive, non-reactive* and notably *do not require the cooperation of a respondent* (Smith 1981). To quote from Van Couvering (2004), "*When Google launched in 1997, it was said to be unspammable ...*". One could potentially interpret this as potential robustness to Goodhart-like effects. However, for scientific citations, in an empirical study, it was found that non-scientific factors play a role in deciding to cite a work, even though the authors stop short of questioning the value of citation analysis (Bornmann and Daniel 2008). While citation analysis has been subject to much critique, citation analysis (and metrics based on it) is notably widely used within contemporary society to inform important decisions (McKiernan, Alperin, and Fleerackers 2019), and we find it reasonable to assert that they are perceived in scientific circles as the most robust among available metrics of scientific achievement. Link-based importance, however, notably differs from citation analysis in one aspect in that it is not just the number of inward links that determine the importance of a web page, but also the importance of the pages that link. It has been observed (Mayer 2009) that this weighted formulation can be substantively traced back to sociometric work from 1953 (Katz 1953) where the intent was to develop a *status index* for social groups, a task with obvious political valence. It is interesting to note that there has been recent research into quantifying the importance of citations (Chakraborty and Narayanam 2016), a formulation that resonates with the link-importance structure of PageRank.

The link-based paradigm, as pointed out critically in Carr (2009), is a manifestation of the idea that each link represents a small bit of human intelligence, that of recommendation/endorsement (first assumption of hyperlink analysis (Henzinger 2001)) and that its aggregation would create immense value. Within the broader backdrop of the politics of search engines, the focus on links creates new forms of exclusion, where web pages not having sufficient backlinks could be excluded from search engines due to not being in the crawl path (Introna and Nissenbaum 2000) (Ref: Table 1 in Introna and Nissenbaum (2000)). In a sharp critique of Google's PageRank, Pasquinelli's (2009) work interprets PageRank as a model of cognitive capitalism. This is so since PageRank, by using links to make judgements of the value of web pages, is appropriating cognitive labour into network value, and thus functions as a global rentier of the

common intellect. The operation of link-based search has also been found to be aligned with capitalism in corresponding to a form of exploitation enabled by the connexionist world (Mager 2012), dubbed to be part of the new spirit of capitalism (Boltanski and Chiapello 2005). These could also be read within the context of Google being observed to be facilitating exploitative and unpaid digital labour (Fuchs 2014).

3.2.3. Myriad Critical Observations on Link-based Search

Ever since the inception of the link-based search paradigm, there have been critical observations pointing out systemic issues. We summarise a representative sample of some such observations herein. First, the **bias of web search to commercial sites** was noted, perhaps for the first time, in Van Couvering (2004), based on several Computer Science Studies at the time. The study speculates on the possibility of link-based search being the determinant, but concludes noting that "*we remain as yet unclear as to why some sites are favoured over others*". On similar lines, it was observed, in 2013, that "Google Search Favors Big Brands over Small Businesses"[12]. The article puts the responsibility of this behaviour to the construction of PageRank in the context of reasons that are characterised as being one of stronger backlink profiles, durable websites and better domain names. Second, **myriad pernicious biases** have been noted as embedded within search results. These were brought into much popular attention through the work on algorithms of oppression (Noble 2018, Rahman 2020) that focuses on the racism in search results. There have been several instances where major search engines, especially Google search, have been criticised for implicitly advancing racist, sexist and otherwise offensive values that manifest as biases. There have also been extant observations of *extreme-right sources appearing in top result positions in web search* (Norocel and Lewandowski 2023). It has been alleged that Google has often adopted a *whack-a-mole* approach where it reacts only when criticism is popular enough to constitute a problem for the brand (Sundin, Lewandowski, and Haider 2022). Fairness in retrieval has evolved to an active research field (Zehlike, Yang, and Stoyanovich 2021) and seeks to address the problem of biased search results through involving technological constraints; yet, as often argued about technological efforts on AI ethics (Katz 2020), the research often focuses on using constraints and similar approaches towards ensuring diversity in the output while not being enthusiastic about identifying or addressing how unfairness comes about in the first place. Third, **algorithmically embodied emissions**, as introduced in Haider, Rödl, and Joosse (2022) outlines how algorithmic information processing systems, particularly search engines, routinely make choices that "*contribute to the climate crisis and other forms of environmental destruction*". We use an example from the authors' blog post[13] to illustrate the point succinctly. A web search query '*summer clothes*' could be the manifestation of multiple possible user intents, such as understanding contemporary or historical summer clothing trends, buying new summer clothes, finding clothing that produce a cooling effect, or buying pre-owned summer clothes from a second-hand shop. Haider, Rödl, and Joosse (2022) point out that search engines consistently favour very high-carbon interpretations of search queries. In fact, an appraisal of the search results for the query may consistently point to an interpretation of *buying branded summer clothes*, an extremely high emission interpretation of the query. This trend, as the authors point out, is consistent across various types of queries; for example, a query '*paris stockholm'* is interpreted as a search for flight options between the cities with

---

[12] https://www.sbwebcenter.com/google-search-favors-big-brands-over-small-businesses/
[13] https://bit.ly/algorithmically-embodied-emissions

such results filling up most slots in the first page of search results. Fourth, there have been general observations on the **reduced utility** of search engines such as Google. A recent article from a popular blog on search engines was titled '*Google search is dying*'[14]; a relevant excerpt says "*... Google search results have gone to shit. You would have already noticed that the first few non-ad results are SEO optimized sites filled with affiliate links and ads*". The article places significant emphasis on SEO actions. Another article, titled '*The Tragedy of Google Search*'[15] talks about Google "*creaking under the weight of its enormous size*" and that "*Google Search is now bloated and overmonetised*". It is notable that these arguments often externalise the reasons for the failures; in other words, they sound like link-based search may have continued well if not for 'extrinsic' factors such as SEO actions, enormous size, and overmonetisation.

3.2.4. Towards an Understanding of the Political Economy of Link-based Search

The critical observations on link-based search, such as those noted above, often tend to be viewed as separate from one another. Well-intended and valid solutions are often proposed, such as the suggestion of extending environmental risk assessments to include algorithmic harm (Haider, Rödl, and Joosse 2022), giving more weighting to source credibility (Sundin, Lewandowski, and Haider 2022), encouraging small businesses to invest in SEO to compete better with big ones[16], exhorting users to specialize queries to include source names (e.g., using the 'site:' operator[17]) and promoting algorithmic capacity for search result diversification[18]. The above strategies differ on the extent to which they are technocratic, qualitative or quantitative. While they may all be judged to be important in their own ways, we have not noted any substantive suggestion on how these myriad phenomena are underpinned by the centrality of links within the search paradigm. We posit that it would be beneficial to develop an understanding of how the myriad issues of web search are systemic issues that have their roots in the *totality of the socio-technical ecosystem and social relations embedded in the paradigm of link-based web search*. This would help understand the core phenomena of which the observed issues are myriad fallouts and manifestations. This requires developing an understanding of the Political Economy of link-based search. The nature of the design choices, historical context, philosophy and ideology, all play vital roles in understanding the Political Economy of link-based search.
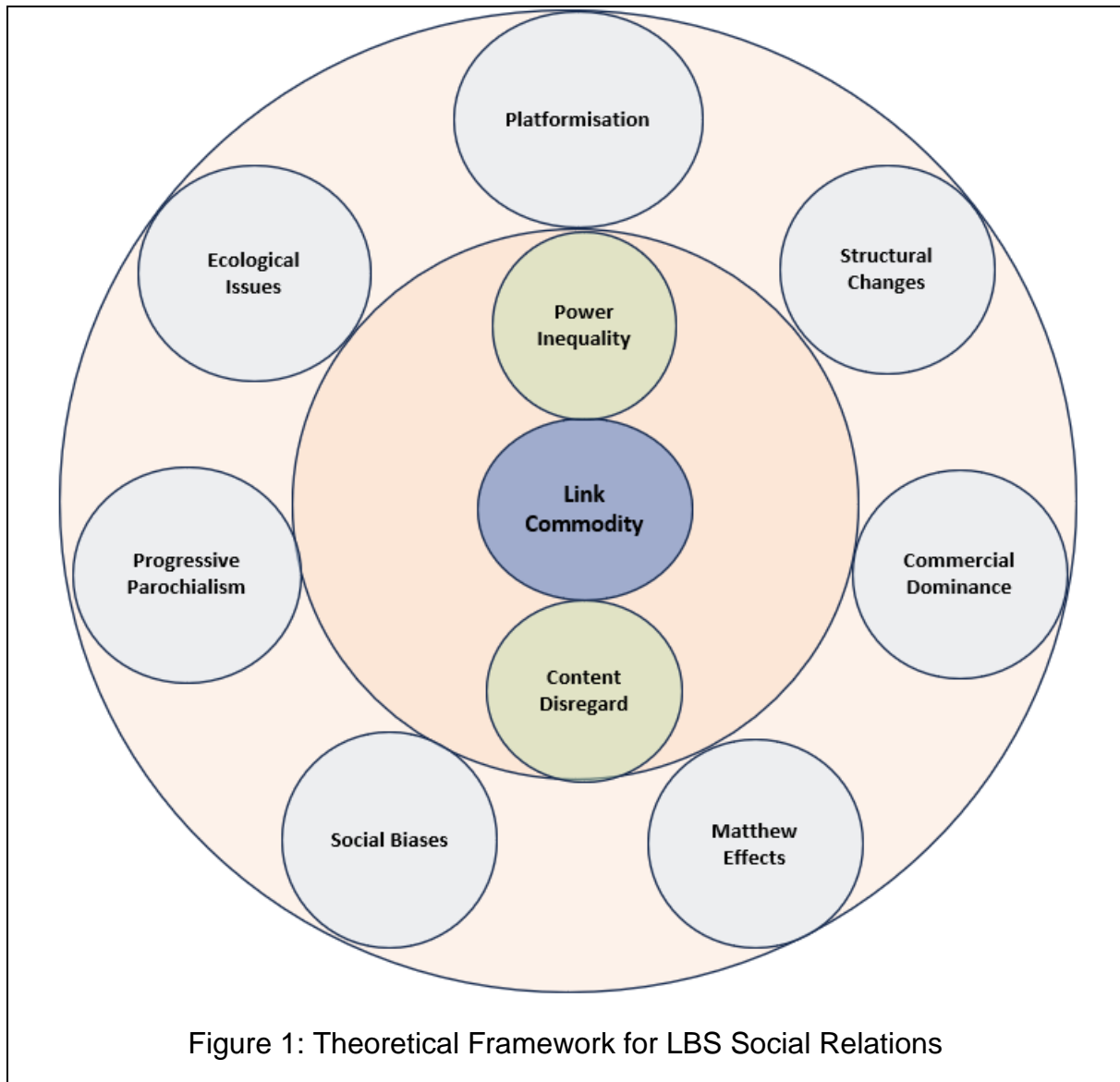
---

[14] https://dkb.blog/p/google-search-is-dying
[15] https://bit.ly/tragedy-of-google-search
[16] https://www.sbwebcenter.com/google-search-favors-big-brands-over-small-businesses/
[17] https://dkb.blog/p/google-search-is-dying
[18] https://fair-trec.github.io/

Figure 1: Theoretical Framework for LBS Social Relations

4. The Political Economy of Link-based Web Search

We now delve into a few steps towards the aforementioned ambitious goal of developing an understanding of the Political Economy of link-based web search. We outline our conceptualisation of the Political Economy of link-based search in Figure 1. The social relations of link-based search (abbreviated LBS hereon) are the focus of our theoretical framework. At the centre of our framework is the commodification of links within LBS. Our framework is focused on understanding how the commodity form of links relates to the myriad kinds of social relations fostered within LBS. We partition the LBS social relations into *first order consequences* (inner circle in Figure 1) and *second order consequences* (outer circle). In what follows, we trace the pathways from LBS to the first order consequences and on to the second order ones. We note that there are synergistic, complementary and other kinds of relations across the various consequences we consider in our theoretical framework; thus, our structuring of consequences as first and second order are to be seen as an effort towards narrative clarity rather than one towards a holistic representation.

Marxist political economy, as a critique of capitalist production, is centred on the commodity, which Marx asserts is a "unity formed of use-value and value" (Marx 1867, 293). The production and circulation of commodities on the market is the means by which capital extracts surplus-value from exploited labour. Any useful object can have use-value, but only a commodity also has value: a portion of "congealed quantities of homogenous human labour" (Marx 1867, 128). Value in this sense is added to a use-value by its production within the specific exploitative social relations of capital wherein generalized commodity exchange—including that of labour-power—is presupposed. In other words, value is a social relation which is grafted onto a particular use-value by capital; the "objective character" of value is "purely social" (Marx 1867 138-139). The commodity is fundamental to Marx's analysis because its social form is the structuring principle of capital and because of this, it acts also to structure the "social totality" (Heinrich 2007) of a society which operates a capitalist mode of production. Over time capitalist relations of production and exchange "dominate the social fabric and pro-gressively come to penetrate to its very ground" (Sohn-Rethel 2019) rendering increas-ing aspects of these relations amenable to capital. Marx aimed to grasp this phenom-enon by referring to the commodity as having the status of "fetish" (Marx 1867, 165). Like a religious artefact endowed with causal efficacy, the commodity, within a system of generalised commodity production, appears to take on autonomous functions, alt-hough in reality it is propelled by the social relations of humans. Commodities thus appear to have a particular value as an inherent property, but this is in fact, Marx holds, an expression of the social relations of production.

By converting social relations into what appear to be properties of commodities, capital gives to commodities the appearance of autonomy and in so doing generates societies which are built around the maintenance and enhancement of "self-moving economic forces that assert themselves behind the backs of the acting subjects, indif-ferent and indeed hostile to their needs" (Bonefeld 2014, 21–22). Social relations and the technologies and institutions that are produced and exist within them then come to be structured in accord with the exigencies of valorisation, a process called "economic form determination" (Kjøsen 2013, 19). Given the primacy of the commodity form within Marxist critiques of political economy, we start our analysis of LBS social relations with the commodity form of most interest in LBS, the link commodity.

4.1. The Link Commodity

We posit that the commodification of links is a key aspect of LBS, and that the com-modity form of the link has shaped the social relations of internet search. We had ob-served earlier, within the context of content-based search, that the presence of web search as an arbiter of user attention creates incentives for webmasters to undertake actions to enhance their visibility. However, in contrast to content-based search where such actions take the form of individual-level content changes that webmasters make over the web pages they own, the dominance of links in LBS requires that pathways to enhance inward links be sought. Unlike content on the web page, inward links to the web page from important[19] web pages *cannot* be created by the webmaster the page in question. Instead, they ought to be created by the webmasters of important web pages. This technical peculiarity makes possible the imposition of the commodity form on links within LBS.

Under LBS, if a webmaster wishes to improve visibility, they need to *publicise* their web page, in perhaps a targeted way to gather eyeballs from webmasters of 'important'

---

[19] Important in accordance to the importance score.

web pages, in order to ensure some inward links to their page, and thus, a high-enough importance score. Given the prime importance of this activity called *link building* in ensuring visibility, guides for link building (such as those from Ahrefs[20]) suggest at least two main options for webmasters: (i) *ask for a link*, or (ii) *purchase a link*. The first option aligns with *marketing*, a tool within a market society, of which, advertising is a popular form. The second option of purchasing links entails the existence of a *marketplace of links*; a visibility-focused market in which links from an external website to that of a buyer can be purchased. While the trading of link commodities between webmasters is, for the most part, opaque to end users, it is visible in now-pervasive practices such as offering incentives for reviews[21]. Review links are said to account significantly for local search in Google[22]; at a conceptual level, this could be understood as Google creating web pages for users and businesses, making each review a link from the former to the latter, enabling the application of LBS logics. The emergence of the link commodity is a substantial aspect of LBS and shapes the social relations within the LBS ecosystem.

4.2. First Order Consequences

We now consider the first order consequences in our theoretical framework and their interrelationships.

4.2.1. Power Inequality

We contend that there are two ways in which LBS social relations engender power inequality across web pages (and thus, webmasters), power interpreted broadly as the quantum of influence they wield. *First*, the link-based formulation of a query/need-agnostic importance score is deeply embedded within LBS, as observed earlier. This creates an a priori and query-agnostic gradient in importance across web pages, one that acts in substantive ways as a determinant of eventual visibility through search engines. For example, across two web pages, the higher importance of one web page can offset any deficit in content-based relevance relative to the other, and thus, the former has a higher propensity to achieve a higher rank. In other words, a web page with a lower query/need relevance than another may be ranked higher than the latter owing to its higher importance score. Under the ethos of relevance-based search (Ref Section 2.1), this could be seen as a violation of equality of opportunity. Similar phenomena are observed in social media where engagement metrics such as the number of *likes* function as an a priori importance measure (Vaidhyanathan 2018); high engagement posts (e.g., sensationalised posts, cat videos), while potentially crowding out relevant posts for the user, are aligned with the interest in maximizing user attention. *Second*, given the formulation of the importance score as an importance weighted aggregate of links, a web page with a high importance score is much better positioned to bestow another with an improvement in importance score through handing out a link. This enables actions of webmasters of high-importance pages to have a higher say in the operation of the web search process. Thus, web pages with possession of higher levels of backlinks (i.e., link commodity) have higher importance scores, and consequently wield much higher power within the socio-technical ecosystem of LBS.

---

[20] https://ahrefs.com/blog/link-building/
[21] https://www.powerreviews.com/what-incentives-generate-reviews/
[22] https://moz.com/learn/seo/local-ranking-factors

4.2.2. Content Disregard

With links and link-based importance scores becoming a dominant determinant of visibility, the role of content is increasingly crowded out from web search. Webmasters increasingly invest time and energy on link building as against creating useful content, running contrary to the conceptualizations of web search discussed in Section 2.1. That LBS cares much more about the *structure or form* than the content of websites makes it well-suited to capitalist production. Capital has a general indifference for content, or one might say that it subordinates content to form. Marx points out that capital is only concerned with specific use-values (i.e. specific content) insofar as they have the general form of a commodity, or in other words, that they can be sold in order to realise an exchange-value (Marx 1867, 461). Whether a commodity actually functions well, or whether it is harmful, are second-order concerns for capital, if they are concerns at all. Consider the long history of capital's unwillingness to confront so-called "negative externalities" such as carbon emissions and exploitation of social stereotypes; such externalities have carried over to the Web[23]. Link-based web search similarly focuses on the formal dimension of link metrics, rather than the specific content of websites. The disregard of content, the bearer of use-value within a web page, is a key aspect of LBS social relations.

4.3. Second Order Consequences

We now consider the second order consequences, their interrelationships as well as their relationships with first order ones.

4.3.1. Platformisation

Search engines are increasingly being viewed as a primary source of information, a function traditionally aligned with media. If so, what kind of medium are search engines? Traditional media, such as newspapers, operate as a two-sided market (Rochet and Tirole 2003), with readers and advertisers. The content is curated and owned by the media operators. It has been observed that search engines (Van Couvering 2017), on the other hand, operationalise a separation of content from delivery, with content being provided by websites outside the control of the search engine. This makes search engines a kind of multi-sided platformised medium. This platformisation creates a new dimension of reciprocal indirect network effects between users and webmasters where the proliferation of websites creates more user demand and vice versa, accelerating the influence of platformised media vis-à-vis two-sided media. This is illustrated schematically in Figure 2 within Van Couvering (2017). We observe that link-based search *adds another dimension* to such network effects, and consequently to enhanced platformisation, with links being a manifestation of network value, and consequently, exchange value. To put it simplistically, proliferation of links among existing websites is capable of catalysing user demand, and vice versa. Thus, the emergence of link-based search enhances platformisation of search engines in substantive ways. Platforms are to be viewed as not just a technological infrastructure, but as a political enabler/catalyst of a multi-sided market (Gillespie 2010). When viewed from the economic theory of conceptualising platforms as markets (Rochet and Tirole 2003), link-based search, through deepening platformisation, enhances the marketisation of

---

[23] https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739

search. Observations such as overmonetisation (Ref: Section 3.2.3) could then be read as fallouts of such consequences.

4.3.2. Structural Changes

It was observed earlier that inward links functioning as a determinant of visibility leads to an entrenchment of the link commodity. This facilitates 'inorganic' or market-driven creation of links through trade in the marketplace of links. Notably, this pathway of *inorganic creation* of links adds to the pace of overall creation of links. This could potentially make link creation a more dominant factor in the evolution of the web, as compared to other factors such as creation of content/pages. It is not obvious to us as to what kind of large-scale effects this could precipitate. Yet, studies point to significant structural changes in the web in the first two decades of this millennium. In 2000, a pioneering work (Broder et al. 2000) on analysing the web structure points out the existence of a single large set of web pages connected through hyperlinks (called a *connected component* in graph terms) and several small components disconnected from the large set; this structure, called the *bow-tie* structure, is not without detractors (Meusel et al. 2014). Two decades later, an analysis of web structure finds a sharp departure from the bow-tie structure, towards multiple local bow-tie structures (Fujita et al. 2019). While it is not clear as to what effect, if any, the emergence of the link commodity may have had on web structures at a macro level, the evolution does suggest the emergence of stronger localised traffic flows. Notably, the enhanced local connectivity is to the benefit of navigational media (Van Couvering 2011), the new ethos of web media which relies on the conceptualisation of traffic distribution as the source of revenue. In other words, it strengthens the hands of new models that rely on traffic (e.g., traffic commodity (Van Couvering 2008)) as opposed to traditional models that rely on audiences as the source of revenue. This could be catalysing the incorporation of navigation-like elements in traditional media, such as QR codes in print media (Nath and Varghese 2020).

4.3.3. Commercial Dominance

We now come to a key aspect of LBS, that of commercial dominance. While commercial dominance is latent in the discussions on platformisation and structural changes, we consider it explicitly here. The emergence of the link commodity as a central factor in LBS naturally makes it a channel of valorisation. Investment by webmasters in amassing link commodities pays off through increased visibility or traffic. This can then be monetised either directly (e.g., an e-commerce site can translate more user visits into more purchases) or through the sale of the 'audience commodity' using advertising; more traffic entails more user data to sell (Fuchs 2014, 95). In either case, the web page functions as an environment for surplus value generation. This makes LBS a key enabler for transforming the web from an information-seeking platform to a commercial platform.

It may further be observed that the commodification of links makes web page visibility directly correlated with their webmasters' monetary spending for link building. Profit-focused content creators who look to translate enhanced visibility into profits are best positioned to spend money on link building (through express spending or offering discounts in exchange for links), since they can translate the enhanced web visibility to monetary outcomes which can then be re-invested in further link commodity purchases. Within the market-based dynamics of link-based search, an academic content creator who may create quality content is ill-positioned to achieve visibility in the link-based era (cf. content-based era), since they are not looking at web visibility as a

source of revenue, and thus cannot bank on enhanced visibility to offset monetary costs incurred in achieving it (and, are not well poised to undertake monetary spending for link building). The commodification of links increasingly crowds out other models of link creation. Marx held that the "immanent laws of capitalist production … assert themselves as the coercive laws of competition" (Marx 1867, 433). LBS forces all web pages to compete on the link commodity market, making even our academic content creator subject to the exigencies of capital.

LBS thus makes commercially oriented web pages well positioned to amass more visibility within the web, bringing about a commercial bias in web search. This could be seen as a particularly narrow and latent manifestation of adversarial Goodhart effects, where link-based search produces an ecosystem that is conducive for a narrow spectrum of adversaries, those oriented commercially, to appropriate value from web search through link building. This, we suggest, functions in a somewhat similar way to the process by which online platforms evolve into digital monopolies, by rendering it near impossible for incumbents to compete with the accrued network value of an established platform (Valente 2021).

The commodification of links thus transforms the web into a platform for commerce, increasingly displacing other kinds of activities on the web by commercial ones. The deepening of commercial dominance can also be read from the prism of the well-studied Matthew effect, as we will see next.

4.3.4. Matthew Effects and Progressive Parochialism

The Matthew effect (Rigney 2010), first outlined by renowned sociologist Robert Merton, is most commonly described using the adage-like statement: *the rich get richer and the poor get poorer*. However, it is also appropriately described as one of *preferential attachment*, where rewards get distributed based on current possessions[24]. The social relations of link-based search are particularly oriented towards the Matthew effect. As an example, a web page (actually, the webmaster behind the web page) that is rich with links commands a higher visibility which further enhances their ability to acquire more links. This can be contrasted with the higher visibility of high use-value web pages in content-based search, which does not produce a feedback loop; that is, visibility does not directly or automatically lead to alteration of content to further enhance visibility. The Matthew effect arguably accentuates every consequence of link-based search. For example, the initial commercial bias progressively leads to steeper skews leading to a big business bias, and an eventual facilitator of monopolisation. Towards understanding this, consider a web-based food delivery platform which is very popular, and thus is bestowed with high page importance scores. They are able to translate their high page importance score to rope in more new businesses (e.g., takeaways) into their platform since their links will ensure higher visibility to the businesses than a link from a less popular food delivery platform. These lead to a progressive parochialism in the operation of link-based search.

4.3.5. Social Biases

We have observed that link-based search moved the heavy lifting in web search from content-matching to link-based scoring. What would this mean for the visibility of social biases within web search? In a somewhat tangential domain, that of social media, it has been found that social media *likes* – arguably a form of endorsement in a manner similar to links on the web – can convey myriad personality traits and political views,

---

[24] Thus the poor needn't get poorer in an absolute sense, only in a relative sense to the rich.

among others (Youyou, Kosinski, and Stillwell 2015). It has been pointed out that likes often convey more than a user would themselves like to[25]. Against this backdrop, it may be reasonable to speculate that links could convey a lot more latent traits than what is willingly offered through the contents of web pages. Racism, a prominent form of social bias, encompasses several types; of which *inadvertent, habitual and explicit* are examples (Brogaard 2020). While the content of web pages may convey explicit racism, inadvertent and habitual racism may not appear in the content but could manifest through links. Thus, link-based search, in contrast to content-based search, could be interpreted as offering new channels of flow of social biases from society to web search results. It has been observed that actual web traffic patterns are more racialised than the hyperlink networks (McIlwain 2017); this may be read in the light of the pervasive usage of search engines for user navigation, making traffic patterns directly dependent on the search engines (and thus, link-based importance scoring). The usage of LBS logics over user data (Ref Sec 3.1.1) could amplify such issues. Further, extant studies explore how reliance on source popularity leads to a far-right politics of exclusion (Norocel and Lewandowski 2023); it is notable that popularity could be correlated with link-based importance scores. Thus, link-based search could be, at least partially, responsible for the observations of widespread social biases in search results.

4.3.6. Ecological Issues

Here, we consider the world of ideas underpinning eco-socialism (Huan 2010). The foundational urge of capital to maximise wealth/value leads to a propensity towards *commodification of everything* (Harvey 2007) so that their values can be realised within markets to enlarge profits. Such extreme commodification spares little, and the natural environment is not immune to the urge. Indeed, as Marx points out, this leads to an "*irreparable rift in the interdependent process of social metabolism, a metabolism prescribed by the natural laws of life itself*" (Marx 1894, 949). This has been described as '*metabolic rift*' (Foster 2000), a breakdown of the metabolic relationship between humans and nature. In other words, advanced forms of capital – embedded in big businesses – predicate an unmistakable anti-nature ethos. Against this backdrop, we may reconsider the various observations made so far. The commodification of links enhances the market ethos on the Web, and the social relations of link-based search engender a commercial bias. The predominance of commercial bias and big business bias in web search translates, somewhat directly, to the privileging of commercially oriented, and thus ecologically unsustainable, interpretations of search queries, as observed in Haider, Rödl, and Joosse (2022) and Haider and Rödl (2023), as discussed in Section 3.2.3. Thus, ecological unsustainability may be seen as being a fallout of the social relations encoded within the socio-technical ecosystem of link-based search, and intricately linked to link commodification.

4.4. Allied Noteworthy Aspects

We now consider some broader and allied aspects herein.

4.4.1. Deterioration of Utility

The discussion so far considered myriad consequences of the social relations and socio-technical ecosystem of link-based search. They all converge on one broad trend,

---

[25] https://www.ted.com/talks/jennifer_golbeck_your_social_media_likes_expose_more_than_you_think/up-next?subtitle=en

that of catalysing a deterioration of the utility of link-based search as a search paradigm. While link-oriented platformisation promotes the marketisation of search, it could be fragmenting the web into silos, as observed in Sec 4.3.2. The myriad social biases and commercial biases could be racialising and commercialising web search, taking web search away from being able to effectively mediate between queries and web pages. Web search could be privileging unsustainable interpretations, aggravating the ecological crisis. The accentuation of the above phenomena through Matthew effects could be creating a subtle but virulent feedback loop parochialising web search. The complementarities, overlaps and synergies among the factors could eventually trigger significant drops in the utility of the link-based search as a paradigm of information seeking on the Web. In light of these observations, we could re-read allegations of futility of link-based web search, as observed in Section 3.2.3.

4.4.2. The Mode of Production in Link-based Web Search

We now attempt to situate link-based search within the concept of automation, an important aspect of Critical Political Economy scholarship. Automation, within traditional manufacturing sectors, is often described as an increase in the *organic composition* (Shaikh 1990) of the mode of production, which increases the proportion of machinery to labour in the production process. Marxist Economics sees automation as an intrinsic tendency of capital (Rowthorn and Harris 1985), which works in lockstep with its drive to enhance surplus value extraction (Marx 1857). Machinery is itself produced using manual labour in the past, and thus, often described as dead labour (Marx 1867), i.e., an ossified manifestation of living labour from the past. We posit that link-based search sits in a somewhat uneasy way within the broader understanding of automation.

Link-based search, as observed earlier, restructures the retrieval process into an importance computation step and a content-based matching step. Let us consider the *nature of work* involved in these separate 'production' processes. The page importance computation process has a very simple technical form but relies heavily on the existence of an up-to-date *link infrastructure* within the Web. The link infrastructure, as we have seen before, is the product of webmaster actions, and thus, may be described as ossified historical and globally distributed digital labour. The matching process, on the other hand, is a query-dependent process and could be potentially complex, given the need to ensure that web pages that satisfy the user's need are ranked sufficiently highly. This is performed by matching algorithms, which are themselves ossified labour of programmers. Under this lens, the importance computation and matching process are both heavily aided by ossified labour, the first by labour ossified primarily within the *link infrastructure* and the second by labour ossified within *matching algorithms*.

A crucial difference between the above two forms of ossified labour is notable. The *link infrastructure* is labour by webmasters who are not paid by the search engines. By creating each link, the webmaster enriches the link infrastructure through undertaking latent and unpaid digital labour, in ways similar to observations of unpaid digital labour elsewhere (Fuchs 2014). On the other hand, the *matching algorithms* are produced by the digital labour of programmers and software architects within the search engine company, which is paid digital labour. Thus, the shift from content-based search to link-based search changes the organic composition of web search from being dominated by the fruits of paid digital labour to one being increasingly the product of unpaid digital labour. This is heavily aligned with the profit urge of capital since unpaid digital labour is intrinsically more profit oriented than paid labour insofar as it does not receive a wage.

Further, it is notable that the quality of web search relies heavily on the link infrastructure being kept up to date, to serve the changing needs of web search users. On the other hand, matching algorithms may only need to be updated much less frequently. Thus, in the long run, the maintenance of the link infrastructure could be viewed as having characteristics of living labour. Thus, the shift from content-based to link-based search may be seen as a shift towards living labour. Given our focus on link-based search, we keep the labour of web page creation outside our remit.

5. Discussion

Having discussed the foundational role of the link commodity within LBS, and the consequent facilitation it provides for advancement of capital on the web, we now turn our attention towards some alternative possibilities. In particular, we consider some of the nuances within the web ecosystem and extant building blocks that could point towards alternatives, against the backdrop of socio-technical challenges.

5.1. Web Pages as Use Values: Challenges

In the beginning of this paper, we considered web pages as use values and that they could help satisfy information needs. The commodification of links within LBS engenders a conceptualisation of web pages first of all as realisation points of link exchange values. This is so because the additional visibility that a web page acquires through a new link is monetised when users visit the web page. As seen earlier, this could be in one of at least two ways; through the web page promoting a transaction such as in e-commerce, or through ad-driven monetisation of attention through models such as Google AdSense[26]. Against this backdrop, let us consider the challenges in conceptualising web pages, as they stand today, as use values.

Web pages are containers of 'knowledge', and thus, different from durable use-values that take a physical form. In contrast to oil, soap or foods that are exhausted upon consumption, web pages, like all information commodities, are not subject to physical limits to concurrent or repetitive consumption. A web page can be consumed by a plurality of people at the same time or at different times, and is thus a 'non-rival' good. The same is true of scientific knowledge that manifest as research papers (a form of intellectual property), a realm that is bound to strict normative constraints such as Mertonian norms (Merton 1973) that mandate common ownership; such mandates strive to keep scientific knowledge outside the remit of commodification. Software code is another similar domain, where there exist mechanisms, such as GNU licenses[27], that resist commercialisation and commodification; a huge community of Free and Open Source Software (FOSS) has been fostered through such mechanisms. While there exist licenses such as Creative Commons[28] that could be used to protect content within web pages more generally, the global and distributed nature of the web makes it hard to operationalise them pervasively. As such, content on the Web remains loosely regulated and bound by few normative frameworks, unlike scientific literature or free software. As early as 1995, the pervasive copying of content and resultant redundancy was noted sharply in an article[29] which notes the Web having been reduced to '*a redundant, ever-multiplying and increasingly chaotic mass of documents*' in that '*it permits infinite redundancy and encourages maximum confusion*'.

---

[26] https://adsense.google.com/start/
[27] https://www.gnu.org/licenses/gpl-3.0.html
[28] https://creativecommons.org/
[29] https://www.wired.com/1995/06/xanadu/

Given that web page content can be copied with few restrictions and that users generally tend to not go beyond the first few results from a search engine[30], what entails is a visibility competition among webmasters for the top spots on the result page. If there exists a content on a topic which is very informative or interesting, webmasters may copy it on to their own web pages hoping that such actions would help them attain higher rankings on search engine result pages. It is notable that such a competition exists even within content-based search, and is not a phenomenon created by LBS. However, LBS skews and amplifies this competition, in that webmasters who have pocketed high importance scores (through amassing link commodity) are offered an a priori advantage; if one creates an informative or interesting web page, the visibility due to that is lost once a webmaster of a higher PageRank web page copies the content on to them. Among two copies of content, the one with higher PageRank would inevitably be favoured by LBS. This engenders a quasi-capitalist social relation, since this amounts to the fruits of the labour of one webmaster being pocketed by another for surplus value. Thus, it appears that the current combination of circumstances, including at least factors such as the model of Web hypertext and the linear presentation of results, are aligned with the logics of competition and thus ill-suited towards conceptualising a use-value based model for web search. We thus consider alternative content models and alternative search result presentation models, and discuss how they may align with use-value based operation.

5.2. Alternative Models of Content on the Web

The Web model of organising content is not the only possibility of operationalising hypertext. We look at an alternative conceptual model from the early history of the Web, and a concrete content model adopted by Wikipedia.

Founded in 1960, Xanadu, 'the original hypertext project'[31], predates the Web by as much as three decades. While Xanadu was a conceptual model, the founder, Ted Nelson, intended it to be a universal, democratic hypertext library that would be most liberating and empowering (Stevens 2023). Documents in Xanadu's hypertext model are designed based on non-sequential writing, where each document is in the form of a virtual file, which has a list of contents, and details of how to put them together. Some of the contents may be text contained within the document, whereas some others could be text within *another document*, connected through links. It may be reasonable to say that a reader could follow such a Xanadu hypertext document by reading the document from the start, following links as appropriate to read parts from linked documents, each time returning to the document itself, and so on. Nelson refers to Xanadu's linking functionality as *transclusion* (Pam 1995), the process of including something by reference rather than by copying. The transclusion functionality is key in the vision of Xanadu as a '*docuverse*' with no redundancy; every document will exist as the original, and any author seeking to include any part will simply reference it through transclusion. Transclusion is notably enabled by the nature of Xanadu's hyperlinks (or, Xanalinks) as being bivisible (visible from both the source and destination page) and bifollowable (can be traversed forward and backward); this is in sharp contrast to the nature of hyperlinks on the web that are univisible (can only be seen from the originating end) and unifollowable (can only be followed in one direction). This enables building upon existing content on the web without copying it. The express prevention of copying, however, must be still enforced through an elaborate intellectual property scheme,

---

[30] https://www.searchenginejournal.com/google-first-page-clicks/374516/
[31] https://www.museumofmediahistory.com/xanadu

which has been laid out to some detail and also described as bold and innovative (Samuelson and Giushko 1992). Thus, if a webmaster seeks to achieve high visibility by creating a new document on Xanadu, they are forbidden from copying content. Yet, they have at least two options. First, they could create new content by undertaking their own inquiry and research; this fosters the creation of new use values, possibly enhancing diversity and pluralism on the (Xana)web. Second, they could create a document that transclusively links to a mashup of useful content already existing on the docuverse; in this case, they would need to transfer the visibility to such existing documents, since transclusion systemically forces them to do so. Thus, the Xanadu docuverse model may be seen as offering some resistance to the coercive laws of competition through systemically restricting specific commodification-oriented webmaster actions. Yet, since Xanalinks are distinct from today's hyperlinks, it is not clear how the logic of LBS could work within Xanadu, or whether there could be models of search that orient towards commodification of Xanalinks.

We now turn our attention to Wikipedia, an alternative and flourishing content model within today's Web. Wikipedia[32] is a free content online encyclopaedia written and maintained by a community of volunteers, hosted by the non-profit Wikimedia foundation, and funded mainly by donations from readers. LBS search engines such as Google often rank Wikipedia pages highly, while also using Wikipedia content to form knowledge panels to augment the result pages[33]. Wikipedia's content model imposes a strict restriction on one Wikipedia page per named topic through enforcing that the web page URL bears the topic itself. For example, there cannot be two Wikipedia pages associated with Marxism, since there can only be one web page at http://en.wikipedia.org/en/Marxism. The single web page model is governed by strong policies, such as enforcing that all content should be '*written from a neutral point of view, representing significant views fairly, proportionately and without bias*'[34]. The policy also forbids any original research requiring that all content should be attributable to a reliable, published source; this guideline restricts the kind of links that can be created on Wikipedia. These policies are enforced through Wikipedia's volunteerism-based moderator community. It is notable that such policies have been observed to be problematic when it comes to representations of several contested geographic territories such as Crimea and Jerusalem (Graham and Dittus 2022). Yet, these guidelines provide for interesting dynamics that resist commercial appropriation. For example, the Wikipedia page for a corporate entity such as Google[35] is also bound to contain a section titled '*Criticism and controversies*' with links to reliable news sources around the controversies; thus, a corporate entity cannot prevent users visiting its Wikipedia web page from being exposed to negative viewpoints about itself. Such strong restrictions on content and link structuring within Wikipedia makes it interestingly different from the general Web and allow it to resist commercialisation in interesting ways.

We saw two models that diverge from the current model on the Web where anybody can create web pages with any content and linking to any other web page. While Xanadu disallows redundancy while allowing for multiple documents for a given topic, Wikipedia is more concerned about providing an encyclopaedic entry point for each topic and less worried about redundancy. Xanadu requires that links are content-focused (for transclusion), whereas Wikipedia encourages links to reliable media (reliability estimated externally, and not intrinsically through links). Consequently, these

---

[32] https://en.wikipedia.org/wiki/Main_Page
[33] https://themathergroupllc.com/how-google-uses-wikipedia-to-improve-search-results/
[34] https://en.wikipedia.org/wiki/Wikipedia:Policies_and_guidelines#%C2%ADContent
[35] https://en.wikipedia.org/wiki/Google

models are both potent to resist link commodification and commercialisation, since the restrictions they pose on links would likely undermine the LBS-based link commodification logic that rests on the assumption that backlinks possess intrinsic value.

5.3. Alternative Models of Search Result Presentation

Among the catalysts for the competition for the top spots on search results is arguably the linear presentation of search results. This advantages the top spot highly, making it almost a 'default' choice, an aspect further facilitated by features such as Google's 'I'm feeling lucky'[36] which enables a user to bypass the results page to directly go to the first result. This is often characterised as position bias (Biega, Gummadi, and Weikum 2018), which leads to disproportionately less attention being paid to low-ranked results. The linear presentation, which leads to a high degree of position bias, disincentivises creation of diverse use values on the web each need, since only the first few would gather visibility anyway.

Other forms of search result presentations which have a lesser degree of position bias could help mitigate the intensity of competition. These could include tiled presentation of results as often used within image search engines and e-commerce results; tile sizes therein can be tailored in accordance with result relevance. Stochastically controlling the search results so that the second result gets flipped to the first position to help achieve occasional good visibility is another approach (Biega, Gummadi, and Weikum 2018). Apart from position bias, other forms of cognitive biases in retrieval as outlined in Azzopardi (2021) could relate to facilitating competition or commodification in ways that would require development of bespoke result presentation schemes.

6. Conclusion

We have proposed a theoretical framework for understanding the Political Economy of link-based web search. We have shown that extant criticisms of link-based search derive from its social relations. We assert that the emergence of link-based search is favourable for the entrenchment of capital valorisation and the accentuation of a libertarian ethos on the Web. The socio-technical environment posited by link-based search appears highly correlated with the interests of the better studied relations of data-intensive, or surveillance, capital in that it enables the expansion of the capital relations to new information domains.

The discussion here lays bare the interdisciplinary and complex nature of challenges towards realising a use-value based model of search operation on the web. It also emphasises the heavily entangled relationship between the web, and the primary model to access it viz. web search. The challenges include sociological, media, political, technological and cognitive aspects, and calls for interdisciplinary inquiry. We suggest that inquiries towards a non-link-based web search may benefit from being guided by one or more of the following three theories viz., the Mertonian norm of knowledge communism (Merton 1973), the Rochdale principle of democratic control (Fairbairn 1994) and insights from Elinor Ostrom's work on institutional arrangements for managing the commons (Ostrom 1990). The Mertonian norm of knowledge communism asserts the need for common ownership of scientific discoveries, and the need for scientists to publicly share their discoveries. Within the web case, we may similarly want to assert common ownership of information on the web; this could have several consequences, one of which could be an interpretation of disallowing ad-based monetisation of web page traffic since that is a mechanism for individual-level monetisation of

---

[36] https://www.howtogeek.com/847170/googles-im-feeling-lucky-explained/

the commons. In case centralised models are practically feasible, such an ad-based visibility incentive (a determinant of the propensity for amassing link commodity) could be trivially dissolved if web pages are predominantly accessed through non-profit digital libraries such as the Internet Archive[37]. The Rochdale principle of democratic control is a norm for cooperatives, which is often realised through the 'one member, one vote' principle[38], one that ensures that the governance model stays truthful to the interests of the majority of participants. In simple terms, a member with a $1m share gets exactly the same voting power as one with a $1 share, so influence is decoupled from investment. While there have been no extant models of web search structured as cooperatives of users or cooperatives of webmasters, there appears to be some emerging interest, as outlined in a cooperative web search project[39]. If explorations for alternative models of web search are directed to decentralised models, institutional arrangements would emerge as extremely important, one where Elinor Ostrom's work could offer interesting directions. Ostrom's work identified defining principles in the management of common pool resources, so that the resources are not overused by individuals at the peril of the commons. The principles touch upon myriad aspects of governance viz., boundaries, operational rules, monitoring, sanctions and so on. Managing the web and search as a common pool resource would need similar mechanisms to ensure sustainability of the model.

References

Azzopardi, Leif. 2021. Cognitive biases in search: a review and reflection of cognitive biases in Information Retrieval. In *Proceedings of the 2021 conference on human information interaction and retrieval,* edited by Falk Scholer and Paul Thomas, New York, 27-37. Association of Computing Machinery. https://doi.org/10.1145/3406522.3446023

Berners-Lee, Tim. 2019. 30 Years On, What's Next #ForTheWeb? Accessed 31 July 2024. https://webfoundation.org/2019/03/web-birthday-30/.

Berners-Lee, Timothy J. 1989. Information Management: A Proposal. https://www.w3.org/History/1989/proposal.html.

Biega, Asia J, Krishna P Gummadi, and Gerhard Weikum. 2018. Equity of Attention: Amortizing Individual Fairness in Rankings. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, edited by Kevyn Collins-Thompson, Qiaozhu Mei, Brian D. Davison, Yiqun Liu, and Emine Yilmaz, New York, 405-414. Association for Computing Machinery. https://doi.org/10.1145/3209978.3210063.

Birch, Kean, D. T. Cochrane, and Callum Ward. 2021. Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech. *Big Data & Society* 8 (1): 1-15. https://doi.org/10.1177/20539517211017308

Boltanski, Luc, and Eve Chiapello. 2005. The New Spirit of Capitalism. *International Journal of Politics, Culture, and Society* 18: 161-188. https://doi.org/10.1007/s10767-006-9006-9.

Bonefeld, Werner. 2014. *Critical Theory and the Critique of Political Economy: On Subversion and Negative Reason*. New York: Bloomsbury.

Bornmann, Lutz, and Hans-Dieter Daniel. 2008. What Do Citation Counts Measure? A Review of Studies on Citing Behavior. *Journal of Documentation* 64 (1): 45-80. https://doi.org/10.1108/00220410810844150.

Brin, Sergey, and Lawrence Page. 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems* 30 (1–7): 107-117. https://doi.org/10.1016/S0169-7552(98)00110-X.

---

[37] https://archive.org/about/

[38] https://cooperativesfirst.com/blog/2020/01/24/why-one-member-one-vote/

[39] http://www.cooperativeweb.org/

Broder, Andrei, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. 2000. Graph Structure in the Web. *Computer Networks* 33 (1-6): 309-320. https://doi.org/10.1016/S1389-1286(00)00083-9.

Brogaard, Berit. 2020. The Many Faces of Racism. Accessed 31 July 2024. https://www.psychologytoday.com/gb/blog/the-superhuman-mind/202012/the-many-faces-racism.

Campbell, Donald T. 1979. Assessing the Impact of Planned Social Change. *Evaluation and Program Planning* 2 (1): 67-90. https://doi.org/10.1016/0149-7189(79)90048-X.

Carr, Nicholas. 2009. *The Big Switch: Rewiring the World, from Edison to Google*. New York: WW Norton & Company.

Chakraborty, Tanmoy, and Ramasuri Narayanam. 2016. All Fingers are not Equal: Intensity of References in Scientific Articles. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, edited by Xavier Carreras Jian Su, Kevin Duh, Austin, 1348-1358. Association for Computational Linguistics. https://doi.org/10.18653/v1/D16-1142.

Chierichetti, Flavio, Ravi Kumar, and Prabhakar Raghavan. 2011. Optimizing Two-Dimensional Search Results Presentation. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, edited by Irwin King, Wolfgang Nejdl, and Hang Li, New York, 257–266. Association for Computing Machinery. https://doi.org/10.1145/1935826.1935873.

Chung, Fan. 2014. A Brief Survey of PageRank Algorithms. *IEEE Transactions on Network Science and Engineering* 1 (1): 38-42. https://doi.org/10.1109/TNSE.2014.2380315.

Davis, Harold, and Iwanow, David. 2009. *Google Advertising Tools: Cashing in with AdSense and AdWords*. O'Reilly Media: Boston.

Fairbairn, Brett. 1994. The meaning of Rochdale: The Rochdale pioneers and the co-operative principles. *Occasional Paper Series*. Centre for the study of Co-operatives, University of Saskatchewan

Figueroa, Alejandro. 2015. Exploring Effective Features for Recognizing the User Intent behind Web Queries. *Computers in Industry* 68: 162-169. https://doi.org/10.1016/j.compind.2015.01.005.

Foster, John Bellamy. 2000. *Marx's Ecology: Materialism and Nature*. New York: Monthly Review Press.

Franceschet, Massimo. 2011. PageRank: Standing on the Shoulders of Giants. *Communications of the ACM* 54 (6): 92-101. https://doi.org/10.1145/1953122.1953146.

Fuchs, Christian. 2014. *Digital Labour and Karl Marx*. New York: Routledge.

Fujita, Yuji, Yuichi Kichikawa, Yoshi Fujiwara, Wataru Souma, and Hiroshi Iyetomi. 2019. Local Bow-Tie Structure of the Web. *Applied Network Science* 4 (15). https://doi.org/10.1007/s41109-019-0127-2.

Garrabrant, Scott. 2017. Goodhart Taxonomy. Accessed 31 July 2024. https://www.lesswrong.com/posts/EbFABnst8LsidYs5Y/goodhart-taxonomy.

Germano, Fabrizio, and Francesco Sobbrio. 2020. Opinion Dynamics Via Search Engines (and Other Algorithmic Gatekeepers). *Journal of Public Economics* 187: 104188. https://doi.org/10.1016/j.jpubeco.2020.104188.

Gigerenzer, Gerd. 2023. Psychological AI: Designing Algorithms Informed by Human Psychology. *Perspectives on Psychological Science* 0 (0). https://doi.org/10.1177/17456916231180597.

Gillespie, Tarleton. 2010. The Politics of 'Platforms'. *New Media & Society* 12 (3): 347-364. https://doi.org/10.1177/1461444809342738.

Goodhart, Charles. 1975. Problems of Monetary Management: The UK Experience. In *Inflation, Depression, and Economic Policy in the West*, edited by Anthony Courakis, 116. New Jersey: Barnes and Noble.

Goodhart, Charles Albert Eric. 1984. *Monetary Theory and Practice: The UK Experience*. London: Macmillan.

Graham, Mark and Martin Dittus. 2022. *Geographies of digital exclusion: Data and inequality*. London: Pluto.

Gross, Paul LK, and Edward M Gross. 1927. College Libraries and Chemical Education. *Science* 66 (1713): 385-389. https://doi.org/10.1126/science.66.1713.385.

Guo, Jiafeng, Yinqiong Cai, Yixing Fan, Fei Sun, Ruqing Zhang, and Xueqi Cheng. 2022. Semantic Models for the First-Stage Retrieval: A Comprehensive Review. *ACM Transactions on Information Systems (TOIS)* 40 (4): 1-42. https://doi.org/10.1145/3486250.

Haider, Jutta, and Malte Rödl. 2023. Google Search and the Creation of Ignorance: The Case of the Climate Crisis. *Big Data & Society* 10 (1). https://doi.org/10.1177/20539517231158997.

Haider, Jutta, Malte Rödl, and Sofie Joosse. 2022. Algorithmically Embodied Emissions: The Environmental Harm of Everyday Life Information in Digital Culture. *Proceedings of CoLIS, the 11th International Conference on Conceptions of Library and Information Science* 27 (Special Issue): colis2224. https://doi.org/10.47989/colis2224.

Harvey, David. 2007. *A Brief History of Neoliberalism*. Cary: Oxford University Press.

Heinrich, M., 2007. Invaders from marx: On the uses of marxian theory, & the difficulties of a contemporary reading. *Left Curve* 31: 83-88.

Henzinger, Monika R. 2001. Hyperlink Analysis for the Web. *IEEE Internet Computing* 5 (1): 45–50. https://doi.org/10.1109/4236.895141.

Huan, Qingzhi. 2010. Eco-Socialism in an Era of Capitalist Globalisation: Bridging the West and the East. In *Eco-Socialism as Politics: Rebuilding the Basis of Our Modern Civilisation*, edited by Qingzhi Huan, 3–12. Dordrecht: Springer. https://doi.org/10.1007/978-90-481-3745-9_1.

Introna, Lucas D, and Helen Nissenbaum. 2000. Shaping the Web: Why the Politics of Search Engines Matters. *The Information Society* 16 (3): 169-185. https://doi.org/10.1080/01972240050133634.

Katz, Leo. 1953. A New Status Index Derived from Sociometric Analysis. *Psychometrika* 18 (1): 39-43. https://doi.org/10.1007/BF02289026.

Katz, Yarden. 2020. *Artificial Whiteness: Politics and Ideology in Artificial Intelligence*. New York: Columbia University Press.

Kang, Hyunjin, and Matthew P. McAllister. 2011. Selling you and your clicks: examining the audience commodification of Google. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society* 9 (2): 141-153. https://doi.org/10.31269/triplec.v9i2.255

Kitchin, Rob. 2019. Thinking critically about and researching algorithms. In *The social power of algorithms*, edited by David Beer, 14-29. Routledge: New York.

Kjøsen, A.M., 2013. Human Material in the Communication of Capital. *communication+1 2*(1): 1-42. https://doi.org/10.7275/R5JS9NCN

Lewandowski, Dirk, Jutta Haider, and Olof Sundin. 2024. JASIST Special Issue Editorial: Re-orienting Search Engine Research in Information Science. *Journal of the Association for Information Science and Technology* 75 (5): 503-511. https://doi.org/10.1002/asi.24845.

Mager, Astrid. 2012. Algorithmic Ideology: How Capitalist Society Shapes Search Engines. *Information, Communication & Society* 15 (5): 769-787. https://doi.org/10.1080/1369118X.2012.676056.

Mager, Astrid, Ov Cristian Norocel, and Richard Rogers. 2023. Advancing Search Engine Studies: The Evolution of Google Critique and Intervention. *Big Data & Society* 10 (2). https://doi.org/10.1177/20539517231191528.

Manheim, David, and Scott Garrabrant. 2018. Categorizing Variants of Goodhart's Law. https://doi.org/10.48550/arXiv.1803.04585.

Marx, Karl. 1857. *Grundrisse: Foundations of the Critique of Political Economy*. London: Penguin Books.

Marx, Karl. 1859. *A Contribution to the Critique of Political Economy*. Moscow: Progress Publishers.

Marx, Karl. 1867. *Capital*. Vol. 1. Moscow: Progress Publishers.

Marx, Karl. 1894. *Capital*. Vol. 3. New York: International Publishers.

Mayer, Katja. 2009. On the Sociometry of Search Engines. In *Deep Search: The Politics of Search Beyond Google*, edited by Konrad Becker and Felix Stalder, 54-72. Innsbruck: Studienverlag.

McIlwain, Charlton. 2017. Racial Formation, Inequality and the Political Economy of Web Traffic. *Information, Communication & Society* 20 (7): 1073-1089. https://doi.org/10.1080/1369118X.2016.1206137.

McKiernan, Erin, Juan Pablo Alperin, and Alice Fleerackers. 2019. The "Impact" of the Journal Impact Factor in the Review, Tenure, and Promotion Process. Accessed 31 July 2024. https://blogs.lse.ac.uk/impactofsocialsciences/2019/04/26/the-impact-of-the-journal-impact-factor-in-the-review-tenure-and-promotion-process/.

Merton, Robert K. 1973. *The sociology of science: Theoretical and empirical investigations*. Chicago: The University of Chicago.

Meusel, Robert, Sebastiano Vigna, Oliver Lehmberg, and Christian Bizer. 2014. Graph Structure in the Web — Revisited: A Trick of the Heavy Tail. In *Proceedings of the 23rd International Conference on World Wide Web*, edited by Chin-Wan Chung, Andrei Z. Broder, Kyuseok Shim, and Torsten Suel, New York, 427-432. Association for Computing Machinery. https://doi.org/10.1145/2567948.2576928.

Nath, Athma B, and Titto Varghese. 2020. Factors Influencing the Effective Usage of QR Code: A Study Among Readers of Trivandrum City. *IJRAR – International Journal of Research and Analytical Reviews (IJRAR)* 7 (1): 822-870. https://ijrar.org/viewfull.php?&p_id=IJRAR19K7814.

Nielsen, Jakob. 2006. F-Shaped Pattern for Reading Web Content. Accessed 31 July 2024. https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content-discovered/.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression*. New York: New York University Press.

Norocel, Ov Cristian, and Dirk Lewandowski. 2023. Google, Data Voids, and the Dynamics of the Politics of Exclusion. *Big Data & Society* 10 (1). https://doi.org/10.1177/20539517221149099.

Ostrom, Elinor. 1990. *Governing the commons: The evolution of institutions for collective action*. Cambridge: Cambridge University Press.

Pam, A. 1995. Where world wide web went wrong. In *AUUG Conference Proceedings 1995*, edited by Philip Tsang, 27-32. AUUG.

Pasquinelli, Matteo. 2009. Google's PageRank Algorithm: A Diagram of Cognitive Capitalism and the Rentier of the Common Intellect. In *Deep Search: The Politics of Search Beyond Google*, edited by Konrad Becker and Felix Stalder, 152-162. London: Transaction Publishers.

Rahman, Anis. 2020. Algorithms of Oppression: How Search Engines Reinforce Racism. *New Media & Society* 22 (3): 575-577. https://doi.org/10.1177/1461444819876115.

Rieder, B. 2022. Towards a political economy of technical systems: The case of Google. *Big Data & Society*, *9* (2): 1-5. https://doi.org/10.1177/20539517221135162

Rigney, Daniel. 2010. *The Matthew Effect: How Advantage Begets Further Advantage*. New York: Columbia University Press.

Rochet, Jean-Charles, and Jean Tirole. 2003. Platform Competition in Two-Sided Markets. *Journal of the European Economic Association* 1 (4): 990-1029. https://doi.org/10.1162/154247603322493212.

Rowthorn, Bob, and Donald J Harris. 1985. The Organic Composition of Capital and Capitalist Development. In *Rethinking Marxism: Struggles in Marxist Theory: Essays for Harry Magdoff & Paul Sweezy*, edited by Stephen Resnick and Richard Wolff, 345-357. New York: Autonomedia.

Samuelson, Pamela and Robert Giushko. 1992. Intellectual property rights for digital library and hypertext publishing systems. *Harv. JL & Tech. 6(1)*: 237-262.

Shaikh, Anwar. 1990. Organic Composition of Capital. In *Marxian Economics*, edited by John Eatwell, Murray Milgate, and Peter Newman, 304-309. London: Palgrave Macmillan. https://doi.org/10.1007/978-1-349-20572-1_45.

Siebert, Horst. 2001. *The Cobra Effect: How to Avoid Wrong Paths in Economic Policy*. Stuttgart: DVA.

Silverstein, Craig, Hannes Marais, Monika Henzinger, and Michael Moricz. 1999. Analysis of a Very Large Web Search Engine Query Log. *ACM SIGIR Forum* 33 (1): 6-12. https://doi.org/10.1145/331403.331405.

Smith, Linda C. 1981. Citation Analysis. *Library Trends* 30 (1): 83-106. https://hdl.handle.net/2142/7190.

Sohn-Rethel, Alfred. 2019/1974. The Formal Characteristics of Second Nature. *Selva: A Journal of the History of Art*. https://selvajournal.org/the-formal-characteristics-of-second-nature/

Song, Yang, Xiaolin Shi, and Xin Fu. 2013. Evaluating and Predicting User Engagement Change with Degraded Search Relevance. In *Proceedings of the 22nd International Conference on World Wide Web*, edited by Daniel Schwabe, Virgílio A. F. Almeida, Hartmut Glaser, Ricardo Baeza-Yates, and Sue B. Moon, New York, 1213-1224. Association for Computing Machinery. https://doi.org/10.1145/2488388.2488494.

Spink, Amanda, Dietmar Wolfram, Major BJ Jansen, and Tefko Saracevic. 2001. Searching the Web: The Public and Their Queries. *Journal of the American Society for Information Science and Technology* 52 (3): 226-234. https://doi.org/10.1002/1097-4571(2000)9999:9999%3C::AID-ASI1591%3E3.0.CO;2-R.

Stevens, H., 2023. Code and Critique: Ted Nelson's Project Xanadu and the Politics of New Media. *Osiris*, *38* (1): 245-264.

Sundin, Olof, Dirk Lewandowski, and Jutta Haider. 2022. Whose Relevance? Web Search Engines as Multisided Relevance Machines. *Journal of the Association for Information Science and Technology* 73 (5): 637-642. https://doi.org/10.1002/asi.24570.

Taghavi, Mona, Ahmed Patel, Nikita Schmidt, Christopher Wills, and Yiqi Tew. 2012. An Analysis of Web Proxy Logs with Query Distribution Pattern Approach for Search Engines. *Computer Standards & Interfaces* 34 (1): 162-170. https://doi.org/10.1016/j.csi.2011.07.001.

Vaidhyanathan, Siva. 2018. *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy*. New York: Oxford University Press.

Valente, Jonas C. L. 2021. *From Online Platforms to Digital Monopolies: Technology, Information and Power*. Leiden: Brill.

Van Couvering, Elizabeth. 2004. New Media? The Political Economy of Internet Search Engines. In *Annual Conference of the International Association of Media & Communications Researchers*, 7-14.

Van Couvering, Elizabeth. 2008. The History of the Internet Search Engine: Navigational Media and the Traffic Commodity. In *Web Search: Multidisciplinary Perspectives*, edited by Amanda Spink and Michael Zimmer, 177-206. Berlin: Springer.

Van Couvering, Elizabeth. 2011. Navigational Media: The Political Economy of Online Traffic. In *The Political Economics of Media: The Transformation of the Global Media Industries*, edited by Dwayne Roy Winseck and Dal Yong Jin, 183-200. London: Bloomsbury Academic.

Van Couvering, Elizabeth. 2017. The Political Economy of New Media Revisited. In *Proceedings of the 50th Annual Hawaii International Conference on System Sciences*, edited by Tung X. Bui and Ralph Sprague, JR., Honolulu, 1812-1819. UH Manoa. http://hdl.handle.net/10125/41374.

Wittgenstein, Ludwig, G. H. von Wright, and G. E. M. Anscombe. 1964. Notebooks, 1914-1916. *Mind* 73 (289): 132-141. https://doi.org/10.1093/mind/LXXIII.289.132.

Wood, Jordan. 2008. Filter and Refine Strategy. In *Encyclopedia of GIS*, edited by Shashi Shekhar and Hui Xiong, 320. Boston: Springer. https://doi.org/10.1007/978-0-387-35973-1_415.

Xu, Zhen, and James Miller. 2016. Identifying Semantic Blocks in Web Pages using Gestalt Laws of Grouping. *World Wide Web* 19: 957-978. https://doi.org/10.1007/s11280-015-0370-0.

Yin, Yuyu, Nan Zhang, Zulong Chen, Mingxiao Li, Yu Li, Honghao Gao, and Lu He. 2022. Exploiting User Preferences for Multiscenarios in Query-Less Search. *IEEE Transactions on Computational Social Systems* 9 (6): 1794-1806. https://doi.org/10.1109/TCSS.2022.3181271.

Youyou, Wu, Michal Kosinski, and David Stillwell. 2015. Computer-Based Personality Judgments Are More Accurate than Those Made by Humans. *Proceedings of the National Academy of Sciences* 112 (4): 1036-1040. https://doi.org/10.1073/pnas.1418680112.

Zehlike, Meike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed, and Ricardo Baeza-Yates. 2017. FA*IR: A Fair Top-k Ranking Algorithm. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, edited by Ee-Peng Lim, Marianne Winslett, Mark Sanderson, Ada Wai-Chee Fu, Jimeng Sun, J. Shane Culpepper, Eric Lo, Joyce C. Ho, Debora Donato, Rakesh Agrawal, Yu Zheng, Carlos Castillo, Aixin Sun, Vincent S. Tseng, and Chenliang Li, New York, 1569-1578. Association for Computing Machinery. https://doi.org/10.1145/3132847.3132938.

Zehlike, Meike, Ke Yang, and Julia Stoyanovich. 2021. Fairness in Ranking: A Survey. https://doi.org/10.48550/arXiv.2103.14000.

Zehlike, Meike, Ke Yang, and Julia Stoyanovich. 2022. Fairness in Ranking, Part I: Score-Based Ranking. *ACM Computing Surveys* 55 (6): 1-36. https://doi.org/10.1145/3533379.

Zhang, Ying, Bernard J Jansen, and Amanda Spink. 2009. Time Series Analysis of a Web Search Engine Transaction Log. *Information Processing & Management* 45 (2): 230-245. https://doi.org/10.1016/j.ipm.2008.07.003.

About the Authors

Deepak P
Deepak P is an Associate Professor in the School of Electronics, Electrical Engineering and Computer Science at Queen's University Belfast. He has a background in data analytics and artificial intelligence and has published over 100 research articles across top avenues in those disciplines. His current research focuses on analysing artificial intelligence and data science through the prism of Marxist political economy.

James Steinhoff
James Steinhoff is an Assistant Professor in the School of Information and Communication Studies at University College Dublin. His research focuses on automation and the political economy of data and artificial intelligence. He is author of Automation and Autonomy: Labour, Capital and Machines in the AI Industry (Palgrave Macmillan 2021) and co-author of Inhuman Power: Artificial Intelligence and the Future of Capitalism (Pluto 2019).

Stanley Simoes
Stanley Simoes is a Research Assistant and PhD student at the School of Electronics, Electrical Engineering, and Computer Science at Queen's University Belfast. He is also associated with the Centre for Public Policy and Administration at Queen's. His research interests lie in AI ethics, currently focussed on fairness in unsupervised learning algorithms.