Algorithmic neutrality [1,2]

Milo Phillips-Brown
University of Edinburgh
Draft as of December, 2023

Abstract. Bias infects the algorithms that wield increasing control over our lives. Predictive policing systems overestimate crime in communities of color; hiring algorithms dock qualified female candidates; and facial recognition software struggles to recognize dark-skinned faces. Algorithmic bias has received significant attention. Algorithmic *neutrality*, in contrast, has been largely neglected. Algorithmic neutrality is my topic. I take up three questions. What is algorithmic neutrality? Is algorithmic neutrality possible? When we have algorithmic neutrality in mind, what can we learn about algorithmic bias? To answer these questions in concrete terms, I work with a case study: search engines. Drawing on work about neutrality in science, I say that a search engine is neutral only if certain values—like political ideologies or the financial interests of the search engine operator—play no role in how the search engine ranks pages. Search neutrality, I argue, is impossible. Its impossibility seems to threaten the significance of search bias: if no search engine is neutral, then every search engine is biased. To defuse this threat, I distinguish two forms of bias—*failing-on-its-own-terms bias* and *exogenous-values bias*. This distinction allows us to make sense of search bias, and capture its normative complexion, despite the impossibility of neutrality.

1 Introduction

In 2005, Adam and Shivaun Raff started a small business: Foundem, a comparison-shopping site similar to Google Shopping. Foundem showed promise. At one point, it was hailed the U.K.'s best comparison-shopping site. But on June 26, 2006, Google changed its search algorithm, dropping foundem.com from the top three search results to the 70s. By all indications, Foundem's drop in Google's rankings was not due to a drop in quality. Foundem.com still held a top place in Yahoo's and Microsoft's search rankings. But in the search engine optimization industry, it's said that if you want to

[1]For many conversations and comments on earlier drafts, thank you to Laura Alvarez Jubete, Serena Booth, David Boylan, Carol Brown, Thomas Byrne, Medb Corcoran, Ray Eitel-Porter, Sina Fazelpour, Nilanjan Das, Kevin Dorst, Dalia Gala, David Grant, Lyndal Grant, Sally Haslanger, Abby Jaques, Matt Mandelkern, Silvia Milano, Carina Prunkl, Bernhard Salow, Jen Semler, Kieran Setiya, Jack Spencer, Charlotte Unruh, Kate Vredenburgh, and audiences at Northeastern University, the Massachusetts Institute of Technology, the Philosophy in an Inclusive Key Summer Institute, the Annual Meeting of the PPE Society, the University of Oxford, the University of Washington, and the University of Wisconsin. Thank you to Ginger Schultheis for extensive and transformative feedback. Thank you especially to Marion Boulicault, Quinn White, and the members of the Jain Family Institute's Digital Ethics and Governance team for years of inspiration, guidance, and encouragement.

[2]This project was supported by a gift from Accenture.

bury a body, you put it on the second page of Google. Foundem was no exception. It would not recover from the loss of traffic from Google (Manthorpe, 2018).

In one way, Foundem's story is unremarkable: Foundem alleged that they were victims of *algorithmic bias*, and it's well documented that algorithmic bias is pervasive in search engines and algorithms more generally. For example, in 2017 the European Union found that Google's search engine was biased in its own favor; Google Shopping undeservedly enjoyed higher search rankings than rival comparison shopping services, including Foundem (European Commission, 2017). (The result was a €2.42 billion fine.) Search engines are biased in other ways, too. Introna and Nissenbaum (2000) argued that the technical architecture of search engines excludes the voices of the less powerful and less wealthy. Noble (2012, 2019) revealed how search engines perpetuate sexism and racism by returning highly sexualized results for queries like 'Black girls'. We find bias infecting algorithmic systems of all kinds—for example, predictive policing systems that overestimate crime in communities of color (Lum and Isaac, 2016); hiring algorithms that dock qualified female candidates (Barocas and Selbst, 2016); and facial recognition software that struggles to recognize dark-skinned female faces (Buolamwini and Gebru, 2018).

In another way, though, Foundem's story *is* remarkable: following Foundem's demotion in Google's search rankings, its founders initiated the search neutrality movement, which calls for search engines to be, well, neutral. *Algorithmic neutrality* has received little attention, despite the considerable work that's been devoted to algorithmic bias. (Algorithmic *fairness* has received significant attention—see e.g. (Castro, 2019) and (Hedden, 2021)—but how fairness and neutrality relate to one another is, quite generally, far from clear. For example, neutrality, as I'll characterize, is a descriptive notion, while fairness is a normative one.)

Algorithmic neutrality is the subject of this paper. I take up three questions. What is algorithmic neutrality? Is algorithmic neutrality possible? When we have algorithmic neutrality in mind, what can we learn about algorithmic bias?

To answer these questions in concrete terms, I will work with a case study: search engines. Search engines warrant special attention because they themselves are remarkable in discussions of algorithmic bias. Search neutrality, in addition to being a particularly rich topic, is one of only two sorts of algorithmic neutrality to receive sustained public, academic, and legal attention.[3] (For work on search neutrality, see e.g. (Grimmelmann, 2010), (Crane, 2012), and (Gillespie, 2014).)

The paper takes in turn each of the three questions, as applied to search engines. §2 asks: "what is search neutrality?" In answering, I draw on work about neutrality in science. §3 asks: "is search neutrality in possible?" I answer *no*. §4 and §5 ask: "when we have search neutrality in mind, what can we learn about search bias?" I consider, for example, how to make sense of search bias given that search neutrality is impossible. My accounts of search neutrality and bias are stated in terms of the *aim* of a given search engine; §6, explores the notion of an aim. §7 generalizes my discussion of search engines to algorithmic systems of all kinds.

[3]The other is net neutrality.

2 What is search neutrality?

To characterize search neutrality, I will first characterize a more general kind of neutrality—algorithmic or otherwise—of which I will argue search neutrality is an instance.

Imagine a governmental agency in charge of distributing vaccines among a country's provinces. The law mandates that vaccines be distributed to provinces so as to maximize lives saved. But, motivated by political allegiances, the agency withholds vaccines from provinces that are governed by a certain party. In so doing, it does not distribute vaccines neutrally. Or imagine that you are teaching a course and assigning grades for participation. You don't keep a record of participation, instead going off of memory. You happen to know which students in your class have parents that donate to your university. Despite your best efforts to assign grades simply on the basis of participation you subconsciously inflate the participation grades of students whose parents donate. In so doing, you do not assign grades neutrally. Or imagine a scientist investigating the effects of smoking on cancer. Her research is funded by a tobacco company, and she selectively ignores evidence that would be unfavorable to her funders' financial interests. In so doing, she does not carry out her research neutrally.

To articulate the kind of neutrality common to these cases, it will help to look at the literature on a certain kind of neutrality: neutrality in science.[4] (Or, as it's often put, "value-freedom" in science.) This literature offers a characterization of neutrality in science—stated below—that explains why the cancer researcher violates neutrality. I argue that this characterization can generalize to cover cases of all kinds.

> **Neutrality in science**[5]
> Science is neutral only if non-epistemic values play no role in how core scientific practices are conducted.

For example, science is not neutral in the case of the cancer researcher because a non-epistemic value (her funder's financial interests) plays a role in how she conducts a core scientific practice (gathering evidence).

Just how should this characterization of scientific neutrality be understood? For my purposes, the key notion is the distinction between epistemic and non-epistemic values. Non-epistemic values are most easily understood in contrast to epistemic values, which are usually understood as values that aim at truth (Steele, 2012). Common examples include empirical adequacy and internal consistency. Non-epistemic values do not aim at truth. Examples include financial interests, political ideologies, and the preservation of human life.[6]

There is much more to be said about scientific neutrality. We could, for example, delve into debates about what counts as a *core* scientific practice. (The basic idea is that scientific neutrality does not seem threatened by non-epistemic values playing a role in how scientists choose research projects, but it does seem threatened by such values

[4]In connecting questions of algorithmic bias and neutrality to the literature on neutrality in science, I follow Dotan (2020), Fazelpour and Danks (2021), and Johnson (fc), among others.

[5]See e.g. (Douglas, 2009).

[6]Some—e.g. Longino (1990)—deny that there is a clean distinction to draw here at all. Whether they're right is orthogonal to my purposes.

playing a role in evidencing gathering. Gathering evidence is in the relevant sense a core scientific practice; choosing a research project is not.) We could also discuss whether *ideal* science is neutral: theorists debate whether it's even possible for science to be neutral—and if so, whether it should be.

To shed light on algorithmic neutrality, we need not address these debates. Rather, what's of interest is the basic notion of neutrality that debates are about: why should scientific neutrality require that non-epistemic values play no role in core scientific practices? I suggest that it's because science, as traditionally understood, *aims at truth* (Reiss and Sprenger, 2020). Truth is science's north star. Scientific neutrality therefore demands that core scientific practices are guided only by that star. If non-epistemic values—values that do not aim at truth—play a role in core scientific practices, those practices deviate from science's aim. Science is thereby not neutral.

The idea, put another way, is that behind this standard characterization of scientific neutrality lies the thesis that epistemic values are *endogenous* to the aim of science, and that non-epistemic values are *exogenous* to it. This suggests a more general characterization notion of neutrality that extends beyond the case of science—a characterization in terms of the aim of a system or practice and values that are either endogenous or exogenous to that aim. This general characterization, I maintain, is the kind of neutrality at issue in the examples I gave above. In the vaccine case, the law aims to distribute vaccines to maximize saving lives. When values that are exogenous to this aim—e.g. political values—play a role in how the agency distributes vaccines, the agency's distribution practice is thereby not neutral. Likewise, your grading policy aims to distribute grades on the basis of participation. When values that are exogenous to this aim—e.g. financial values—play a role in how you give grades, your grading practice is thereby not neutral.

I propose to characterize neutrality in search engines in terms of their aim and the values that are endogenous and exogenous to it. We must then ask: what do search engines aim at? The standard answer is that search engines aim to give *relevant* results (Munton, 2022). As Google CEO Sunday Pichai puts it: 'We provide users with... the most relevant information. And that's our true north' (C-SPAN, 2018). Why think that search engines aim to give relevant results? Because the Internet is vast. When we want to find information or explore a topic online, we often cannot do it on our own, or at least not easily. We often don't know which websites have the information that we're looking for, or which pages on a website would be of help. And so we turn to search engines. We give them queries and, if all goes well, they give us relevant results (exactly what this amounts to I explore in §3).

Some search engines aim not to give results that are simply relevant, but that also meet some further criteria—such as results that are useful or that have true content. From now through §4, I will focus on search engines that aim simply to give relevant results (and to reduce clutter, will often write 'search engines' instead of 'search engines that aim to give relevant results'). Everything I will say generalizes to search engines with other aims. In §5, I discuss such search engines, and in §6, I discuss what it is for a search engine to have an aim. (I should note also that the kinds of search engines that I discuss in this paper are of the traditional style that returns a ranked list of webpages when given a query. However, what I say applies equally well to a newer brand of search engine that uses generative artificial intelligence, directly delivering a response to the query without necessarily providing a list of links.)

Search engines that aim simply at relevance are of special interest because of the pride of place that relevance holds in common understandings of search engines (as we've just seen) and both search neutrality and search bias (as you'll see just below and then in §4.1). For such search engines, we can characterize neutrality like this:

> **Neutrality in search**
> A search engine that aims at relevance is neutral only if values other than relevance play no role in how the search engine ranks pages.

Relevance is endogenous to a search engine that aims to give relevant results; all other values are exogenous.

Consider some examples. Imagine a search engine operator that intentionally ranks pages of its own products above those of its competitors' even when its competitors' pages are more relevant. (This is exactly what the EU fined Google for doing.) A value other than relevance—the exogenous value of the operator's financial interests—plays a role in how the search engine ranks pages. The search engine is not neutral. Or imagine that an exposé has been published about a scandal involving a certain politician, Ms. F. The operators of a search engine are partial to Ms. F's political party and so engineer their search engine to rank the webpage on which the exposé is published far down in the search results for the query 'Ms. F scandal'. A value other than relevance—the exogenous value of a political ideology—plays a role in how the search engine ranks pages. The search engine is not neutral.

Compare how I've characterized search neutrality with how it's normally characterized by search engine operators, their critics, and scholars alike. For example:

> We [i.e. Google] do get concerns [about political bias] across both sides of the aisle. I can assure you we do this [i.e. deliver search results] in a neutral way. And we do this based on a specific keyword, what we are able to assess as the most relevant information. (Google CEO Sundar Pichai in (C-SPAN, 2018))

> Search Neutrality can be defined as the principle that search engines… should have no editorial policies other than that their results be… based solely on relevance. (Search Neutrality, 2009)

> Search neutrality… at its heart is some idea that Internet search engines ought… [to] employ "neutral" search algorithms that determine search result rankings based on some "objective" metric of relevance. (Crane, 2012, p. 1199)

The similarity between these characterizations and mine is no accident. Implicit in them is what I have made explicit: when we take relevance as the aim of search engines, we can—in keeping with the general, aim-based understanding of neutrality—characterize search neutrality in terms of relevance.

3 Is search neutrality possible?

Search neutrality is not possible. This I argue in §3.1. I then develop and reply to objections to my argument in §3.2 and §3.3.

3.1 Search neutrality is impossible

Let me begin by introducing the notion of a *multidimensional* concept. Consider intelligence. Jack, imagine, is quicker-witted than Nashid but worse at solving mathematical problems (Kamp, 1975). Is Jack more intelligent than Nashid? It seems that there may be no good answer to this question. Certainly, along one *dimension* of intelligence, quick-wittedness, he is more intelligent. Along another dimension, ability to solve mathematical problems, he is not. But is he more intelligent, full stop? To answer this question, we must have some way to compare Jack's superior quick-wittedness to Nashid's superior mathematical problem-solving ability—we must have some way to weight the dimensions of intelligence against one another. How, then, should these dimensions be weighted?

Many theorists—such as Kamp (1975) and Sen (1997)—argue that for intelligence, or other multi-dimensional concepts, there is no way that these dimensions should be weighted. In other words, there is no privileged weighting. These theorists conclude that multidimensional concepts therefore generate *incomparability*. Jack is neither more nor less intelligent than Nashid, nor are they equally intelligent. They are incomparable with respect to intelligence. This move—from multidimensionality to incomparability—is made by many, although it's questioned by some,[7] and I will assume that it's legitimate without further argument.

Relevance, I maintain, is a multidimensional concept. To see why, consider a case analogous to Jack and Nashid's. Imagine that you enter 'hurricane' into a search engine. What might pages relevant to this query discuss? There are many candidates. For example: what a hurricane is; whether human-caused global warming has exacerbated the frequency or intensity of hurricanes; natural disasters similar to hurricanes, like tsunamis or tornadoes; the human toll of hurricanes, or how that toll is unequal across racial and socioeconomic lines; hurricanes that are often discussed (for example, in the United States, Hurricane Katrina of 2005); how governments in different parts of the world respond differently to hurricanes; stories about particular people or communities who have been affected by hurricanes. The list could go on.

Now imagine two pages, P_1 and P_2. P_1 discusses in detail how hurricanes form, the nature of forest fires, and the history of Hurricane Hortense of 1984 (a hurricane that had a low death toll and no other particularly remarkable features). P_2 only briefly notes how hurricanes form, the nature of tornadoes, and the history of Hurricane Katrina. I maintain that P_1 is more relevant than P_2 along some dimensions of relevance but not along others.

What are these dimensions? One is a certain kind of *aboutness*. Facts about how hurricanes form, for example, are about hurricanes in a way that facts about, say, the

[7] See e.g. (Dorr et al., fc).

nutritional value of mushrooms are not. Another dimension is a certain kind of popularity. Hurricane Katrina is more often discussed—it's a more popular topic—than Hurricane Hortense, and so pages that discuss Katrina rather than Hortense are, along the dimension of popularity, more relevant. Another dimension still is similarity; pages that discuss tornadoes are more relevant along this dimension than pages that discuss forest fires. The final dimension I'll discuss—there are no doubt more—is a certain informativeness (Roberts, 2012). In particular, the amount of information concerning topics that are about the search term, or that are popular, or that are similar. A given webpage might carry more information concerning hurricanes than another, and in so doing be more relevant than the other along this dimension.

Is P_1 more, less, or equally as relevant as P_2, not merely along a dimension, but full stop? To answer this question, we must have some way to weight the dimensions of relevance against one another. But there is no privileged weighting of these dimensions, just as there is no privileged weighting of the dimensions of intelligence. And so P_1 and P_2 are incomparable with respect to relevance, just as Jack and Nashid are incomparable with respect to intelligence.

We can now give a first-pass statement of my argument. A search engine that aims to rank pages on the basis of relevance cannot rank one page above another if those pages are incomparable with respect to relevance. The best it can do is to rank one page over another on the basis of a given weighting of the dimensions of relevance. But which weighting? One cannot advert to the aim of relevance itself to answer this question since there is no privileged weighting. In other words, the aim of relevance *underdetermines* which weighting to use in ranking pages, and so underdetermines how to rank pages. Values other than relevance must play a role in determining what this weighting is. Values other than relevance must then play a role in ranking pages. Search neutrality is impossible.[8] (My argument here parallels arguments that scientific neutrality is impossible on the grounds that the aim of truth underdetermines how to conduct core scientific practices.[9])

3.2 Objection: the searcher's purposes

An objector might contest my claim that the aim of relevance underdetermines how to weight relevance's dimensions. The objector's reasoning would begin with an observation with which I agree: using search engines is a certain way of inquiring, and in general, what is relevant to an inquiry varies with its purpose (Anderson, 1995). This fact is evocatively evinced by Garfinkel (1981):

> When [infamous bank robber] Willie Sutton was in prison, a priest who was trying to reform him asked him why he robbed banks. 'Well,' Sutton replied, 'that's where the money is.' Clearly there are different... *purposes* shaping the question and answer. [The priest and Sutton] take different things... to stand in need of explanation.(Garfinkel, 1981, p. 21, emphasis mine)

[8] For other arguments that neutral search is somehow impossible or incoherent see e.g. (Grimmelmann, 2010) and (Gillespie, 2014).

[9] See e.g. (Rudner, 1953).

In the case of search, what is relevant to a search query varies with the purpose of the searcher making the query. Imagine two people. The purpose of one aligns with Sutton's; the purpose of the other aligns with the priest's. Each person enters 'Why did Sutton rob banks?' into a search engine. What's relevant to this very same search query differs by the searcher's purpose. For a purpose like Sutton's, the most relevant pages concern issues like the financial windfall, degree of difficulty, and risk of imprisonment in robbing banks rather than, say, grocery stores or movie theaters. For a purpose like the priest's, the most relevant pages concern issues like Sutton's motives, character, or religious background.

We should not simply talk, then, of whether one page is more relevant than another. We should rather talk of whether one page is more relevant than another *given a purpose*. It's for this reason that the statement of my underdetermination argument in §3.1 was only a first pass. I had made a key claim about underdetermination—call it my 'underdetermination claim'—that made no mention of purposes. My underdetermination claim, as stated, was simply that the aim of relevance underdetermines how to weight the dimensions of relevance. This claim, properly stated, is *for at least some purposes*, the aim of relevance underdetermines how to weight the dimensions of relevance given those purposes. On this much, my objector and I agree.

But the objector goes further. She says the *searcher's purposes determine* how to weight the dimensions of relevance, and indeed that every purpose determines how to weight these dimensions. For every purpose, then, there is no underdetermination. Neutral search may be possible after all.

It might be that some purposes do determine how to weight the dimensions of relevance. But not all purposes do. Imagine that a schoolchild searches 'hurricanes' with the purpose of *exploring* the topic of hurricanes. She has just heard about hurricanes for the first time and is simply curious to know more about them; she has no particular aspects or facts about hurricanes in mind. All of the considerations that I appealed to in §3.1—in support of the thesis that the aim of relevance underdetermines how to weight its dimensions for the query 'hurricanes'—apply in the schoolchild's case. They all apply for exploratory purposes.

We have thus established my underdetermination claim, properly stated: for at least some purposes—exploratory purposes—the aim of relevance underdetermines how to weight the dimensions of relevance given those purposes. (I suspect that the purposes of most searches have an exploratory element, but whether they do is inessential to my argument.) Search neutrality remains impossible.

3.3 Objection: randomness

An objector could go along with my underdetermination claim, but resist my conclusion that search neutrality is therefore impossible. Specifically, the objector might say that if one picks *randomly* among the possible weightings, then neutrality may well be possible. It's natural to think that as a general fact, one is neutral if one picks randomly, and not neutral if one does not.

The objection fails. To see why, consider an analogy. Imagine that there is an upcoming election in the U.S. There are two Democratic candidates and four Republican candidates. A certain news outlet can interview only one candidate and it commits to choosing randomly in determining who to interview. But randomly among what? It could pick randomly among individual candidates, giving each candidate an equal—one-in-six—chance to be interviewed. Or it could pick randomly along party lines, giving each party an equal—one-in-two—chance to have a candidate of theirs interviewed. Or it could pick randomly along any other number of lines—gender lines, racial lines, etc.

Let us adopt the objector's view about the connection between randomness and neutrality. Then we conclude that if the news outlet picks randomly among individual candidates, it's neutral among individual candidates but *not* neutral along party lines. (There is a two-in-six chance that a Democratic candidate is interviewed, but a four-in-six chance that a Republican candidate is interviewed.) If the news outlet picks randomly along party lines, it's neutral along party lines but not among individual candidates. (Each Democratic candidate has a one-in-four chance of being interviewed, but each Republican candidate has a one-in-eight chance.[10]) So, in selecting which candidate to interview, the news outlet can be neutral among individual candidates or it can be neutral along party lines, but it's impossible to be neutral in both ways at once.

It will help to state the point in terms of sets. Consider the set of candidates. There are different ways of *partitioning* this set into cells (or subsets). In one partition there are two cells—one is the set of Democratic candidates and the other is the set of Republican candidates. In another partition, each candidate is a member of her own cell. In selecting which candidate to interview, the news outlet can at best be neutral *with respect to a partition*, by randomizing among the cells of the partition. But even if the news outlet is neutral with respect to any given partition, it's impossible for it to be neutral with respect to every partition all at once.

As for the news outlet, so too for search engines. Consider the set of the dimensions of relevance. Just as there are different ways of partitioning the set of candidates, there are different ways of partitioning the set of weightings. Just as the news outlet can be neutral with respect to a partition on the set of candidates by randomizing among the partition's cells, a search engine can be neutral with respect to a partition on the set of weightings by randomizing among the partition's cells. Just as it's impossible for the news outlet to be neutral with respect to every partition on the set of candidates all at once, it's impossible for a search engine to be neutral with respect to every partition on the set of weightings all at once. Search neutrality remains impossible.

4 What can we learn about bias? Two forms of bias

The impossibility of search neutrality seems to threaten the significance of search bias. After all, if no search engine is neutral, then every search engine is biased. And as Antony (1993) provocatively asks when discussing bias in epistemology: "If bias is ubiquitous and ineliminable... what are we complaining about?" (p. 136). Do complaints of search bias rest on a mistake? The answer is no. Below, I distinguish two forms

[10] If we assume, for simplicity, that the Democrats have the same chance as each other and the Republicans have the same chance as one another.

of search bias—*failing-on-its-own-terms bias* and *exogenous-values bias*—and show that complaints about either form are on firm footing.

4.1 Failing-on-its-own-terms bias

To get a grip on failing-on-its-own terms bias, consider an algorithm that aims to rank pages according to how recently they were updated. More recently updated pages are to be ranked above less recently updated ones. This algorithm *fails on its own terms* if how it ranks pages deviates from how recently they were updated. It fails on its own terms if, for example, it ranks a Google Shopping page that was updated yesterday above a page of Foundem that was updated today.

The system is *biased in failing on its own terms* if it fails on its own terms *systematically*. Specifically, the system is biased in failing on its own terms if how it ranks certain kinds of pages systematically deviates from how recently they were updated. (This is an instance of a more general understanding of bias—that we find in other domains, most prominently statistics—according to which bias is a matter of systematic deviation with respect to a baseline (Fazelpour and Danks, 2021).) The system exhibits such bias if, for example, it tends to rank Google Shopping pages above pages of competitors' shopping services even when Google Shopping's pages were less recently updated than its competitors'.

The same goes for a search engine that aims at relevance. Such a search engine fails on its own terms if how it ranks pages deviates from how relevant those pages are. There is subtlety here, since some pages are incomparable with respect to relevance. Recall the two pages, P_1 and P_2, that are incomparable with respect to relevance (for the query 'hurricanes'). If a search engine ranks P_1 above P_2 or *vice versa*, its ranking does not deviate from relevance. Some pages, though, are comparable with respect to relevance. Some pages are simply more relevant than others. (Put one way, some pages are more relevant than others according to every way of weighting the dimensions of relevance.) It's with such pages that we can define deviating from relevance. Take the query 'Foundem'. Foundem.com is more relevant to this query than a Google Shopping page. If a search engine ranks foundem.com below a Google Shopping page for this query, how the search engine ranks those pages deviates from how relevant they are. If such deviations are systematic, they amount to bias:

> **Failing-on-its-own-terms bias**
> A search engine that aims at relevance is biased in failing on its own terms if how it ranks certain kinds of pages deviates systematically from how relevant those pages are.

Failing-on-its-own-terms bias is, I argue, the kind of bias at issue in the prominent accusations of bias I discussed in §1—accusations from the EU, Noble, and Introna and Nissenbaum (2000) (but for the sake of space, I won't discuss Introna and Nissenbaum's concerns here).

When the EU levied its €2.42 billion fine against Google, part of the rationale was that Google systematically placed Google Shopping at the top of the shopping-related search results, while systematically placing rival comparison shopping services on the fourth page and below. Google Shopping may well be more relevant than these other services. But whatever gap there is in relevance between Google Shopping and its competitors, that gap is presumably not so big that it warrants systematically awarding Google Shopping the top search result and relegating competitors to the fourth page and below (European Commission, 2017). How Google's search engine ranks pages of Google's products and those of its competitors systematically deviates from how relevant those pages are. In other words, Google's search engine was biased in failing on its own terms.

Turn now to Noble's concerns about queries like 'Filipina girls'. On July 19, 2022, for example, entering this query (in a "private" browser with no search history) returned sexualized results from all of the most-used search engines in the English-speaking world—Google's, Microsoft's, Yahoo's, and DuckDuckGo's. In one search, for example, the following were three of the four top-ranked pages: "What Philippines Girls like in BED," "Dating a Filipino Woman - Russian brides," "Dating a Filipina Woman: The Complete Guide for Men."

How do such results relate to bias? Recall that which results are relevant to a given query partly depends on the purpose for which the query is made. Imagine two searchers who enter the query, 'Filipina girls'. The first searcher is interested in dating Filipina women; the second is a high school student researching the role of Filipina girls in civil rights movements (to adapt a case from (Noble, 2012)). The relevance of sexualized results differs by the purposes of these searchers. Sexualized results may well be highly relevant given the purpose of the first searcher. But they are patently not highly relevant (if relevant at all) given the purposes of the second searcher (Noble 2012, 2019). So, a search engine that aims at relevance fails on its own terms if it ranks sexualized pages highly for the student's purpose. Such a search engine might return results systematically. If it does so, it's biased in failing on its own terms.

The inevitable non-neutrality of search engines threatens to rob complaints of search bias of normative force. Complaints of failing-on-its-own-terms bias are not subject to this threat. While it's inevitable that search engines are not neutral, it's not inevitable that search engines are biased in failing on their own terms. It's not inevitable, for example, that a search engine returns sexualized results for the query 'Filipina girls'.

We now know that complaints of failing-on-its-own-terms bias have normative force. What is that force? The answer is that such complaints pack a powerful punch. For example, when Noble called the world's attention to the results that search engines deliver for queries like 'Filipina girls', she did not "merely" show that those results reinforced sexism. She also showed, as she herself emphasizes, that the search results reinforced sexism *because* the search engines systematically failed at what they themselves were trying to do—in other words, because the search engines were biased in failing on their own terms.

This fact precludes a certain defense of bias that might be made on behalf of search engines. We can get a grip on the defense by considering the following passage from Grimmelmann (2010):

> Every search result requires both a user to contribute a search query, and

> websites to contribute the content to be ranked. Neither users nor websites are passive participants; both can be wildly, profoundly biased... Some bias is going to leak through [into search results] as long as search engines help users find what they want. And helping users find what they want is such a profound good that one should be skeptical of trying to inhibit it. (pp. 446–7)

While Grimmelmann does not discuss results of the kind Noble discusses, one might defend these results by adapting his well-articulated thought to concern them: "It's well documented that women are perniciously sexualized. That sexualization is a certain form of bias, which manifests itself both in the websites that Google ranks high for queries like 'Filipina girls' and in the users who are interested in those sites. That bias, though, resides entirely in the *users and the websites*, not in the search engine itself. For Google to try to combat this kind of bias by changing its search results would inhibit the profound good that its search engine provides."

This Grimmelmann-inspired defense of bias is a non-starter with failing-on-its-own-terms bias. One can concede that when search engines yield sexualized results for 'Filipina girls' *for purposes relative to which such sites are relevant*, the search engines does provide the good that it promises: giving relevant results. And one can concede that to change search results for such purposes would inhibit that good. The same does not apply *for purposes relative to which such sites are not relevant*. In delivering the results that it does for such purposes, a search engine that aims at relevance fails on its own terms—it does not provide the good that it promises. To change the search results would not inhibit that good. It would promote it.

The Grimmelmann-inspired defense attempts to pass the normative buck to users and websites, locating the bias entirely in them, and not at all in the search engine itself. But the buck cannot be wholly passed. Users and websites are biased, yes, but failing-on-its-own-terms bias is due to what *the search engine itself* is doing—systematically failing on its own terms—not in what searchers or websites are doing. The search engine itself, or rather its operator, must answer for how its results reinforce sexism.

Before concluding §4.1, let me note that failing-on-its-own-terms bias is a purely descriptive notion. Certain instances of failing-on-its-own-terms bias are cause for concern, as we've just seen. But other forms may be harmless or even desirable, as I discuss in §5 and §6.

4.2 Exogenous-values bias

A search engine that aims at relevance is neutral only if values other than relevance play no role in how the search engine ranks pages. Correspondingly:

> **Exogenous-values bias**
> A search engine that aims at relevance is exogenous-values biased if values other than relevance play a role in how the search engine ranks pages.

Because values other than relevance must play a role in how a search engine ranks pages, every search engine is inevitably exogenous-values biased.

What does this mean from the normative point of view? The inevitability of exogenous-values bias is not in and of itself a problem. Exogenous-values bias—like failing-on-its-own-terms bias—is a purely descriptive notion. It's simply a matter of whether certain values play a role in how a search engine ranks pages, not of whether such values playing that role is cause for concern or celebration or neither. But the inevitability of exogenous-values bias does prompt us to ask whether complaints of bias then have any normative force. To answer, we must distinguish two roles that exogenous values may play: they can *override* relevance or they can *complement* it.

Imagine that a search engine ranks a less relevant page above a more relevant one (for some search given some purpose). If it does so because an exogenous value is at play, then that other value overrides relevance. For example, call to mind the case from §2 of the politician Ms. F, who is embroiled in a scandal about which an exposé has been written. To protect Ms. F, a search engine operator that favors her party ranks the page on which the exposé was published far down in search results for the query 'Ms. F scandal'. An exogenous value—a political ideology—overrides relevance. (If the search engine were to rank pages in this way systematically, it would be failing-on-its-own-terms biased. Generally, when exogenous value values systematically override relevance, there is failing-on-its-own-terms bias.)

It's *not* inevitable that exogenous values override relevance. In the case of Ms. F, for example, the search engine could simply rank the highly relevant page, on which the story about Ms. F was published, high in its search rankings. It's no mystery, then, how complaints about values overriding relevance can have normative force.

Because the aim of relevance underdetermines how to weight the dimensions of relevance, exogenous values must play a role in determining how to weight these dimensions. When they do so, they complement relevance. Given that it's inevitable that exogenous values complement relevance, what are we complaining about when we complain about exogenous values complementing relevance? The complaint cannot be *that* exogenous values play this role, but rather *which* values these are. Such a complaint can have normative bite, since it's not inevitable that one value rather than another plays this role. And so the question to be answered in adjudicating complaints of inevitable exogenous-values bias is simply: which values should complement relevance?

To answer this question would take us beyond the scope of this paper, but I will point to where we might look in developing an answer. First, to those, such as Anderson (1995), who are concerned with the analogous question of which values should complement epistemic values in conducting scientific practices. Second, to authors, such as Mager (2023), who discuss what values, other than relevance, should play a role in how search engine deliver results. And third, to the various fields—such as value-sensitive design, responsible research and innovation, and design justice—that concern the role that values should play in building technologies in general.[11]

[11] See (Fazelpour and Danks, 2021) for similar points.

5 What can we learn about bias? On the normative significance of a search engine's aim

As I noted in §2, some search engines aim to give results that are not only relevant but that also meet some further criteria. A search engine might aim to give results that are not just relevant but that also have true content; or are useful; or that satisfy the searcher's preferences; or some combination of these; or some combination of relevance with other notions still. The fact that search engines differ in their aims brings into view something else we can learn about bias: the normative significance of certain forms of bias for a given search engine is beholden to the normative significance of the search engine's aim. To see why this is so, consider another case study, a search engine that aims to give results that are both relevant and have true content. (To reduce clutter, call such search engines 'relevance-truth engines' and call search engines that aim simply at relevance 'relevance engines'.)

The idea that search engines should give results that are not only relevant, but also have true content, has gained traction in response to the online spread of misinformation. This is because search engines that simply give relevant results accelerate that spread (Bush and Zaheer, 2019). Consider, for example, the question 'did the Holocaust happen?' A theory that the Holocaust did not happen is relevant to this question, just as the actual history of the Holocaust is. (In general, both p and not-p are relevant to the question of whether p (Roberts, 2012).) So, a relevance engine will rank highly pages that discuss Holocaust denial theories. This is in fact exactly what Google's search engine used to do. In December 2016, for example, the top search result for the query 'did the Holocaust happen?' was a page—from the site of the American neo-Nazi group Stormfront—titled 'Top 10 reasons why the holocaust didn't happen' (Roberts, 2016b). The relevance of this page was exactly what Google appealed to in explaining its search results:

> "The fact that hate sites appear in Search results does not mean that Google endorses these views," said the spokesperson in a statement. According to the company, a site's ranking in search results is determined by computer algorithms using hundreds of factors to calculate a page's relevance to a given query. (Roberts, 2016b)

Because Stormfront's page was relevant to the query, its high place in search results was the correct result, if a disturbing one, for a relevance engine.

In contrast, a relevance-truth engine would not rank Stormfront's site so high, since it's shot-through with falsehoods. Google's search engine no longer ranks Stormfront's site so high (or ranks it at all) for this very reason (Roberts, 2016a).

Relevance-truth engines differ from relevance engines in their criteria for bias and neutrality. Consider failing-on-its-own-terms bias, for example. As we've said, a relevance engine is biased in failing on its own terms if how it ranks certain kinds of pages deviates systematically from how relevant those pages are. Likewise, a relevance-truth engine is biased in failing on its own terms if how it ranks certain kinds of pages deviates systematically from how relevant *and true* the content of those pages is.

How a search engine ranks pages may amount to bias if the search engine has one aim but not if it has another. Imagine a search engine that *systematically* ranks pages lowly that carry misinformation, even when those pages are relevant to a query. If the search engine is a relevance engine, then it's biased in failing on its own terms: how it ranks relevant pages that carry misinformation systematically deviates from how relevant those pages are. But if the search engine is a relevance-truth engine, then it's not biased in failing on its own terms. As we've just said, assigning a low rank to pages that are relevant but carry misinformation need not deviate from how relevant and true the content of those pages is.

Because how a search engine ranks pages may amount to bias if the search engine has one aim but not if it has another, some questions about whether bias is worth avoiding amount to questions about what aim a search engine should have. For example, would it be cause for concern that a relevance engine is biased in failing on its own terms if it systematically assigns a low rank to pages that are relevant but carry misinformation? This is a question, ultimately, of the extent to which the aim of relevance alone is worth pursuing in the first place. It is a question of whether the terms of a relevance engine are worth succeeding on.

6 What's in an aim?

The aim of a given search engine is central in my accounts of search neutrality and bias. So far, I have worked with a rough and ready understanding of a search engine's aim. This understanding needs refining. We should not talk simply of *the* aim of a search engine, as I have been doing. We can, for example, distinguish between a search engine's *intended* aim, what the search engine's operator intends the search engine to do, and its a *stated* aim, what the operator says that the search engine is doing. Imagine a search engine operator that claims that its search engine is designed to return relevant results, but which secretly serves the interests of a political party, the S Party. The system's stated aim is to return relevant results. Its intended aim is to return relevant results except when relevant results would harm the S Party.

We can characterize neutrality and bias for each of these aims, and in turn, ask and answer all of the same questions I've posed in the preceding sections about neutrality and bias. Take, for example, my characterization of failing-on-its-own-terms-bias. A search engine that aims at relevance fails on its own terms if how it ranks pages systematically deviates from how relevant those pages are. This characterization—and likewise for my characterizations of neutrality and exogenous-values bias—can apply if the aim under consideration is the intended aim or its stated aim. And then if the aim, either intended or stated, is something other than simply relevance, then neutrality and bias can be characterized in terms of that aim. For example, a search engine that aims to give results that are relevant except when relevant results would hurt the interest of the S Party is biased in failing on its own terms if how it ranks pages deviates from how relevant those pages are except when it would harm the S Party.

Whether a complaint of bias is apt differs depending on which aim is at issue. Imagine that the secretly politicized search systematically assigns very low rankings to pages that are highly relevant but that would harm the S Party. In so doing, it systematically fails on the terms of its stated aim. A complaint of failing-on-its-own-terms bias with respect to its stated aim is therefore warranted: members of the public, governments, or corporate competitors can rightfully complain that the search engine is biased. Not so for a complaint of bias with respect to its intended aim. In systematically assigning very low rankings to pages that are highly relevant but that would harm the S Party, the search engine *succeeds* on the terms of its intended aim; with respect to its intended aim, it's not failing-on-its-own-terms biased.

In other cases, a complaint of bias can be apt with respect to the intended aim but not with respect to the stated aim. Imagine that the secretly politicized search engine systematically, and unintentionally, does *not* assign low rankings to pages that are highly relevant but would harm the female members of the S Party. Such a search engine systematically fails at its intended aim. It's thereby failing-on-its-own-terms biased with respect to that aim. And this bias can engender complaints. Female complaints of the S Party, it's natural to think, would have standing to complain the search engine systematically does not afford them the protection that it does for male members of the party. The search engine, though, is not thereby biased with respect to its stated aim. How ranks highly relevant pages that would harm female candidates of the S Party does not systematically deviate (or deviate at all) from how relevant those pages are.

A search engine's intended aim and its stated aim may not be the only aims we can attribute to it. If there were a law, for example, that required search engines to give simply relevant results, the *legally-required aim* of a search engine subject to that law would be to deliver simply relevant results. My ambition is not to catalogue what kinds of aims a given search engine might have. Rather, I am concerned to point out that different kinds of bias are indexed to different aims, and so that in evaluating claims of bias, we must have in view the aim to which that bias is indexed. The same goes not just for failing-on-its-own-terms bias, but also exogenous-values bias, and indeed also for claims about neutrality.

I what follows, I will for simplicity return to talking simply of *the* aim of a given algorithmic system, with the understanding that what I say applies to the system's intended and stated aims, or any other aim that one might attribute to it.

7 Generalizing

This paper is animated by three questions: What is algorithmic neutrality? Is algorithmic neutrality possible? When we have algorithmic neutrality in mind, what can we learn about algorithmic bias? In §2–6, I discussed these questions in the setting of search engines. Here in §7, I show how my discussion generalizes to algorithmic systems of all kinds.

For a given search engine, I used the aim of that search as the central notion with which to characterize neutrality and bias. For any given algorithmic system, I propose to likewise use the aim of that system as the central notion with which to characterize neutrality and bias.

Consider an algorithm for college admissions. Assume that the algorithm aims to rank candidates on the basis of merit. *What is neutrality for the admissions algorithm?* The algorithm is neutral only if values other than merit play no role in how the algorithm ranks candidates. *Is neutrality possible for the admissions algorithm?* No. Merit underdetermines how to rank candidates, since merit—like relevance—is (one would think) a multidimensional concept. And so values other than merit must play a role in ranking candidates. Neutrality is impossible. *When we have neutrality for the admissions algorithm in mind, what can we learn about bias?* Bias comes in at least two kinds, each with its own normative complexion. The algorithm is biased in failing on its own terms if how it ranks certain kinds of candidates deviates systematically from how much merit those candidates have. (For example, the algorithm might systematically rank female candidates lower than male candidates who have less merit.) The algorithm is exogenous-values biased if values other than merit play a role in how the algorithm ranks candidates.

More generally, take any given algorithmic system. *What is neutrality for the system?* The system is neutral only if no values—other than the system's aim—play a role in how the system delivers its results.

Is it possible for the system to be neutral? The answer is "no" if the system's aim underdetermines how to deliver the system's results. We have seen that underdetermination arises if the system's aim is multidimensional. Underdetermination has many other sources too, of which I will canvas some. Because underdetermination is pervasive, neutrality is impossible for many—if not most—algorithmic systems.

Underdetermination may arise if a system's aim involves an *arbitrary threshold*.[12] Consider an algorithm—similar to the Family Screening Tool used by the Department of Human Services in Allegheny, Pennsylvania (Allegheny County, 2022)—for use in a foster care system. The algorithm, imagine, is used to identify whether it's safe for a child in foster care to return to their original family or guardians. In particular, the aim of the system is to categorize children as at a low, medium, or high risk of being abused if they were to return. How likely must abuse be for a child to be categorized as at high risk? 10%? 20%? 21%? 50%? In other words, what is the threshold of high risk? Likewise, what are the thresholds for low risk and medium risk? The aim of categorizing children at low, medium, or high risk itself underdetermines what these thresholds are, and so underdetermines how to categorize children.

Underdetermination may arise if the system has *multiple aims*.[13] Imagine an algorithm for use in pre-trial detention decisions in the U.S. judicial system, along the lines of those that are in fact widely used (Angwin et al., 2016). Such decisions are supposed to be based on two factors: if the defendant is released, whether they will commit a crime (likelihood of recidivism) and whether they will fail to appear for a future court appearance (likelihood of flight). Our algorithm assigns a defendant a single score—based on likelihood of recidivism and likelihood of flight—that is supposed to represent their aptness for pretrial detention. In other words, the algorithm aims to assign scores on the basis of *both* the likelihood of recidivism and the likelihood of flight. These two

[12] Johnson (fc) makes a similar point.

[13] See (Fazelpour and Danks, 2021) for a similar point.

aims underdetermine how to assign scores to candidates. Imagine that one defendant has slightly higher risk of recidivism than another while having a slightly lower risk of flight. Which defendant should receive a higher risk score, or should they receive the same score? This is a matter of how to weight likelihood of recidivism and likelihood of risk against each other. To resolve it, we cannot appeal to the dual aims of predicting recidivism risk and predicting flight risk, since these two aims "disagree" with one another. The aims of predicting recidivism risk and flight risk therefore underdetermine how to assign scores to defendants.

Underdetermination may arise from other sources still. Dotan (2020) and Johnson (fc), among others, show how various scientific practices have direct analogues in algorithmic systems. (Dotan, for example, discusses how questions of scientific theory choice also arise in developing machine learning algorithms.) After all, algorithmic systems often aim to get at the truth—for example, they aim to predict whether someone will commit a crime or fail to appear for a court date. In §2, I noted that many claim that the aim of truth underdetermines how to conduct certain scientific practices. Suppose such claims are true. Then, Dotan and Johnson show, also true will be analogous claims that the aim of truth in an analogous algorithmic system underdetermines how that system delivers its results.

Turn now to our third question: *What can neutrality teach us about bias in a given algorithmic system?* The system can be exogenous-values biased—with values other than the system's aim either overriding that aim or complementing it. The system can also be biased in failing on its own terms—such bias is only worth avoiding if those terms are worth succeeding on.

In closing, I will recap and synthesize what we gain from my characterizations of algorithmic neutrality and bias. For contrast, consider authors, like Dotan and Johnson, who are concerned the distinction between epistemic and non-epistemic values in algorithmic systems. Such authors are usually concerned with the important question of whether non-epistemic values play certain roles in algorithmic systems (just as many in the philosophy of science literature are concerned with the question of whether non-epistemic values play a role in core scientific practices).

My concern is different. I am not concerned with the epistemic–non-epistemic distinction *per se*, but rather with the endogenous–exogenous distinction. These two distinctions are not the same. For example, consider my argument in §5 that if truth plays a role in how pages are ranked by a search engine that aims at relevance, then that search engine is thereby not neutral. For such a system, an epistemic value (truth) is exogenous. However, when the aim of an algorithmic system is to get at the truth, the two distinctions perfectly align. For such a system, epistemic values are endogenous and non-epistemic values are exogenous.

The endogenous–exogenous distinction does important work, as we've seen throughout the paper. As I argued in §2, the distinction is central to a general characterization of neutrality, of which scientific neutrality and algorithmic neutrality are instances. And, as we also saw in §2, the public, legal, and academic debates over search neutrality are about whether values other than relevance—exogenous values—play a role in how a search engine ranks pages.

The endogenous–exogenous distinction plays a starring role in my characterizations of two of the three notions I've introduced—algorithmic neutrality and exogenous-values bias. In closing, I'll reflect on the third notion, failing-on-its-own-terms bias. By identifying certain forms of bias as a failure on the algorithm's own terms, we can do various things. We can, for example, better respond to *defenses* of algorithmic bias. Some defend bias in a given algorithm—see e.g. the "Grimmelmann-style defense" in §4.1—by maintaining that it's not the algorithm itself that is biased, but rather the world that surrounds it. So, the defense continues, it's not the algorithm itself—or its developer—that's at fault, but rather the biased world. When we identify an algorithm as failing-on-its-own-terms biased, though, it's clear that the algorithm can *criticized on its own terms*. The algorithm, or its developers, unequivocally are at fault: it's not merely the world that is biased, but also the algorithm itself. The notion of failing-on-its-own-terms bias has further use. It makes vivid that certain forms of bias in an algorithmic system are only worth avoiding if the aim of the system is worth pursuing.

References

Allegheny County (2022). The Allegheny family screening tool. https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx.

Anderson, E. (1995). Knowledge, human interests, and objectivity in feminist epistemology. *Philosophical Topics*, 23(2):27–58.

Angwin, J., Larson, J., Matthu, S., and Kirchner, L. (2016). Machine bias. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. *ProPublica*.

Antony, L. M. (1993). Quine as feminist: The radical import of naturalized epistemology. In Antony, L. M. and Witt, C., editors, *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, pages 110–153. Westview Press.

Barocas, S. and Selbst, A. (2016). Big data's disparate impact. *California Law Review*, 104(671):671—732.

Buolamwini, J. and Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability, and Transparency*, 9:77–91.

Bush, D. and Zaheer, A. (2019). Bing's top search results contain an alarming amount of disinformation. https://cyber.fsi.stanford.edu/io/news/bing-search-disinformation. *Stanford Internet Observatory*.

C-SPAN (2018). Google CEO Sundar Pichai testifies on data collection. https://www.youtube. com/watch?v=WfbTbPEEJxI.

Castro, C. (2019). What's wrong with machine bias. *Ergo: An Open Access Journal of Philosophy*, 6.

Crane, D. (2012). Search neutrality as an antitrust principle. *George Mason Law Review*, 19(5):1199–1209.

Dorr, C., Nebel, J. M., and Zuehl, J. (fc). The case for comparability. *Noûs*.

Dotan, R. (2020). Theory choice, non-epistemic values, and machine learning. *Synthese*, (11):1–21.

Douglas, H. (2009). *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press.

European Commission (2017). Antitrust: Commission fines Google 2.42 billion euro. https://ec.europa.eu/commission/presscorner/detail/en/IP_17_1784.

Fazelpour, S. and Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass*, 16(8).

Garfinkel, A. (1981). *Forms of Explanation: Rethinking the Questions of Social Theory*. Yale University Press.

Gillespie, T. (2014). The relevance of algorithms. In Gillespie, T., Boczkowski, P., and Foot, K., editors, *Media Technologies: Essays on Communication, Materiality, and Society*, pages 167–193. MIT Press.

Grimmelmann, J. (2010). Some skepticism about search neutrality. In Szoka, B. and Marcus, A., editors, *The Next Digital Decade: Essays on the Future of the Internet*, pages 435–459. TechFreedom.

Hedden, B. (2021). On statistical criteria of algorithmic fairness. *Philosophy and Public Affairs*, 49(2):209–231.

Introna, L. and Nissenbaum, H. (2000). Shaping the web: Why the politics of search engines matter. *The Information Society*, 16(3):98–122.

Johnson, G. (fc). Are algorithms value-free? Feminist theoretical virtues in machine learning. *Journal of Moral Philosophy*.

Kamp, J. W. (1975). Two theories about adjectives. In Keenan, E. L., editor, *Formal semantics of natural language*. Cambridge University Press.

Longino, H. E. (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press.

Lum, K. and Isaac, W. (2016). To predict and serve? *Significance*, 13(5):14–19.

Mager, A. (2023). European search? How to counter-imagine and counteract hegemonic search with european search projects. *Big Data & Society*, pages 1–13.

Manthorpe, R. (2018). Google's nemesis: meet the British couple who took on a giant, won... and cost it €2.1 billion. https://www.wired.co.uk/article/fine-google-competition-eu-shivaun-adam-raff. *Wired*.

Munton, J. (2022). Answering machines: How to (epistemically) evaluate a search engine. *Inquiry*.

Noble, S. (2012). Missed connections: What search engines say about women. *Bitch Media*, (54):37–41.

Noble, S. (2019). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.

Reiss, J. and Sprenger, J. (2020). Scientific Objectivity. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*.

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69.

Roberts, J. J. (2016a). Google demotes Holocaust denial and hate sites in update to algorithm. https://fortune.com/2016/12/20/google-algorithm-update/. *Fortune*.

Roberts, J. J. (2016b). A top Google result for the Holocaust is now a white supremacist site. https://fortune.com/2016/12/12/google-holocaust/. *Fortune*.

Rudner, R. (1953). The scientist qua scientist makes value judgments. *Philosophy of Science*, 20(1):1–6.

Search Neutrality (2009). Making the case for search neutrality. http://www.searchneutrality.org/search-neutrality.

Sen, A. (1997). *On Economic Inequality*. Clarendon Press.

Steele, K. S. (2012). The scientist qua policy advisor makes value judgments. *Philosophy of Science*, 79(5):893–904.