Article

Superintelligence Skepticism as a Political Tool

Seth D. Baum [iD]

Global Catastrophic Risk Institute, P.O. Box 40364, Washington, DC 20016, USA; seth@gcrinstitute.org

check for updates

Abstract: This paper explores the potential for skepticism about artificial superintelligence to be used as a tool for political ends. Superintelligence is AI that is much smarter than humans. Superintelligence does not currently exist, but it has been proposed that it could someday be built, with massive and potentially catastrophic consequences. There is substantial skepticism about superintelligence, including whether it will be built, whether it would be catastrophic, and whether it is worth current attention. To date, superintelligence skepticism appears to be mostly honest intellectual debate, though some of it may be politicized. This paper finds substantial potential for superintelligence skepticism to be (further) politicized, due mainly to the potential for major corporations to have a strong profit motive to downplay concerns about superintelligence and avoid government regulation. Furthermore, politicized superintelligence skepticism is likely to be quite successful, due to several factors including the inherent uncertainty of the topic and the abundance of skeptics. The paper's analysis is based on characteristics of superintelligence and the broader AI sector, as well as the history and ongoing practice of politicized skepticism on other science and technology issues, including tobacco, global warming, and industrial chemicals. The paper contributes to literatures on politicized skepticism and superintelligence governance.

Keywords: artificial intelligence; superintelligence; skepticism

1. Introduction

The purpose of this paper is to explore the potential for skepticism about artificial superintelligence to be used for political ends. Artificial superintelligence (for brevity, henceforth just *superintelligence*) refers to AI that is much smarter than humans. Current AI is not superintelligent, but the prospect of superintelligence is a topic of much discussion in scholarly and public spheres. Some believe that superintelligence could someday be built, and that, if it is built, it would have massive and potentially catastrophic consequences. Others are skeptical of these beliefs. While much of the existing skepticism appears to be honest intellectual debate, there is potential for it to be politicized for other purposes.

In simple terms (to be refined below), *politicized skepticism* can be defined as public articulation of skepticism that is intended to achieve some outcome other than an improved understanding of the topic at hand. Politicized skepticism can be contrasted with *intellectual skepticism*, which seeks an improved understanding. Intellectual skepticism is essential to scholarly inquiry; politicized skepticism is not. The distinction between the two is not always clear; statements of skepticism may have both intellectual and political motivations. The two concepts can nonetheless be useful for understanding debates over issues such as superintelligence.

There is substantial precedent for politicized skepticism. Of particular relevance for superintelligence is politicized skepticism about technologies and products that are risky but profitable, henceforth *risk–profit politicized skepticism*. This practice dates to 1950s debates over the link between tobacco and cancer and has since been dubbed the *tobacco strategy* [1]. More recently, the strategy has been applied to other issues including the link between fossil fuels and acid rain, the link between fossil fuels and global warming, and the link between industrial chemicals and neurological disease [1,2]. The essence of the strategy is to

promote the idea that the science underlying certain risks is unresolved, and therefore the implicated technologies should not be regulated. The strategy is typically employed by an interconnected mix of industry interests and ideological opponents of regulation. The target audience is typically a mix of government officials and the general public, and not the scientific community.

As is discussed in more detail below, certain factors suggest the potential for superintelligence to be a focus of risk–profit politicized skepticism. First and foremost, superintelligence could be developed by major corporations with a strong financial incentive to avoid regulation. Second, there already exists a lot of skepticism about superintelligence, which could be exploited for political purposes. Third, as an unprecedented class of technology, it is inherently uncertain, which suggests that superintelligence skepticism may be especially durable, even within apolitical scholarly communities. These and other factors do not guarantee that superintelligence skepticism will be politicized, or that its politicization would follow the same risk–profit patterns as the tobacco strategy. However, these factors are at least suggestive of the possibility.

Superintelligence skepticism may also be politicized in a different way: to protect the reputations and funding of the broader AI field. This form of politicized skepticism is less well-documented than the tobacco strategy, and appears to be less common. However, there are at least hints of it for fields of technology involving both grandiose future predictions and more mundane near-term work. AI is one such field of technology, in which grandiose predictions of superintelligence and other future AI breakthroughs contrast with more modest forms of near-term AI. Another example is nanotechnology, in which grandiose predictions of molecular machines contrast with near-term nanoscale science and technology [3].

The basis of the paper's analysis is twofold. First, the paper draws on the long history of risk–profit politicized skepticism. This history suggests certain general themes that may also apply to superintelligence. Second, the paper examines characteristics of superintelligence development to assesses the prospect of skepticism being used politically in this context. To that end, the paper draws on the current state of affairs in the AI sector, especially for artificial general intelligence, which is a type of AI closely related to superintelligence. The paper further seeks to inform efforts to avoid any potential harmful effects from politicized superintelligence skepticism. The effects would not necessarily be harmful, but the history of risk–profit politicized skepticism suggests that they could be.

This paper contributes to literatures on politicized skepticism and superintelligence governance. Whereas most literature on politicized skepticism (and similar concepts such as denial) is backward-looking, consisting of historical analysis of skepticisms that have already occurred [1,2,4–7], this paper is largely (but not exclusively) forward-looking, consisting of prospective analysis of skepticisms that could occur at some point in the future. Meanwhile, the superintelligence governance literature has looked mainly at institutional regulations to prevent research groups from building dangerous superintelligence and support for research on safety measures [8–11]. This paper contributes to a smaller literature on the role of corporations in superintelligence development [12] and on social and psychological aspects of superintelligence governance [13].

This paper does not intend to take sides on which beliefs about superintelligence are most likely to be correct. Its interest is in the potential political implications of superintelligence skepticism, not in the underlying merits of the skepticism. The sole claim here is that the possibility of politicized superintelligence skepticism is a worthy topic of study. It is worth studying due to: (1) the potential for large consequences if superintelligence is built; and (2) the potential for superintelligence to be an important political phenomenon regardless of whether it is built. Finally, the topic is also of inherent intellectual interest as an exercise in prospective socio-political analysis on a possible future technology.

The paper is organized as follows. Section 2 presents a brief overview of superintelligence concerns and skepticisms. Section 3 further develops the concept of politicized skepticism and surveys the history of risk–profit politicized skepticism, from its roots in tobacco to the present day. Section 4 discusses prospects for politicized superintelligence skepticism. Section 5 discusses opportunities for constructive action. Section 6 concludes.

2. Superintelligence and Its Skeptics

The idea of humans being supplanted by their machines dates to at least the 1863 work of Butler [14]. In 1965, Good presented an early exposition on the topic within the modern field of computer science [15]. Good specifically proposed an "intelligence explosion" in which intelligent machines make successively more intelligent machines until they are much smarter than humans, which would be "the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control" [15] (p. 33). This intelligence explosion is one use of the term *technological singularity*, though the term can also refer to wider forms of radical technological change [16]. The term *superintelligence* refers specifically to AI that is much more intelligent than humans and dates to at least the 1998 work of Bostrom [17]. A related term is *artificial general intelligence*, which is AI capable of reasoning across many intellectual domains. A superintelligent AI is likely to have general intelligence, and the development of artificial general intelligence could be a major precursor to superintelligence. Artificial general intelligence is also an active subfield of AI [18,19].

Superintelligence is notable as a potential technological accomplishment with massive societal implications. The effects of superintelligence could include anything from solving a significant portion of the world's problems (if superintelligence is designed well) to causing the extinction of humans and other species (if it is designed poorly). Much of the interest in superintelligence derives from these high stakes. Superintelligence is also of intellectual interest as perhaps the ultimate accomplishment within the field of AI, sometimes referred to as the "grand dream" of AI [20] (p. 125).

Currently, most AI research is on *narrow AI* that is not oriented towards this grand dream. The focus on narrow AI dates to early struggles in the field to make progress towards general AI or superintelligence. After an initial period of hype fell short, the field went through an "AI winter" marked by diminished interest and more modest expectations [21,22] This prompted a focus on smaller, incremental progress on narrow AI. It should be noted that the term AI winter most commonly refers to a lull in AI in the mid-to-late 1980s and early 1990s. A similar lull occurred in the 1970s, and concerns about a new winter can be found as recently as 2008 [23].

With most of the field focused on narrow AI, artificial general intelligence has persisted only as a small subfield of AI [18]. The AI winter also caused many AI computer scientists to be skeptical of superintelligence, on grounds that superintelligence has turned out to be much more difficult than initially expected, and likewise to be averse to attention to superintelligence, on grounds that such hype could again fall short and induce another AI winter. This is an important historical note because it indicates that superintelligence skepticism has wide salience across the AI computer science community and may already be politicized towards the goal of protecting the reputation of and funding for AI. (More on this below.)

Traces of superintelligence skepticism predate AI winter. Early AI skepticism dates to 1965 work by Dreyfus [24]. Dreyfus [24] critiqued the overall field of AI, with some attention to human-level AI though not to superintelligence. Dreyfus traced this skepticism of machines matching human intelligence to a passage in Descartes' 1637 *Discourse On Method* [25]: "it must be morally impossible that there should exist in any machine a diversity of organs sufficient to enable it to act in all the occurrences of life, in the way in which our reason enables us to act."

In recent years, superintelligence has attracted considerable attention. This has likely been prompted by several factors, including a growing scholarly literature (e.g., [9,19,26–29]), highly publicized remarks by several major science and technology celebrities (e.g., Bill Gates [30], Stephen Hawking [31], and Elon Musk [32]), and breakthroughs in the broader field of AI, which draw attention to AI and may make the prospect of superintelligence seem more plausible (e.g., [33,34]). This attention to superintelligence has likewise prompted some more outspoken skepticism. The following is a brief overview of the debate, including both the arguments of the debate and some biographical information about the debaters. (Biographical details are taken from personal and institutional webpages and are accurate as of the time of this writing, May 2018; they are not necessarily accurate as of the time of the publication of the cited literature.) The biographies can be

politically significant because, in public debates, some people's words carry more weight than others'. The examples presented below are intended to be illustrative and at least moderately representative of the arguments made in existing superintelligence skepticism (some additional examples are presented in Section 4). A comprehensive survey of superintelligence skepticism is beyond the scope of this paper.

2.1. Superintelligence Cannot Be Built

Bringsjord [35] argued that superintelligence cannot be built based on reasoning from computational theory. Essentially, the argument is that superintelligence requires a more advanced class of computing, which cannot be produced by humans or existing AI. Bringsjord is Professor of Cognitive Science at Rensselaer Polytechnic University and Director of the Rensselaer AI and Reasoning Lab. Chalmers [36] countered that superintelligence does not necessarily require a more advanced class of computing. Chalmers is University Professor of Philosophy and Neural Science at New York University and co-director of the NYU Center for Mind, Brain, and Consciousness.

McDermott [37] argued that advances in hardware and algorithms may be sufficient to exceed human intelligence, but not to massively exceed it. McDermott is Professor of Computer Science at Yale University. Chalmers [36] countered that, while there may be limits to the potential advances in hardware and software, these limits may not be so restrictive as to preclude superintelligence.

2.2. Superintelligence Is Not Imminent Enough to Merit Attention

Crawford [38] argued that superintelligence is a distraction from issues with existing AI, especially AI that worsens inequalities. Crawford is co-founder and co-director of the AI Now Research Institute at New York University, a Senior Fellow at the NYU Information Law Institute, and a Principal Researcher at Microsoft Research.

Ng argued that superintelligence may be possible, but it is premature to worry about, in particular because it is too different from existing AI systems. Ng memorably likened worrying about superintelligence to worrying about "overpopulation on Mars" [39]. Ng is Vice President and Chief Scientist of Baidu, Co-Chairman and Co-Founder of Coursera, and an Adjunct Professor of Computer Science at Stanford University.

Etzioni [40] argued that superintelligence is unlikely to be built within the next 25 years and is thus not worth current attention. Etzioni is Chief Executive Officer of the Allen Institute for Artificial Intelligence and Professor of Computer Science at University of Washington. Dafoe and Russell [41] countered that superintelligence is worth current attention even if it would take more than 25 years to build. Dafoe is Assistant Professor of Political Science at Yale University and Co-Director of the Governance of AI Program at the University of Oxford. Russell is Professor of Computer Science at University of California, Berkeley. (An alternative counter is that some measures to improve AI outcomes apply to both near-term AI and superintelligence, and thus it is not essential to debate which of the two types of AI should be prioritized [42].)

2.3. Superintelligence Would (Probably) Not Be Catastrophic

Goertzel [43] argued that superintelligence could be built and is worth paying attention to, but also that superintelligence is less likely to result in catastrophe than is sometimes suggested. Specifically, Goertzel argued that it may be somewhat difficult, but very difficult, to build superintelligence with values that are considered desirable, and that the human builders of superintelligence would have good opportunities to check that the superintelligence has the right values. Goertzel is the lead for the OpenCog and SingularityNET projects for developing artificial general intelligence. Goertzel [43] wrote in response to Bostrom [28], who suggested that, if built, superintelligence is likely to result in catastrophe. Bostrom is Professor of Applied Ethics at University of Oxford and Director of the Oxford Future of Humanity Institute. (For a more detailed analysis of this debate, see [44].)

Views similar to Goertzel [43] were also presented by Bieger et al. [45], in particular that the AI that is the precursor to superintelligence could be trained by its human developers to have safe

and desirable values. Co-authors Bieger and Thórisson are Ph.D. student and Professor of Computer Science at Reykjavik University; co-author Wang is Associate Professor of Computer and Information Sciences at Temple University.

Searle [46] argued that superintelligence is unlikely to be catastrophic, because it would be an unconscious machine incapable of deciding for itself to attack humanity, and thus humans would need to explicitly program it to cause harm. Searle is Professor Emeritus of the Philosophy of Mind and Language at the University of California, Berkeley. Searle [46] wrote in response to Bostrom [28], who arqued that superintelligence could be dangerous to humans regardless of whether it is conscious.

3. Skepticism as a Political Tool

3.1. The Concept of Politicized Skepticism

There is a sense in which any stated skepticism can be political, insofar as it seeks to achieve certain desired changes within a group. Even the most honest intellectual skepticism can be said to achieve the political aim of advancing a certain form of intellectual inquiry. However, this paper uses the term "politicized skepticism" more narrowly to refer to skepticism with other, non-intellectual aims.

Even with this narrower conception, the distinction between intellectual and politicized skepticism can in practice be blurry. The same skeptical remark can serve both intellectual and (non-intellectual) political aims. People can also have intellectual skepticism that is shaped, perhaps subconsciously, by political factors, as well as politicized skepticism that is rooted in honest intellectual beliefs. For example, intellectuals (academics and the like) commonly have both intellectual and non-intellectual aims, the latter including advancing their careers or making the world a better place per whatever notion of "better" they subscribe to. This can be significant for superintelligence skepticism aimed at protecting the reputations and funding of AI researchers.

It should be stressed that the entanglement of intellectual inquiry and (non-intellectual) political aims does not destroy the merits of intellectual inquiry. This is important to bear in mind at a time when trust in science and other forms of expertise is dangerously low [47,48]. Scholarship can be a social and political process, but, when performed well, it can nonetheless deliver important insights about the world. For all people, scholars included, improving one's understanding of the world takes mental effort, especially when one is predisposed to believe otherwise. Unfortunately, many people are not inclined to make the effort, and other people are making efforts to manipulate ideas for their own aims. An understanding of politicized skepticism is essential for addressing major issues in this rather less-than-ideal epistemic era.

Much of this paper is focused on risk–profit politicized skepticism, i.e., skepticism about concerns about risky and profitable technologies and products. Risk–profit politicized skepticism is a major social force, as discussed throughout this paper, although it is not the only form of politicized skepticism. Other forms include politicized skepticism by concerned citizens, such as skepticism about scientific claims that vaccines or nuclear power plants are safe; by religious activists and institutions, expressing skepticism about claims that humans evolved from other species; by politicians and governments, expressing skepticism about events that cast them in an unfavorable light; and by intellectuals as discussed above. Thus, while this paper largely focuses on skepticism aimed at casting doubt about concerns about risky and profitable technologies and products, it should be understood that this is not the only type of politicized skepticism.

3.2. Tobacco Roots

As mentioned above, risk–profit politicized skepticism traces to 1950s debates on the link between tobacco and cancer. Specifically, in 1954, the tobacco industry formed the Tobacco Industry Research Committee, an "effort to foster the impression of debate, primarily by promoting the work of scientists whose views might be useful to the industry" [1] (p. 17). The committee was led by C. C. Little,

who was a decorated genetics researcher and past president of the University of Michigan, as well as a eugenics advocate who believed cancer was due to genetic weakness and not to smoking.

In the 1950s, there was substantial evidence linking tobacco to cancer, but it was not as conclusive of a link as is now available. The tobacco industry exploited this uncertainty in public discussions of the issue. It succeeded in getting major media to often present the issue as a debate between scientists who agreed vs. disagreed in the tobacco–cancer link. Among the media figures to do this was the acclaimed journalist Edward Murrow, himself a smoker who, in tragic irony, later died from lung cancer. Oreskes and Conway speculated that, "Perhaps, being a smoker, he was reluctant to admit that his daily habit was deadly and reassured to hear that the allegations were unproven" [1] (pp. 19–20).

Over subsequent decades, the tobacco industry continued to fund work that questioned the tobacco–cancer link, enabling it to dodge lawsuits and regulations. Then, in 1999, the United States Department of Justice filed a lawsuit against nine tobacco companies and two tobacco trade organizations (United States v. Philip Morris). The US argued that the tobacco industry conspired over several decades to deceive the public, in violation of the Racketeer Influenced and Corrupt Organizations (RICO) Act, which covers organized crime. In 2006, the US District Court for the District of Columbia found the tobacco industry guilty, upheld unanimously in 2009 by the US Court of Appeals. This ruling and other measures have helped to protect people from lung cancer, but many more could have also avoided lung cancer were it not for the tobacco industry's politicized skepticism.

3.3. The Character and Methods of Risk–Profit Politicized Skepticism

The tobacco case provided a blueprint for risk–profit politicized skepticism that has since been used for other issues. Writing in the context of politicized environmental skepticism, Jacques et al. [4] (pp. 353–354) listed four overarching themes: (1) rejection of scientific findings of environmental problems; (2) de-prioritization of environmental problems relative to other issues; (3) rejection of government regulation of corporations and corporate liability; and (4) portrayal of environmentalism as a threat to progress and development. The net effect is to reduce interest in government regulation of corporate activities that may pose harms to society.

The two primary motivations of risk–profit politicized skepticism are the protection of corporate profits and the advancement of anti-regulatory political ideology. The protection of profits is straightforward: from the corporation's financial perspective, the investment in politicized skepticism can bring a substantial return. The anti-regulatory ideology is only slightly subtler. Risk–profit politicized skepticism is often associated with pro-capitalist, anti-socialist, and anti-communist politics. For example, some political skeptics liken environmentalists to watermelons: "green on the outside, red on the inside" [1] (p. 248), while one feared that the Earth Summit was a socialist plot to establish a "World Government with central planning by the United Nations" [1] (p. 252). For these people, politicized skepticism is a way to counter discourses that could harm their political agenda.

Notably, both the financial and the ideological motivations are not inherently about science. Instead, the science is manipulated towards other ends. This indicates that the skepticism is primarily political and not intellectual. It may still be intellectually honest in the sense that the people stating the skepticism are actually skeptical. That would be consistent with author Upton Sinclair's saying that "It is difficult to get a man to understand something when his salary depends upon his not understanding it." The skepticism may nonetheless violate that essential intellectual virtue of letting conclusions follow from analysis, and not the other way around. For risk–profit politicized skepticism, the desired conclusion is typically the avoidance of government regulation of corporate activity, and the skepticism is crafted accordingly.

To achieve this end, the skeptics will often engage in tactics that clearly go beyond honest intellectual skepticism and ordinary intellectual exchange. For example, ExxonMobil has been found to express extensive skepticism about climate change in its public communications (such as newspaper advertisements), but much less skepticism in its internal communications and peer-reviewed publications [7]. This finding suggests that ExxonMobil was aware of the risks of climate change and

misled the public about the risks. ExxonMobil reportedly used its peer-reviewed publications for "the credentials required to speak with authority in this area", including in its conversations with government officials [7] (p. 15), even though these communications may have presented climate change risk differently than the peer-reviewed publications did. (As an aside, it may be noted that the ExxonMobil study [7], published in 2017, has already attracted a skeptic critique by Stirling [49]. Stirling is Communications Manager of the Canadian nonprofit Friends of Science. Both Stirling and Friends of Science are frequent climate change skeptics [50].)

While the skeptics do not publicly confess dishonesty, there are reports that some of them have privately done so. For example, Marshall [51] (p. 180) described five energy corporation presidents who believed that climate change was a problem and "admitted, off the record, that the competitive environment forced them to suppress the truth about climate change" to avoid government regulations. Similarly, US Senator Sheldon Whitehouse, an advocate of climate policy to reduce greenhouse gas emissions, reported that some of his colleagues publicly oppose climate policy but privately support it, with one even saying "Let's keep talking—but don't tell my staff. Nobody else can know" [52] (p. 176). Needless to say, any instance in which skepticism is professed by someone who is not actually skeptical is a clear break from the intellectual skepticism of ordinary scholarly inquiry.

One particularly distasteful tactic is to target individual scientists, seeking to discredit their work or even intimidate them. For example, Philippe Grandjean, a distinguished environmental health researcher, reported that the tuna industry once waged a $25 million advertising campaign criticizing work by himself and others who have documented links between tuna, mercury, and neurological disease. Grandjean noted that $25 million is a small sum for the tuna industry but more than the entire sum of grant funding he received for mercury research over his career, indicating a highly uneven financial playing field [2] (pp. 119–120). In another example, climate scientists accused a climate skeptic of bullying and intimidation and reported receiving "a torrent of abusive and threatening e-mails after being featured on" the skeptic's blog, which calls for climate scientists "to be publicly flogged" [51] (p. 151).

Much of the work, however, is far subtler than this. Often, it involves placing select individuals in conferences, committees, or hearings, where they can ensure that the skeptical message is heard in the right places. For example, Grandjean [2] (p. 129) recounted a conference sponsored by the Electric Power Research Institute, which gave disproportionate floor time to research questioning the health effects of mercury. In another episode, the tobacco industry hired a recently retired World Health Organization committee chair to "volunteer" as an advisor to the same committee, which then concluded to not restrict use of a tobacco pesticide [2] (p. 125).

Another common tactic is to use outside organizations as the public face of the messaging. This tactic is accused of conveying the impression that the skepticism is done in the interest of the public and not of private industry. Grandjean [2] (p. 121) wrote that "organizations, such as the Center for Science and Public Policy the Center for Indoor Air Research or the Citizens for Fire Safety Institute, may sound like neutral and honest establishments, but they turned out to be 'front groups' for financial interests." Often, the work is done by think tanks. Jacques et al. [4] found that over 90% of books exhibiting environmental skepticism are linked to conservative think tanks, and 90% of conservative think tanks are active in environmental skepticism. This finding is consistent with recent emphasis in US conservatism on unregulated markets. (Earlier strands of US conservatism were more supportive of environmental protection, such as the pioneering American conservative Russell Kirk, who wrote that "There is nothing more conservative than conservation" [53].)

3.4. The Effectiveness of Politicized Skepticism

Several broader phenomena help make politicized skepticism so potent, especially for risk–profit politicized skepticism. One is the enormous amounts of corporate money at stake with certain government regulations. When corporations use even a tiny fraction of this for politicized skepticism, it can easily dwarf other efforts. Similarly, US campaign finance laws are highly permissive.

Whitehouse [52] traced the decline in bipartisan Congressional support for climate change policy to the Supreme Court's 2010 *Citizens United* ruling, which allows unlimited corporate spending in elections. However, even without election spending, corporate assets tilt the playing field substantially in the skeptics' favor.

Another important factor is the common journalistic norm of balance, in which journalists seek to present "both sides" of an issue. This can put partisan voices on equal footing with independent science, as seen in early media coverage of tobacco. It can also amplify a small minority of dissenting voices, seen more recently in media coverage of climate change. Whereas the scientific community has overwhelming consensus that climate change is happening, that it is caused primarily by human activity, and that the effects will be mainly harmful, public media features climate change skepticism much more than its scientific salience would suggest [54]. (For an overview of the scientific issues related to climate change skepticism, see [55]; for documentation of the scientific consensus, see [56].)

A third factor is the tendency of scientists to be cautious with respect to uncertainty. Scientists often aspire to avoid stating anything incorrect and to focus on what can be rigorously established instead of discussing more speculative possibilities. Scientists will also often highlight remaining uncertainties even when basic trends are clear. "More research is needed" is likely the most ubiquitous conclusion of any scientific research. This tendency makes it easier for other parties to make the state of the science appear less certain than it actually is. Speaking to this point in a report on climate change and national security, former US Army Chief of Staff Gordon Sullivan states "We seem to be standing by and, frankly, asking for perfectness in science . . . We never have 100 percent certainty. We never have it. If you wait until you have 100 percent certainty, something bad is going to happen on the battlefield" [57] (p. 10).

A fourth factor is the standard, found in some (but not all) policy contexts, of requiring robust evidence of harm before pursuing regulation. In other words, the burden of proof is on those who wish to regulate, and the potentially harmful product is presumed innocent until proven guilty. Grandjean [2] cited this as the most important factor preventing the regulation of toxic chemicals in the US. Such a protocol makes regulation very difficult, especially for complex risks that resist precise characterization. In these policy contexts, the amplification of uncertainty can be particularly impactful.

To sum up, risk–profit politicized skepticism is a longstanding and significant tool used to promote certain political goals. It has been used heavily by corporations seeking to protect profits and people with anti-regulatory ideologies, and it has proven to be a powerful tool. In at least one case, the skeptics were found guilty in a court of law of conspiracy to deceive the public. The skeptics use a range of tactics that deviate from standard intellectual practice, and they exploit several broader societal phenomena that make the skepticism more potent.

4. Politicized Superintelligence Skepticism

4.1. Is Superintelligence Skepticism Already Politicized?

At this time, there does not appear to be any superintelligence skepticism that has been politicized to the extent that has occurred for other issues such as tobacco–cancer and fossil fuels–global warming. Superintelligence skeptics are not running ad campaigns or other major dollar operations. For the most part, they are not attacking the scholars who express concern about superintelligence. Much of the discussion appears in peer-reviewed journals, and has the tone of constructive intellectual discourse. An exception that proves the rule is Etzioni [40], who included a quotation comparing Nick Bostrom (who is concerned about superintelligence) to Donald Trump. In a postscript on the matter, Etzioni [40] wrote that "we should refrain from ad hominem attacks. Here, I have to offer an apology". In contrast, the character attacks of the most heated politicized skepticism are made without apology.

However, there are already at least some hints of politicized superintelligence skepticism. Perhaps the most significant comes from AI academics downplaying hype to protect their field's reputation and funding. The early field of AI made some rather grandiose predictions, which soon

fell flat, fueling criticisms as early as 1965 [24]. Some of these criticisms prompted major funding cuts, such as the 1973 Lighthill report [58], which prompted the British Science Research Council to slash its support for AI. Similarly, Menzies [59] described AI as going through a "peak of inflated expectations" in the 1980s followed by a "trough of disillusionment" in the late 1980s and early 1990s. Most recently, writing in 2018, Bentley [60] (p. 11) derided beliefs about superintelligence and instead urges: "Do not be fearful of AI—marvel at the persistence and skill of those human specialists who are dedicating their lives to help create it. And appreciate that AI is helping to improve our lives every day." (For criticism of Bentley [60], see [61].) This suggests that some superintelligence skepticism may serve the political goal of protecting the broader field of AI.

Superintelligence skepticism that is aimed at protecting the field of AI may be less of a factor during the current period of intense interest in AI. At least for now, the field of AI does not need to defend its value—its value is rather obvious, and AI researchers are not lacking for job security. Importantly, the current AI boom is largely based on actual accomplishments, not hype. Therefore, while today's AI researchers may view superintelligence as a distraction, they are less likely to view it as a threat to their livelihood. However, some may nonetheless view superintelligence in this way, especially those who have been in the field long enough to witness previous boom-and-bust cycles. Likewise, the present situation could change if the current AI boom eventually cycles into another bust—another winter. Despite the success of current AI, there are arguments that it is fundamentally limited [62]. The prospect of a new AI winter could be a significant factor in politicized superintelligence skepticism.

A different type of example comes from public intellectuals who profess superintelligence skepticism based on questionable reasoning. A notable case of this is the psychologist and public intellectual Steven Pinker. Pinker recently articulated a superintelligence skepticism that some observers have likened to politicized climate skepticism [63,64]. Pinker does resemble some notable political skeptics: a senior scholar with an academic background in an unrelated topic who is able to use his (and it is typically a *he*) platform to advance his skeptical views. Additionally, a close analysis of Pinker's comments on superintelligence finds them to be flawed and poorly informed by existing research [65]. Pinker's superintelligence skepticism appears to be advancing a broader narrative of human progress, and may be making the intellectual sin of putting this conclusion before the analysis of superintelligence. However, his particular motivations are, to the present author's knowledge, not documented (It would be especially ironic for Pinker to politicize skepticism based on flawed intellectual reasoning, since he otherwise preaches a message intellectual virtue).

A third type of example of potential politicized superintelligence skepticism comes from the corporate sector. Several people in leadership positions at technology corporations have expressed superintelligence skepticism, including Eric Schmidt (Executive Chairman of Alphabet, the parent company of Google) [66] and Mark Zuckerberg (CEO of Facebook) [67]. Since this skepticism comes the corporate sector, it has some resemblance to risk–profit politicized skepticism and may likewise have the most potential to shape public discourse and policy. One observer postulated that Zuckerberg professes superintelligence skepticism to project the idea that "software is always friendly and tame" and avoid the idea "that computers are intrinsically risky", the latter of which "has potentially dire consequences for Zuckerberg's business and personal future" [67]. While this may just be conjecture, it does come at a time in which Facebook is under considerable public pressure for its role in propagating fake news and influencing elections, which, although unrelated to superintelligence, nonetheless provides an antiregulatory motivation to downplay risks associated with computers.

To summarize, there may already be some politicized superintelligence skepticism, coming from AI academics seeking to protect their field, public intellectuals seeking to advance a certain narrative about the world, and corporate leaders seeking to avoid regulation. However, it is not clear how much superintelligence skepticism is already politicized, and there are indications that it may be limited, especially compared to what has occurred for other issues. On the other hand, superintelligence

is a relatively new public issue (with a longer history in academia), so perhaps its politicization is just beginning.

Finally, it is worth noting that while superintelligence has not been politicized to the extent that climate change has, there is at least one instance of superintelligence being cited in the context of climate skepticism. Cass [68,69] cited the prospect of superintelligence as a reason to not be concerned about climate change. A counter to this argument is that, even if superintelligence is a larger risk, addressing climate change can still reduce the overall risk faced by humanity. Superintelligence could also be a solution to climate change, and thus may be worth building despite the risks it poses. At the same time, if climate change has been addressed independently, then this reduces the need to take risks in building superintelligence [70].

4.2. Prospects for Politicized Superintelligence Skepticism

Will superintelligence skepticism be (further) politicized? Noting the close historical association between politicized skepticism and corporate profits—at least for risk–profit politicized skepticism—an important question is whether superintelligence could prompt profit-threatening regulations. AI is now being developed by some of the largest corporations in the world. Furthermore, a recent survey found artificial general intelligence projects at several large corporations, including Baidu, Facebook, Google, Microsoft, Tencent, and Uber [19]. These corporations have the assets to conduct politicized skepticism that is every bit as large as that of the tobacco, fossil fuel, and industrial chemicals industries.

It should be noted that the artificial general intelligence projects at these corporations were not found to indicate substantial skepticism. Indeed, some of them are outspoken in concern about superintelligence. Moreover, out of 45 artificial general intelligence projects surveyed, only two were found to be dismissive of concerns about the risks posed by the technology [19]. However, even if the AI projects themselves do not exhibit skepticism, the corporations that host them still could. Such a scenario would be comparable to that of ExxonMobil, whose scientists confirmed the science of climate change even while corporate publicity campaigns professed skepticism [7].

The history shows that risk–profit politicized skepticism is not inherent to corporate activity—it is generally only found when profits are at stake. The preponderance of corporate research on artificial general intelligence suggests at least a degree of profitability, but, at this time, it is unclear how profitable it will be. If it is profitable, then corporations are likely to become highly motivated to protect it against outside restrictions. This is an important factor to monitor as the technology progresses.

In public corporations, the pressure to maximize shareholder returns can motivate risk–profit politicized skepticism. However, this may be less of a factor for some corporations in the AI sector. In particular, voting shares constituting a majority of voting power at both Facebook and Alphabet (the parent company of Google) are controlled by the companies' founders: Mark Zuckerberg at Facebook [71] and Larry Page and Sergey Brin at Alphabet [72]. Given their majority stakes, the founders may be able to resist shareholder pressure for politicized skepticism, although it is not certain that they would, especially since leadership at both companies already display superintelligence skepticism.

Another factor is the political ideologies of those involved in superintelligence. As discussed above, risk–profit politicized skepticism of other issues is commonly driven by people with pro-capitalist, anti-socialist, and anti-communist political ideologies. Superintelligence skepticism may be more likely to be politicized by people with similar ideologies. Some insight into this matter can be obtained from a recent survey of 600 technology entrepreneurs [73], which is a highly relevant demographic. The study finds that, contrary to some conventional wisdom, this demographic tends not to hold libertarian ideologies. Instead, technology entrepreneurs tend to hold views consistent with American liberalism, but with one important exception: technology entrepreneurs tend to oppose government regulation. This finding suggests some prospect for politicizing superintelligence skepticism, although perhaps not as much as may exist in other industries.

Further insight can be found from the current political activities of AI corporations. In the US, the corporations' employees donate mainly to the Democratic Party, which is the predominant party of American liberalism and is more pro-regulation. However, the corporations themselves have recently shifted donations to the Republican Party, which is the predominant party of American conservatism and is more anti-regulation. Edsall [74] proposed that this divergence between employees and employers is rooted in corporations' pursuit of financial self-interest. A potential implication of this is that, even if the individuals who develop AI oppose risk–profit politicized skepticism, the corporations that they work for may support it. Additionally, the corporations have recently been accused of using their assets to influence academic and think tank research on regulations that the corporations could face [75,76], although at least some of the accusations have been disputed [77]. While the veracity of these accusations is beyond the scope of this paper, they are at least suggestive of the potential for these corporations to politicize superintelligence skepticism.

AI corporations would not necessarily politicize superintelligence skepticism, even if profits may be at stake. Alternatively, they could express concern about superintelligence to portray themselves as responsible actors and likewise avoid regulation. This would be analogous to the strategy of "greenwashing" employed by companies seeking to bolster their reputation for environmental stewardship [78]. Indeed, there have already been some expressions of concern about superintelligence by AI technologists, and likewise some suspicion that the stated concern has this sort of ulterior motive [79].

To the extent that corporations do politicize superintelligence skepticism, they are likely to mainly emphasize doubt about the risks of superintelligence. Insofar as superintelligence could be beneficial, corporations may promote this, just as they promote the benefits of fossil fuels (for transportation, heating, etc.) and other risky products. Or, AI corporations may promote the benefits of their own safety design and sow doubt about the safety of their rivals' designs, analogous to the marketing of products whose riskiness can vary from company to company, such as automobiles. Alternatively, AI corporations may seek to sow doubt about the possibility of superintelligence, calculating that this would be their best play for avoiding regulation. As with politicized skepticism about other technologies and products, there is no one standard formula that every company always adopts.

For their part, academic superintelligence skeptics may be more likely to emphasize doubt about the mere possibility of superintelligence, regardless of whether it would be beneficial or harmful, due to reputational concerns. Or, they could focus skepticism on the risks, for similar reasons as corporations: academic research can also be regulated, and researchers do not always welcome this. Of course, there are also academics who do not exhibit superintelligence skepticism. Again, there is no one standard formula.

4.3. Potential Effectiveness of Politicized Superintelligence Skepticism

If superintelligence skepticism is politicized, several factors point to it being highly effective, even more so than for the other issues in which skepticism has been politicized.

First, some of the experts best positioned to resolve the debate are also deeply implicated in it. To the extent that superintelligence is a risk, the risk is driven by the computer scientists who would build superintelligence. These individuals have intimate knowledge of the technology and thus have an essential voice in the public debate (though not the only essential voice). This is distinct from issues such as tobacco or climate change, in which the risk is mainly assessed by outside experts. It would be as if the effect of tobacco on cancer was studied by the agronomists who cultivate tobacco crops, or if the science of climate change was studied by the geologists who map deposits of fossil fuels. With superintelligence, a substantial portion of the relevant experts have a direct incentive to avoid any restrictions on the technology, as do their employers. This could create a deep and enduring pool of highly persuasive skeptics.

Second, superintelligence skepticism has deep roots in the mainstream AI computer science community. As noted above, this dates to the days of AI winter. Thus, skeptics may be abundant

even where they are not funded by industry. Indeed, most of the skeptics described above do not appear to be speaking out of any industry ties, and thus would not have an industry conflict of interest. They could still have a conflict of interest from their desire in protect the reputation of their field, but this is a subtler matter. Insofar as they are perceived to not have a conflict of interest, they could be especially persuasive. Furthermore, even if their skepticism is honest and not intended for any political purposes, it could be used by others in dishonest and political ways.

Third, superintelligence is a topic for which the uncertainty is inherently difficult to resolve. It is a hypothetical future technology that is qualitatively different from anything that currently exists. Furthermore, there is concern that its mere existence could be catastrophic, which could preclude certain forms of safety testing. It is thus a risk that defies normal scientific study. In this regard, it is similar to climate change: moderate climate change can already be observed, as can moderate forms of AI, but the potentially catastrophic forms have not yet materialized and possibly never will. However, climate projections can rely on some relatively simple physics—at its core, climate change largely reduces to basic physical chemistry and thermodynamics. (The physical chemistry covers the nature of greenhouse gasses, which are more transparent to some wavelengths of electromagnetic radiation than to others. The thermodynamics covers the heat transfer expected from greenhouse gas buildup. Both effects can be demonstrated in simple laboratory experiments. Climate change also involves indirect feedback effects on much of the Earth system, including clouds, ice, oceans, and ecosystems, which are often more complex and difficult to resolve and contribute to ongoing scientific uncertainty.) In contrast, AI projections must rely on notions of intelligence, which is not so simple at all. For this reason, it is less likely that scholarly communities will converge on any consensus position on superintelligence in the way that they have on other risks such as climate change.

Fourth, some corporations that could develop superintelligence may be uniquely well positioned to influence public opinion. The corporations currently involved in artificial general intelligence research include some corporations that also play major roles in public media. As a leading social media platform, Facebook in particular has been found to be especially consequential for public opinion [80]. Corporations that serve as information gateways, such as Baidu, Google, and Microsoft, also have unusual potential for influence. These corporations have opportunities to shape public opinion in ways that the tobacco, fossil fuel, and industrial chemicals industries cannot. While the AI corporations would not necessarily exploit these opportunities, it is an important factor to track.

In summary, while it remains to be seen whether superintelligence skepticism will be politicized, there are some reasons for believing it will be, and that superintelligence would be an especially potent case of politicized skepticism.

5. Opportunities for Constructive Action

Politicized superintelligence skepticism would not necessarily be harmful. As far as this paper is concerned, it is possible that, for superintelligence, skepticism is the correct view, meaning that superintelligence may not be built, may not be dangerous, or may not merit certain forms of imminent attention. (The paper of course assumes that superintelligence is worth some imminent attention, or otherwise it would not have been written.) It is also possible that, even if superintelligence is a major risk, government regulations could nonetheless be counterproductive, and politicized skepticism could help avoid that. That said, the history of politicized skepticism (especially risk–profit politicized skepticism) shows a tendency for harm, which suggests that politicized superintelligence skepticism could be harmful as well.

With this in mind, one basic opportunity is to raise awareness about politicized skepticism within communities that discuss superintelligence. Superintelligence skeptics who are motivated by honest intellectual norms may not wish for their skepticism to be used politically. They can likewise be cautious about how to engage with potential political skeptics, such as by avoiding certain speaking opportunities in which their remarks would be used as a political tool instead of as a constructive intellectual contribution. Additionally, all people involved in superintelligence debates can insist on

basic intellectual standards, above all by putting analysis before conclusions and not the other way around. These are the sorts of things that an awareness of politicized skepticism can help with.

Another opportunity is to redouble efforts to build scientific consensus on superintelligence, and then to draw attention to it. Currently, there is no consensus. As noted above, superintelligence is an inherently uncertain topic and difficult to build consensus on. However, with some effort, it should be possible to at least make progress towards consensus. Of course, scientific consensus does not preclude politicized skepticism—ongoing climate skepticism attests to this. However, it can at least dampen the politicized skepticism. Indeed, recent research has found that the perception of scientific consensus increases acceptance of the underlying science [81].

A third opportunity is to engage with AI corporations to encourage them to avoid politicizing skepticism about superintelligence or other forms of AI. Politicized skepticism is not inevitable, and while corporate leaders may sometimes feel as though they have no choice, there may nonetheless be options. Furthermore, the options may be especially effective at this early stage in superintelligence research, in which corporations may have not yet established internal policy or practices.

A fourth opportunity is to follow best practices in debunking misinformation in the event that superintelligence skepticism is politicized. There is a substantial literature on the psychology of debunking [81–83]. A debunking handbook written for a general readership [82] recommends: (1) focusing on the correct information to avoid cognitively reinforcing the false information; (2) preceding any discussion of the false information with a warning that it is false; and (3) when debunking false information, also give the correct information so that people are not left with a gap in their understanding of the topic. The handbook further cautions against using the *information deficit model* of human cognition, which proposes that mistaken beliefs can be corrected simply by providing the correct information. The information deficit model is widely used in science communication, but it has been repeatedly found to work poorly, especially in situations of contested science. This sort of advice could be helpful to efforts to counter superintelligence misinformation.

Finally, the entire AI community should insist that policy be made based on an honest and balanced read of the current state of knowledge. Burden of proof requirements should not be abused for private gain. As with climate change and other global risks, the world cannot afford to prove that superintelligence would be catastrophic. By the time uncertainty is eliminated, it could be too late.

6. Conclusions

Some people believe that superintelligence could be a highly consequential technology, potentially even a transformative event in the course of human history, with either profoundly beneficial or extremely catastrophic effects. Insofar as this belief is plausible, superintelligence may be worth careful advance consideration, to ensure that the technology is handled successfully. Importantly, this advance attention should include social science and policy analysis, and not just computer science. Furthermore, even if belief in superintelligence is mistaken, it can nonetheless be significant as a social and political phenomenon. This is another reason for social science and policy analysis. This paper is a contribution to the social science and policy analysis of superintelligence. Furthermore, despite the unprecedented nature of superintelligence, this paper shows that there are important historical and contemporary analogs that can shed light on the issue. Much of what could occur for the development of superintelligence has already occurred for other technologies. Politicized skepticism is one example of this.

One topic not covered in this paper is the prospect of beliefs that superintelligence will occur and/or will be harmful to be politicized. Such a phenomenon could be analogous to, for example, belief in large medical harms from nuclear power, or, phrased differently, skepticism about claims that nuclear power plants are medically safe. The scientific literature on nuclear power finds medical harms to be substantially lower than is commonly believed [84]. Overstated concern (or "alarmism") about nuclear power can likewise be harmful, for example by increasing use of fossil fuels. Similarly, the fossil fuel industry could politicize this belief for its own benefit. By the same logic, belief in

superintelligence could also be politicized. This prospect is left for future research, although much of this paper's analysis may be applicable.

Perhaps the most important lesson of this paper is that the development of superintelligence could be a contentious political process. It could involve aggressive efforts by powerful actors—efforts that not only are inconsistent with basic intellectual ideals, but that also actively subvert those ideals for narrow, self-interested gain. This poses a fundamental challenge to those who seek to advance a constructive study of superintelligence.

References

1. Oreskes, N.; Conway, E.M. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*; Bloomsbury: New York, NY, USA, 2010.
2. Grandjean, P. *Only One Chance: How Environmental Pollution Impairs Brain Development—And How to Protect the Brains of the Next Generation*; Oxford University Press: Oxford, UK, 2013.
3. Selin, C. Expectations and the emergence of nanotechnology. *Sci. Technol. Hum. Values* **2007**, *32*, 196–220. [CrossRef]
4. Jacques, P.J.; Dunlap, R.E.; Freeman, M. The organisation of denial: Conservative think tanks and environmental skepticism. *Environ. Politics* **2008**, *17*, 349–385. [CrossRef]
5. Lewandowsky, S.; Oberauer, K. Motivated rejection of science. *Curr. Dir. Psychol. Sci.* **2016**, *25*, 217–222. [CrossRef]
6. Lewandowsky, S.; Mann, M.E.; Brown, N.J.; Friedman, H. Science and the public: Debate, denial, and skepticism. *J. Soc. Polit. Psychol.* **2016**, *4*, 537–553. [CrossRef]
7. Supran, G.; Oreskes, N. Assessing ExxonMobil's climate change communications (1977–2014). *Environ. Res. Lett.* **2017**, *12*, 084019. [CrossRef]
8. McGinnis, J.O. Accelerating Ai. *Northwest. Univ. Law Rev.* **2010**, *104*, 366–381. [CrossRef]
9. Sotala, K.; Yampolskiy, R.V. Responses to catastrophic AGI risk: A survey. *Phys. Scr.* **2014**, *90*, 018001. [CrossRef]
10. Wilson, G. Minimizing global catastrophic and existential risks from emerging technologies through international law. *VA Environ. Law J.* **2013**, *31*, 307–364.
11. Yampolskiy, R.; Fox, J. Safety engineering for artificial general intelligence. *Topoi* **2013**, *32*, 217–226. [CrossRef]
12. Goertzel, B. The Corporatization of AI Is a Major Threat to Humanity. *H+ Magazine*, 21 July 2017.
13. Baum, S.D. On the promotion of safe and socially beneficial artificial intelligence. *AI Soc.* **2017**, *32*, 543–551. [CrossRef]
14. Butler, S. Darwin among the Machines. *The Press*, 13 June 1863.
15. Good, I.J. Speculations concerning the first ultraintelligent machine. *Adv. Comput.* **1965**, *6*, 31–88.
16. Sandberg, A. An overview of models of technological singularity. In *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*; More, M., Vita-More, N., Eds.; Wiley: New York, NY, USA, 2010; pp. 376–394.
17. Bostrom, N. How Long before Superintelligence? 1998. Available online: https://nickbostrom.com/superintelligence.html (accessed on 18 August 2018).
18. Goertzel, B. Artificial general intelligence: Concept, state of the art, and future prospects. *J. Artif. Gen. Intell.* **2014**, *5*, 1–48. [CrossRef]
19. Baum, S.D. A Survey of Artificial General Intelligence Projects for Ethics, Risk, and Policy. Global Catastrophic Risk Institute Working Paper 17-1, 2017. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3070741 (accessed on 18 August 2018).

20. Legg, S. Machine Super Intelligence. Ph.D. Thesis, University of Lugano, Lugano, Switzerland, 2008.

21. Crevier, D. *AI: The Tumultuous History of the Search for Artificial Intelligence*; Basic Books: New York, NY, USA, 1993.

22. McCorduck, P. *Machines Who Think: 25th Anniversary Edition*; A.K. Peters: Natick, MA, USA, 2004.

23. Hendler, J. Avoiding another AI winter. *IEEE Intell. Syst.* **2008**, *23*, 2–4. [CrossRef]

24. Dreyfus, H. Alchemy and AI. RAND Corporation Document P-3244, 1965. Available online: https://www.rand.org/pubs/papers/P3244.html (accessed on 18 August 2018).

25. Descartes, R. *A Discourse on Method*; Project Gutenberg eBook, 1637. Available online: http://www.gutenberg.org/files/59/59-h/59-h.htm (accessed on 18 August 2018).

26. Chalmers, D.J. The singularity: A philosophical analysis. *J. Conscious. Stud.* **2010**, *17*, 7–65.

27. Eden, A.H.; Moor, J.H.; Soraker, J.H.; Steinhart, E. (Eds.) *Singularity Hypotheses: A Scientific and Philosophical Assessment*; Springer: Berlin, Germany, 2013.

28. Bostrom, N. *Superintelligence: Paths, Dangers, Strategies*; Oxford University Press: Oxford, UK, 2014.

29. Callaghan, V.; Miller, J.; Yampolskiy, R.; Armstrong, S. (Eds.) *The Technological Singularity: Managing the Journey*; Springer: Berlin, Germany, 2017.

30. Rawlinson, K. Microsoft's Bill Gates Insists AI Is a Threat. *BBC*, 29 January 2015.

31. Cellan-Jones, R. Stephen Hawking Warns Artificial Intelligence Could End Mankind. *BBC*, 2 December 2014.

32. Dowd, M. Elon Musk's Billion-Dollar Crusade to Stop the A.I. Apocalypse. *Vanity Fair*, 26 March 2017.

33. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]

34. Silver, D.; Huang, A.; Maddison, C.J.; Guez, A.; Sifre, L.; Van Den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **2016**, *529*, 484–489. [CrossRef] [PubMed]

35. Bringsjord, S. Belief in the singularity is logically brittle. *J. Conscious. Stud.* **2012**, *19*, 14–20.

36. Chalmers, D. The Singularity: A reply. *J. Conscious. Stud.* **2012**, *19*, 141–167.

37. McDermott, D. Response to the singularity by David Chalmers. *J. Conscious. Stud.* **2012**, *19*, 167–172.

38. Crawford, K. Artificial Intelligence's White Guy Problem. *New York Times*, 25 June 2016.

39. Garling, C. Andrew Ng: Why 'Deep Learning' Is a Mandate for Humans, not just Machines. *Wired*, May 2015.

40. Etzioni, O. No, the Experts Don't Think Superintelligent AI Is a Threat to Humanity. *MIT Technology Review*, 20 September 2016.

41. Dafoe, A.; Russell, S. Yes, We Are Worried about the Existential Risk of Artificial Intelligence. *MIT Technology Review*, 2 November 2016.

42. Baum, S.D. Reconciliation between factions focused on near-term and long-term artificial intelligence. *AI Soc.* **2017**. [CrossRef]

43. Goertzel, B. Superintelligence: Fears, promises and potentials. *J. Evol. Technol.* **2015**, *25*, 55–87.

44. Baum, S.D.; Barrett, A.M.; Yampolskiy, R.V. Modeling and interpreting expert disagreement about artificial superintelligence. *Informatica* **2017**, *41*, 419–428.

45. Bieger, J.; Thórisson, K.R.; Wang, P. Safe baby AGI. In Proceedings of the 8th International Conference on Artificial General Intelligence (AGI), Berlin, Germany, 22–25 July 2015; Bieger, J., Goertzel, B., Potapov, A., Eds.; Springer: Cham, Switzerland, 2015; pp. 46–49.

46. Searle, J.R. What your computer can't know. *The New York Review of Books*, 9 October 2014.

47. Nichols, T. *The Death of Expertise: The Campaign against Established Knowledge and Why It Matters*; Oxford University Press: New York, NY, USA, 2017.

48. De Vrieze, J. 'Science wars' veteran has a new mission. *Science* **2017**, *358*, 159. [CrossRef] [PubMed]

49. Stirling, M. Merchants of Consensus: A Public Battle against Exxon. 2017. Available online: https://ssrn.com/abstract=3029939 (accessed on 18 August 2018).

50. Hampshire, G. Alberta Government Cool on Controversial Climate Change Speaker. *CBC News*, 19 January 2018.

51. Marshall, G. *Don't Even Think About It: Why Our Brains Are Wired to Ignore Climate Change*; Bloomsbury: New York, NY, USA, 2014.

52. Whitehouse, S. *Captured: The Corporate Infiltration of American Democracy*; The New Press: New York, NY, USA, 2017.

53. Kirk, R. Conservation activism is a healthy sign. *Baltimore Sun*, 4 May 1970.

54. Boykoff, M.T.; Boykoff, J.M. Balance as bias: Global warming and the US prestige press. *Glob. Environ. Chang.* **2004**, *14*, 125–136. [CrossRef]

55. Baum, S.D.; Haqq-Misra, J.D.; Karmosky, C. Climate change: Evidence of human causes and arguments for emissions reduction. *Sci. Eng. Ethics* **2012**, *18*, 393–410. [CrossRef] [PubMed]

56. Oreskes, N. The scientific consensus on climate change. *Science* **2004**, *306*, 1686. [CrossRef] [PubMed]

57. CNA Military Advisory Board. *National Security and the Threat of Climate Change*; The CNA Corporation: Alexandria, VA, USA, 2007.

58. Lighthill, J. *Artificial Intelligence: A Paper Symposium*; Science Research Council: Swindon, UK, 1973.

59. Menzies, T. 21st-century AI: Proud, not smug. *IEEE Intell. Syst.* **2003**, *18*, 18–24. [CrossRef]

60. Bentley, P.J. The three laws of artificial intelligence: Dispelling common myths. In *Should We Fear Artificial Intelligence? In-Depth Analysis*; Boucher, P., Ed.; European Parliamentary Research Service, Strategic Foresight Unit: Brussels, Belgium, 2018; pp. 6–12.

61. Häggström, O. A spectacularly uneven AI report. *Häggström Hävdar*, 30 March 2018.

62. Marcus, G. Artificial intelligence is stuck. Here's how to move it forward. *New York Times*, 29 July 2017.

63. Bengtsson, B. Pinker is dangerous. *Jag är Här*, 22 October 2017.

64. Häggström, O. The AI meeting in Brussels last week. *Häggström Hävdar*, 23 October 2017.

65. Torres, P. A Detailed Critique of One Section of Steven Pinker's Chapter "Existential Threats" in Enlightenment Now. Project for Future Human Flourishing Technical Report 2, Version 1.2, 2018. Available online: https://docs.wixstatic.com/ugd/d9aaad_8b76c6c86f314d0288161ae8a47a9821.pdf (accessed on 21 August 2018).

66. Clifford, C. Google billionaire Eric Schmidt: Elon Musk is 'exactly wrong' about A.I. because he 'doesn't understand'. *CNBC*, 29 May 2018.

67. Bogost, I. Why Zuckerberg and Musk are fighting about the robot future. *The Atlantic*, 27 July 2017.

68. Cass, O. The problem with climate catastrophizing. *Foreign Affairs*, 21 March 2017.

69. Cass, O. How to worry about climate change. *National Affairs*, Winter 2017; pp. 115–131.

70. Baum, S.D. The great downside dilemma for risky emerging technologies. *Phys. Scr.* **2014**, *89*, 128004. [CrossRef]

71. Heath, A. Mark Zuckerberg's plan to create non-voting Facebook shares is going to trial in September. *Business Insider*, 4 May 2017.

72. Ingram, M. At Alphabet, there are only two shareholders who matter. *Fortune*, 7 June 2017.

73. Broockman, D.; Ferenstein, G.F.; Malhotra, N. *The Political Behavior of Wealthy Americans: Evidence from Technology Entrepreneurs*; Stanford Graduate School of Business: Stanford, CA, USA, 2017; Stanford Graduate School of Business Working Paper, No. 3581.

74. Edsall, T.B. Silicon Valley takes a right turn. *New York Times*, 12 January 2017.

75. Mullins, B. Paying professors: Inside Google's academic influence campaign. *Wall Street Journal*, 15 July 2017.

76. Taplinaug, J. Google's disturbing influence over think tanks. *New York Times*, 30 August 2017.

77. Tiku, N. New America chair says Google didn't prompt critic's ouster. *Wired*, 6 September 2017.

78. Marquis, C.; Toffel, M.W.; Zhou, Y. Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organ. Sci.* **2016**, *27*, 483–504. [CrossRef]

79. Mack, E. Why Elon Musk spent $10 million to keep artificial intelligence friendly. *Forbes*, 15 January 2015.

80. Pickard, V. Media failures in the age of Trump. *Political Econ. Commun.* **2017**, *4*, 118–122.

81. Lewandowsky, S.; Gignac, G.E.; Vaughan, S. The pivotal role of perceived scientific consensus in acceptance of science. *Nat. Clim. Chang.* **2013**, *3*, 399–404. [CrossRef]

82. Cook, J.; Lewandowsky, S. The Debunking Handbook. St. Lucia, Australia: University of Queensland, 2011. Available online: https://skepticalscience.com/Debunking-Handbook-now-freely-available-download.html (accessed on 18 August 2018).

83. Chan, M.P.; Jones, C.R.; Hall Jamieson, K.; Albarracín, D. Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychol. Sci.* **2017**, *28*, 1531–1546. [CrossRef] [PubMed]

84. Slovic, P. The perception gap: Radiation and risk. *Bull. At. Sci.* **2012**, *68*, 67–75. [CrossRef]