

A Quantitative Study of Diversity,  
Equity, and Inclusion in STEM  
Subjects at American Universities

by Mason D. Goad & Bruce R. Chartwell

NATIONAL  
ASSOCIATION  
*of* SCHOLARS

# Ideological Intensification

*A Quantitative Study of Diversity,  
Equity, and Inclusion in STEM  
Subjects at American Universities*

A report by the

NATIONAL  
ASSOCIATION  
*of* SCHOLARS

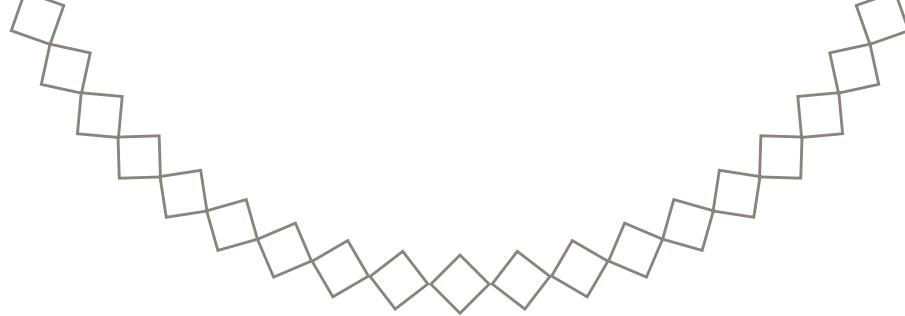
420 Madison Avenue, 7th Floor  
New York, NY 10017

Authors

**Mason Goad**  
Research Associate  
National Association of Scholars

**Bruce R. Chartwell**  
Data Scientist  
National Association of Scholars

Cover Design by Beck&Stone  
Interior Design by Chance Layton  
Published December 2022  
© 2022 National Association of Scholars



# About the National Association of Scholars

## Mission

The National Association of Scholars is an independent membership association of academics and others working to sustain the tradition of reasoned scholarship and civil debate in America's colleges and universities. We uphold the standards of a liberal arts education that fosters intellectual freedom, searches for the truth, and promotes virtuous citizenship.

## What We Do

We publish a quarterly journal, *Academic Questions*, which examines the intellectual controversies and the institutional challenges of contemporary higher education.

We publish studies of current higher education policy and practice with the aim of drawing attention to weaknesses and stimulating improvements.

Our website presents educated opinion and commentary on higher education, and archives our research reports for public access.

NAS engages in public advocacy to pass legislation to advance the cause of higher education reform. We file friend-of-the-court briefs in legal cases defending freedom of speech and conscience and the civil rights of educators and students. We give testimony before congressional and legislative committees and engage public support for worthy reforms.

NAS holds national and regional meetings that focus on important issues and public policy debates in higher education today.

## Membership

NAS membership is open to all who share a commitment to its core principles of fostering intellectual freedom and academic excellence in American higher education. A large majority of our members are current and former faculty members. We also welcome graduate and undergraduate students, teachers, college administrators, and independent scholars, as well as non-academic citizens who care about the future of higher education.

NAS members receive a subscription to our journal *Academic Questions* and access to a network of people who share a commitment to academic freedom and excellence. We offer opportunities to influence key aspects of contemporary higher education.

Visit our website, [www.nas.org](http://www.nas.org), to learn more about NAS and to become a member.

## Our Recent Publications

*Shifting Sands: Flimsy Food Findings*. 2022.

*After Confucius Institutes: China's Enduring Influence on American Higher Education*. 2022.

*Skewed History: Textbook Coverage of Early America and the New Deal*. 2021.

*Climbing Down: How the Next Generation Science Standards Diminish Scientific Literacy*. 2021.

*Priced Out: What College Costs America*. 2021.

*Freedom to Learn: Amending the Higher Education Act*. 2021.

*Rebalancing the Narrative: Higher Education, Border Security, and Immigration*. 2021.

*Disfigured History: How the College Board Demolishes the Past*. 2020.

*Dear Colleague: The Weaponization of Title IX*. 2020.

*Corrupting the College Board: Confucius Institutes and K-12 Education*. 2020.

*Critical Care: Policy Recommendations to Restore American Higher Education after the 2020 Coronavirus Shutdown*. 2020.

# Contents

<b>Abstract</b>	10
<b>Executive Summary</b>	11
<b>Introduction</b>	14
<b>Methodology</b>	15
<b>Dataset Analyses</b>	16
University Websites	16
University Twitter Feeds	23
Academic Associations	26
Scientific Grants and Awards	34
Scientific Literature	42
<b>General Discussion</b>	50
<b>Technical Appendix</b>	51

# Abstract

Diversity, Equity, and Inclusion (DEI) ideology has spread aggressively through institutions of higher education in the United States and throughout the Western world. Arising from within the humanities and social sciences, DEI has begun to influence the natural sciences, where it promises to do great intellectual and economic damage. This report documents and quantifies the growing prevalence of DEI-associated language in science, technology, engineering, and mathematics (STEM) fields in the United States. For this study, five new datasets were assembled—comprising over 30 GB of data in over 280,000 files—focusing on American university webpages, university Twitter accounts, annual programs of academic associations, grants of major scientific research funders, and publications of scientific research, making this the largest study to date of DEI ideology's intensification in STEM fields.

Our findings show that DEI-related language has increased significantly in all STEM sectors over the last decade, and exponentially so within the last few years. DEI indicators linked with STEM have risen 2,600 percent compared to a decade ago on university websites, with similar trends observed in social media content. This increase has been catalyzed, in part, by significant increases in government spending. The two principal funders of scientific research, the National Institutes of Health (NIH) and the National Science Foundation (NSF), increased DEI-related research spending by roughly 300 percent in one year alone (2020–2021). This increase is also visible in scientific publications, with DEI-related language rising up to 4,200 percent between 2010 and 2021.

# Executive Summary

Diversity, Equity, and Inclusion (DEI) ideology has spread rapidly through the universities and colleges of the United States. The most alarming spread has been that which infects the natural sciences, where it undermines the norms of open discourse, objectivity, devotion to evidence, and intellectual independence that constitute robust science. This report is the largest quantitative study of the growth of DEI-related language in the sciences published thus far, with the construction of five original datasets drawn from a variety of online sources. Analyses of these datasets document the growth of DEI-related language over time.

The five datasets are:

1. The websites of the top 100 research universities and colleges in the United States;
2. Twitter feeds of the top 100 research universities and colleges, as well as their affiliated accounts;
3. The annual programs of academic associations in the four major branches of the natural sciences (physics, chemistry, life sciences, and mathematics);
4. Grants by three major funders of scientific research in the United States (the National Institutes of Health, the National Science Foundation, and the Ford Foundation);
5. Scientific publications in four literature databases and preprint repositories (Google Scholar, arXiv, Web of Science, and PubMed).

The analyses and results for each dataset are presented in five sections, along with an explanation of how data were gathered, and the limitations of the individual datasets. We present the results in graphs, tables, and an online animated map, which chart the rise of DEI terminology in the sciences over the last few decades.

To summarize our findings:

- In all datasets, with the exception of the annual programs of academic associations, DEI language has increased significantly over the last decade, accelerating over the last few years in particular.

- Of the 100 university websites we surveyed, the number of webpages that use both STEM and DEI terms grew from 110 in 2010 to 2,891 in 2021. This finding suggests that DEI is being linked with STEM over 26 times more frequently than it was a decade ago.
- DEI ideology appears to be more prominent at elite institutions. As of 2021, Ivy League school webpages linking DEI and STEM were, proportionally, almost double those of non-Ivy League universities.
- DEI or social justice-related language on university Twitter accounts has grown since 2015. Tweets incorporating DEI-related terms were primarily explosions of interest followed by reversion to the mean, rather than the steady growth exhibited in the universities' websites.
- Growth of DEI language in the annual programs of academic associations is mixed, ranging from moderate to no growth, and sometimes negative growth. However, the use of “antiracist”<sup>1</sup> language in annual programs has grown since 2020.
- Both the National Science Foundation and the National Institutes of Health have diverted significantly more funding to DEI-related topics over the last few years. NSF funding to projects with antiracist themes approximately tripled from 2020 to 2021. The number of DEI-related projects receiving NIH funding grew significantly between 2010 and the present, increasing roughly 5 to 1,000 times depending on the DEI term.
- Between 2010 and 2021, scientific publications and preprints that incorporate DEI or antiracist language grew between 3 to 42 times faster than scientific topics in general in the Web of Science. Similar patterns are observed in the data from Google Scholar and PubMed. The number of preprints on arXiv that incorporate DEI or antiracist language has grown significantly in the last two years, with increases varying depending on the DEI or antiracist term.

Some care must be taken when interpreting these results. For instance, while the co-occurrence of DEI and STEM terminology on university websites is 26 times more prevalent than it was a decade ago, the general use of websites also has grown. Some data allowed for appropriate base-rate comparisons, but other forms of data (such as university websites) did not. However, the extremely sharp growth in the co-occurrence of DEI and STEM terms on university websites at a time when DEI was gaining rapid salience in the broader culture suggests that the data still register a strong signal despite this limitation.

---

<sup>1</sup> “Antiracism” is a racist ideology. See: John McWhorter, *Woke Racism: How a New Religion Has Betrayed Black America* (Portfolio, 2021) for specific details. While we do not uniformly put “antiracism” in scare quotes, the readers should note that “antiracism” is a term that means the opposite of what it seems to denote.

Accompanying this report are database and data-mining tools that should allow other scholars and researchers to build upon these results and to foster a broader effort to explore and understand the effect of DEI on the conduct and content of STEM education and research in the United States.<sup>2</sup>

---

<sup>2</sup> Please see the National Association of Scholars' GitHub repository for this project at: <https://github.com/NASorg/quant-dei>; The raw data can be found at: <https://zenodo.org/record/6360904#.Y4HkfMJPZ>; The animated map can be found at: [https://github.com/NASorg/quantdei/blob/master/graphs/dei\\_stem\\_growth\\_map.gif](https://github.com/NASorg/quantdei/blob/master/graphs/dei_stem_growth_map.gif).

# Introduction

Diversity, Equity, and Inclusion (DEI) has become the new trinity of American higher education—and of Western civilization. The new trinity goes by various names: the ironic DIE (Diversity, Inclusion, Equity), the obscure EDI (Equity, Diversity, Inclusion), and sometimes JEDI or EDIJ (the ‘J’ meaning ‘Justice’). No matter the acronym, the ideology is more or less the same, and it is an ideology that undermines science’s core values.

Consider, for example, the consequences of the sciences adopting an ethic of equity. The vice president of the United States, Kamala Harris, has stated plainly that “equitable treatment means we all end up in the same place.”<sup>3</sup> In other words, equity is equality of outcomes. If the sciences were to adopt such an ethic, a scholar who is consistently correct would be considered equal in merit to a scholar who is consistently incorrect. In fact, the very concept of correct or incorrect would have to be discarded, as the terms insinuate that unequal outcomes are possible, and that we will not all end up in the same place.

The scientific enterprise cannot function if it truly adopts this ethic, yet the ethic appears to have spread rapidly throughout the STEM fields. While further critiques and qualitative analyses of DEI, as well as Ibram X. Kendi’s antiracism ideology, may be addressed elsewhere, this report details multiple quantitative analyses of just how far DEI language (as a proxy for DEI ideology) seems to have risen in the STEM fields over the last few decades.

Our study builds on Leif Rasmussen’s *Increasing Politicization and Homogeneity in Scientific Funding: An Analysis of NSF Grants, 1990-2020*. We have broadened the scope of our study to include several other data sources relevant to the sciences.<sup>4</sup> We also build on Zach Goldberg’s studies of what he calls “The Great Awokening.”<sup>5</sup> We examine the same set of topics, but with a specific focus on the STEM fields.

3 Kamala Harris, Twitter Post, 1 November 2020. <https://twitter.com/kamalaharris/status/1322963321994289154?lang=en>.

4 Leif Rasmussen, *Increasing Politicization and Homogeneity in Scientific Funding: An Analysis of NSF Grants, 1990-2020*, The Center for the Study of Partisanship and Ideology (Report No. 4.) 16 November 2021, <https://cspcenter.org/wp-content/uploads/2021/11/CSPI.Report.4.pdf>.

5 Zach Goldberg, “How the Media Led the Great Racial Awakening,” *Tablet Magazine*, 4 August 2020, <https://www.tabletmag.com/sections/news/articles/media-great-racial-awakening>.

# Methodology

For this study, we created and analyzed five new datasets that assess the salience of DEI language in STEM content within the United States. These datasets were analyzed to count the prevalence of terms associated with both DEI and STEM. A general limitation to this study is that quantitative measures of DEI language do not necessarily evaluate the extent to which DEI ideology is influencing the conduct and output of scientific research or education overall. However, it is reasonable to presume that the more these terms come to intermingle with scientific work, the more DEI ideology has influenced the scientific enterprise. The precise nature of DEI's ideological influence on STEM only can be assessed through detailed qualitative study on a case-by-case basis.

Our focus here is almost entirely on quantitative description, such as incidence counts of relevant key words over time. For example, the first dataset comprises HTML files downloaded from nearly 60,000 separate URLs across 100 universities. These files were then scanned for DEI language, and the findings are presented as simple counts. Where possible, we have tried to normalize the data to a relevant correlative measure, such as proportion of incidence on websites compared to total number of websites. Similarly, the number of Tweets using DEI language must be normalized against the growth of universities' Twitter accounts. For grants, we compiled the dollar amounts awarded rather than the number of grants awarded, because dollar amounts are better indicators of institutional priorities than the overall number of grants.

A discussion of each of the datasets follows.

# Dataset Analyses

## University Websites

The first dataset created drew from university websites. Most public organizations now use their websites to communicate with present and future internal and external stakeholders. If DEI ideology was gaining influence in STEM teaching and research at universities in the United States, one would expect to see an increase of DEI language on STEM-related university websites over time. To test the hypothesis that DEI language has increased in STEM, we gathered nearly 60,000 URLs from research university websites. We scanned these for the co-occurrence of at least one STEM term from our dictionary and one DEI term. Our STEM dictionary included the following terms: “astronomy,” “biology,” “chemistry,” “engineering,” “geology,” “mathematics,” “meteorology,” “oceanography,” “physics,” and “science.” Our DEI dictionary included the following terms: “diversity,” “equity,” “inclusion,” “systemic racism,” “white supremacy,” “anti-racism,” “justice,” “privilege,” and “critical race theory.”<sup>6</sup>

## Data Sampling

These scans raised an immediate question: were the data taken from the scans of any actual value in proving that DEI has encroached on STEM? The data only tell us that a DEI term and a STEM term have occurred together on the same university web page, with a single specified date. The best way of measuring the appropriateness of the data was to sample files from the dataset at random and look at what information they truly contain.

Table 1 includes examples of data demonstrating some of the typical dynamics of DEI and STEM language in specific university websites. In each case, the File column contains the university name, the DEI term, and the STEM term (e.g., {universityname\_deiterm\_stemterm}). These files were drawn randomly from the dataset. The Strength column indicates the extent to which STEM and DEI terms intermingled. Strength was assessed as either “Strong” or

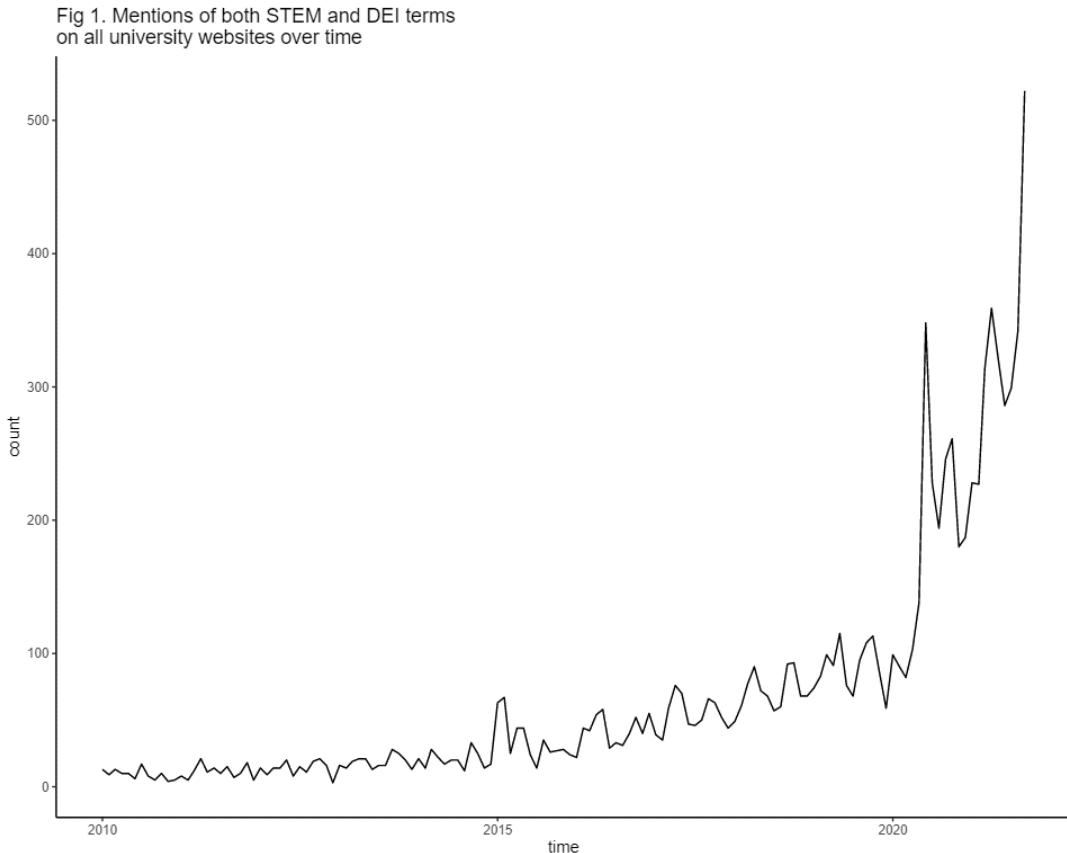
<sup>6</sup> These terms were chosen due to their prevalence within DEI. Naturally, “diversity,” “inclusion,” and “equity” are terms of the DEI trinity itself. “Systemic racism,” “white-supremacy,” “anti-racism,” “privilege,” and “critical race theory” are terms typical of the so-called “antiracism” initiatives. “Justice” may refer to social justice or racial justice.

“Weak.” In those cases where it could not be said that DEI ideology had encroached on STEM, Strength was listed as “None.” The Notes column gives details of assessment for each entry.

**Table 1. Results of sample searches for the 100 universities and colleges.**

File	Notes	Strength
arizona_equity_meteorology	“Equity” appears in the bibliography of a scholar in environmental science.	Weak
duke_critical-race-theory_geology	A blog by a professor discussing both geology and CRT.	Weak
asu_inclusion_engineering	Reference to Apple-funded “Inclusion and Diversity Engineering” scholarships.	Strong
arizona_anti-racism_astronomy	Examines relationship between STEM and the “JEDI” DEI initiative.	Strong
gatech_justice_biology	An employer list.	None
hmc_inclusion_biology	Link to Harvey Mudd College’s Women’s Inclusion in Science, Technology, Engineering & Mathematics application page.	Strong
ucsb_diversity_engineering	Advertisement for a seminar on “Science and Diversity” led by a professor of biomedical engineering.	Strong
tufts_justice_chemistry	A list of STEM research projects based on antiracist or social justice principles, including “Connecting the Dots between Organic Chemistry and Social Justice through Mechanistic Reasoning, and Systemic Racism as a Driver for Math Anxiety.”	Strong
illinois_diversity_meteorology	A landing page for weather data. DEI appears in a link.	None
wisc_anti-racism_engineering	News of a \$5 million award “to fund an interdisciplinary, multi-year project to advance anti-racist practices and pedagogy in science, technology, engineering, mathematics, and medicine.”	Strong

This sample is a telling indication of the kind of data our large-scale web scraping operation surfaced. Six out of ten files we randomly sampled showed strong evidence between DEI initiatives and STEM, while four out of ten files included either weak evidence or no evidence. We discarded nearly 90% of the data (most of which were files with multiple dates) to make the signal as clear as possible. However, noise clearly remains in the dataset. Still, there is no reason to suspect that weak or non-existent evidence increased over time, and even if only 60% of the data in the remaining 10,000 datapoints include strong evidence of DEI’s encroachment on STEM, it would still signal that DEI and STEM topics have been co-occurring with much greater frequency in recent years. Moreover, the findings within the dataset of university websites is congruent with the findings in the other datasets described in the later sections, giving little reason to doubt the conclusions of this particular dataset.



## Results

The growth of correlated DEI and STEM terms was quantified by extracting dates from web pages, excluding every URL which included more than one date. We graphed the remainder on frequency plots. The results show the number of URLs on university websites over time that feature a STEM term, a DEI term, and a single date. Full details are explained in the Technical Appendix. Figure 1 exhibits all mentions of both DEI and STEM terms for all universities and shows relatively steady growth until 2020, when it rapidly increases. This explosion was likely prompted by the Black Lives Matter riots of that summer.

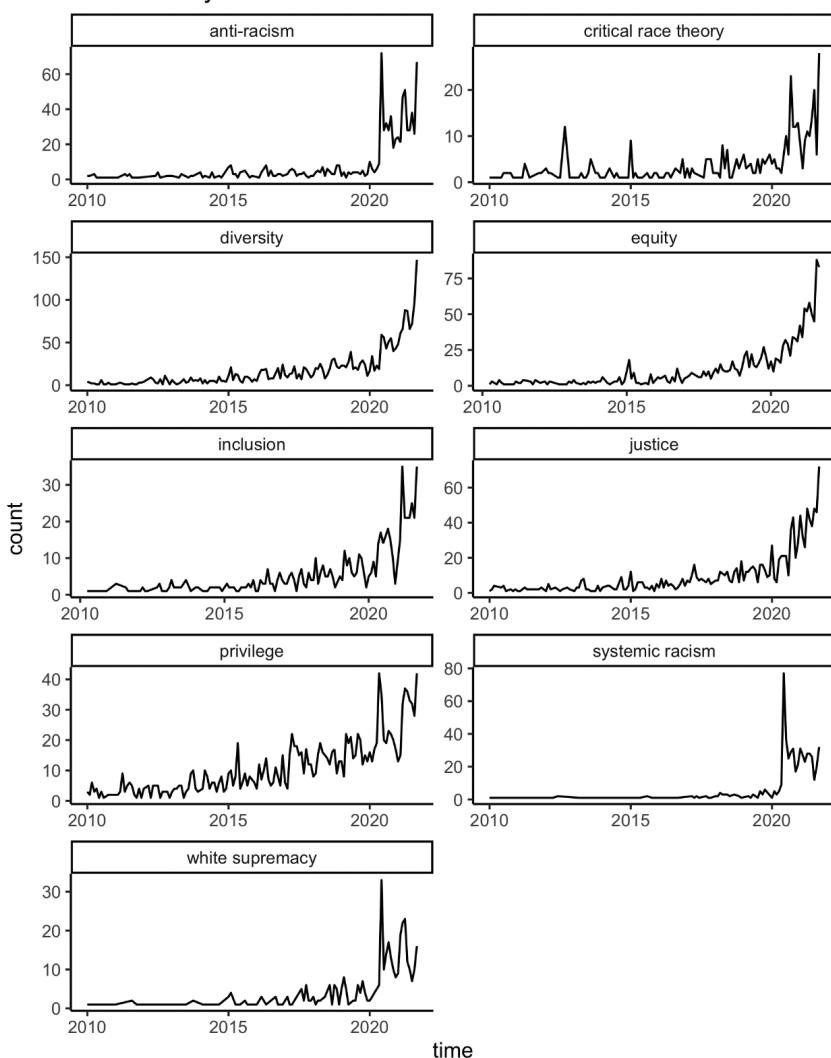
The data for Figure 1 are tallied on a monthly basis. Calculated on a yearly basis, we observe that Ivy League schools,<sup>7</sup> had 409 web pages featuring both DEI and STEM language in 2021, while non-Ivy universities had 2,490. In other words, the co-occurrence of DEI and STEM terms was disproportionately high in the eight Ivy League schools (51 per school) compared with the 92 “non-Ivy” colleges and universities (27 per institution).

<sup>7</sup> The Ivy League schools include: Brown University, Columbia University, Cornell University, Dartmouth College, Harvard University, the University of Pennsylvania, Princeton University, and Yale University.

We then examined how specific DEI and STEM terms operated within these overall trends. Figure 2 shows the co-occurrence of specific DEI terms amid all STEM terms. Figure 3 shows the co-occurrence of specific STEM terms amid all DEI terms. The co-occurrence of specific DEI terms and STEM terms appears to follow a similar pattern to the overall trends: a slow increase from 2010, demarcated by a spike in 2020 leading to a sharp increase.

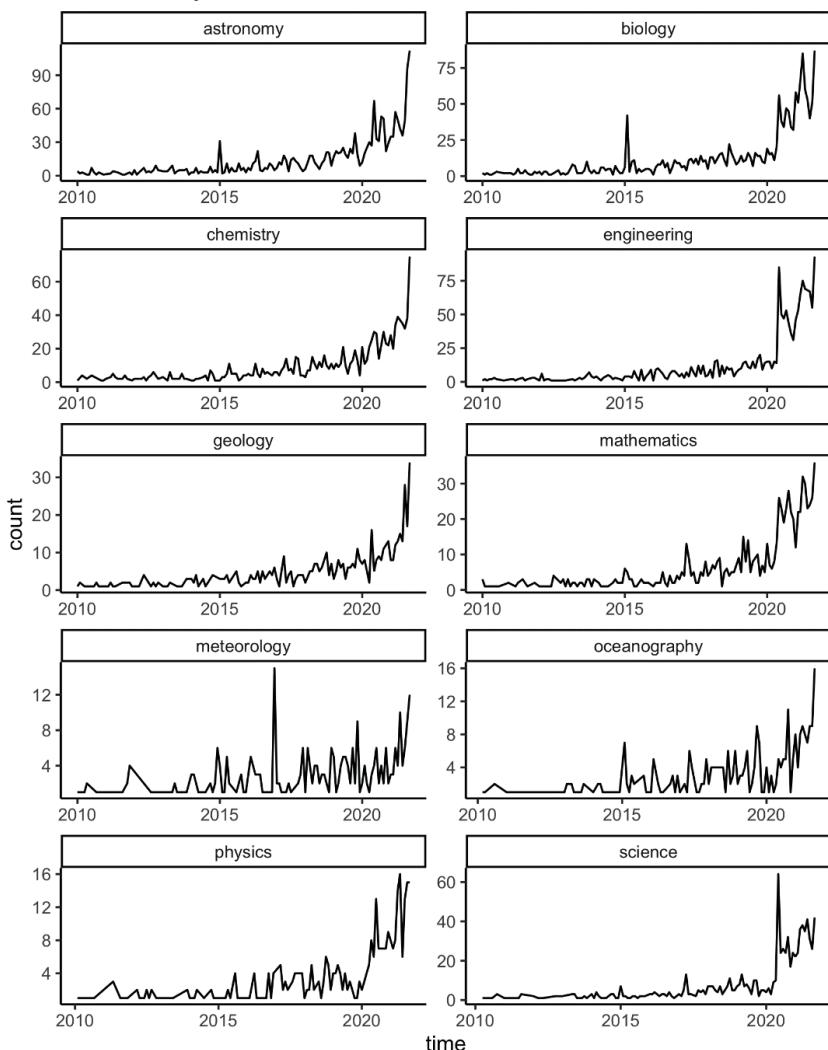
By tracking how the usage of specific DEI terms has spread through all sciences, we observe another interesting pattern. Terms that have long been in use, such as “diversity,” “equity,” or “inclusion,” did not show the 2020 spike, yet nevertheless shared in the dramatic 2020 rise. Terms that have only more recently come into use, such as “critical race theory,”

**Fig 2. Mentions of specific DEI terms alongside all STEM terms on university websites over time**



“systemic racism,” or “white supremacy,” showed the 2020 spike, followed by an increased use. The recent adoption of these terms within the broader DEI umbrella is likely in response to the 2020 Black Lives Matter riots, as seen in the overall trends (Figure 1).<sup>8</sup>

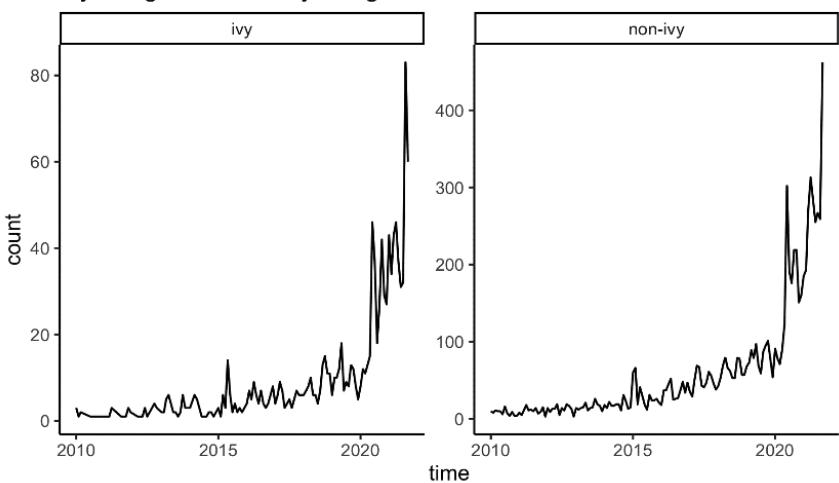
**Fig 3. Mentions of all DEI terms alongside specific STEM terms on university websites over time**



Being that Ivy League schools had a disproportionately high number DEI- and STEM-related webpages compared to non-Ivy universities, we were interested to see if the overall trends differed. The trends do not appear to differ much. Figure 4 shows that each group had a slow increase up to 2020, followed by a rapid rise thereafter.

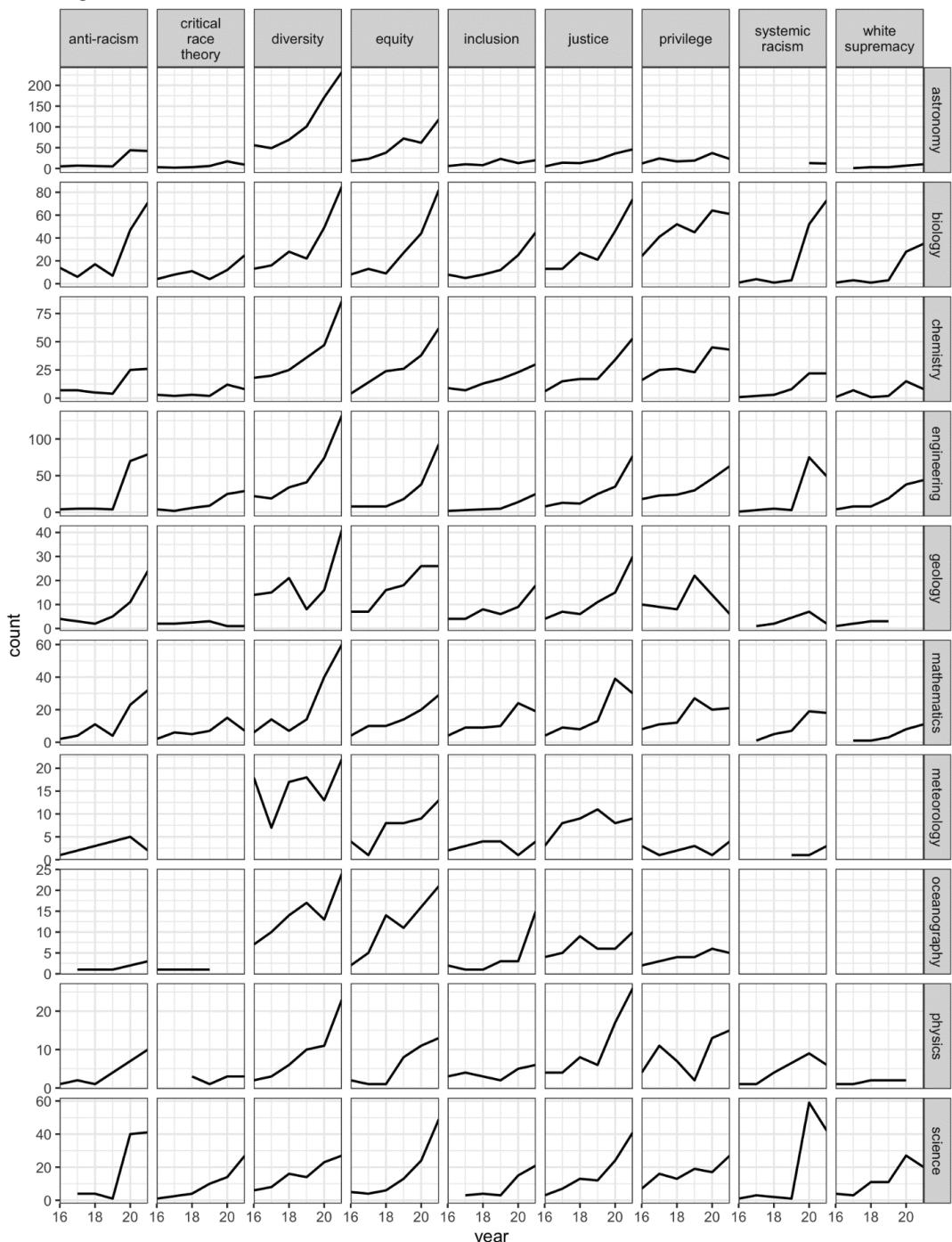
<sup>8</sup> Note: We put the graphs on the same page and let the axes float, meaning that their absolute values differ. This brings to focus what we are truly interested in: the relative trends, rather than the absolute values.

**Fig 4. Mentions of all DEI terms alongside all STEM terms on Ivy League vs non-Ivy League universities over time**



To see if any other patterns might be observed, we then cross-tabulated the STEM terms with the DEI terms. Figure 5 revealed some interesting patterns. In biology, for example, the usage of all DEI terms increased over time. In astronomy, however, the usage of DEI terms increased more selectively: “diversity” and “equity” showed increases similar to other fields, but there was almost no increase in “inclusion,” “justice,” “privilege,” and so on.

Fig 5. Matrix of DEI vs STEM terms at all universities since 2016



Finally, we were interested to see where, geographically, the rise of these co-occurrences had taken place since 1991. Figure 6, presented here, is a static map showing the locations of the 100 universities we surveyed. An animated version of this map, showing where

these co-occurrences have been found over the past 30 years, is available on the National Association of Scholars' GitHub account.<sup>9</sup>

**Fig 6. Geographic spread of co-occurrence of DEI and STEM terms at universities**



## Summary: University Websites

Our large-scale web scraping operation uncovered telling indicators of the growth of DEI language in university webpages. While some noise in the data remains, there is a strong signal that DEI-related language has risen in STEM content on university websites, particularly in recent years. These trends, both in general and specific to certain DEI terms, appear to have been rising steadily until 2020, when we observe a large spike followed by a rapid increase. This spike and subsequent increase are likely in response to the Black Lives Matter riots of 2020.

## University Twitter Feeds

The second dataset we created was from the Twitter feeds of universities, or university-related accounts. Twitter has become a key platform for communication, so we were interested in the degree to which the social media presence of academic institutions has come to adopt DEI-related language. This analysis of Twitter feeds focuses primarily on DEI language in general, not merely the confluence of STEM and DEI. Still, this dataset complements our study of STEM and DEI, as it shows that the increase of DEI in STEM follows a broader trend of increased usage of DEI language by university administrators overall.

<sup>9</sup> For the animated version of Figure 6, please visit: [https://github.com/NASorg/quantdei/blob/master/graphs/dei\\_stem\\_growth\\_map.gif](https://github.com/NASorg/quantdei/blob/master/graphs/dei_stem_growth_map.gif).

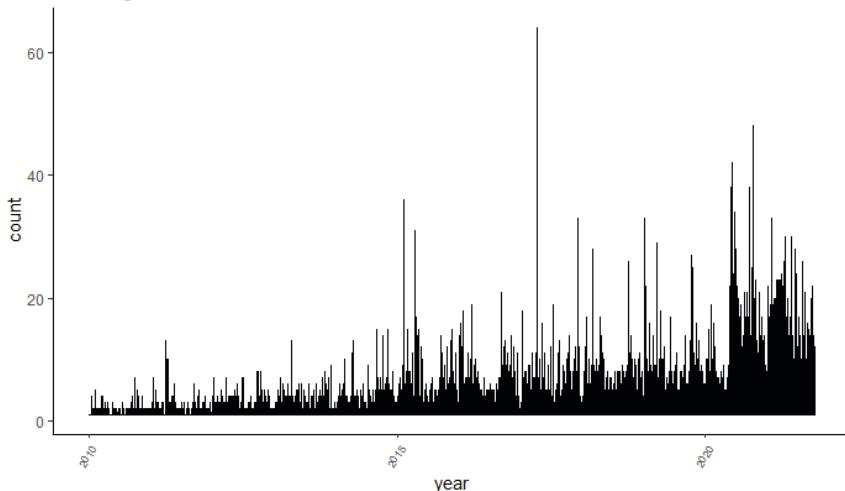
## Methods

We used the Google JSON API<sup>10</sup> to search for the name of the university and “twitter” (i.e., Harvard University twitter, etc.) for the 100 schools. We dumped the raw JSON search results and parsed them for all Twitter accounts identified on the first page of results (a maximum of 10). This allowed us to identify both the primary Twitter account run by the school and up to nine other related accounts. The University of Arizona, for example, has a main university account, alongside related accounts, such as @arizonaalumni, @arizonafball, @dailywildcat, etc. Incorporating related accounts allowed for more detailed statistics and let us see both the depth and breadth of DEI terminology in social media. Parsing out the Twitter accounts from the JSON search results led us to identify 895 unique accounts. We then linked them with their respective schools and combined them with 21 DEI terms. This led to 22,475 individual API calls yielding 151,284 Tweets. These Tweets—each from a school or school-related account, where the Tweet uses a DEI term—formed the dataset.

## Results

Every university primary account we surveyed used DEI terms on its Twitter timeline at some stage during the period under observation. Figure 7 shows all DEI related Tweets over time, and the familiar pattern of a rapid rise beginning in 2020 was evident. DEI-related Tweets show some curious periodic spikes around 2011, 2015, and 2017, which may correlate with political events such as presidential or mid-term elections, but this hypothesis remains untested.

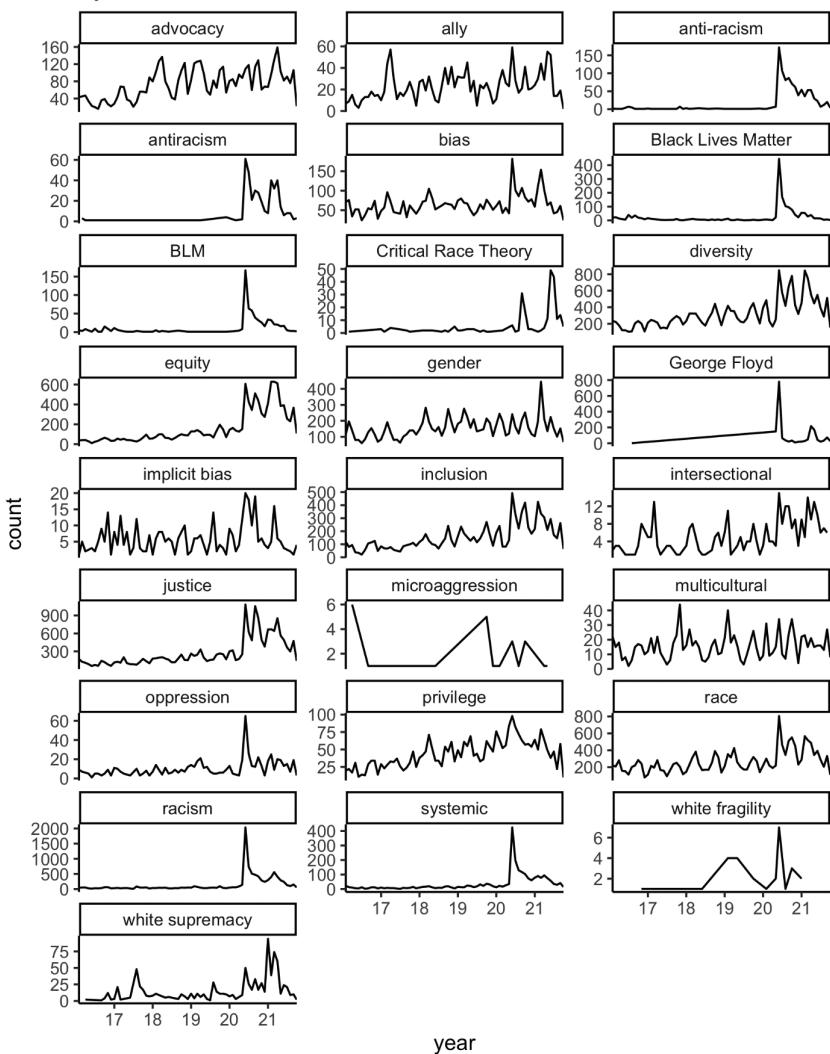
**Fig 7. All DEI-related Tweets from all school-related accounts over time**



<sup>10</sup> The Google JavaScript Object Notation (JSON) Application Programming Interface (API). Please see <https://developers.google.com/custom-search/v1/introduction> for more information.

We were then interested to see the patterns of specific DEI terms amid all university and university-related accounts. Figure 8 shows two specific patterns. The incidence of terms like “advocacy,” “implicit bias,” and “race” was fairly steady over time. For other terms, such as “anti-racism,” “Black Lives Matter,” and “systemic,” the incidence was very low until 2020, with a sharp spike followed by increased usage thereafter.

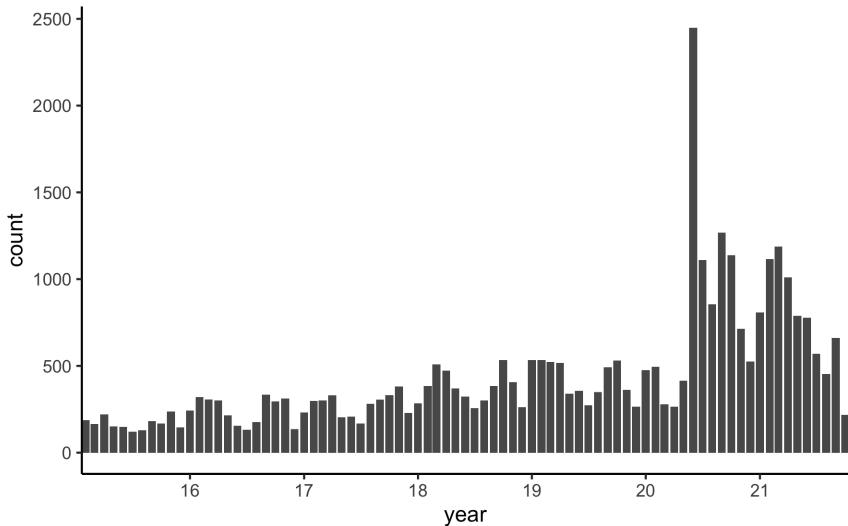
**Fig 8. DEI-related Tweets from all school-related accounts by DEI term**



We were curious whether the use of DEI language on university primary accounts was similar to the use on all accounts (see Figure 7). By excluding these related accounts, we are able to see the trends of DEI language coming directly from the main university

administrations. Figure 9 shows a similar pattern: generally low and steady incidence until the summer of 2020, with a large spike followed by high usage.

**Fig 9. All DEI-related Tweets from university primary accounts**



## Summary: University Twitter Feeds

University Twitter feeds show a similar pattern in the usage of DEI language to universities' web sites: a modest yet steady amount of growth in the usage of DEI language, with a large spike in the summer of 2020. This spike is almost certainly an accompaniment to the Black Lives Matter riots. The dataset did not examine the co-occurrence of DEI language with STEM language on university Twitter accounts, but simply the growth of DEI language over time. Still, the data do reveal where administrators' attention lies on DEI-related topics, with the administrators then serving as a proxy for DEI ideology's encroachment on STEM (as well as every other facet of the university).

## Academic Associations

The growth of DEI language in university websites and Twitter feeds is a proxy for the penetration of DEI ideology into STEM. However, if DEI ideology is truly influencing the output of STEM fields, outside of the universities, we would expect to see it most clearly in three places:

1. the pipeline of new research and thinking as published by academic associations,

2. scientific grants and awards as approved by funders of scientific research, and
3. scientific publications and preprints.

This section details the growth of DEI language in the output of academic associations. Scientific grants and awards are discussed later in Section 4, and scientific publications and preprints are discussed in Section 5.

Regarding academic associations, we chose the four largest we could identify based on membership size and financial resources. The four societies composing our dataset are:

- American Physical Society (APS)
- American Chemical Society (ACS)
- American Mathematical Society (AMS)
- American Society for Biochemistry and Molecular Biology (ASBMB)

## Methods

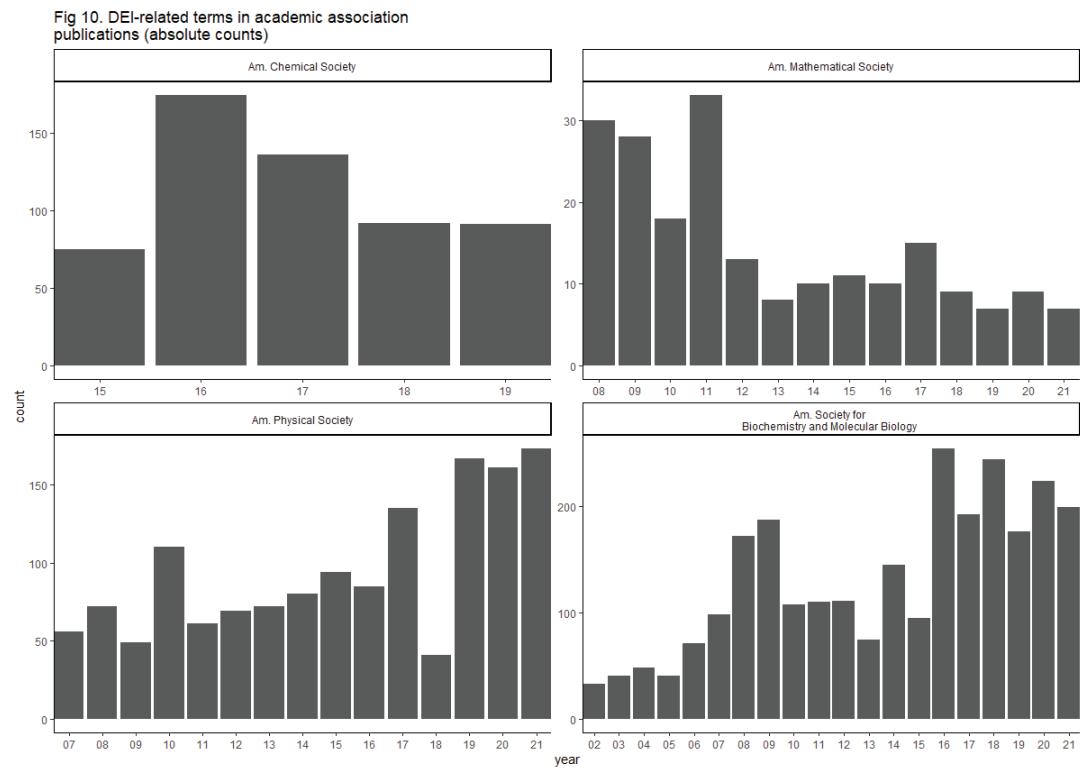
Most scientific associations hold annual conferences where researchers present scientific abstracts of their work. We were able to obtain a collection of annual programs for three of the four societies. In most cases, this collection was of all the years available. In the case of the ACS, however, it was only for ~5 years. In the case of the ASBMB, we used 19 years (~225 PDFs) of their monthly newsletters and conducted text analyses on this archive. Table 2 summarizes the data we collected:

**Table 2. Data obtained from the four professional scientific societies.**

Academic Association	Data Obtained
American Physical Society (APS)	Program of abstracts from the March and April annual scientific meetings from 2007–2021 (30 PDFs)
American Chemical Society (ACS)	Program of abstracts from bi-annual scientific meetings from 2015–2019 (10 PDFs)
American Mathematical Society (AMS)	Annual scientific abstracts from 2008–2021 (14 HTML files)
American Society for Biochemistry and Molecular Biology (ASBMB)	Monthly internal newsletter, ASBMB Today, from April 2002 to October 2021 (225 PDFs)

We conducted a straightforward dictionary analysis, counting the appearance of each DEI term. The dictionary of DEI-related terms used for this analysis included: “1619 project,” “advocacy,” “ally,” “anti-racism,” “antiracism,” “bias,” “Black Lives Matter,” “BLM,” “Critical

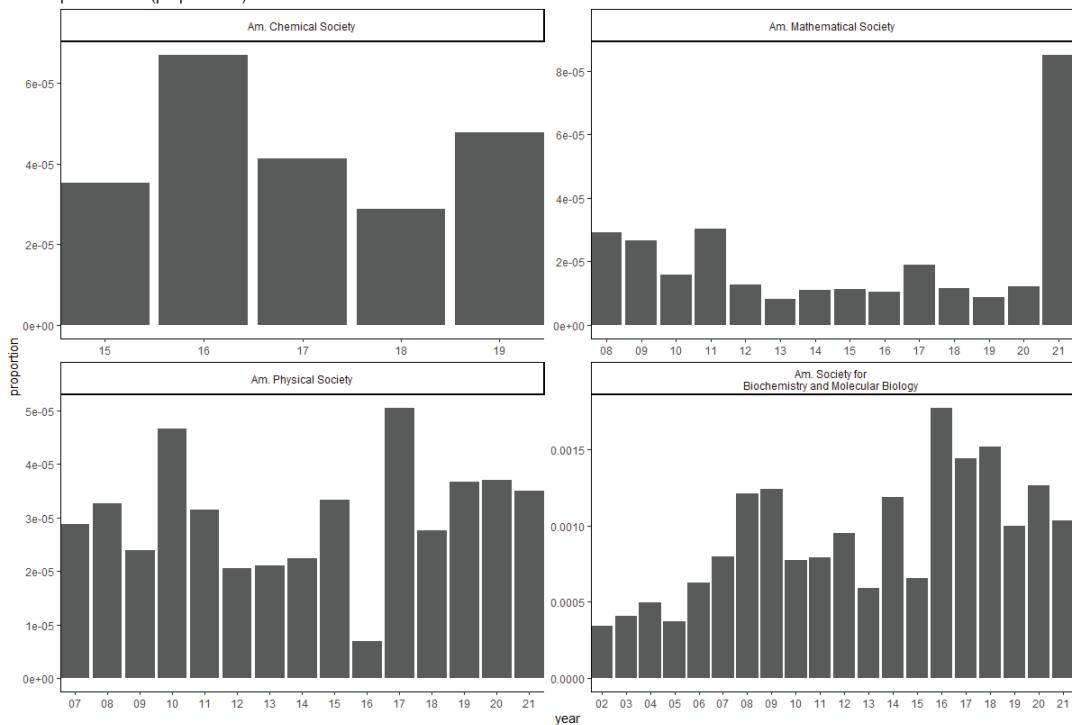
Race Theory,” “discrimination,” “diverse,” “diversity,” “equity,” “gender,” ‘George Floyd,’ “inequality,” “implicit bias,” “indigenous,” “Indigenous,” “inclusion,” “intersectional,” “justice,” “Kendi,” “microaggression,” “multicultural,” “oppression,” “privilege,” “race,” “racism,” “racial,” “racist,” “social justice,” “systemic,” “transgender,” “trans,” “white fragility,” and “white supremacy.” The R files associated with this analysis are described in the Technical Appendix.



## Results

We did not see a consistent pattern of DEI language increasing over time. Figure 10 exhibits the absolute counts of DEI terms, which shows downward trends for the ACS and AMS, but increasing trends for the APS and ASBMB. Figure 11 exhibits the counts as proportional to the total number of words, but the pattern remains mostly inconsistent. Proportionally, the ASBMB appears to have the same steadily increasing trend as seen in the absolute counts, and the AMS has a sudden jump in 2021, but the ACS and APS appear to go flat, neither increasing nor decreasing.

**Fig 11. DEI-related terms in academic association publications (proportional)**

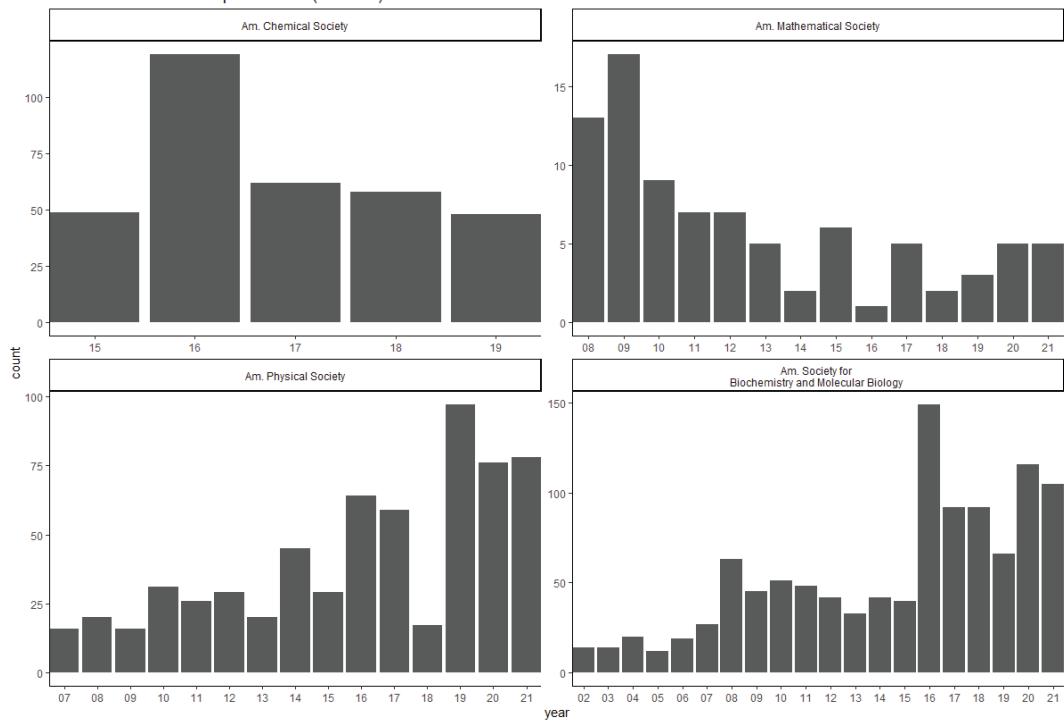


We then performed the same analysis on grouped sets of DEI terms. These were:

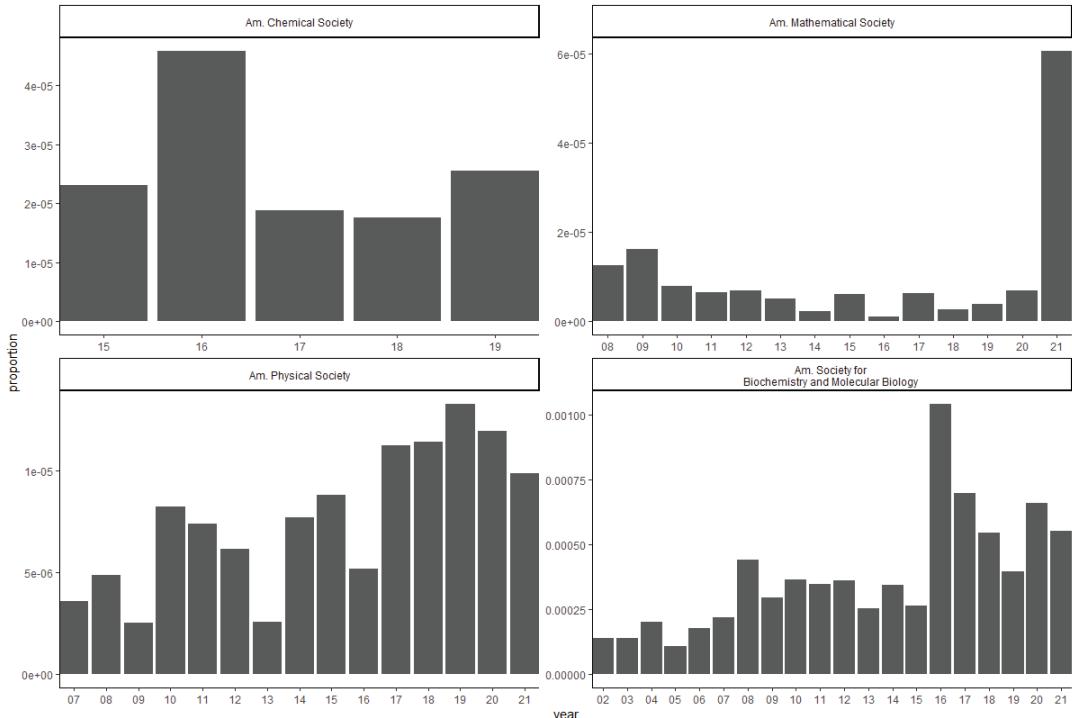
- diversity, equity, inclusion (Figures 12–13);
- racism, antiracism (Figures 14–15);
- gender (Figures 16–17);

Figures 12 and 13 show inconsistent patterns of usage similar to those seen for all DEI terms. Two of the societies did not mention either “racism” or “antiracism” at all in their publications, leaving only the APS and ASBMB to chart. Figures 14 and 15 show that only the ASBMB has an upward trend of antiracist language, with no trend visible amid the APS. Figures 16 and 17 show no trend in any society regarding the term “gender,” whether counted as absolute or as proportional.

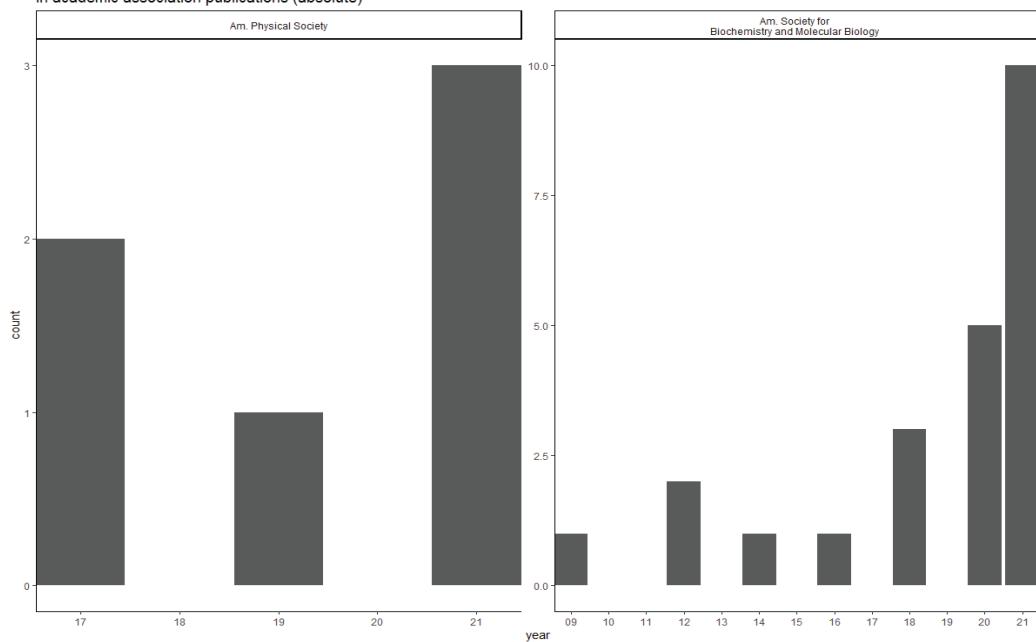
**Fig 12.** Terms 'diversity', 'equity', and 'inclusion' in academic association publications (absolute)



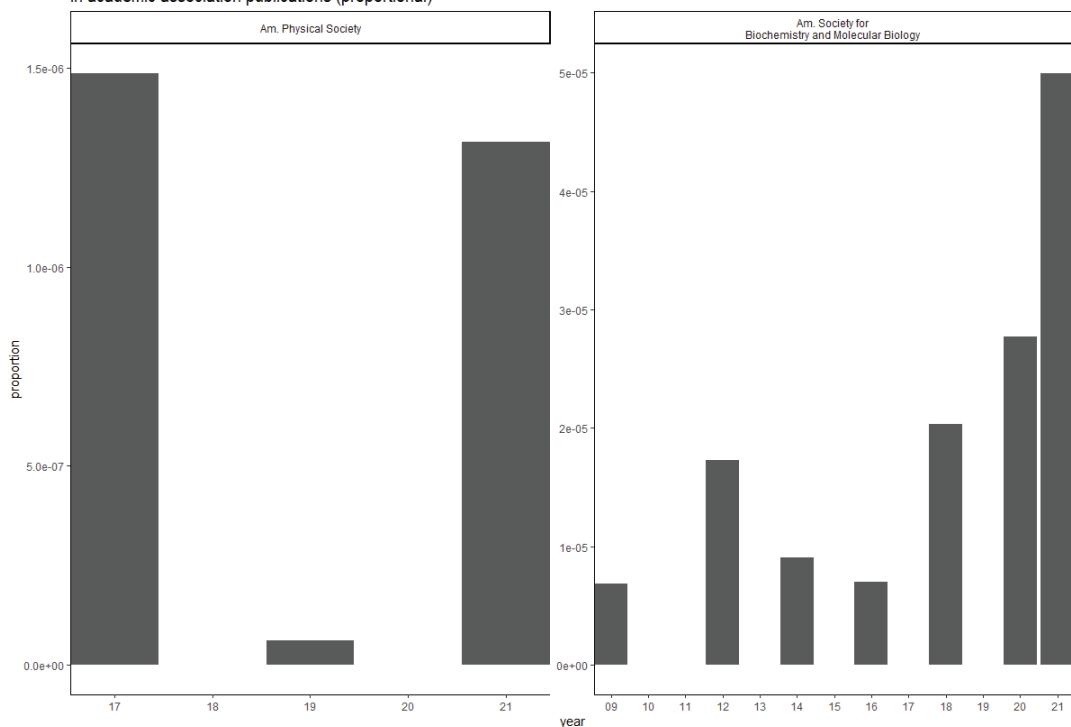
**Fig 13.** Terms 'diversity', 'equity', and 'inclusion' in academic association publications (proportional)



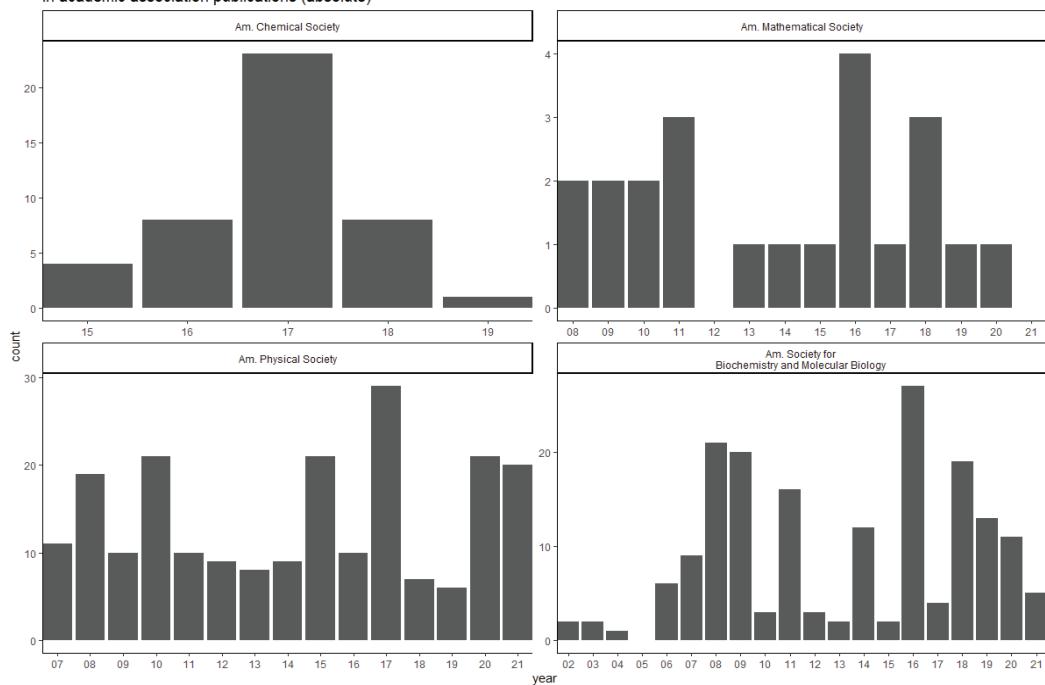
**Fig 14.** Terms 'racism' and 'antiracism'  
in academic association publications (absolute)



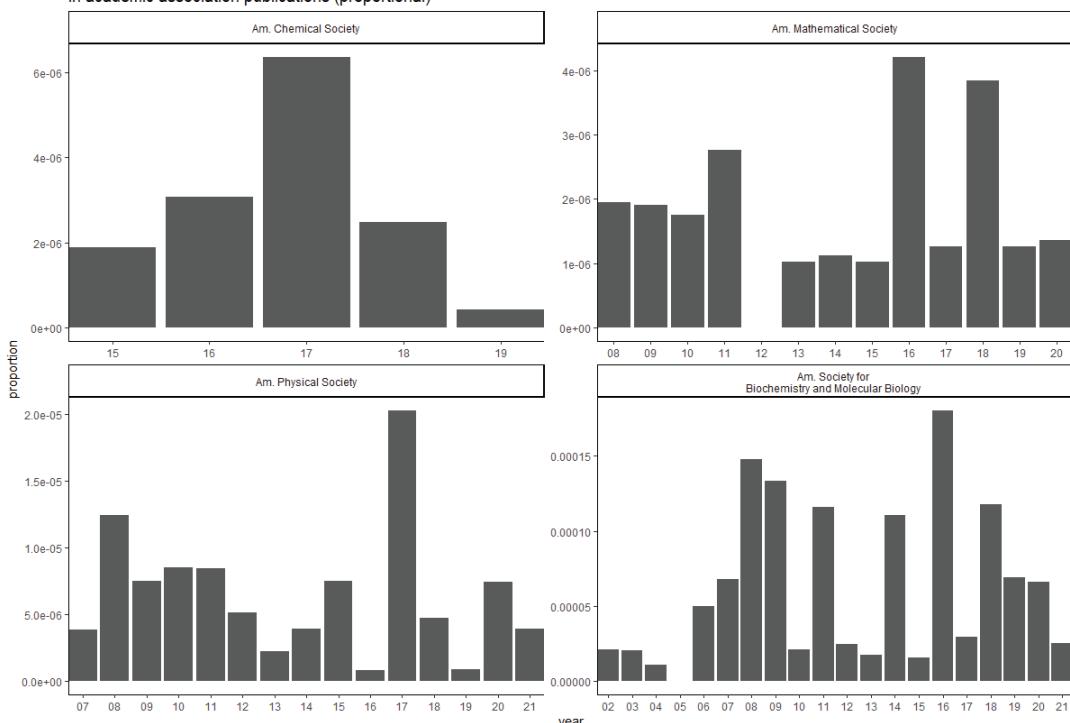
**Fig 15.** Terms 'racism' and 'antiracism'  
in academic association publications (proportional)



**Fig 16.** Term 'gender'  
in academic association publications (absolute)



**Fig 17.** Term 'gender'  
in academic association publications (proportional)



## Summary: Academic Associations

Scientific societies did not show the clear rise of DEI language that university websites and Twitter feeds did. For opponents of DEI, these results would seem to be encouraging: for all the increase of DEI language by universities, its penetration into the scientific societies appears to be modest. If DEI language was being heavily incorporated into the abstracts of scientific conferences (for instance, “feminist glaciology”<sup>11</sup>), this analysis would have shown it. It did not show such a pattern. Two of the four societies showed modest increases in DEI language, while the other two showed a decline.

However, a deeper dive into the societies themselves reveals patterns that should be watched closely. For example, even though the societies’ annual programs did not show a strong skew toward DEI-related content, all of the societies have a section on their websites and a variety of committees devoted to DEI issues – e.g., the American Physical Society’s Committee on the Status of Women in Physics,<sup>12</sup> the American Chemical Society’s Committee on Minority Affairs,<sup>13</sup> etc.<sup>14</sup> Thus, professional associations are possibly promoting DEI ideology through organizational and staffing change. If so, this has not yet led to a significant effect on the content of scientific output the societies evaluate and discuss annually. Even so, heightened vigilance by the members of these societies may be warranted.

There are two limitations to the analysis of this dataset. First, we do not have annual reports of abstracts for the ASBMB, where we relied instead on monthly newsletters. Second, we possess only five years of data for the ACS, from 2015 to 2019. Missing are the years (2020 on) where DEI language is known to have rapidly increased elsewhere. However, the data we did obtain for each of these cases seem appropriate. If the ASBMB had made a hard turn toward embracing DEI ideology, that should be apparent over 225 issues of their internal newsletter, which are generally lighter on scientific content and heavier on social topics. While there is a steady increase within the ASBMB, it is not exactly indicative of a hard turn. The result for the APS is similarly ambiguous given the missing data and wide variance in the estimates; there is no clear pattern.

The quantitative measurements here do not address the extent to which the administrative resources or researcher attention is being diverted to DEI ideology, even if it doesn’t affect the actual science itself. That question could be answered through qualitative analysis of the associations’ activities.

<sup>11</sup> Mark Carey, M. Jackson, Alessandro Antonello, and Jaclyn Rushing, “Glaciers, gender, and science: A feminist glaciology framework for global environmental change research,” *SAGE Journals: Progress in Human Geography*, 10 January 2016. <https://doi.org/10.1177/030913251562336>.

<sup>12</sup> American Physics Society, “Committee on the Status of Women in Physics” <https://www.aps.org/about/governance/committees/cswp/index.cfm>.

<sup>13</sup> American Chemical Society, “Committee on the Status of Minority Affairs” <https://www.acs.org/content/acs/en/about/governance/committees/minority-affairs.html>.

<sup>14</sup> See also the American Mathematical Society’s Equity, Diversity and Inclusion page: <https://www.ams.org/about-us/diversity> as well as the ASBMB’s: <https://www.asbmb.org/diversity>.

## Scientific Grants and Awards

We created a dataset for DEI language in scientific grants and awards. Grantmaking and awards drive research programs, and scientists shape their work to conform to funding opportunities. If DEI ideology is affecting the STEM fields overall, not just in the universities, it should be evident in the allocation of scientific grants and awards.

### Methods

We based our analysis on datasets compiled from three major funders for scientific research. Table 3 details these funders and what data was collected for this dataset.

**Table 3. Data sources for DEI language from three major funders of scientific research.**

Grantmaking Body	Data Obtained
National Science Foundation	Full award data from 2010-2021 (~140,000 awards, 1.5gb)
National Institutes of Health	Awards for nine targeted searches (described below)
Ford Foundation	Full historical award data (~24,000 entries)

The DEI-related dictionary used for this analysis is the same as that used for the analysis of academic associations above. The R files associated with this analysis are described in the Technical Appendix.

### Results

#### National Science Foundation

The NSF data are presented on an annual basis, which smoothed out the effect of bi-annual grant cycles. Figure 18 exhibits the total number of DEI terms in all NSF grants by year, showing a steady increase over time. Figure 19 exhibits the total number of NSF grants with antiracist language by year—such as: “anti-racism,” “antiracism,” “Black Lives Matter,” “BLM,” “Critical Race Theory,” “George Floyd,” “racism,” “racist,” “white fragility,” and “white supremacy”—and shows little to no growth until an abrupt increase occurring in 2020. Figure 20 exhibits the total NSF funding for all DEI-related projects, showing a steady increase over time, with an abrupt increase in 2020.

Fig 18. All DEI terms in all NSF grants by year

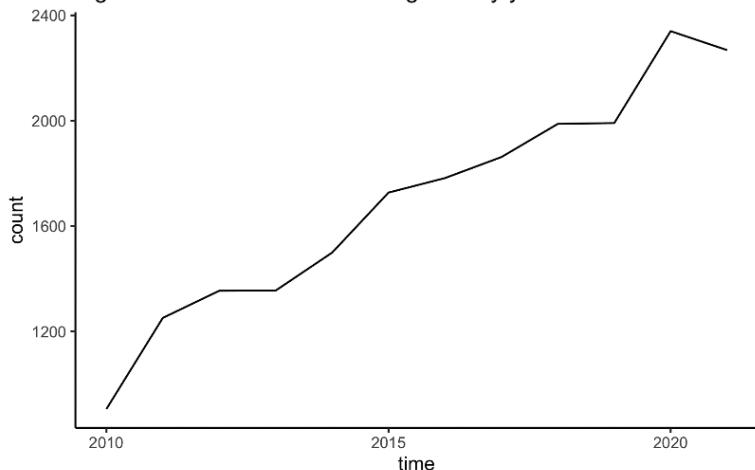


Fig 19. All NSF grants with antiracist terms by year

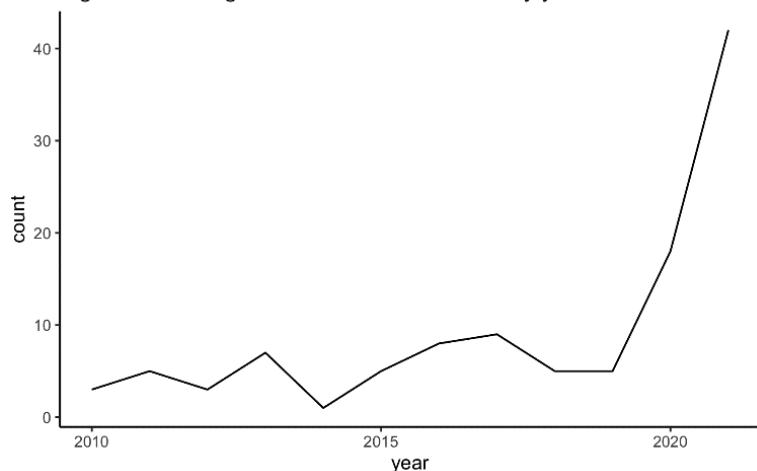
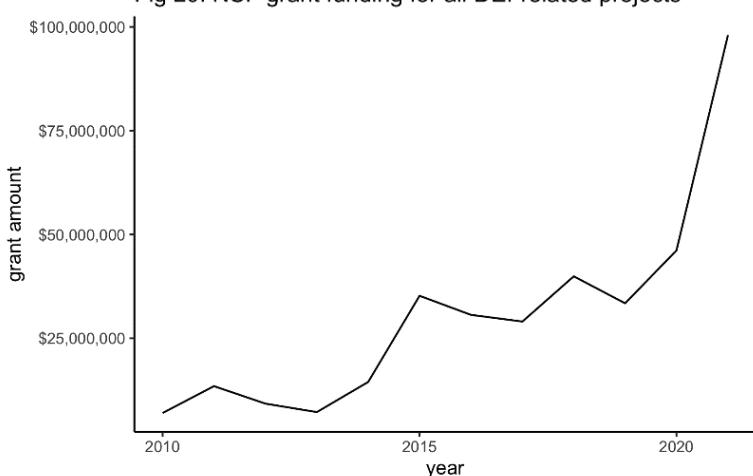
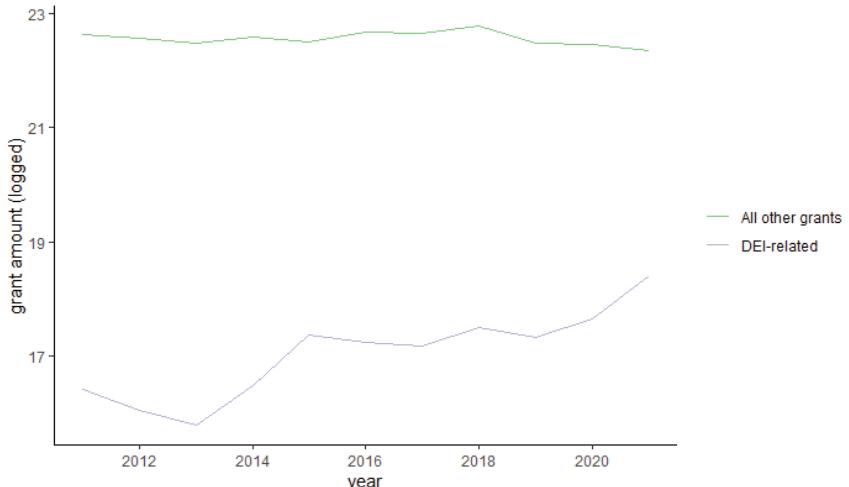


Fig 20. NSF grant funding for all DEI-related projects



However, compared to the annual trends of NSF funding, DEI language is clearly growing. Figure 21 shows grant awards for DEI-related projects compared to all other NSF grants. DEI related grants increased by about an order of magnitude, while funding for all other grants was comparatively steady, or slightly decreased. Note that the y-axis is logarithmic.

**Fig 21. NSF grant funding for all DEI-related projects compared to other projects**



The same patterns are evident for funding related specifically to anti-racist themes. The category of “antiracism-related projects” displayed in the following graphs refers to any grant that includes any of the following terms in its project narrative: “antiracism,” “anti-racism,” “Black Lives Matter,” “BLM,” “Critical Race Theory,” “George Floyd,” “racism,” “racist,” “white fragility,” and “white supremacy.” Figure 22 exhibits levels of funding, showing a dramatic rise beginning in 2020. Figure 23 exhibits the same trend compared to all other grants. Note that the y-axis is logarithmic.

**Fig 22. NSF grant funding for all antiracism-related projects (absolute)**

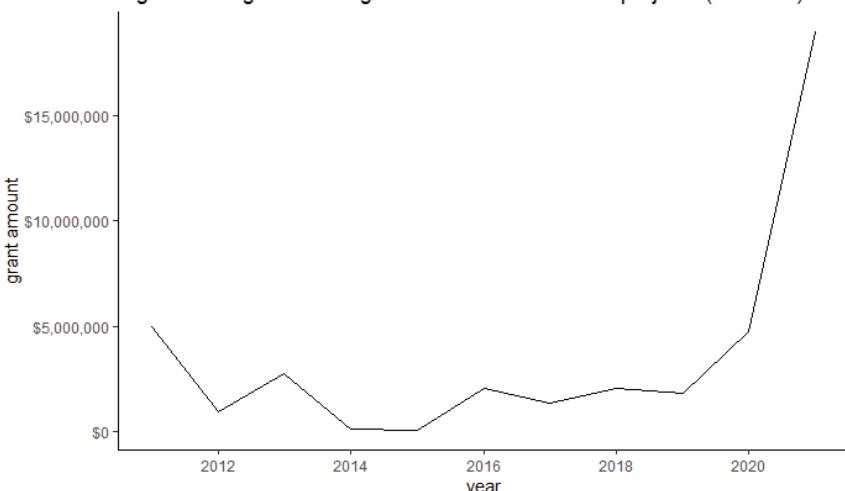


Fig 23. NSF grants for antiracist and all other topics by grant amount (logged)

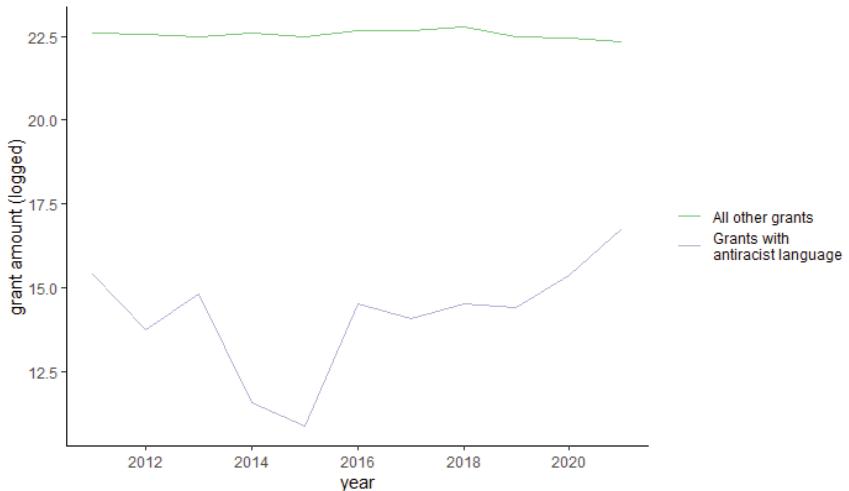
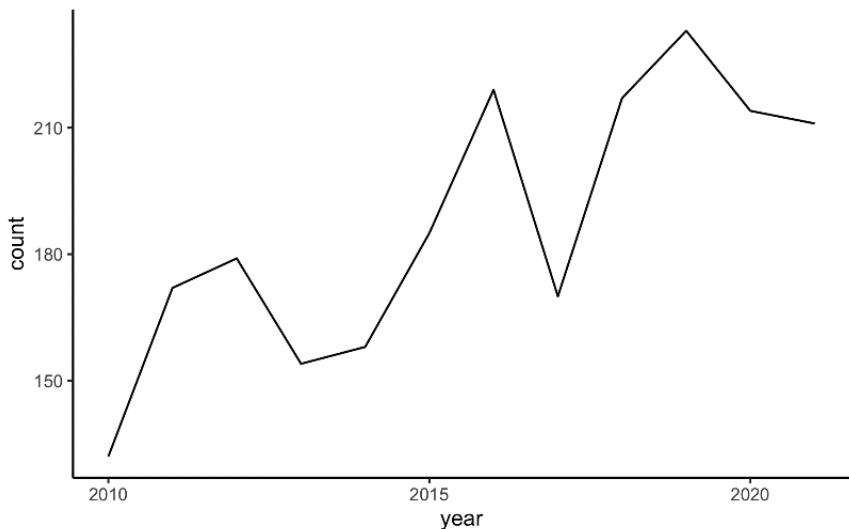


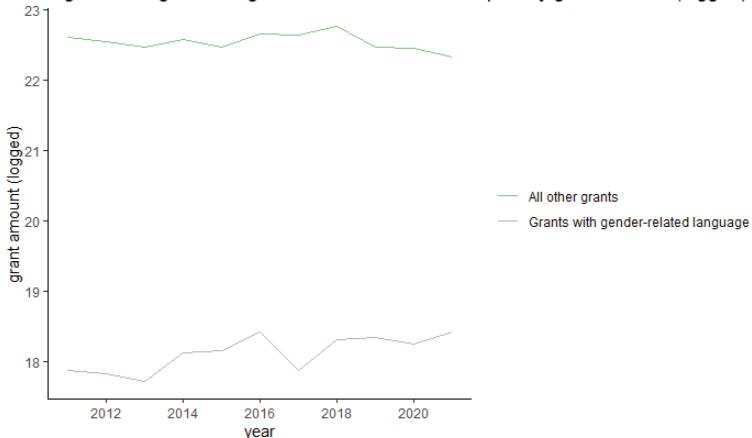
Figure 24 shows a steady increase over time of NSF grants incorporating the terms “gender” and “transgender.” Figure 25 compares the grants with gender-related language to to all other grants funded by the NSF. Increases of gender-related grants were modest, as were decreases of all other grants.

Fig 24. All NSF grants with terms 'gender' and 'transgender' by year

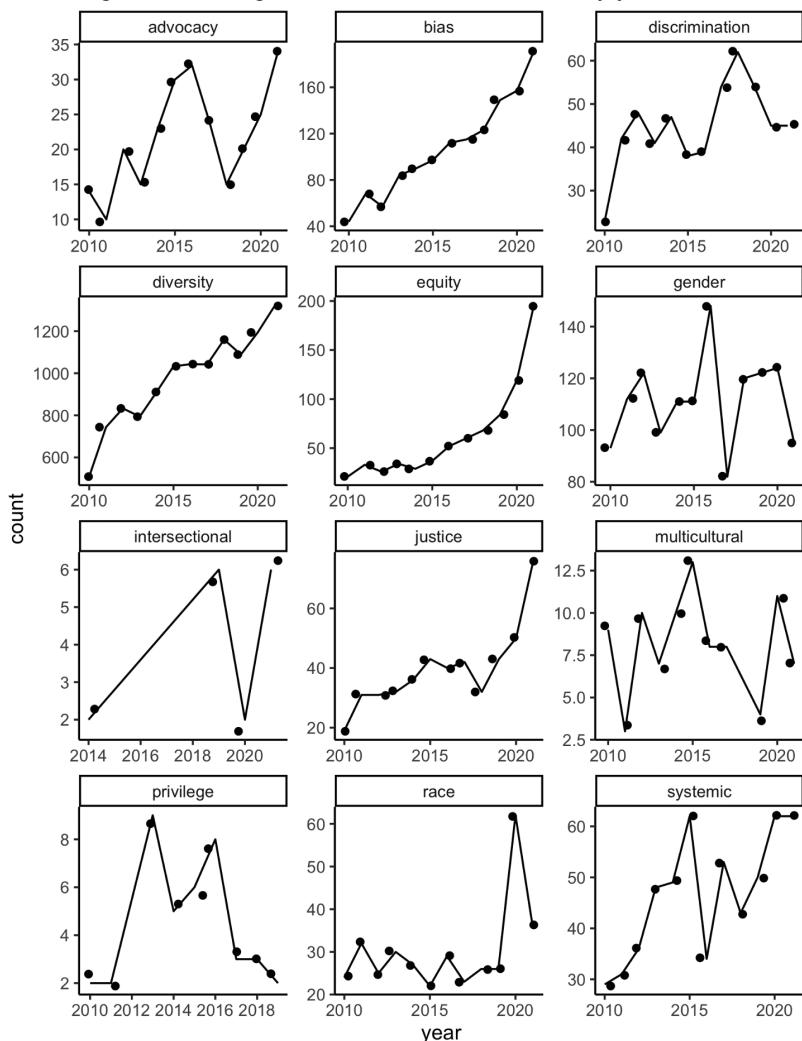


We were then interested to see if any individual DEI-related terms in NSF grants had variations in their trends. The cross-tabulation plotted in Figure 26 shows that some terms—such as “bias,” “diversity,” “equity”—show a steady upward trend, while other terms – such as “advocacy,” “gender,” “multicultural” and “privilege” – appear to fluctuate over time.

**Fig 25. NSF grants for gender-related and other topics by grant amount (logged)**



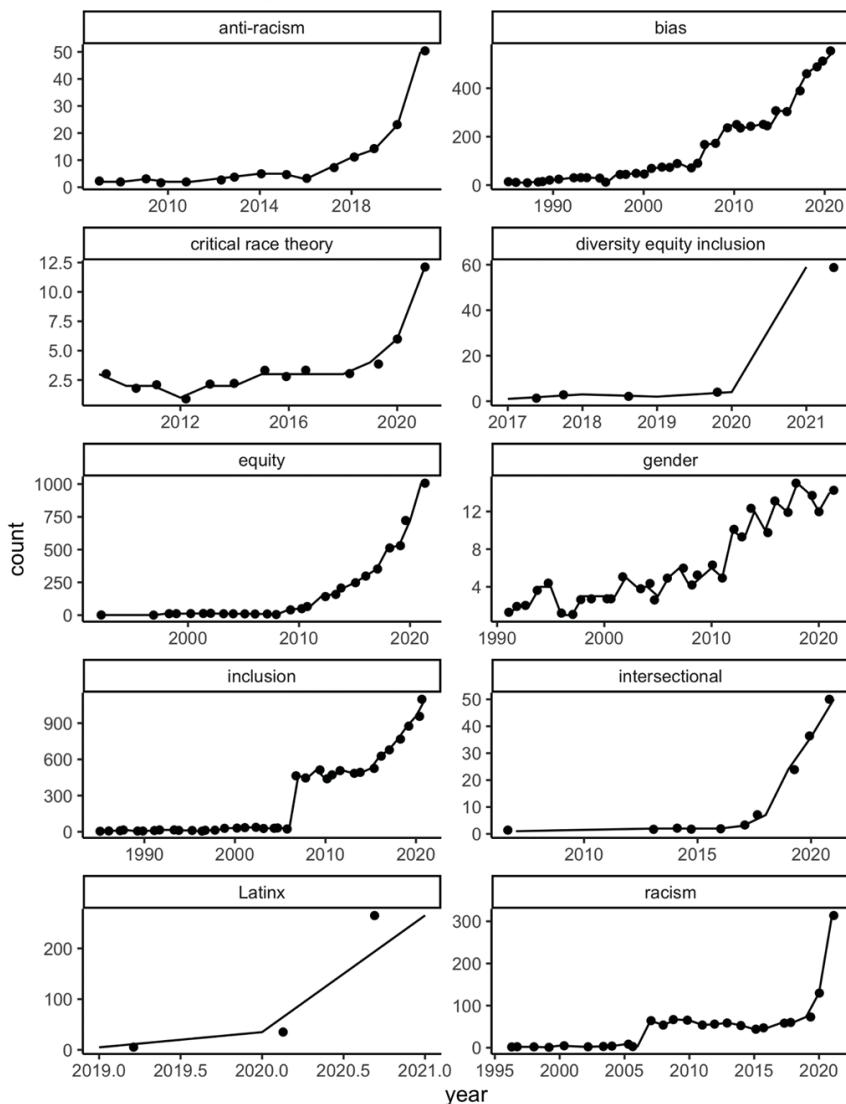
**Fig 26. All NSF grants with DEI-related terms by year**



## National Institutes of Health

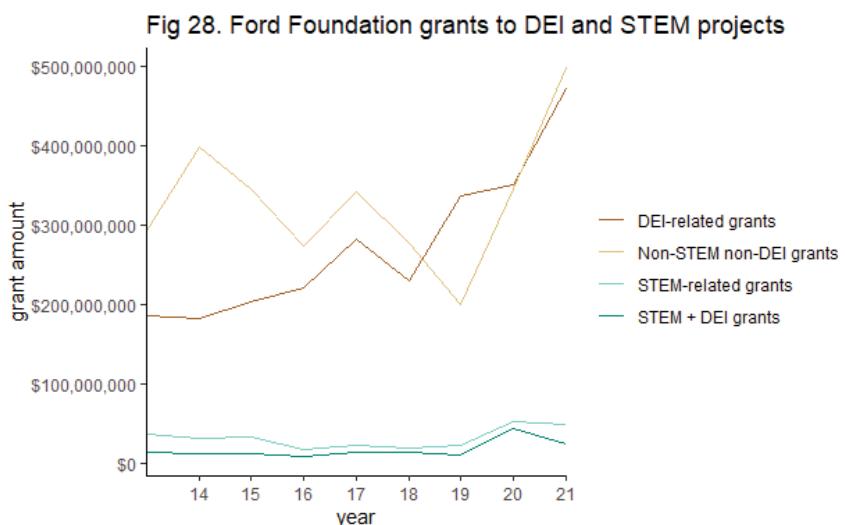
The NIH database was not as versatile or powerful as the NSF database and search engine. We could not query the entire corpus at once, for example. The searches therefore queried the abstracts and headlines only. We were able to obtain the absolute frequency of DEI language over time, but not what portion of total grants included DEI terms, or what portion of total grant funding each year involved such terms. Figure 27 exhibits 10 DEI-related terms, showing abrupt increases in typical anti-racist terms, steady increases in “bias” and “gender,” and a curiously sharp increase of “inclusion” around 2008.

**Fig 27. All NIH grants with DEI-related terms by year**



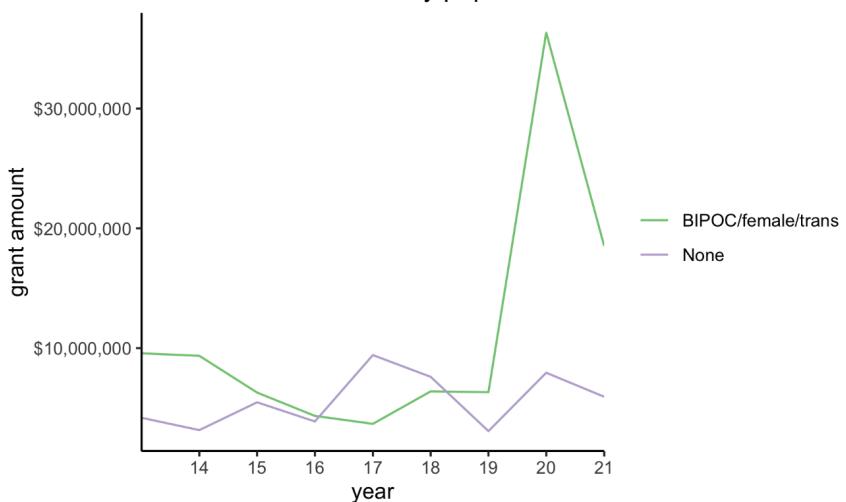
### Ford Foundation

The Ford Foundation publishes their full grant database. We accessed the database between 2010 and 2022, consisting of 16,630 grants. However, only grants between 2013 and 2021 were used, as data for 2012 were absent. Figure 28 shows that the dollar amounts going to all DEI-related grants has an upward trend similar to that found in the NSF data. The amount of funding for STEM and DEI-related grants has remained only slightly less than the funding of STEM-only grants since approximately 2016, followed by a sharp uptick in 2020, and a clear divergence by 2021.



The Ford Foundation's grant database also considers "Benefiting Populations." Of the 352 grants that contain both STEM and DEI-related language, 208 of these refer to "BIPOC" (Black, Indigenous, People of Color), females, or transgender / non-binary populations. Of the remaining 144, only five indicate some other beneficiary population ("United States" race ethnicity three times, "West Africa's race ethnicity," and "Quilombolas"), while the rest do not declare any. As exhibited in Figure 29 below, funding considering "BIPOC/female/trans" as a "benefiting population," showed an abrupt increase of over 20 million dollars (USD) between 2019 and 2020, followed by an abrupt decrease in the following year.

**Fig 29. Ford Foundation grants to DEI and STEM projects with identified beneficiary populations**



## Summary: Scientific Grants and Awards

This dataset clearly indicates that DEI-related language has been gaining ground in the allocation of funding for scientific research. The analysis of approximately 140,000 NSF grants makes clear that DEI language has risen as both absolute numbers of grants, and amount of funding. Anti-racist language in particular has risen sharply in the last few years, as one might expect. Gendered language has also increased, both as raw counts and funding. These trends are visible in the data of the Ford Foundation as well.

The data for the National Institutes of Health are greatly limited. As aforementioned, the NIH database was not as versatile or as powerful as the NSF database or search engine. While we were able to obtain the absolute frequency of DEI language in NIH grants over time, it was not possible to determine what portion of the total grants included DEI terms, nor the portion of funding that involved such terms. Still, the absolute frequency of ten DEI terms increased across the board, with a more sudden rise in specifically anti-racist terms, which is congruent with our findings elsewhere.

## Scientific Literature

The fifth and final dataset created was from the scientific literature itself. DEI terminology on university websites and Twitter feeds, within the annual programs of scientific academic associations, and within research funding, are all solid indicators of DEI language's intrusion and intensification within the STEM fields. However, our analysis would be incomplete if we failed to look at the repositories of scientific literature, of which we used data from four: the Web of Science, Google Scholar, arXiv, and PubMed.<sup>15\*</sup> These are among the largest and most authoritative scientific databases in the world, each with its own focus and characteristics.

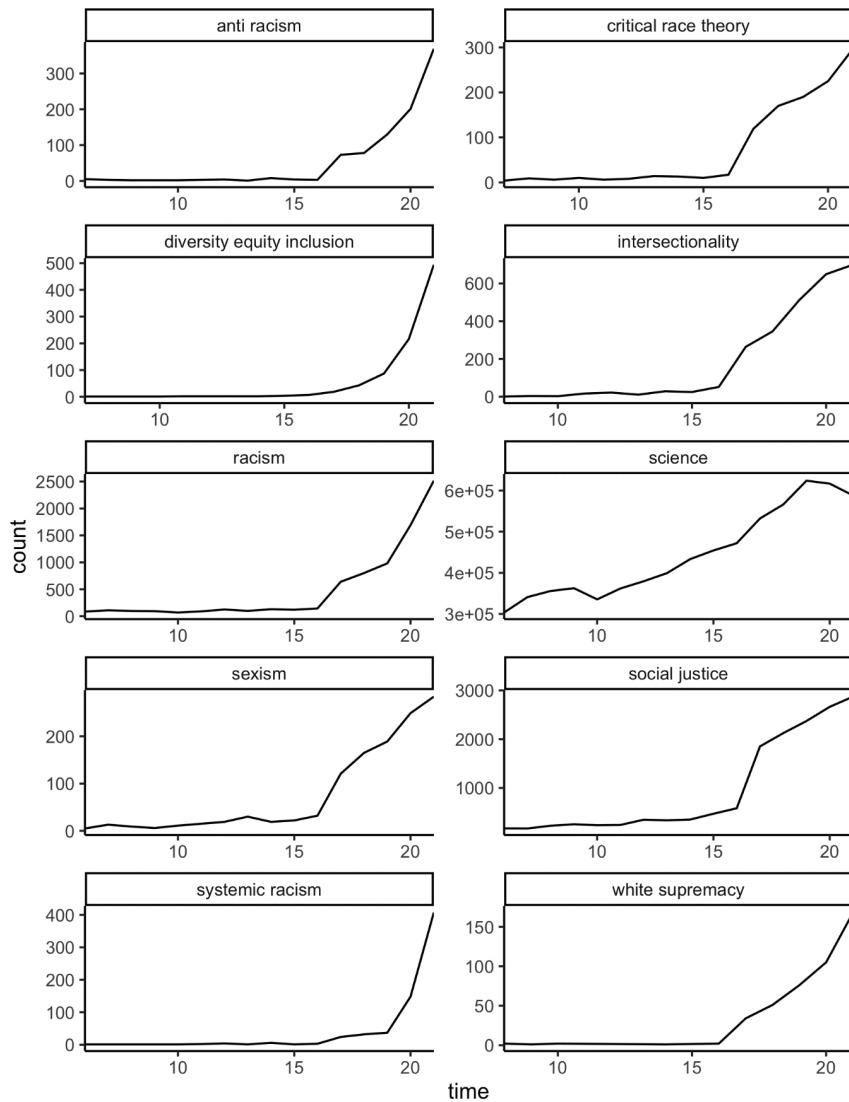
### Web of Science

Web of Science, published by Clarivate, is among the largest commercial citation indices in the world. It contains 1.9 billion cited references from over 171 million records, with several subsidiary databases related to discipline. We searched inside the following citation indices: Science Citation Index Expanded (SCI-EXPANDED), Conference Proceedings Citation Index – Science (CPCI-S), Current Chemical Reactions (CCR-EXPANDED), Index Chemicus (IC), and Emerging Sources Citation Index (ESCI). We excluded the three social science indexes: Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (AHCI), and Conference Proceedings Citation Index – Social Science & Humanities (CPCI-SSH). Our searches were limited to the default “Topic” field. We used the control phrase “science” to ensure that the patterns observed in the DEI terms represented the overall content of the database.

Figure 30 shows obvious growth in DEI-related terms in science-only indices since 2016. However, this chart conceals how rapidly DEI terms have grown and over how many orders of magnitude, because the total number of publications also rose over time. Table 4 shows the average growth rate for all of the years under observation for each of the categories. From 2010 to 2021, reports about antiracist and DEI-related topics have grown between 3 to 42 times faster than scientific topics in general in the Web of Science.

15 \* See: Web of Science, <https://clarivate.com/webofsciencengroup/solutions/web-of-science/>; Google Scholar, <https://scholar.google.com/>; arXiv, <https://arxiv.org/>; and Pub Med, <https://pubmed.ncbi.nlm.nih.gov/>

**Fig 30. Web of Science count of DEI terms over time (with control)**



**Table 4. Differential growth of DEI related terms in the published scientific literature.**

Topic	Growth Rate
science	5%
social justice	15%
racism	32%
sexism	45%
critical race theory	62%
diversity equity inclusion	73%
intersectionality	107%

## Google Scholar

Google Scholar is a prominent repository of academic and peri-academic publications. In comparison to Web of Science, Google Scholar contains a much larger number of preprints or otherwise informally published works. This difference is a helpful distinguishing feature, because it means data from Google Scholar will not be identical to data from Web of Science. Google Scholar does not have an application programming interface (API), and Google limits the manner in which web scraping bots can interact with the site. It is thus difficult to get comprehensive information, making it infeasible to use a reliable control search as in Web of Science above.

Figure 31 shows the growth of publications for a range of DEI and STEM terms. The data were obtained using open-source software that queries Google Scholar and returns a maximum of 1000 results.<sup>16</sup> These results appear to be broadly distributed over all years in which data are available, according to searches for non-DEI terms. Because these searches could not be performed programmatically, we did not satisfy every combination of DEI / STEM terms as in previous cases. Instead, there is only a sample of the total possible combinations. The findings suggest wider applicability than just these cases, however.

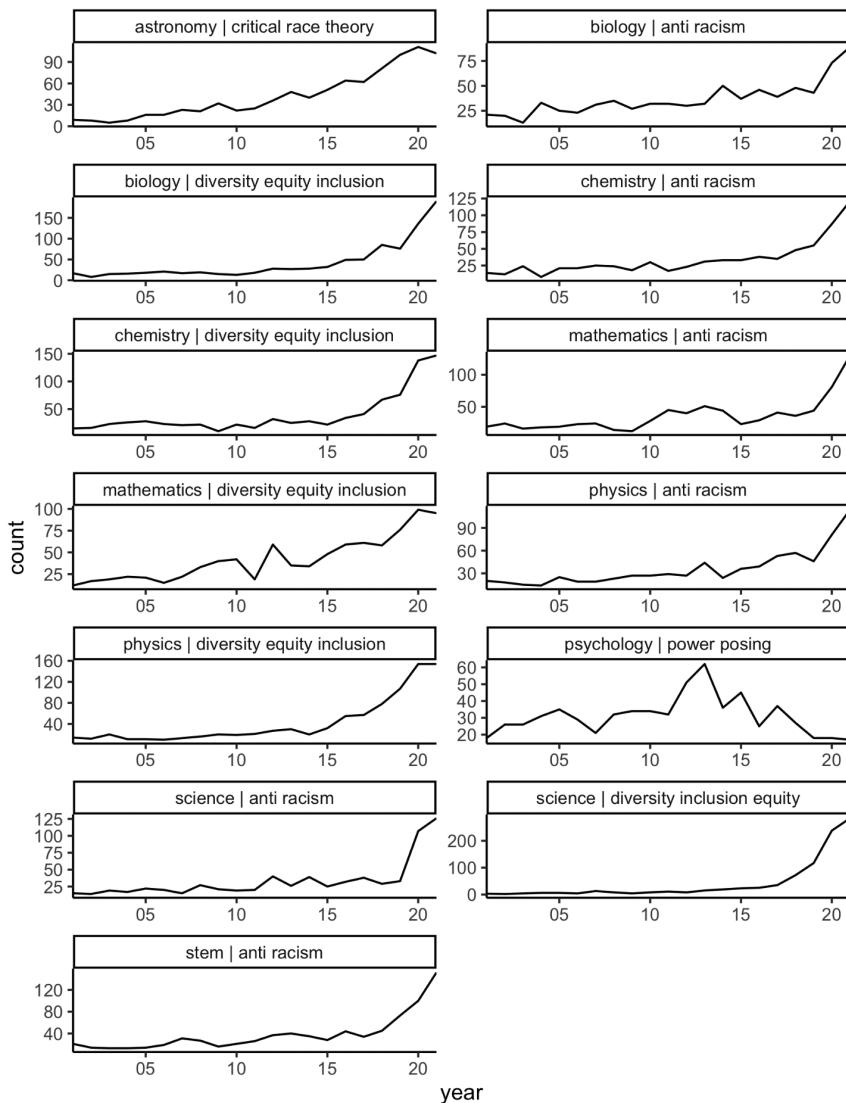
All of the searches below were input as single strings (i.e., “astronomy critical race theory,” or “biology diversity equity inclusion”). The control we used in this case was the search term “psychology power posing.” This is because it is well known in the social psychology literature that the theory of “power posing” went through a rise and fall in influence and popularity.<sup>17</sup> Therefore, if the method of querying Google Scholar we used was roughly representative of the full content of the database (rather than severely oversampling more recent years, which would be a threat to inference), we would expect to see this reflected in the results for “power posing.” This is indeed the case. We thus feel comfortable that the following findings broadly reflect the actual content on these topics in the Google Scholar database. Yet another reason this claim is robust is because the query is limited to 1,000 results, and in none of our searches was the limit reached.

---

<sup>16</sup> This program is called Publish or Perish. For more information, see: Anne-Wil Harzing, “Publish or Perish: Explains the use of Publish or Perish and its metrics,” Harzing.com, 6 February 2016. <https://harzing.com/resources/publish-or-perish>.

<sup>17</sup> Andrew Gelman, “What has happened down here is the winds have changed,” *Statistical Modeling, Causal Inference, and Social Science*, 21 September 2016. <https://statmodeling.stat.columbia.edu/2016/09/21/what-has-happened-down-here-is-the-winds-have-changed/>; Lee Jussim, “Are Most Published Social Psychology Findings False?” *Psychology Today*, 23 February 2016. <https://www.psychologytoday.com/us/blog/rabble-rouser/201602/are-most-published-social-psychology-findings-false>; Joseph P. Simmons and Uri Simonsohn, “Power Posing: P-Curving the Evidence,” *SAGE Journals: Psychological Science*, 20 March 2017. <https://doi.org/10.1177/0956797616658563>.

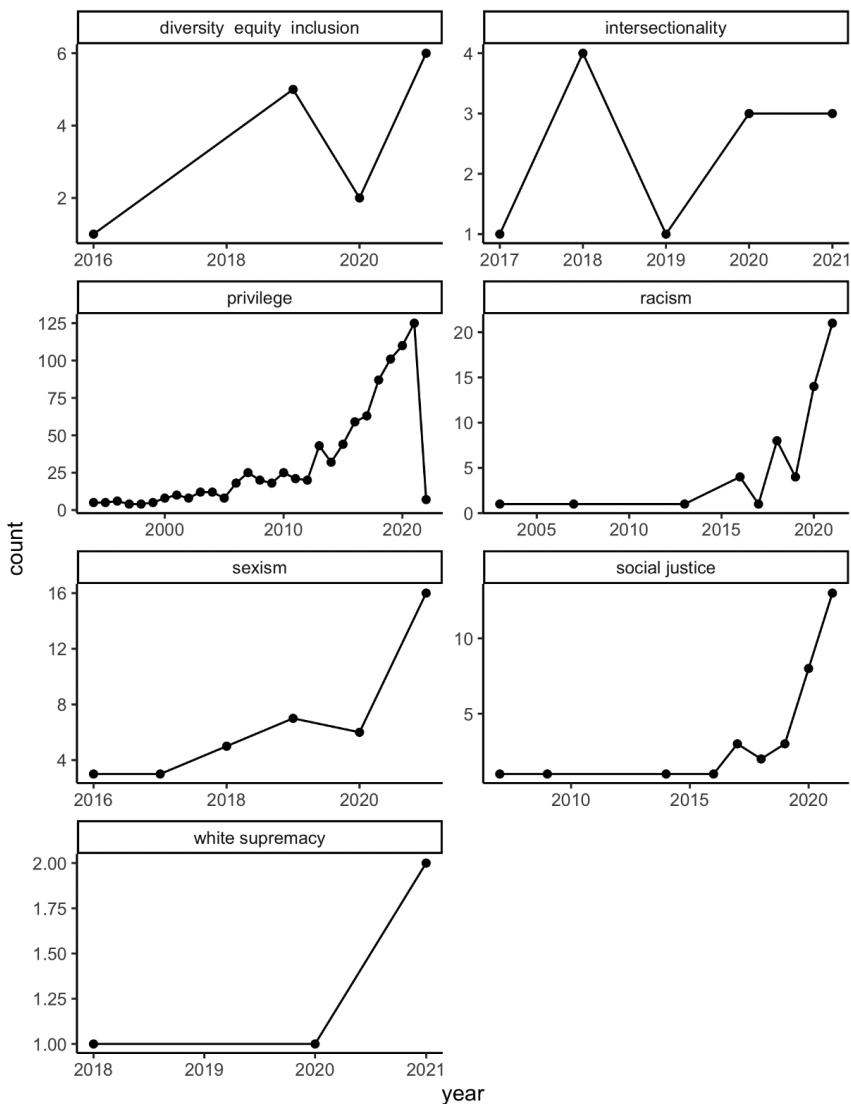
Fig 31. Google Scholar results for DEI and STEM terms



arXiv

arXiv (pronounced “archive”) is the largest and most well-known preprint server for publications in the sciences. We used the arXiv API to conduct searches on DEI-related terms. It was not necessary to search for STEM terms, because the entire database consists of scientific preprints. Figure 32 shows the results of these searches. Three search terms—antiracism, critical race theory, and systemic racism—were dropped because they yielded few or no results.

Fig 32. arXiv results for DEI terms



## PubMed

PubMed, short for PubMed Central, is a leading database of biomedical research hosted by the National Institutes of Health. Substantially everything in PubMed is STEM-related, and so it can be queried directly for DEI-related terms. PubMed does not have an API, so large-scale programmatic searching was impossible. Therefore, we conducted searches directly for DEI-related terms in the front-end interface. PubMed did, however, have the best control data available so far: numbers of citations annually. This attribute is a close

representation of the growth of the overall database, and thus allows for a natural comparison to the growth of DEI-related publications.

Figure 33 exhibits the use of DEI terminology in PubMed over the last decade, and shows that usage has grown rapidly since 2015. Then, we were interested to see what the historical data might show, and pushed the x-axis back to the first time a paper mentioning one of the terms appears. Figure 34 shows these historical trends, highlighting just how rapid the rise of DEI-related language has been.

**Fig 33. PubMed results for DEI terms (recent)**

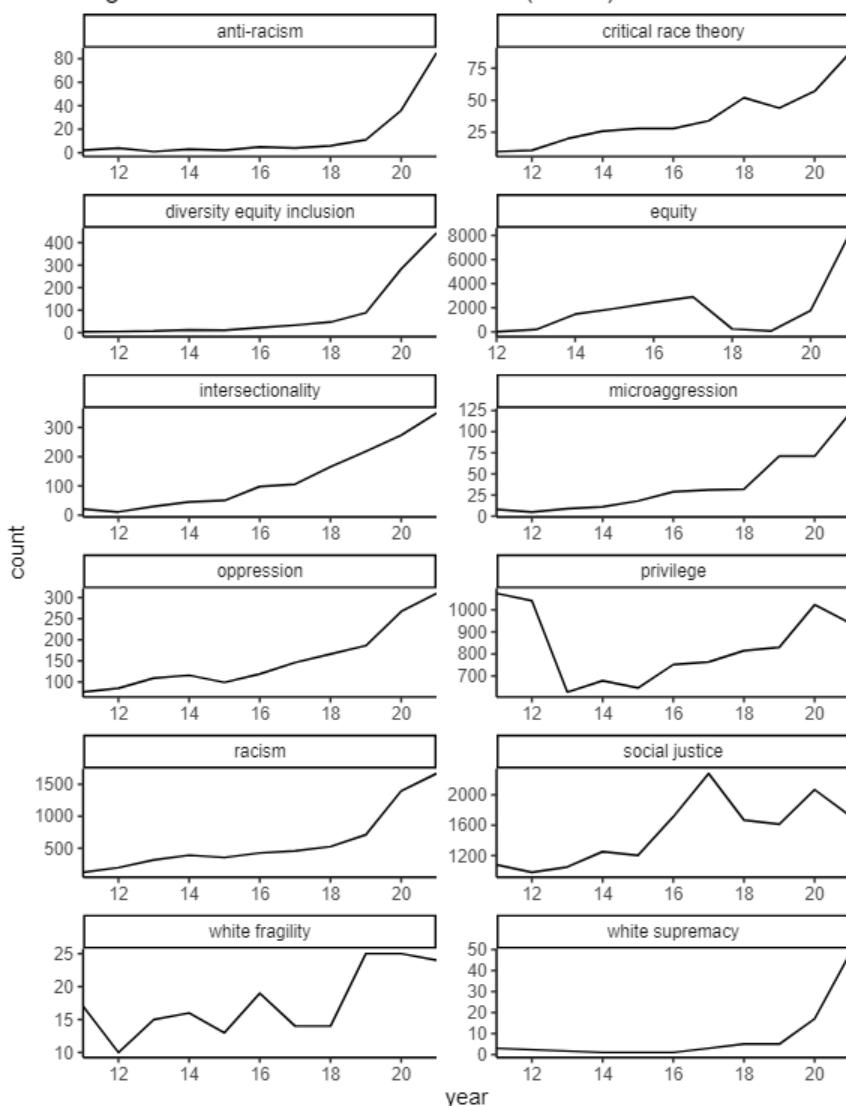
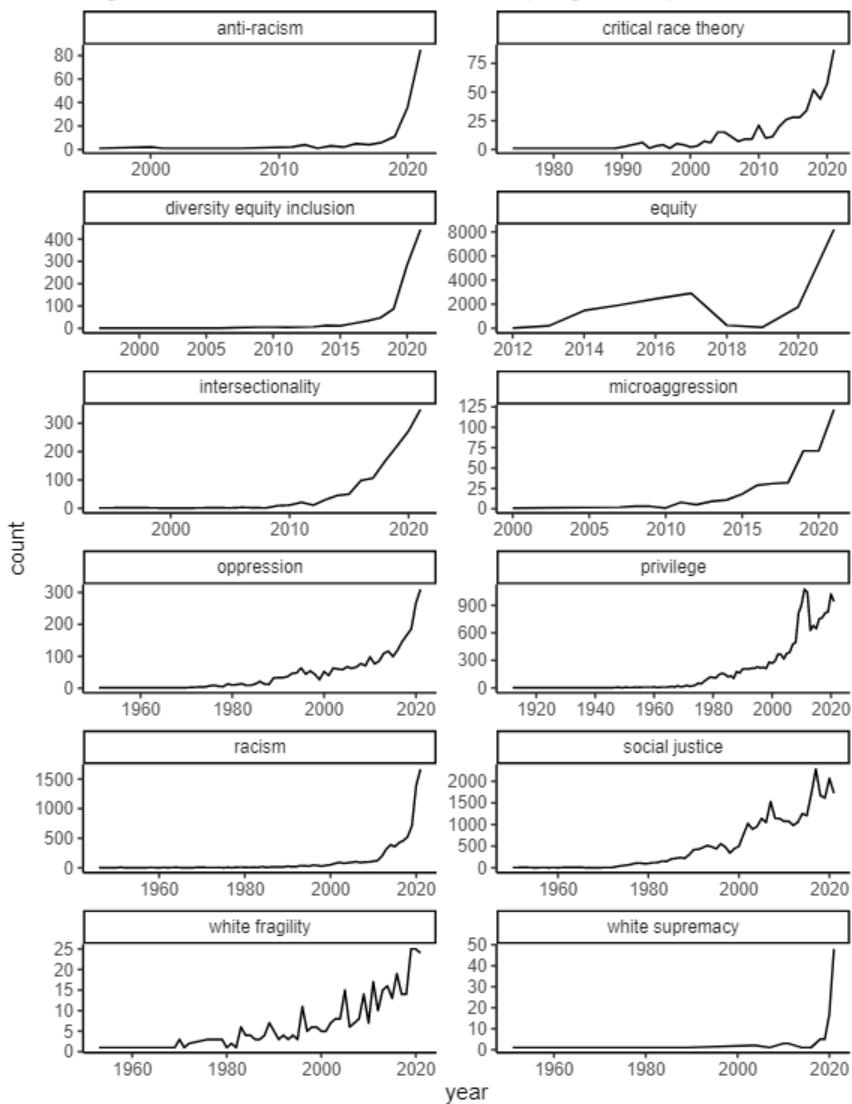
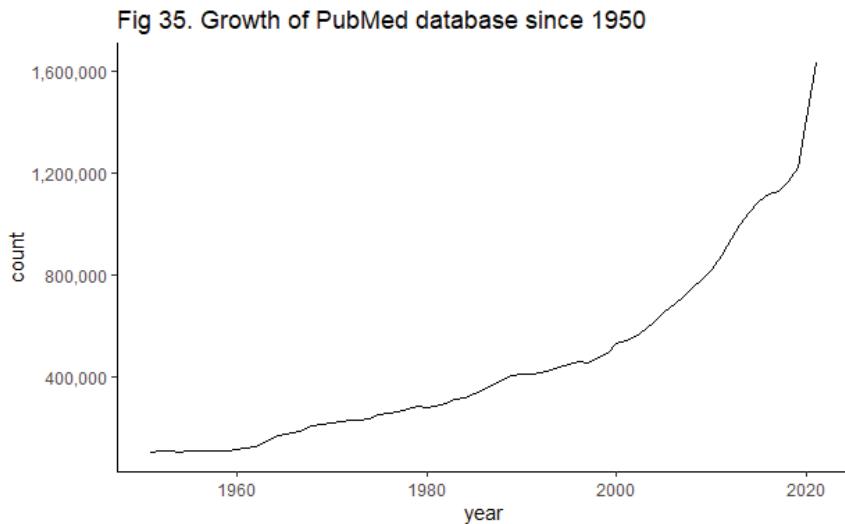


Fig 34. PubMed results for DEI terms (longitudinal)



Additionally, we wanted to compare these trends with the overall growth of the PubMed database, which should better exhibit the dramatic rise of DEI-related language. The data for this final graph were obtained from PubMed's own “PubMed Total Records by Publication Year” dataset.<sup>18</sup> Figure 35 exhibits the growth of PubMed overall, showing once more that DEI-related publications have expanded far more rapidly than the total body of publications in PubMed.

<sup>18</sup> PubMed, “Total Records by Publication Year,” PubMed, <https://datadiscovery.nlm.nih.gov/dataset/PubMed-total-records-by-publication-year/eds5-ig9r>.



## **Summary: Scientific Literature**

It is clear from this analysis that DEI-related publications have risen sharply in the scientific literature, especially in recent years. Between four different repositories—Web of Science, Google Scholar, arXiv, and PubMed—the general pattern holds. When this dataset is compared to those of the university websites and Twitter feeds, the annual programs of academic associations, and research funding, one overall conclusion can be drawn: the empirical evidence clearly shows that the vocabulary of DEI ideology has increased dramatically in frequency in the STEM fields, both within and outside of the universities.

# General Discussion

By nearly every measure, language associated with the Diversity, Equity, and Inclusion ideology has risen dramatically in STEM fields. Over the past decade, and for longer where data were available to assess it, DEI language associated with STEM keywords has become more prevalent on university websites and twitter feeds. Outside the universities, the same trends are visible in science funding sites and funded research proposals, in the scientific literature and literature databases, and, to a lesser extent, scientific professional societies.

The rise of DEI language in the sciences has been driven largely by activists who argue that the sciences systematically exclude and marginalize women, along with designated racial minorities and identity groups defined by sexual orientation. The claim is that science is too white, too male, and too heterosexual, and that aggressive action must be taken to reverse this. Nearly the entire societal infrastructure that supports scientific research, including universities, funding agencies, scientific societies, and accreditation boards, and increasingly scientific journals have accepted this premise, reflected in the increasing use of DEI-associated language we have documented here.

The activists' aim is to radically transform the sciences by making all aspects of scientists' careers conform to the demands of DEI ideology. Training and education in the sciences, hiring of new science educators and researchers, decisions on research grants, accreditation, governance of research universities, and more are now being reshaped to meet the demands of DEI ideology. The ideological intensification over the past few years is now empirically evident. Only the future remains unknown. Yet if the ideology continues to intensify and be institutionalized, then the future of STEM, along with the rest of the academy, is almost certainly imperiled.

# Technical Appendix

Much of the work in this report involves original data scraped from the internet, processed, and displayed in graphs. There are choices at each step of that process, each of which can influence the outcome significantly. The most robust method of ensuring the integrity of such work is a transparent presentation of the exact methods used, along with the code and data, so that other analysts can scrutinize the methods and replicate them. The best way for interested parties to do that is to look at the GitHub repository on which all the code is deposited, download from Zenodo the 30gb of raw data that was used, and reproduce our analyses for themselves.<sup>19\*</sup> This report was written in a combination of English and computer code (using R), and is compiled by a number of pieces of software (R, RStudio, pandoc, etc.). To assist interested parties in reproducing the analysis, some notes follow that explain where to find the files and what to do with them. Given that much of this is obvious once examining the files, these notes only deal with a few of the sections.

## University Websites

Each university name was extracted from the QS Top Universities list of the 100 top universities in the US as of April 2021.<sup>20</sup> A dozen of these were updated with more technically-oriented universities. Each university's domain name (i.e., website) was then obtained from other online resources. These were used in combination with DEI and STEM terms to search the Google Search application programming interface (API) for all web pages on those websites that contained a given combination of terms. The script paginated until no further results were reached (for a maximum of 100 results per DEI + STEM + university combination). All results were stored in raw JavaScript Object Notation (JSON) format.

This resulted in 8,718 JSON files, each containing a maximum of 10 links. Some of these links were dead, duplicates, or otherwise corrupt, and not every file contained the maximum possible 10 links. We were able to download and save to disk 116,324 files. We passed each

<sup>19</sup>\* See the raw data on Zenodo at <https://zenodo.org/record/6360904#Yj4HkufMJPZ> and the NAS GitHub repository at <https://github.com/NASorg/quantdei>

<sup>20</sup> Craig O., "Ranked: The Top 100 Universities In The USA," *QS Top Universities*, (As of April 2021). <https://www.topuniversities.com/where-to-study/north-america/united-states/ranked-top-100-us-universities>.

URL to a hash function (a one-way function that will turn a given string to another string of some fixed length—in this case 32 characters), resulting in files of the form: school-name\_dei-term\_stem-term\_hash.html, or, for example:

wpi\_inclusion\_chemistry\_1446e6d34af76ef78c9edf973aebcd3b.html

(where “wpi” is Worcester Polytechnic Institute).

We then converted these HTML files to text with the Linux utilities `html2text` and `lynx -dump`. We were able to convert substantially all of them to text, though nearly 2,000 resulted in 0, 1, or 2 byte files, meaning the HTML could not be parsed and the result was empty. Hashing the URLs meant that each file could be uniquely linked to a URL in the dataset (because it is unwieldy to use URLs as file names). We then developed and refined a number of regular expressions to extract dates from each of these pages. We filtered out any page that included more than a single date, in order to avoid counting a large number of dates on a single page.

This date duplication often happens with event listings, and adds noise to the data. As we have designed this enquiry, each date pulled from a page becomes a datapoint – an instance of a DEI term and a STEM term being on the same page. When performing this operation, the number of observations in the dataset decreases from over 198,710 to just 10,115. The data was presented as a raw count across all the pages, and then in a series of different combinations (i.e. Ivy League vs non-Ivy, etc.). The data was initially collected in September and October of 2021, and the code reran in January of 2022. The results are current as of mid-January 2022. These are in the `./data/school_websites/json` folder in the project’s replication files.

We used the Google JSON Search API to construct the search calls. Each search call consisted of a combination of a DEI term, STEM term, and school. These were the DEI and STEM terms used:

```
dei_terms ← c("diversity," "equity," "inclusion," "systemic racism," "white supremacy" ,  
"anti-racism," "justice," "privilege," "critical race theory")
```

```
stem_terms ← c("astronomy," "biology," "chemistry," "engineering," "geology," "mathematics,"  
"meteorology," "oceanography," "physics," "science")
```

When mapping all unique combinations of these two terms across the 100 universities, this led to 9,000 unique combinations, and thus 9,000 searches. The R files associated with this task are:

- 1.analyze\_json.R
- 1.analyze\_school\_websites.R
- 1.clean\_school\_website\_data.R
- 1.clean\_schools.R
- 1.download\_school\_websites\_curl.R
- 1.dump\_html\_txt.R

1.parse\_json\_to\_csv.R  
 1.pull\_school\_websites\_json\_api.R  
 1.school\_list\_url.R

All of the data files associated with the task are in /out/school\_websites/.

The regular expressions for pulling the dates out of the text files look like this:

"\\b(3[01][12][0-9]|0?[1-9])/([1|0-2]|0?|[1-9])/([?:[0-9]{2}})?[0-9]{2}\\b" and  
 "\\b(((Jan(uary)?|Feb(ruary)?|Mar(ch)?|Apr(il)?|May|Jun(e)?|Jul(y)?|Aug(ust)?|Sep(tember)?|Oct(ober)?|Nov(ember)?|Dec(ember)?))\\s+\\d{1,2},\\s+\\d{4}))\\b."

Together they capture most date combinations, such as Mar 1, 2011, April 19, 2020, 10/7/17, 01/09/2021, etc.

## University Twitter Feeds

The R files associated with this task are:

2.pull\_twitter\_feeds.R  
 2.parse\_twitter\_data.R

All of the data files associated with the task are in /out/twitter/.

## Academic Associations

The R files associated with this task are:

3.downloads\_learned\_societies.R  
 3.clean\_learned\_societies\_data.R  
 3.convert\_pdf\_text.R

All of the data files associated with the task are in /out/learned\_societies/.

Analysis of the grants and scholarship data (Sections 4 and 5 of this report) follows an identical format: The R files are designated by the “4” and “5” prefix respectively in the /code/PREPROCESSING/ folder, and the relevant data files are appropriately named in the /out/ folder.

In recent years, Diversity, Equity, and Inclusion (DEI) ideology has spread aggressively throughout institutions of higher education in the United States and throughout the Western world. Arising from within the humanities and social sciences, DEI has begun to influence the natural sciences.

The spread of DEI has torn through barriers established to protect students and professors from ideological intrusion into their learning and research. By systemically dismantling academic freedom and freedom of association, DEI's followers have established themselves as the preeminent ideologues on America's college campuses today.

This rapid and radical transformation of the American academy has left many outside these institutions bewildered. In an attempt to document the natural sciences' conversion from seeking truth to forwarding ideological goals, the National Association of Scholars commissioned this study, *Ideological Intensification*.

*Ideological Intensification* documents the dramatic takeover of the natural sciences by DEI ideologues over the last two decades.<sup>1</sup> This study contains five new datasets—comprising over 30 GB of data in over 280,000 files—focusing on American college and university webpages, Twitter accounts, annual programs of academic associations, grants of major scientific research funders, and publications of scientific research. These datasets make this one of the largest such attempts to document DEI's ideological takeover of America's natural sciences.