Show me the money! A roadmap to biomedical grants

Project 4 Final Write-up Report

Team: Quarto_quartet.qmd

Name	JH Email
Ding Ding	dding15@jh.edu
Lo-Yu Chang	lchang36@jhmi.edu
Hui Yao	hyao21@jh.edu
Linda Ye	lye25@jh.edu

Introduction

Federal funding is the leading source of scientific research, accounting for nearly 40% of research and development expenditures. In the current sociopolitical circumstances, securing research funding is becoming less certain and even less predictable in the United States. This differentially impacts academic institutions and researchers, particularly students and junior to mid-level faculty at large, elite research institutions such as Johns Hopkins, which receives the most federal grant dollars in the country.

Compared to 2024, federal spending on scientific research has decreased by 53% across multiple agencies since the new administration took office in January 2025, and at least $5 billion has been frozen at some point at multiple elite research institutions (Malakoff and Brainard 2025). These federal funding agencies have publicly available search tools to identify active projects and previously awarded grants, however these websites are maintained separately by each department, and the advanced search criteria are too clunky, specific, and impractical for the average user to utilize efficiently. The National Institutes of Health (NIH) RePORTER website has a visual dashboard which provides a broad overview of grants under its own purview, however granular details are lacking. The National Science Foundation (NSF) does not maintain a dashboard or provide detailed statistics. Thus, we aim to fill this gap by providing an user-friendly dashboard built on aggregated data to allow users to explore biomedical research funding and understand the current funding landscape.

As we near the end of 2025, it is apt to look back on the year to evaluate how funding dynamics have changed in order to prepare for the future and set expectations. Our primary aim was to aggregate and analyze available public data from the separate federal funding agencies on currently and previously awarded biomedical, clinical, and human research topics and describe the trends in allocation of federal research grant awards. Our original goal was to use data from separate sources (NIH, NSF, Department of Defense [DoD], Department of Energy [DoE], Department of Health and Human Services [HHS], and ClinicalTrials.gov) to generate an aggregated, digestible data visualization dashboard and create an interactive analytic tool incorporating machine learning methods to predict the most appropriate funding mechanism based on the user's proposed research domain and subdomain, university affiliation, among other details. The intended audience is students, junior faculty, and any researcher interested in current trends in research activity and funding allocation.

Methods

Search Strategy and Data Collection

The NIH and NSF were the agencies chosen for this data analysis because they are the funding agencies most germane to our target audience. We examined the time period after 01/01/2021 in order to analyze the current trends after allowing washout of the effects of the COVID-19 pandemic, which was associated with many spurious societal and financial events in 2020. The date of last data collection was 12/01/2025.

NIH RePORTER

All new or renewed, active and expired grants awarded after 01/01/2021 from all available funding institutions relating to biological and biomedical research were included. Subprojects were excluded. We excluded grants from centers whose primary mission was unrelated to human health research: the National Center for Health Marketing (NCHM), National Center for Research Resources (NCRR), NIH Center for Scientific Review (CSR), NIH Office of the Director (OD), Health Insurance Reform Program (HIRP), Laboratory Science, Policy, and Practice Program Office (LSPPPO), Office of the Chief of Public Health Practice (OCPHP), Office of Workforce and Career Development (OWCD), Public Health Informatics and Technology Program Office (PHITPO), Public Health Program Office (PHPPO), and Scientific Education and Professional Development Program Office (SEPDPO).

There were >70,000 grants fulfilling this inclusion criteria. Data retrieval via API was attempted, however due to data transfer limits, the data were systematically exported in chunks to CSV format from advanced search results. The following relevant fields were included: project abstract, title, public health relevance, administering institution/center, application ID, award notice date, project number, grant type, activity category, project start and end dates, organization name and location, fiscal year, and total cost.

NSF

All active and expired grants awarded after 01/01/2021 from the Division of Molecular and Cellular Biosciences (BIO/MCB) and Behavioral and Cognitive Sciences (SBE/BCS) were included, as these were the organizations most related to human and biomedical research. The remainder of the NSF institutions funded grants that were predominantly non-human-related and excluded from data collection. Similar to the NIH RePORTER, there were limitations on data retrieval via API. The data were exported to CSV format from the advanced search results. The following relevant fields were included: award number, title, abstract, funding office, award instrument, organization name and location, project start and end dates, last application amendment date, and awarded amount to date.

The full raw dataset can be accessed through here:
https://drive.google.com/drive/folders/1yXcOL0Kl7UDMljaD_2O-Ss07FjI0UcX6?usp=drive_link

Data Wrangling

After obtaining the raw data from both sources, we column-matched both datasets and merged them. To account for the non-exact naming for institution names (eg. Johns Hopkins University and Johns Hopkins School of Medicine), we used fuzzy string matching to collapse non-exact matched institute names. For efficient computing, we reasoned that the similarly named institutions should only be considered in close geographical locations. Thus, we stratified the institutions by state and city name, fuzzy matched for individual city level, then merged the data. Parallel computing was used to increase speed.

Given the large number of variables in each dataset, a composite data dictionary was created to clarify variable definitions across the NSF and NIH datasets and to organize variables in preparation for merging during data aggregation.

Data Cleaning

Following data wrangling, additional data cleaning steps were performed to ensure variable consistency and analytical validity. These steps included resolving inconsistencies introduced during fuzzy institution name matching. For example, standardizing punctuation in institution names for schools with multiple campuses. Across NSF and NIH records, funding IDs were standardized, composite variables were identified and renamed, and formats and data types were coordinated. Duplicate and incomplete records generated during data aggregation were removed. The final cleaned dataset was saved as an RDS file and deposited in a GitHub repository to support reproducibility.

The GitHub repository can be accessed here:
https://github.com/jhu-statprogramming-fall-2025/project04-quarto_quartet-qmd

Outcomes

Since many grants are awarded at least several months prior to the project start date, we chose to use the date of award notification to analyze funding by year. The date and year of award notification was felt to be more closely reflective of up-to-date trends in funding. The date of award notification was readily available in the NIH dataset, however the NSF dataset did not have an equivalent variable. Thus, in the NSF dataset, the date of last application amendment was used as a surrogate date thought to be most closely reflective of the time of award notification.

In order to best quantify the total number of distinct, funded awards, we performed data wrangling to exclude duplicate projects and subprojects. The total cost variable, which is a sum of the direct and indirect costs of a grant, was used as the funding amount.

Version Control

GitHub (San Francisco, CA) was the platform used for version control. Local and remote repositories were maintained. Git programming was performed locally using the command line interface. Due to the GitHub policy regarding large raw files, the unprocessed raw data was stored separately as a publicly available file in Google Drive, allowing users to access the raw data.

Machine Learning

We aimed to use grant types, institution, funding agency, state, and years as predictors to develop an interpretable machine learning model, not only to help predict award amounts, but to analyze the importance of predictors to understand the potential trend behind award amounts. The machine learning was performed using `tidymodels`. The dataset was first randomly split to 80% training and 20% validation sets. To account for the uneven distribution of individual predictors, random sampling of the training set was done in a stratified manner to ensure a relatively similar distribution of individual covariates between the training and validation sets. Linear regression and random forest models were used on both crude award amount and log-transformed award amount. R-squared and root mean-squared error (RMSE) were calculated to evaluate model fit and accuracy.

Text Analysis

In order to evaluate recent trends and different agencies' missions, we analyzed titles and abstracts using a 'tidytext' workflow parallel to the data wrangling tools. We first tokenized the text into unigrams and bigrams using 'unnest_tokens()' and removed English stop words as well as self-defined low-information terms (e.g., project, study, grant). For each agency, we computed relative frequencies of tokens. The count of tokens after standardization was plotted as comparative barplots of frequent terms and bigrams.

To study how language evolved over time, we aggregated bigram counts by Source × Fiscal year and again converted them to proportions of all bigrams in that stratum, effectively performing a time-series analysis of categorical text features. Finally, we grouped selected bigrams into three interpretable "buzzword sets" (scientific content, narrative language, and training/education) and plotted their yearly proportions, comparing proportions across groups and over time.

Interactive Dashboard

The final product was an interactive dashboard deployed as a website. R (version 4.5.2, R-project.org) was the predominant programming language used in addition to CSS. Positron (Posit, Boston, MA) and Visual Studio Code (Microsoft, Redmond, WA) were the main integrated development environments (IDEs) for R. The programming was performed using Quarto (Posit, Boston, MA) dashboards. The final website was deployed using Shiny Apps (Posit, Boston, MA). Both static and interactive pages were designed. The following R packages were employed: `rsconnect`, `shiny`, `tidyverse`, `leaflet`, `bslib`, `bsicons`, `fontawesome`, `quarto`, `lubridate`, `sf`, `tigris`, `knitr`, `here`, `ggiraph`, `scales`, `flexdashboard`, `tidytext`, `tibble`, `DT`, and `readr`.

Graphs and plots were generated using the `ggplot2` package within `tidyverse`, and interactivity was additionally applied using the `ggiraph` package. The interactive map was generated using the `leaflet`, `sf`, and `tigris` packages. Value boxes were designed to illustrate digestible summary statistics using the `bslib`, `bsicons`, `fontawesome` packages. The interactive website was then deployed using `rsconnect` and the `shiny` server.

Programming Paradigms

Multiple statistical programming paradigms were used to generate the final product. Parallel computing was used to perform fuzzy string matching to resolve non-exact institution names recorded in the databases. For example, a single, large university may have multiple schools or names under the main institutional umbrella.

Machine learning models were created from the dataset, albeit the models were ultimately not informative. Using total award amount as the response variable and institution, grant type, and funding year as the predictors, linear regression and random forest models were generated with and without log transformation of the total award amount.

Many of our exploratory analyses and results are prominently displayed on our website with interactivity. Functional programming paradigms were employed to generate interactive plots in a Shiny dashboard. We generated static graphs and tables, an interactive U.S. map illustrating the funding distribution per state by source, an interactive histogram of funding distribution by year and agency, interactive heatmaps of the award count, dollar award funded, and percent change in total award amount by state, funding institute/center, and grant types and subtypes, and an interactive dashboard visualizing the results from an exploratory text analysis.

Finally, throughout the process, we used version control to maintain a shared repository and reproducible document using git in the command line and GitHub.

Results

Exploratory data analysis

Exploratory data analysis was conducted to examine trends in research funding across states, award years, and award types, with the goal of characterizing recent patterns in the biomedical funding landscape.
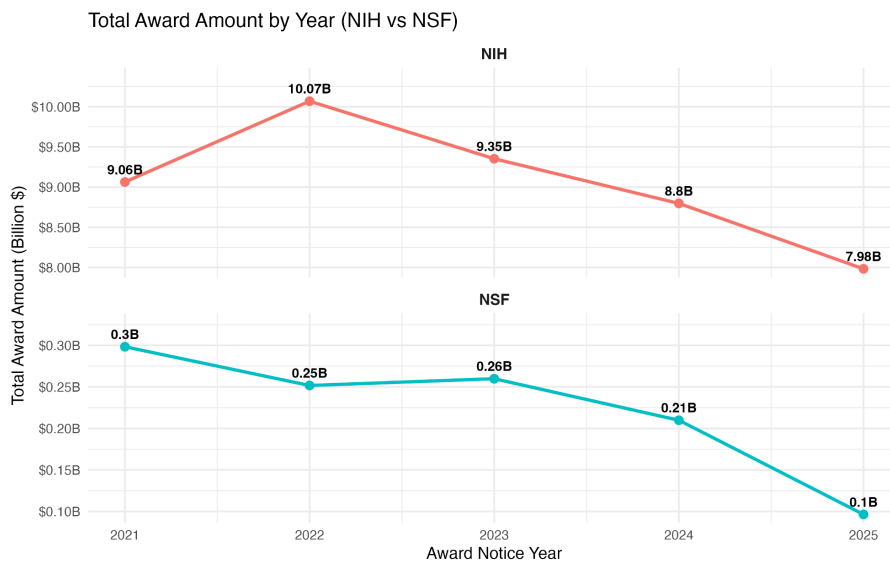
Total Award Amount by Year (NIH vs NSF)

Figure 1. Total Award Amounts for NIH and NSF by year (2021–2025)

Grant funding has declined consistently since 2022, with 2025 experiencing the most pronounced reductions over the past five years. Compared with 2024, total funding in 2025 dropped by 10.2%, alongside a substantial 31.4% decrease in award counts. As shown in Figure 1, total annual award amounts from both the NIH and NSF peaked in 2022 and have steadily decreased in each subsequent year, with the sharpest decline occurring between 2024 and 2025.
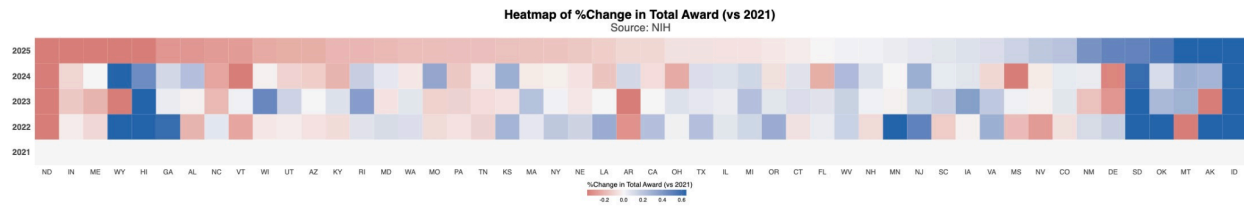
Figure 2. Heatmap of Percent Change in Total NIH Award Amounts Relative to 2021, by State
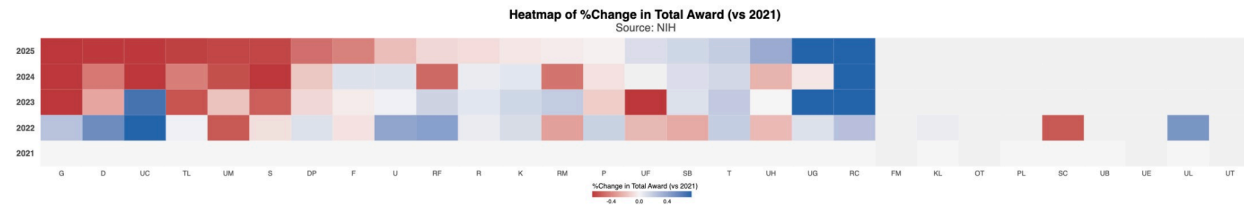

Figure 3. Heatmap of Percent Change in Total NIH Award Amounts Relative to 2021, by Award Type
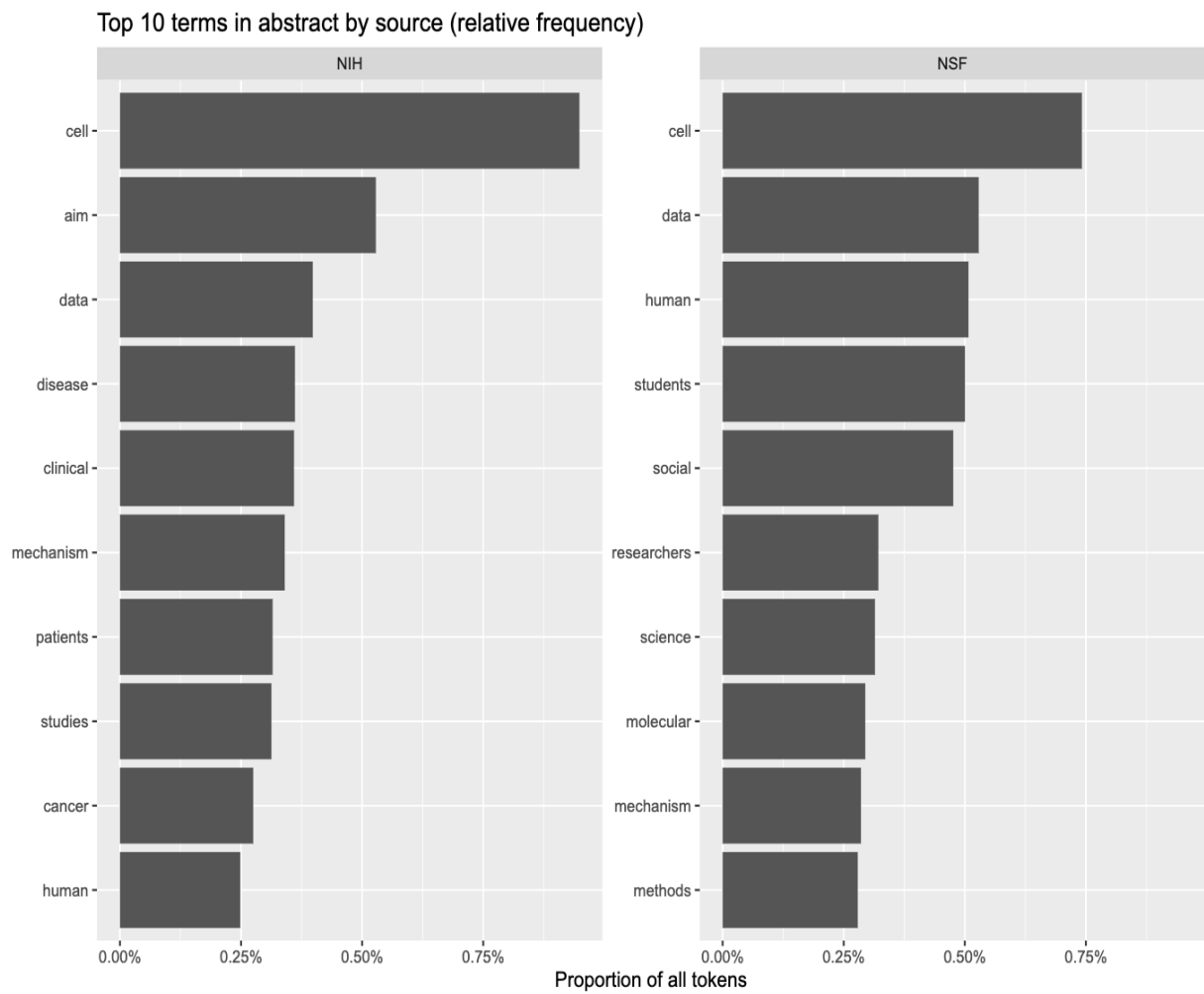
Figures 2 and 3 display heatmaps of the percent change in total NIH award amounts relative to the 2021 baseline, stratified by state and award type, respectively. These visualizations highlight substantial heterogeneity in funding trajectories across both geographic regions and grant types. As shown in Figure 2, by 2025, Idaho, Alaska, and Montana experienced the largest relative increases in NIH funding compared with 2021, whereas North Dakota, Indiana, and Maine exhibited the greatest relative declines. These patterns show uneven geographic shifts in funding distribution over time. It does appear that over 2022-2025, Idaho and South Dakota had already been receiving increased grant funding, and North Dakota, Indiana, and Maine had already been receiving decreased grant funding. Figure 3 illustrates the variation in funding trends across NIH award types. Relative to 2021, RC (high-impact research), UG/UH (cooperative agreements), and T (institutional training) -type grants had the largest increases in total award amounts by 2025. In contrast, G (infrastructure support), D (drug development), and UC (university cooperative) grants experienced the most pronounced decreases in funding over the same period. Similar to the change in grant funding by state, it appears that certain grant subtypes were already experiencing a trend toward relative increased or decreased funding over the previous years.

Text Analysis

When we compare the top unigrams (Figure 4A), both agencies commonly emphasized basic biology (e.g. *cell, data, molecular, mechanism*), but other contexts were different. The NIH more commonly funded abstracts mentioning topics relating to disease and patients (e.g. *disease, clinical, patients, cancer*), whereas the NSF tended to fund topics relating to people and training (*students, social, researchers, science, methods*). This suggests that the NIH

prioritized topics relating to diseases and clinical questions, while the NSF prioritized the scientific community and training.
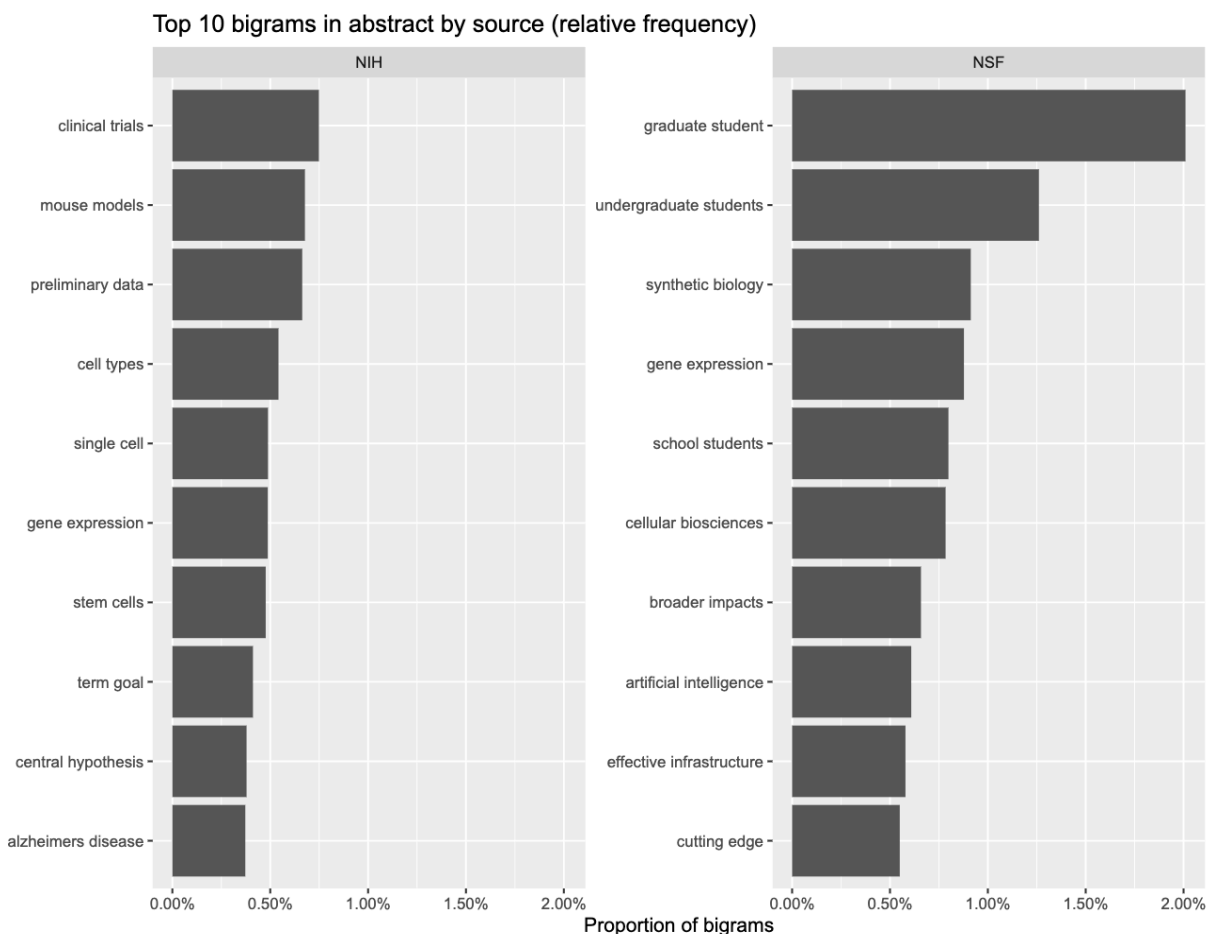
A.

Top 10 terms in abstract by source (relative frequency)

B.

Top 10 bigrams in abstract by source (relative frequency)

Figure 4. Top 10 single terms (A) and bigrams (B) for NIH and NSF (2021–2025)

The bigram barplots (Figure4) and the log-ratio of bigram frequency figure (Figure 5) further emphasized the difference in funding priorities by keywords. The NIH-funded grants included keywords that were almost entirely disease or clinic oriented (e.g., *clinical trials, breast cancer, lung cancer, heart failure, chronic pain, cognitive impairment, clinical outcomes, anti-tumor*). In contrast, NSF-funded grants referred to funding programs and broader-impact framing (e.g., *graduate student, undergraduate students, broader impacts, cellular biosciences, synthetic cells, national science, science foundation, effective infrastructure, local communities, participant observation*). The text found in funded NIH projects tended to be related to diseases and clinical outcomes, while the text from NSF-funded grants prioritized basic biosciences, training, infrastructure, and broader societal impacts.
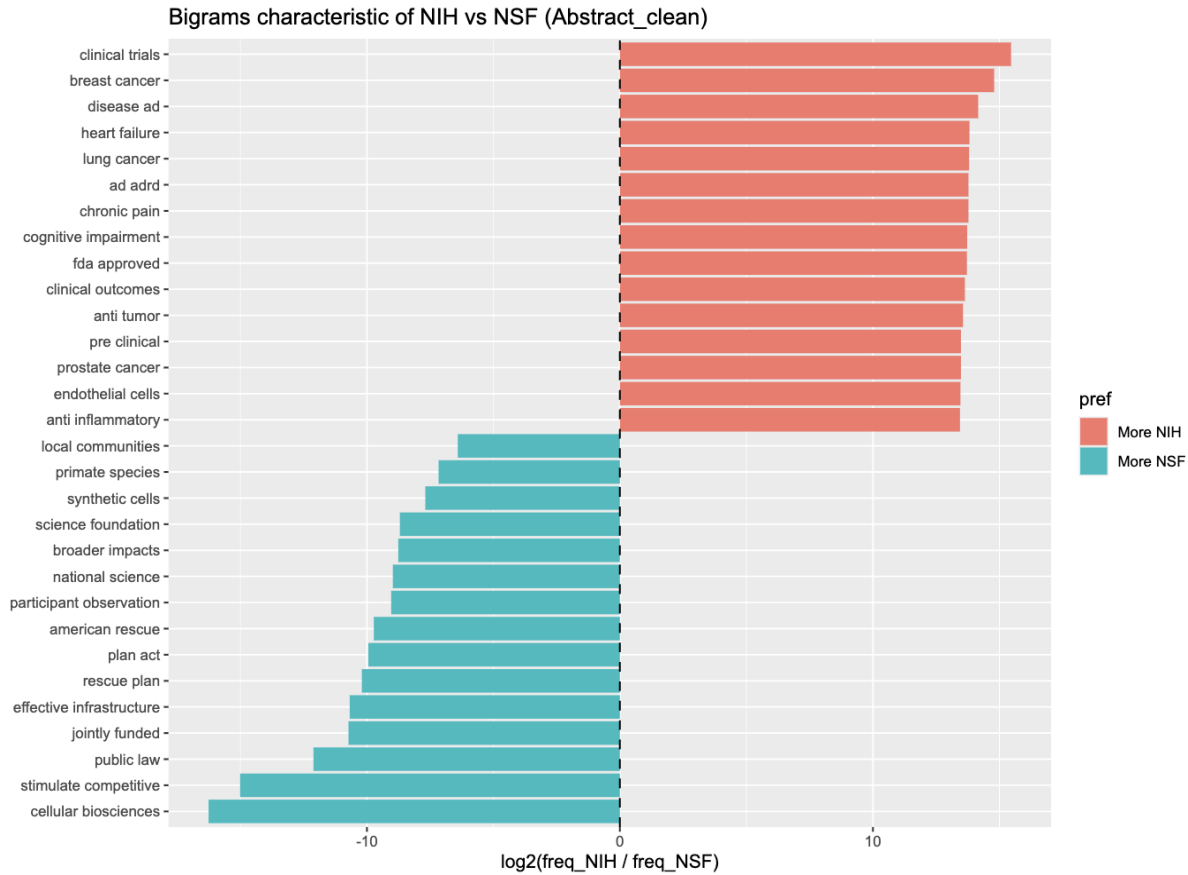
Figure 5. Distribution of Log-Ratio of Bigram Frequency Between the NIH and NSF Abstracts (2021–2025)

The faceted time-series panels (Figure 6) showed that the proportion of funded grants containing language relating to COVID (i.e. *SARS-CoV*) declined after 2021 from both agencies, while more general mechanistic topics were stable or increased (e.g. *cell types, mouse models, gene expression*). The NSF carried the most signal for '*gene expression*', '*molecular mechanisms*', and '*single cell*', and showed a marked rise for '*single cell*' in 2025, consistent with a growing interest in -omics and cell-level biology. The NIH, in contrast, funded grants consistently mentioning use of the animal model (e.g. *animal models, mouse models*) and disease-specific phrases, such as '*breast cancer*' and '*Alzheimer's disease*', although with some decline after 2023. Overall, the trends suggest a shift away from prioritization of COVID toward mechanistic and cell-based science, with the NSF prioritizing upstream basic sciences and the NIH on downstream disease and clinical translation.
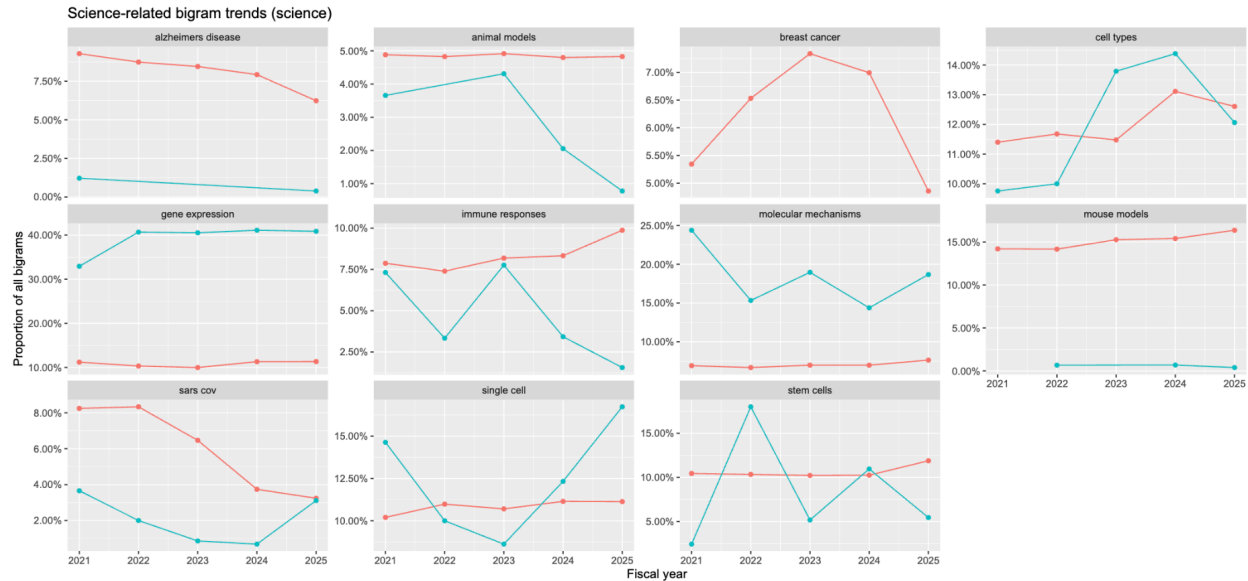
Figure 6. Bigram Trends of Select Science-Related Keywords (2021–2025)

Machine Learning

Due to the skewed data distribution, we log-transformed the award amount to better fit the data. Based on the R-squared, log-transformation did help the model fit, as seen by the increase from 0.005 (not displayed) to 0.07 (Figure 7). However, even with transformation of the response variable, the R-squared was still very low, suggesting a large amount of variation unaccounted for by the model. Thus, we conclude that our current model can not reliably predict award amount based on the current dataset.
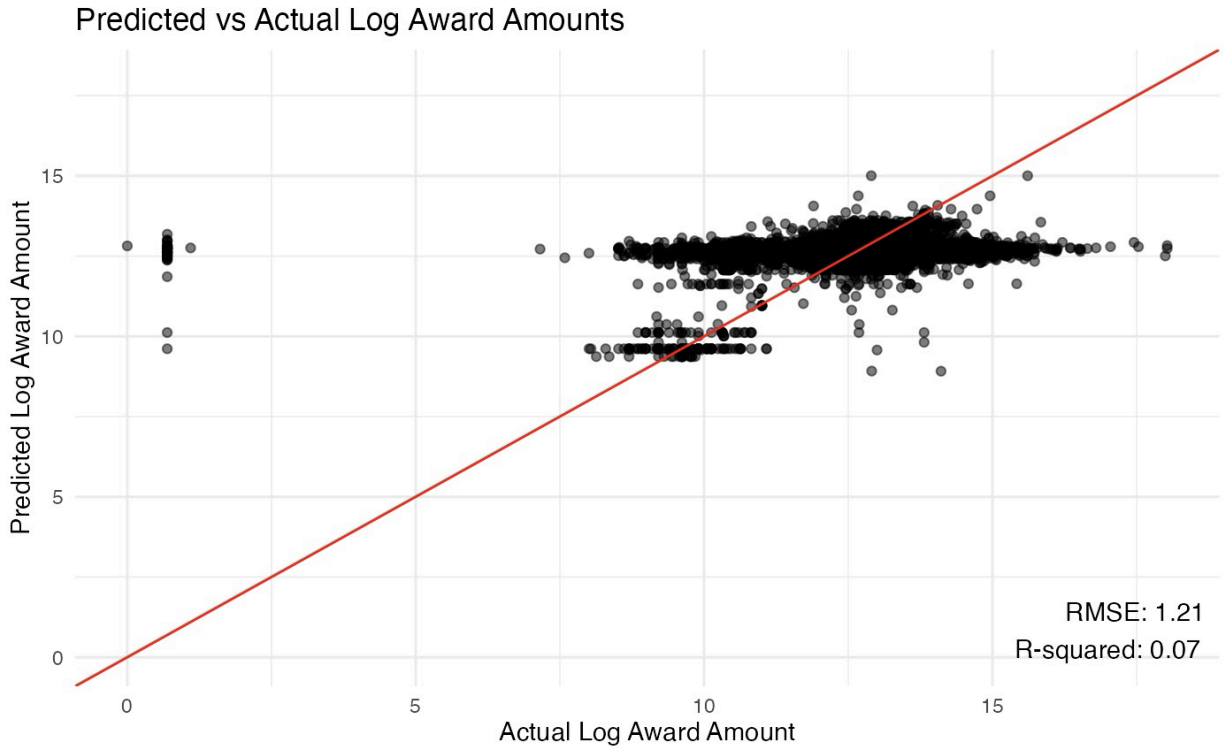
Figure 7. Distribution of random forest model predicted and actual award amount (post log-transformation)

Table 1. Metrics evaluating the accuracy of the crude and log-transformed award amount in the random forest model.

	RF	RF+Log
RMSE	1239285	1.21
RSQ	0.00579	0.0708
MAE	388156	0.842

RF: random forest; Log: log transformation; RMSE: root mean squared error; RSQ: R-squared; MAE: mean absolute error

Functionality

Overall, the Shiny dashboard works as intended and the four pages are tightly integrated around a single objective: helping users understand the current biomedical research funding landscape, specifically for the NIH and NSF.

The **Introduction** page introduces the data source and time frame and displays the raw dataset of funded awards as well as a data dictionary, so users can view the raw records and interpret each variable before interacting with the dashboard.

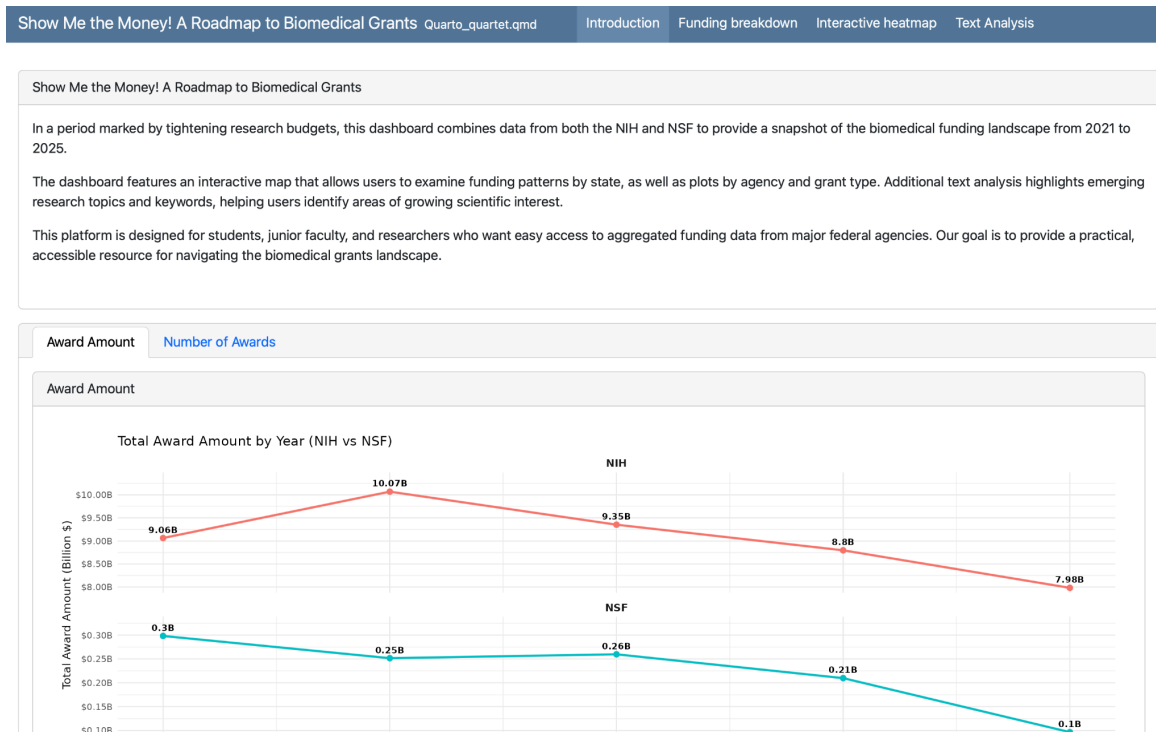Here are the links to [raw dataset](#), [GitHub repository](#), and deployed [dashboard](#).

Figure 8. Interactive Dashboard: Introduction

The **Funding breakdown** page summarizes core statistics (change in award counts and numbers, total award amounts, total new & renewed awards) and provides an interactive choropleth map of funding distribution by state and interactive agency-year histogram. Input filters for the year and funding source updates the map, allowing users to quickly search for a specific funding source and year to see the distribution of award amounts across the United States. Additionally, hovering over each state reveals specific funding amounts of that year.
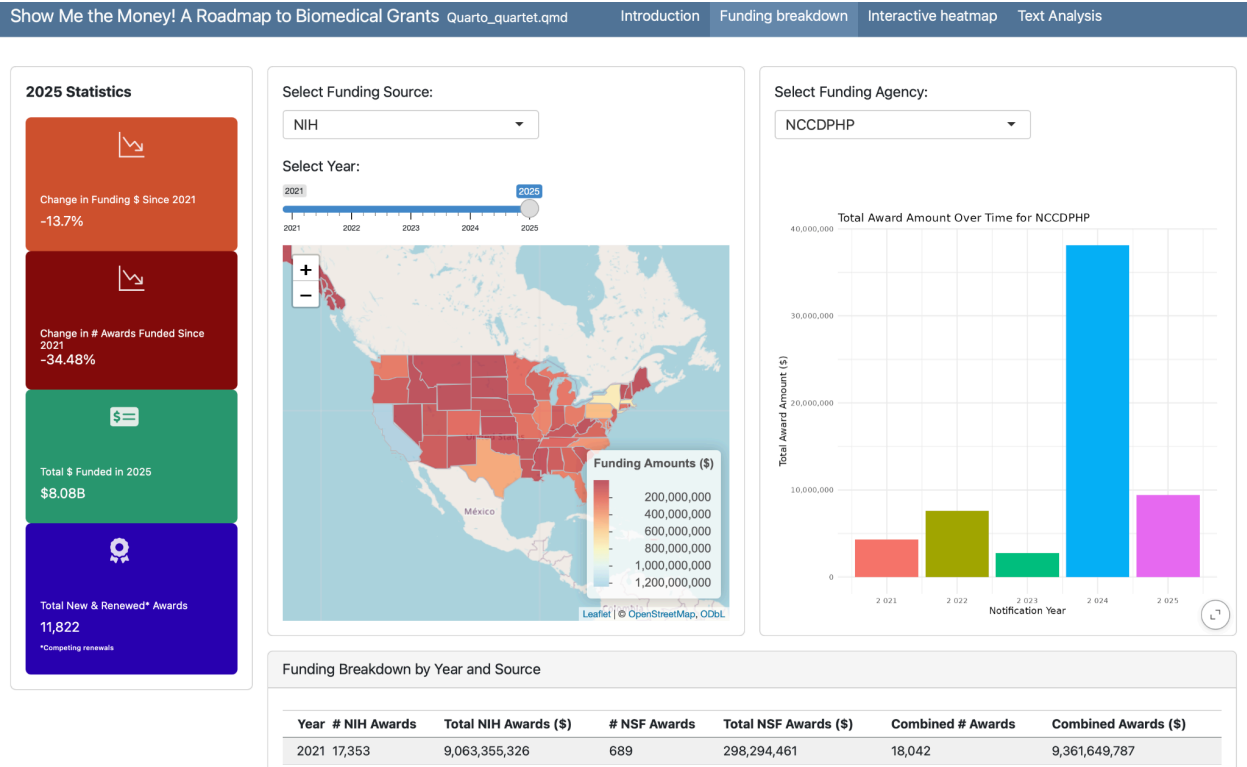
Figure 9. Interactive Dashboard: Funding breakdown

The **Interactive heatmap** page further elucidates funding trends and details by allowing users to select the X-axis (state, organization, grant subtype, grant type) and the color metric (mean/median award, total funding, counts, or percent change vs. 2021). These input functions are linked so that a single choice of group and metric produces a consistent, comparative view across categories.
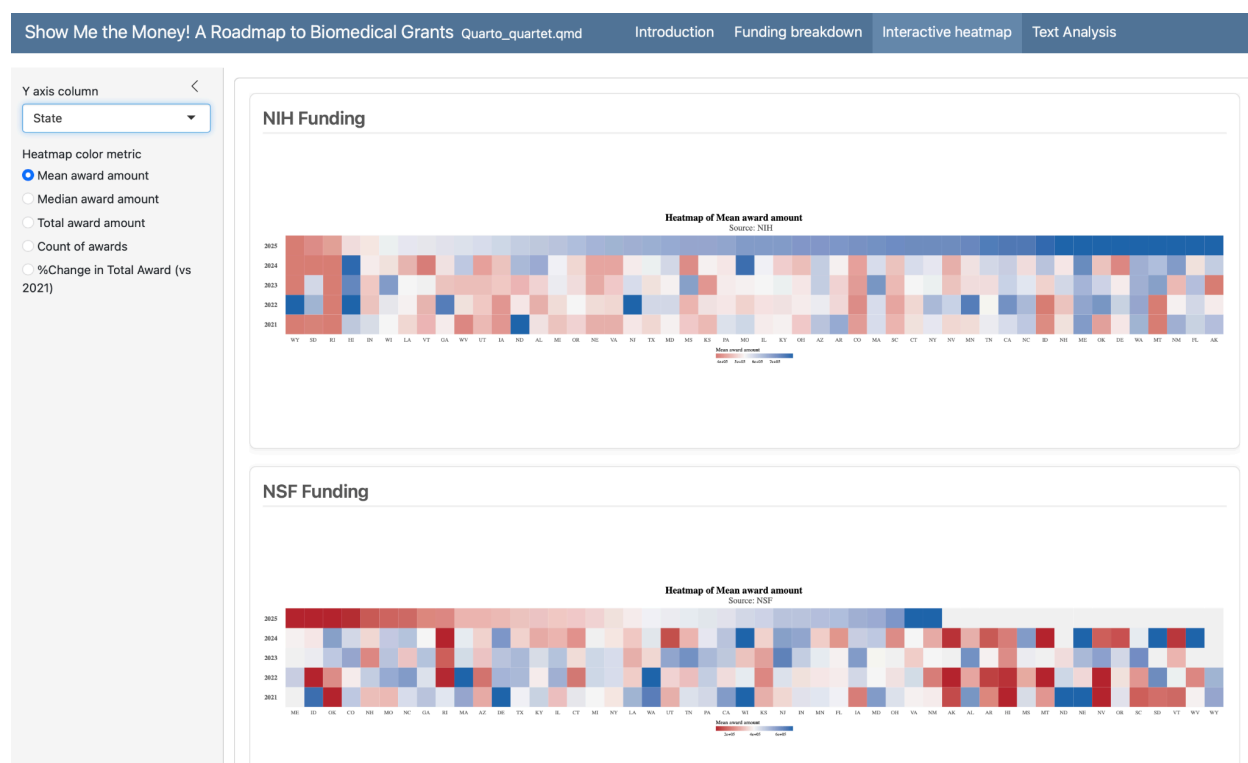
Figure 10. Interactive Dashboard: Interactive heatmap

Finally, the **Text analysis** page applies the agency and year filters to the tidy-text outputs. Users can select NIH vs NSF, titles vs abstracts, and the number of top bigrams to display, and can also view time trends for pre-defined "buzzword sets" (narrative, science, or training).
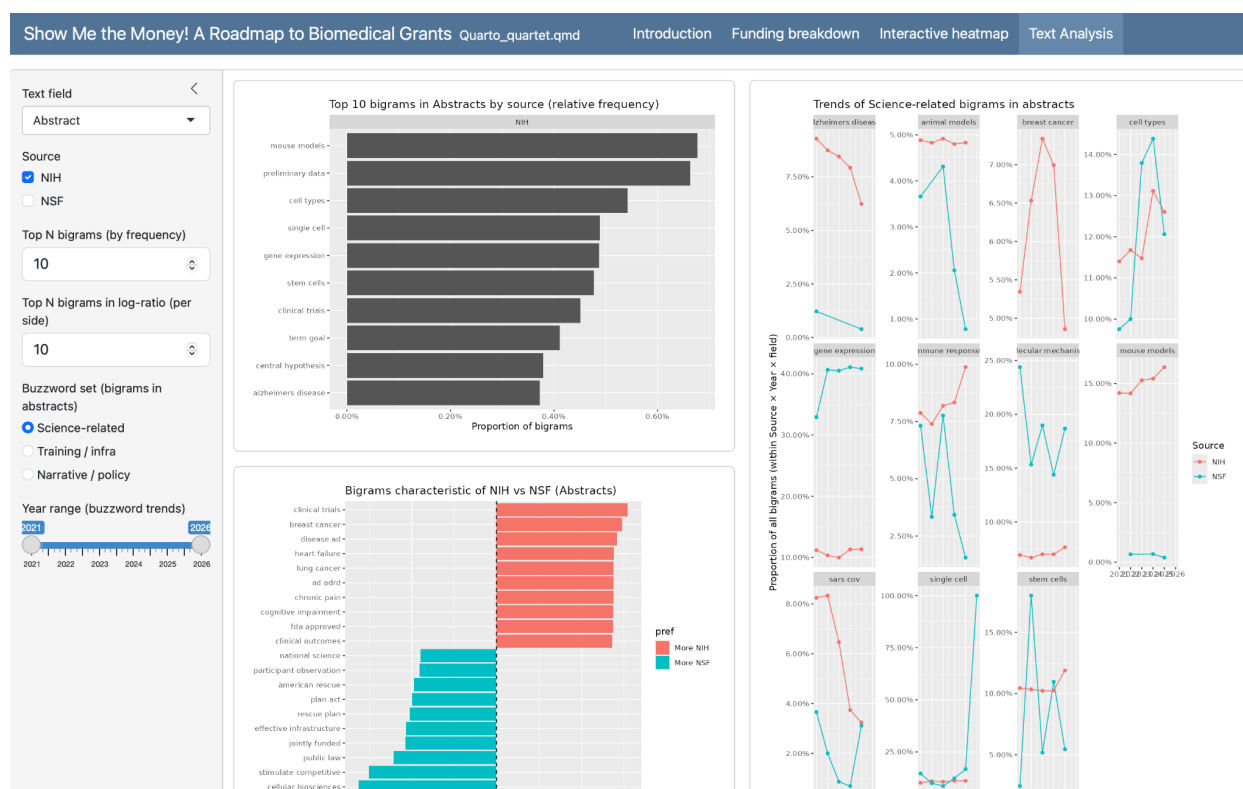
Figure 11. Interactive Dashboard: Text analysis

Because all pages share the same underlying dataset, the user can move from an overview of funding patterns, to organization- or state-level summaries, and then to the language of the corresponding abstracts without breaking context. Taken together, these design choices indicate that the components are well integrated and that the data analytic product supports the intended exploratory tasks.

Discussion

Data Synthesis

Based on our exploratory data analysis, there has been a consistent downtrend in federal grant funding by both the NSF and NIH after an increase in funding in 2022 compared to 2021. The increase in funding in 2022 may potentially be associated with the high inflation of that year following COVID, however this is purely speculative. Every subsequent year has resulted in a decrease in funding compared to the year prior in both total funding dollars as well as number of awards. Compared to 2024, 2025 had the greatest decrease in grant funding compared to years prior. Of note, we acknowledge that there was a government shutdown lasting 43 days between October 1, 2025 and November 12, 2025, which may have contributed to the decreased funding in 2025, during which time the NIH was closed and grant review

committees did not convene. In addition, the date of last data extraction was December 1, 2025, which would not include the remaining 30 days of the fiscal year. We view this as only a minor limitation of the study design, as significant numbers of grant reviews would not have occurred during the winter holidays.

Analysis of funding trends by state and NIH grant subtype showed state-specific increases and decreases in funding. Of note, certain states had already been experiencing similar trends in the years prior to 2025. Overall, it seems that the notable funding trends of 2025 appear to follow the same general patterns as years prior albeit more dramatically.

Text analysis demonstrated overall that the text found in funded NIH project abstracts generally tended to be related to diseases and clinical outcomes, while the text from NSF-funded grant abstracts prioritized basic biosciences, training, infrastructure, and broader societal impacts. Notably, there was a clear increase in frequency of the term '*single cell*' in NSF-funded grants in 2025, likely relating to recent growing interest in -omics and cell-based research.

Finally, machine learning models predicting award amount based on grant type, institution, and year were uninformative, likely due to unadjusted confounding. The award amount is strongly correlated with the grant subtype, by knowledge of funding mechanisms. However, without grant subtype, the only predictors would have been year and institution, which are too broad for meaningful use. The overall conclusion is that this dataset was not the most effective to use in machine learning modeling.

Learning Points

One of the greatest takeaways from this project is the appreciation of how accessible it is to create an online interactive dashboard for data visualization. Having already had experience designing and publishing a Quarto website, the design of a Quarto dashboard incorporating interactive visual elements was a natural extension of those skills and did not mandate extensive experience in CSS or HTML programming. It was refreshing to learn how many different ways data can be displayed and organized such that it is digestible and interesting to the user. While programming static data dashboards was more straightforward, incorporating interactivity was not significantly more complicated, provided additional layers of data presentation and details, and enhanced user experience by customizing the displays to his or her specific interests.

Additionally, we learned that it is essential to be familiar with the data and data sources, as the data clearly dictates the ability and appropriateness to perform certain analyses. As an example, machine learning models were in retrospect not the most appropriate methods to apply to our specific data. We were vaguely aware of our data sources, but until we were already deep in the process of data collection, we were not expecting the volume of data and degree of data heterogeneity. We then were required to slightly modify the types of data

analyses we could perform. In the same vein, we became aware that not having access to rejected or unfunded grants would severely limit our ability to perform predictive modeling. In the real world, we would anticipate being more involved and involved earlier in the formulation of the study question, database design, and data acquisition process to answer that question.

Similarly, it was also helpful to know of different methods for data collection. API appeared to be the most available and straightforward method to extract data, however we also learned there are limitations as to how much data can be collected in that manner, even when using iterative loop functions. We circumvented this by exporting the data directly to CSV files from the sites' advanced search tool.

The computational resource limitation not only existed for processing the raw data but also when we attempted to use fuzzy string matching to resolve non-exact institution names. Initially, we tried all-by-all fuzzy matching to the entire dataset, which required exceedingly large physical memory to achieve. We learned that the way to efficiently divide the computational work was critical. By stratifying the institutions by state and city level, we were able to achieve a much more efficient computation task.

Finally, working with grant titles and abstracts underscored the special role of text data in this analysis. Text analysis ultimately became one of the more novel and elucidative areas of this project. Language encodes agencies' priorities, methods, and target populations in ways that are not captured by structured variables such as state or grant mechanisms. On the other hand, text data is also noisy, high-dimensional, and context-dependent. This experience showed us that text features are both challenging and promising: with more sophisticated natural language processing methods and richer labels, they could substantially improve study design, portfolio monitoring, and future machine learning tools.

Technical Challenges

While our original intent was to aggregate data from these multiple funding sources to generate a centralized database of actively funded projects, we realized early in the process that there was an extremely high volume of data from these separate sources. For example, the data collected from the NIH alone resulted in over 70,000 funded grants and large raw data files. In addition, the variables and details available from each source varied significantly, thus the aggregation of these data was not practical or productive to answer our question. Thus, we proceeded to limit our data sources to the NIH and NSF, as we felt those agencies were the most relevant to the pre-specified target audience.

We approached data extraction intending to use APIs. However, as mentioned, there was a large volume of data that fulfilled our inclusion criteria. In addition, given our data sources were federally maintained websites, there were limitations on the amount of data able to be extracted over a short period via API, even when functional programming paradigms using `purrr` were used in order to create iterative functions. We navigated this by performing detailed

searches using the built-in advanced search tools for each website and downloaded the data of interest into CSV files.

The final component of our original goal was to create a predictive tool utilizing machine learning to inform potential grant applicants of the likelihood of success in funding based on predictive variables, such as project topic or subdomain, state, institution, grant subtype, and proposed funding agency. In carrying out the design of this model, it became clear that there was significant data bias in that we only had data on projects that were funded. The lack of information on unfunded or rejected projects made the original purpose of this tool impossible to execute, as we were unable to train the data on any negative outcomes. Furthermore, there are structural limitations posed on individual grant types for award amount, thus machine learning was unable to yield meaningful results.

The NIH RePORTER did have aggregate data available on percent funding success based on grant type, however without individual-level data, we were unable to utilize this data. Therefore, we attempted to use the total award amount as a different response variable and institution, grant type, and year to execute this model. Despite modeling with linear regression and random forest methodologies and performing log-transformation of the response variable, the machine learning model performed poorly with an R-squared of 0.07 at best. This was consistent with our background knowledge of the data, which is that federal grants typically have a predetermined allocation annually, thus there was significant collinearity between our predictor and response variables.

Conclusion

Funding patterns varied by state, funding agency, and grant mechanism, reflecting heterogeneity in the U.S. research funding landscape. While overall funding levels differed across agencies, certain states and grant types at the extremes of relative funding change exhibited consistent trends across multiple years. In addition, keyword analyses revealed agency-specific and time-varying thematic priorities. Our analysis revealed that similar funding changes had already begun happening after a peak in funding in 2022. This could potentially be reflective of economic inflection points following the COVID-19 pandemic, such as a multifactorial response to inflation. However, in 2025, the trend in funding by both the NIH and NSF has decreased even further compared to prior years.

This project addresses the data integration and analytical challenge of heterogeneous, agency-specific datasets that are not typically available in standardized format. Existing federal funding agencies maintain separate publicly available search tools for identifying active projects and previously awarded grants. However, these tools rely on advanced search criteria that are overly specific, and impractical for the user to process funding information efficiently. By constructing a composite dataset and performing cross-agency exploratory analyses, this work enables a more comprehensive and integrated examination of funding trends in the biomedical field.

Beyond its analytical contributions, this work provides a reproducible and accessible resource for researchers, trainees, and institutions seeking to understand shifts in the funding landscape, particularly during periods of federal disruptions. In this context, the analysis also serves as a practical tool for identifying funding opportunities and mitigating uncertainty periods of constrained government funding.

References

Malakoff, David, and Jeffrey Brainard. 2025. "How Trump Upended Science." Science, 577.
U.S. National Institutes of Health. 2025a. NIH RePORTER - Grants and Awards Portal. https://reporter.nih.gov/.
———. 2025b. NIH RePORTER API Documentation. https://api.reporter.nih.gov/.
U.S. National Science Foundation. 2025a. NSF Award Search. https://www.nsf.gov/awardsearch/simple-search/.
———. 2025b. NSF Developer Resources - Open Government. https://www.nsf.gov/digital/developer.