

ADSA Data Science Community Newsletter

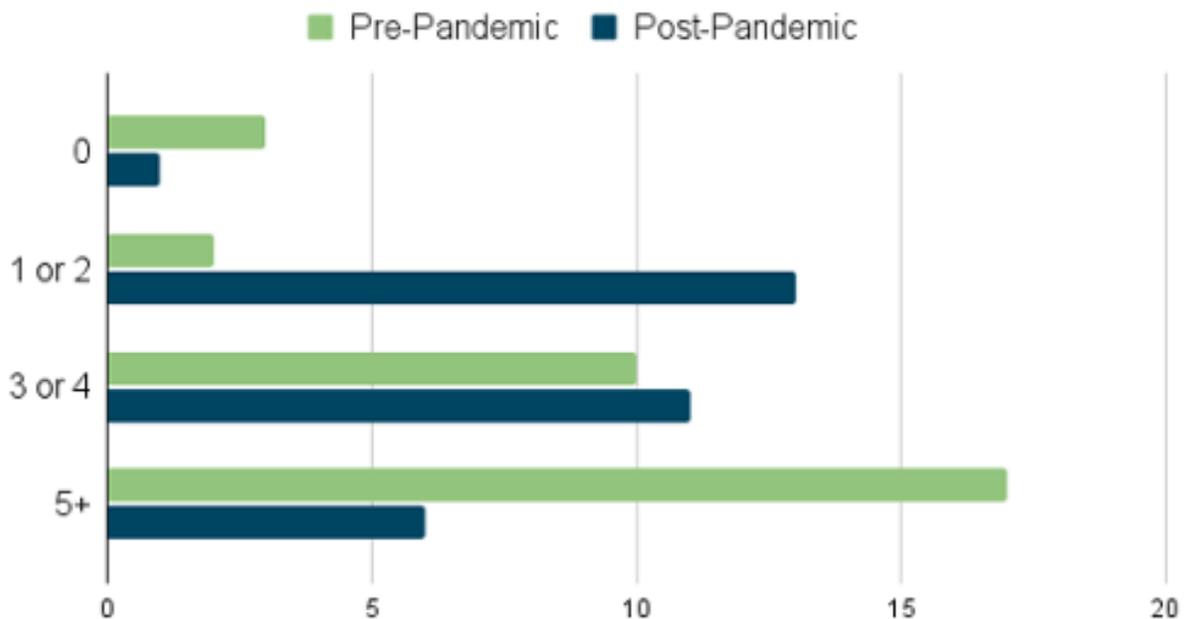
Data Science Community Newsletter features journalism, research papers and tools/software for January 27, 2022.

Please let us ([Micaela Parker](#), [Catherine Cramer](#), [Brad Stenger](#), [Laura Norén](#)) know if you have something to add to next week's newsletter. We are grateful for the generous financial support from the [Academic Data Science Alliance](#).

CAMPUS VS. WFH SURVEY RESULTS

In a non-representative poll of 34 readers the global trend toward more flexible work will impact academia.

How many days do you plan to spend on campus each week?



About a third of the people surveyed who had been going to campus 5+ days per week before the pandemic expect to maintain that level of presence on-campus post-pandemic. Most people pre-pandemic were on campus 4-5 days per week, but that shifts down to 2-3 days per week post-pandemic with the largest number of respondents marking that they're expecting a 1-2 days per week cadence post-pandemic. (If you want to contribute to this survey, [do it.](#))

This week's super short survey builds on these results to ask how these changes may impact the work experience and student experience of higher edu life going forward. Does work from home flexibility make academia a better place to work? A better or worse place to get an undergraduate degree? A graduate degree?

[>> Take Campus Quality Survey.](#)

ADSA In-Person Spring Meeting



ADSA MENTIONED IN THE FEDERAL REGISTER

For the first time, ADSA was [mentioned](#) in the *Federal Register*. Why? Through ADSA, **Kristen Eschenfelder** and **Bret Largent** at **University of Wisconsin-Madison** organized an 85-person letter writing campaign to explain to the **Department of Homeland Security** that Data Science is a STEM degree. The STEM designation allows foreign Data Science degree holders with F-1 visas to participate in the [OPT visa program](#) that allows them to work in the US.

Featured Events

See the [ADSA Events Page](#) for more details and more opportunities.

UTSA Draper Data Science Business Plan Competition.

Open to university students globally. Event is virtual & in-person: April 13-15, 2022. Teams can win prizes totaling \$100,000.

Application Deadline: 11 February 2022.

SCHMIDT FUTURES' \$40M VIRTUAL INSTITUTE OF SCIENTIFIC SOFTWARE

Two of the four inaugural recipients of the **Schmidt Futures [funding](#)** — **University of Washington** and **Georgia Tech** — are also founding members of ADSA. They are joined by **University of Cambridge** and **Johns Hopkins** to support research software positions. At UW, the funding will bolster the **eScience Institute's** 14 staff data science positions with 7 new **[software engineering positions](#)**.

ACADEMIC MERGERS AND ACQUISITIONS

Academia is not typically associated with merger and acquisition activity, but it has happened at least two times over the past decade. During the pandemic, **Boston College [acquired](#) Pine Manor College**, a 2-year community college that was unlikely to make it out of the pandemic without a big windfall such as BC's \$100 million purchase. Now Pine Manor College will be called Messina and will offer a residential 2-year experience to 100 students who can subsequently apply to Boston College (or any other college) if they decide to complete a 4-year bachelor's degree. In 2014, **New York University [acquired](#) Polytechnic Institute of New York** which became the **NYU Tandon School of Engineering**. This gave NYU an engineering school which it sorely needed, an opportunity to sell naming rights to the engineering school (also, \$100 million), and a fantastic piece of real estate in downtown Brooklyn. Though it is barely a trend, we expect to see more academic mergers and acquisitions as the entire sector faces lower enrollment. The bigger, more powerful institutions will selectively acquire smaller programs that fill gaps in their educational offerings and/or provide excellent campus locations for expansion.

GRAD STUDENTS GET A 25% RAISE

Princeton University grad students will be getting a 25% bump in their stipends, bringing them up to about \$40,000 during the 10-month academic year. One tweeter reports their net effective bump will go to \$45,600 and another reports the 12-month effective will be \$48,000. It's hard to tell if this has anything to do with the recent **[successful](#)** organizing action by **Columbia University's** grad students who will be getting quite a bit less. Interesting moment for salary transparency.

Princeton's **[Official announcement](#)**

[Tweet 1](#) (\$45,600) from a Princeton Humanities PhD student

[Tweet 2](#) (\$48,000) from Kelsey Hatzell, an assistant professor of chemistry at Princeton

Also, remember **[most places aren't Princeton](#)**

SAME-GENDER MARRIED MEN MAKE THE MOST MONEY

U.S. Census Bureau data collection practices recently became more inclusive of same-gender couples with the **American Community Survey 2020 (ACS)**. The **Brookings Institution [took a look](#)** at the data and found that same-gender male couples make more than same-gender female couples. Same-gender male couples also make more than heterosexual couples. Hello, gender wage gap! Brookings also found that being married is correlated with higher incomes for all three groups compared to their gender-matched cohabiting pairs and compared to singletons. It's quite unique to have nationally representative data on same-sex couples. It's illegal to even collect data on sexual orientation in some countries so this is a smashing breakthrough.

MUSTAFA SULEYMAN LEAVES DEEPMIND, JOINS GREYLOCK

Mustafa Suleyman, founder of AI powerhouse **DeepMind** that we featured twice in our best of 2021 issue for protein folding and natural language models, has **[joined](#)** venture capital firm **Greylock**. Suleyman has been on administrative leave since 2019 after accusations of bullying and boorish behavior for which he apologized in a **[podcast](#)**, to **Reid Hoffman** of Greylock (it's unclear if he

apologized directly to the people he managed).

IBM SOLD WATSON HEALTH TO A PRIVATE EQUITY FIRM

The heady days of Watson have *been* over, but now it's official. Watson Health was [sold](#) to Menlo Park-based private equity firm **Francisco Partners**.

INTEL TO BUILD 2 CHIP FACTORIES IN OHIO. \$20 BILLION INVESTMENT.

With Watson put out to pasture, **Intel** has [announced](#) it will build two new chip factories in Ohio just outside Columbus. Never underestimate the importance of hardware.

IRS WANTS TAXPAYERS TO USE THIRD-PARTY FACIAL RECOGNITION

In a move [met by](#) "incredulity" in the security and privacy world, the IRS is requesting that all taxpayers use **ID.me** to verify their identity by uploading selfies and additional identity verifying documentation. ID.me has been working with other agencies to prevent unemployment fraud. Their ability to handle the onslaught of Americans — almost all! — who must pay taxes is better than the IRS's ability to upgrade their identity verification capacity by tax season (which has already started). While I am generally skeptical of facial recognition, ID.me has had no breaches, often uses humans to help with verifications, exists only for the purpose of providing identity verification, and is likely to do a much better job than the IRS at preventing fraud. We don't want to turn into dogmatic commenters over here at the DSCN. That's no fun. So in this case, we're going to split with the privacy community and support this move. We would like to see some additional language about how ID.me will respond to law enforcement requests, one of the issues raised by critic **Caitlin Seeley George**.

US FALLS BEHIND CHINA IN SCIENCE

The National Science Foundation [released a report](#) that, "confirms China has overtaken the United States as the world's leader in several key scientific metrics, including the overall number of papers published and patents awarded." **Julia Phillips**, a retired physicist and key contributor, notes that the distrust between the international scientific community and U.S. government intelligence agencies is not helping. In related [news](#), another scientist — **MIT's Gang Chen** — who had been charged in connection with the China Initiative has now been cleared of any wrongdoing. He reports his trust is "shattered" and he will no longer accept any U.S. government funding.

Derek Thompson [notes](#) that one strategy federal funders could pursue is to give more small grants to researchers under 35 years old who currently get less than 5% of it. He cites **Thomas Kuhn's** research that found scientific revolutions tended to come from young scientists (not big corporations or seasoned faculty).

SCANDAL PLAGUES UNIVERSITY OF MICHIGAN, AGAIN

The **University of Michigan** [fired](#) its President, **Mark Schlissel**, because he was having an inappropriate romantic relationship with a subordinate while married and while ushering the school through a sexual abuse scandal. (See: [nearly 1200 rapes by University of Michigan athletic doctor.](#)) The evidence cited in *The New York Times* in the Schlissel case involves a lot of email correspondence, a good reminder that your university-provided email is available to internal and external investigators.

JUDGE FLEXES HIS VERBAL SKILLS AGAINST UNIVERSITY OF FLORIDA

University of Florida faculty had a [linguistically resplendent](#) win against the administration. Recall that earlier, the UF administration attempted to prevent faculty in political science from offering expert testimony in voting rights cases and prevented medical school faculty from offering expert testimony about the protective value of children wearing masks in school (which was not the maskless situation **Florida Governor Ron DeSantis** wanted). The school's position was that in cases where

faculty testimony opposed the policies of Governor DeSantis, it was disallowed. The judge's ruling used phrases such as the: "defendants' side of the scale sits empty; plaintiffs' overflows." The report turned up plenty of evidence that UF administration attempted to enforce the will of the Governor's administration including requesting that the words "critical" and "race" never appear in the same sentence.

MYTH-BUSTING THE NO-SUGAR DIET

Think you should stop eating sugar because it's going to cause all sorts of problems? Tired of hearing about the keto diet but not sure how to silence the chattering keto cheerleaders? You'll like this [article](#) by **Jennifer Rooke** at the **Morehouse School of Medicine**.

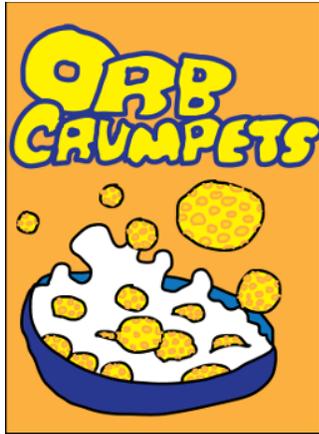
NEW PROGRAMS, FOLLOW THE MONEY

Click through to access [a structured spreadsheet](#) of New Programs and money moving around in academic data science.

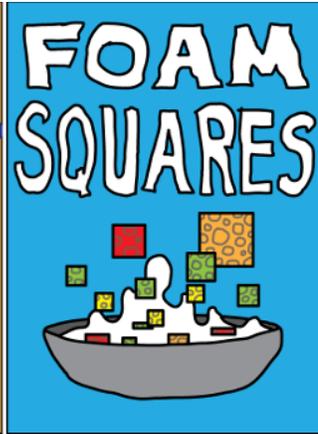
DATA VISUALIZATION OF THE WEEK

Twitter, Janelle Shane from January 21, 2022

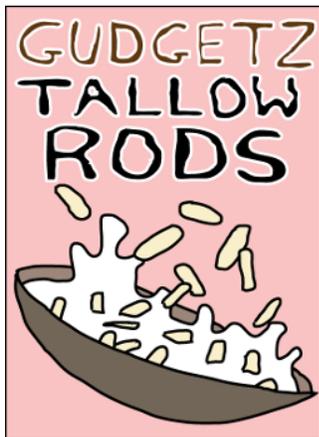
"I used GPT-3 to generate new breakfast cereals and descriptions. The cereals from Ada, the smallest GPT-3 model, were a bit questionable."



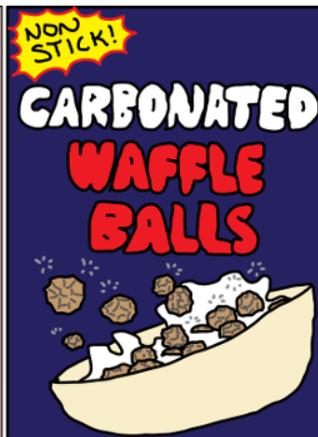
Bright yellow biscuit made with only high quality, raw ingredients.



Chewy and colorful fun fun fun foam squares deliver the fun flavor of muffins and covered train cars!!!



A science developed recipe.



Gold standard for fun waffle ball innovation using wheat-based flavors and crispy, aluminum foil.

Deadlines

Conferences

[Consider joining our CHI workshop \(to be held virtually :\) on Investigating Data Work Across Domains: New Perspectives on the Work of Creating Data.](#)

"WS website w call for participation and more info is here:

<https://web.asu.edu/dataworkworkshopchi2022/home> @najaholten @MilagrosMiceli @nicksu28 @ginasue" Deadline for submissions is February 21.

RFPs

[Call for Proposals: New Program Focuses on Methodologies, Tools, Infrastructure for Internet Measurements](#)

"The Internet Measurement Research (IMR) program, headed by the **National Science Foundation's Directorate for Computer and Information Science and Engineering** in partnership with the **Directorate of Mathematical and Physical Sciences**, is a new focused program to support methodologies, tools, and research infrastructure for internet measurements." Initial proposal deadline is February 15.

Tools & Resources

[The State of Web Scraping 2022](#)

ScrapeOps from January 12, 2022

"After consistent growth in interest every year over the last 10 years, 2021 seems to have been a year when "web scraping" became uncool. Google searches for the term "web scraping" have dropped 30-40% compared to 2020 volumes."

[Working with JSON data in BigQuery](#)

Medium, Google Cloud - Community, Lak Lakshmanan from January 24, 2021

"JSON support in BigQuery is not new. For example, BigQuery has had JSON querying and JSON generation functions for a long time. ... What's new is that JSON is a data type (plus a set of functions to work with that data type) — think of how Geography is a type, and how that opened up the entire universe of ST_* functions instead of us having to muck around with latitude and longitude calculations ourselves."

Featured Jobs

See the [ADSA Jobs Page](#) for more opportunities.

About Us: The Data Science Community Newsletter was founded in 2015 in the Moore-Sloan Data Science Environment at NYU's Center for Data Science. We continue to be supported by the Gordon and Betty Moore Foundation and the Alfred P. Sloan Foundation through the [Academic Data Science Alliance](#). Our archive of newsletters is at <https://academicdatascience.org/resources/newsletter>. Our mailing address is [1037 NE 65th St #316; Seattle, WA 98115](#).