

ADSA Data Science Community Newsletter

Data Science Community Newsletter features journalism, research papers and tools/software for October 21, 2021.

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](#).

ADSA ANNOUNCEMENTS

[ADSA Leadership Summit & Annual Meeting Virtual Series Kick-off](#)

Title: Data Science for Social Impact in University-based Programs

Date: November 10 at 12:30 p.m. Pacific (3:30 p.m. Eastern).

Description: University-based initiatives seeking to make a positive social impact with data-intensive technologies require partnerships across a range of stakeholders. This session will have short "roadmap" presentations on two current programs at the University of Washington and the University of Pennsylvania, followed by a panel session featuring leaders of additional programs.

[REGISTER!](#)

FREEDOM OF EXPRESSION VS. DEI

It was inevitable and has now happened on the front stage of academia: ideology about inclusivity crashed directly into ideology about freedom of expression. Professor **Dorian Abbot, University of Chicago**, was invited to give a prestigious, off-site public lecture about his research on climate change models at the invitation of **MIT's Department of Earth, Atmospheric and Planetary Sciences** (EAPS/Course 12). MIT students objected to giving Professor Abbot a public stage because of his [co-authored op-ed](#) objecting to the pervasive "lens of DEI" in academic admissions and hiring: "The underlying premise of DEI is that any statistical difference between group representation on campus and national averages reflects systemic injustice and discrimination by the university itself...[and] compromises the university's mission" to pursue academic and scientific excellence. Abbot had come under scrutiny for similar reasons in 2020 when students at U of C demanded that he be removed from a Committee position. They didn't get their wish, but Abbot [did respond](#) by defending his right to voice a dissenting opinion about the impacts and scoping of DEI initiatives without being "assumed to have racist or otherwise bigoted motives."

This time around, the lecture Abbot was invited to give at MIT was demoted from a fancy public event to an invitation to give an on-campus lecture. **Princeton** got in on the action by inviting Professor Abbot to give an in-person lecture on the same day as the original MIT invitation, and

professor **David Romps** tried to invite Professor Abbot to his center at the **University of California, Berkeley**, but met strong opposition from other center faculty. Romps ended up resigning his directorship of the **Berkeley Atmospheric Sciences Center** [over Twitter](#) because "it was never my intention to lead an organization that is political even ambiguously so." That's fairly dramatic collateral damage.

In terms of the ideological collision here — DEI versus freedom of expression — there is a strong argument to be made that the scope, outcomes, and processes of DEI initiatives should be open for debate. There are plenty of social science studies and theoretical work available to foster such conversations. Having the same conversations about DEI over and over can be exhausting. The type of objection to DEI raised by Professor Abbot is not new and in fact a very recent [court ruling](#) over **University of North Carolina's** use of affirmative action upheld the school's practices due to "substantial evidence demonstrating its good faith in pursuing the educational benefits that flow from diversity."

ERODING TENURE

Georgia's Board of Regents has taken a step toward eroding the protection tenure affords to academics' freedom of expression. As [reported](#) in the *The New York Times*, "Georgia's public university system will now let its colleges' administrations remove a tenured professor with little to no faculty input."

THE GREAT RESIGNATION HITS YALE

Thanks in part to a cushy severance package offered to faculty over 70 years old, **Yale University** saw [3x more retirements](#) last year than in the previous year. Have you seen headlines like this elsewhere? If so, please alert us.

SPONSORED CONTENT

*Data for Good,
Ethically*

November 2-3 | tdai.osu.edu

THE OHIO STATE UNIVERSITY
TRANSLATIONAL DATA ANALYTICS INSTITUTE

RESPONSIBLE DATA SCIENCE
AI & HEALTHCARE
ENVIRONMENT & SUSTAINABILITY
ART & DATA
SMART MOBILITY
FOUNDATIONS OF AI & ML

UNIVERSITY VS. BOOTCAMP FOR CS TRAINING

Bootcamps for computer science training may be starting to pull students away from traditional

academic degrees. In terms of job placement rate, there is [some evidence](#) that top-rated coding bootcamps have comparable job placement to four-year Computer Science degrees. More research is required to investigate salary, benefits, and job permanence. But this news comes at a time when Midwestern colleges and universities are already seeing the start of a predicted [downturn in enrollment](#) due largely to lower birth rates and COVID. Bootcamps are more affordable, so it is reasonable to expect that if they continue to demonstrate high job placement rates, they will pull cost-conscious students away from traditional 4-year degrees and the associated student loan debt.

REFORESTATION OVERSTATED AS A CARBON OFFSET

Reality check: [new analysis](#) published in *Nature* suggests that using reforestation as a carbon offset tactic has a number of difficulties that haven't been properly account for, partly because that type of accounting is hard and partly because it leads to unpopular revelations. "Our results show major errors in carbon accounting in reforestation projects if they (1) ignore the time needed for trees to reach their carbon capture potential; (2) ignore the GHG emissions involved in setting up a plot; (3) report the carbon capture potential per tree planted, thereby ignoring limitations at the forest ecosystem level; or (4) under-estimate tree losses due to inevitable human and climatic disturbances."

SPORTS EQUINOX!

The sports equinox is [upon us](#). October 21 will see all four major professional American sports in action: football, baseball, basketball, hockey. Alas, no major women's sports are included this equinox but that should improve in the future with [ongoing research](#) into the issue and events like **ESPN's [Women + Sport Summit](#)**. In fact, this Halloween, October 31, will be another sports equinox and it will include **NWSL** women's professional soccer.

Luke Bornn, a former statistics professor at **Harvard** and **Simon Fraser** universities, is a reluctant entrepreneur, according to the recent [online profile](#) at *SportTechie*. Bornn came across optical player tracking data from NBA basketball games and thought it was "the richest space-time data I'd ever seen." He is currently a technical co-founder of **Zelus Analytics** (with our former **NYU Center for Data Science** colleague **Dan Cervone**). Zelus, in addition to being a second-tier Greek god, is a sports data advisory and investment consulting firm with projects in soccer, basketball, baseball and football. Other current and former academics who lead technology development in this industry: **Rajiv Maheswaran** of **Second Spectrum** and **University of Southern California** ([@rajivmaheswaran](#)), **Patrick Lucey** of **Stats Perform** and **Carnegie Mellon** ([recent interview](#)), **Mehrsan Javan** of **Sportlogiq** and **McGill University** ([Google Scholar](#)).

This season the **National Hockey League** has [inserted a tracking chip](#) in the hockey puck. The league is looking to turn the data into fan interaction through gambling — "How far does the puck travel in a game? Who skates the fastest in the second period? Who has the hardest shot?" This is at a time when the **State of Connecticut** has just [legalized online gambling](#) and **New Jersey** has now [earned its first \\$1 billion](#) in legal gambling revenue. **Iowa** [hailed in](#) \$210 million during this past September alone. **ESPN** and **FOX Sports** are [meeting the public interest](#) in gambling head-on with more programming. [Some see](#) sports gambling operations taking a Wall Street-like approach to wagering. The downside risk, however, falls disproportionately on universities where

dismal earning opportunities for college athletes make them targets for illegal match fixing. Journalist **Will Leitch** [calls it](#) a "disaster waiting to happen."

INKJECT PRINTERS + ML = AI CRAFTWORK

A team from **Google Research** and **California Institute of Technology** have been [using inkjet printers](#) loaded up with various metal oxides to lay down many permutations of oxide combinations. There are roughly 63,000 possible combinations that might have promising properties in the context of renewable energy storage. So, after the inkjet printed hundreds of oxide tracks rated for desirability by human experts, an ML model was trained to predict which would have the best combination of properties for use in renewables. That [model is available](#) to other teams who want to contribute to the materials science hunting mission.

400 RESEARCHERS, 17 PAPERS: MAJOR ADVANCE FOR BRAIN ATLAS

In 17 papers written by 400 researchers, "millions of human, marmoset, and mouse cells extracted from a brain region involved in coordinating movement" have been analyzed and mapped. This kind of [disciplinary coordination](#), historically more common in physics, is an incredibly powerful strategy for moving entire fields forward by leaps at a time.

NEW KING OF NATURAL LANGUAGE MODELS

NVIDIA and **Microsoft** joined forces to develop the ["Megatron-Turing Natural Language Generation model"](#) which has 105 layers and 530 billion parameters. It is trained to deliver "completion prediction, reading comprehension, commonsense reasoning, natural language inferences and word sense disambiguation." GPT-3 has met a formidable match.

SPONSORED CONTENT



IUPUI
SCHOOL OF INFORMATICS AND COMPUTING
INDIANA UNIVERSITY

**Seeking tenure-track faculty:
Assistant/Associate
Professor of Data Science**

Department of Human-Centered Computing
Apply by January 3, 2022

FOLLOW THE MONEY

\$50,000,000 Anonymous -> **University of California, Berkeley** and **University of California, San Francisco** to create [a joint program](#) in Computational Precision Health.

\$19,000,000 **National Science Foundation** -> **University of Idaho** for the [Deep Soil Ecotron](#) project that will study soil 1-2 meters below the surface.

\$14,000,000 **Daniel J. Epstein** -> **University of Southern California, Viterbi School of Engineering** to “[grow recruitment](#) of expert faculty members...with a focus on interdisciplinary areas such as computational systems, machine learning and artificial intelligence.”

\$11,700,000 **National Institutes of Health** -> **University of California, San Diego, School of Medicine** to study [precision medicine and genomics](#) among people who are not predominantly of European descent.

\$5,000,000 **National Science Foundation** -> **Georgia State University** to establish the “[Center for Dynamic Multiscale and Multimodal Brain Mapping Over the Lifespan](#), which will focus on brain development, structure and connectivity from childhood onward”

\$2,400,000 **University of Tennessee at Chattanooga** -> **UTC College of Engineering and Computer Science, Innovation-Advanced Manufacturing Center (I-AMAC)** for a new [Engineering Barn](#)

\$2,000,000 **The Bob & Renee Parsons Foundation** -> **Arizona State University Foundation** to [expand](#) ASU's Math, Computer Science and Statistics Accelerator-style education to middle school students

\$1,200,000 **National Science Foundation** -> **University of Illinois at Urbana-Champaign** for developing techniques to [accurately label image data](#) with less human input.

\$1,000,000 **AAAI Squirrel AI Award** -> **Cynthia Rudin** for [her work](#) “advocating for and developing ‘interpretable’ machine learning”.

\$1,000,000 **National Science Foundation** -> **Virginia Tech** to use “physics-guided machine learning” in an attempt to [predict the mechanics](#) of living cells.

\$1,000,000 **National Science Foundation** -> **University of Hawai'i** for [Cyberinfrastructure Training to Advance Climate Science](#)

\$358,578 **National Science Foundation** -> **Aaron Naeger** at **University of Alabama in Huntsville** and **Guanyu Huang** at **Spelman College** for [new machine learning models](#) that monitor air quality and integrate new satellite data with existing terrestrial data sources

negative \$5,000,000 **University of North Carolina at Chapel Hill** <- **UNC Chapel Hill Libraries collections budget** as part of a [7.5% operating budget shortfall](#) exacerbated by COVID.

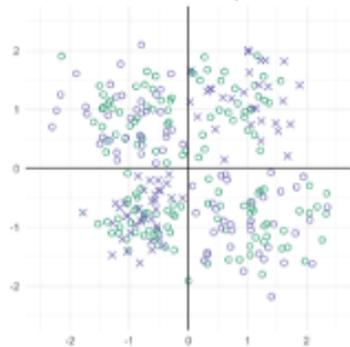
NEW PROGRAMS

[Data Science Major](#) at **Albion College** ... also, *Detroit Free Press* education reporter **David Jesse** wrote [this recent feature article](#) on Albion's efforts to enroll students post-pandemic.

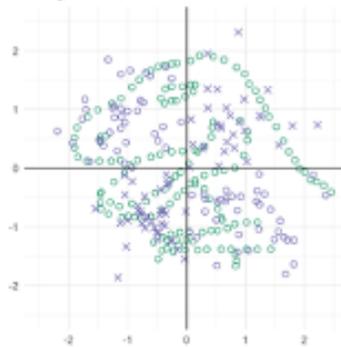
DATA VISUALIZATION OF THE WEEK

Twitter, Steve Franconeri, from October 17, 2021

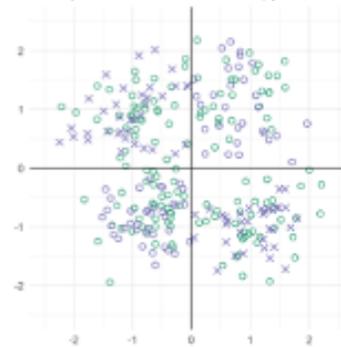
Jurassic Mark: Intentional Blindness for a Datasaurus Reveals that Visualizations are Explored, not Seen



Boger, Most, & Franconeri (2021), IEEE VIS



Info: <https://datasaurus-vis.herokuapp.com/>



"Data visualizations are not 'seen' in a glance – they require active exploration, across a series of visual filtering operations" ... "93% of people to miss a *dinosaur* within the ignored data values"

Deadlines

Studies/Surveys

[Diffusion MRI Data for Best Practices in Image Preprocessing](#)

"Your contributions will allow the Diffusion Study Group to evaluate how preprocessing pipelines and tools affect the reproducibility of diffusion MR analyses." Sponsor: **National Institutes of Health, Center for Advanced Imaging Innovation and Research.**

Education Opportunities

[Applications Now Open for the MGB-SIAM Early-Career Fellowship](#)

"The MSEC Fellowship reflects a joint commitment by **Mathematically Gifted & Black (MGB)** and **SIAM** to promote long-term engagement of MSEC Fellows within SIAM and continued success within the wider applied mathematics and computational sciences community. Apply now through November 15, 2021."

RFPs

[National Science Foundation accepting proposals for new set of National Artificial Intelligence Research Institutes](#)

"The **National Science Foundation (NSF)** is currently accepting proposals for their National Artificial Intelligence (AI) Research Institutes program." Deadline for proposals is January 14, 2022.

Tools & Resources

[General-Purpose Question-Answering with Macaw](#)

Medium, Ai2 blog, Peterc from October 15, 2021

"While there are a few other pretrained QA systems available, none has quite matched GPT-3's few-shot QA performance — until now. **AI2** has just released Macaw (multi-angle question-answering), a versatile, generative question-answering (QA) system that exhibits strong zero-shot performance on a wide range of question types."

Events

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

[The MaD seminar features leading specialists at the interface of Applied Mathematics, Statistics and Machine Learning.](#)

New York, NY October 21, starting at 2 p.m. Eastern, **NYU Center for Data Science**, 60 Fifth Ave.
Speaker: **Tim Roughgarden** from **Columbia University**."

[American Statistical Association Webinar, Funding Opportunities at NSF and NIH: An Introduction](#)

Online October 22, starting at 1 p.m. Eastern. "Statisticians, learn about funding opportunities at NIH & NSF at this free 10/22 webinar." Featuring **Judy Wang** from **National Science Foundation** and **Karen Messer** from **University of California-San Diego**.

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.