Data Science Community Newsletter

6 March 2020   Data Science Community Newsletter Issue 192

Please let us (Laura Norén, Brad Stenger) know if you have something to add to the newsletter. We are grateful for financial support from the Academic Data Science Alliance.

The coronavirus is the news this week. **UW-Seattle** is the **first university in the country to suspend all in-person activities on campus**, including classes. Click **here** if you want to contribute unused compute resources to protein folding computations that will hopefully lead to drug discovery. The March meeting of **American Physical Society** was canceled in Denver due to coronavirus fears and in a bizarre game of epidemiological chicken, **Austin** finally pulled the plug on **SXSW**. The American application of laissez faire to public health will be interesting to watch, kind of like a horror movie.

As universities roll into spring break season, please **contribute to this crowd-sourced data collection effort** to track what universities are advising faculty, staff, and students to do. Are students, faculty, and staff banned from international travel at your university like they are at NYU? What about travel to Seattle? NYU hasn't said anything about Seattle, but it appears to be the hardest hit US area. What happens to students, faculty, and staff who choose to travel into hot zones on their upcoming spring breaks? Will they be expected to quarantine themselves for 14 days? Has any school offered incentives for self-quarantining? In my experience, getting people to commit to quarantine is hard.

As social scientists like **Erik Klinenberg** have **pointed out**, natural disasters rip the window dressing off society to reveal the rot of underlying social inequalities. America's lack of universal health care and growing economic inequality are reflected in academia as well, where many classes are taught by contingent adjuncts and precarious graduate students who may feel they will be punished if they actually practice self-quarantine.

Be smart, stay healthy.

-Laura and Brad

DATA SCIENCE NEWS
University Data Science News

Yoshua Bengio is advocating the adoption of Slow Science, or at least the **adoption of slower more thoughtful publication cycles**. He's concerned that the flurry of activity leading up to conference paper deadlines results in "at least 50% more papers...that may contain errors, lack in rigour or simply be rather incremental." I'm with Bengio. The current situation leads to a mind-numbing avalanche of articles and less, ahem, deep learning. ;)

The **University of California System** has had a week of highs and lows. **Fifty-four graduate students** at **UC-Santa Cruz** have been *terminated* because they continued to strike after UC

Chancellor **Janet Napolitano** told them she would can them if they continued striking. She was good to her word, in a bad way. Graduate students at UC-Santa Cruz continue to strike and have been joined by graduate students at **UC-Santa Barbara** who are now **striking for a cost of living adjustment increase** in their pay. Graduate students at **UC-Davis**, **UCLA**, and the **Department of African American Studies** at **Berkeley** are protesting. Update: **Berkeley students staged a walkout strike** Thursday.

Academics who wish to show support for the striking graduate students can **sign a petition** pledging solidarity.

In more positive news — though the seesawing here is jarring enough to metaphorically knock some teeth loose — **UC-Berkeley** has announced that an anonymous donor is putting up $252 million to build a physical home for the new **Division of Computing, Data Science, and Society**. They're calling the building the Data Hub. The new building will require another $300 million for which the school is currently fundraising. This is the **"single largest gift in Berkeley's history"** and marks a home run for new Associate Provost **Jennifer Chayes**.

And just to add color on life in California, the **City of Berkeley** had its first coronavirus case, while Santa Clara County — which contains much of Silicon Valley including Palo Alto, Cupertino, and San Jose — **had 11 cases** but then ominously stopped putting out press releases.

Anshumali Shrivastava led a group of computer scientists at **Rice University** to **speed up deep learning technology without resorting to GPUs**. The team used an algorithm called the "sub-linear deep learning engine (SLIDE)" in place of the typical matrix multiplication found in back propagation. SLIDE relies on hash tables. Teammate **Tharun Medini** explains, "you don't need to train all the neurons on every case. We thought, 'If we only want to pick the neurons that are relevant, then it's a search problem.' So, algorithmically, the idea was to use locality-sensitive hashing to get away from matrix multiplication." The team is now working with researchers at **Intel** who have helped them cut time by half.

Biologists have a few options for sharing their giant datasets (and many options for sharing smaller data sets). As a big fan of data infrastructure, it's with pleasure that I point biology researchers to Electron Microscopy Public Image Archive (**EMPIAR**) and the Image Data Resource (**IDR) which are going to be merged into a new BioImage Archive** hosted in the UK.

In another data commons story, **Northwestern University** has announced a **data commons partnership** with the **Food Allergy Research & Education** (FARE) organization.

The **University of North Texas** will offer a **masters degree in artificial intelligence**.

University of Toledo in Ohio will offer a new **bachelor of science degree in data science**.

Two researchers at the **Cleveland Clinic** found that **Apple** watches **aren't accurate enough to be**

used to detect atrial fibrillation, one of the most common heart problems. These results have not stopped Apple from partnering with **Johnson and Johnson** to see if they can use their watches **to predict risk of stroke** (and get more seniors using their watches?). I wish them luck, but it looks like we are years away from the kind of accuracy we expect with health care devices. By then, a new cohort of people will be old...and already accustomed to wearing Apple watches. [Full disclosure: I do not wear a watch of any kind.]

Google announced its new cohort of **150 faculty fellows** which include **Cristina Savin** and **Kyunghyun Cho**, two **NYU** computational neuroscientists.

Michael Evans, **Michael Reed** and additional researchers from **Tufts University** used 30 years of satellite data to prove that **federally protected lands do, indeed, protect habitat and support species diversity** compared to private non-protected lands in the U.S.

Company Data Science News

Two new real-time facial recognition technologies are being sold to law enforcement officers. **Wolfcom** is a paintball retailer and personal spying corporation **turned live facial recognition purveyor**. Their technology is an Android app that will return predicted age, gender, and facial expression while returning more granular results if the person is a match for outstanding warrants or a missing person case. I will tell you this, readers; they always get you with the missing persons argument. 'How will we find the lost children if we don't subject everyone to persistent public surveillance?!' Be wary of common security show ponies like that. They are often trotted out to distract us from civil rights infringement. We found this story in OneZero, a new all digital critical technology publication on Medium that's worth putting on your radar.

Wolfcom looks amateurish - though still objectionable - compared to **Banjo**, a company also **selling live artificial intelligence data**. The **State of Utah** has authorized them to run their algorithm against all data coming off government-owned surveillance cameras and traffic cameras in the state. They enrich their data with data pulled from "social media, satellites, and other apps" which is difficult, but incredibly powerful. It's not clear what the "other apps" are or how they are getting access to satellite data. The technology is also installed across the **University of Utah** campus and is likely privy to the university's student, faculty and staff directory. It is about 6 months past the time when privacy mavens should have taken all their photos off social media, but it's still a reasonable idea.

Another form of surveillance (remember: survey and surveillance have the same root) that is increasingly celebrated in the data science world is the use of personality testing in hiring processes. This seems as good a time as any to refer to existing research that suggests that despite the widespread use of big 4 personality testing, that type of personality testing shows **no correlation with job placement success**. **Whitney Martin** cites data from **Frank L. Schmidt's** 2013 address to the Personnel Testing Council Metropolitan Washington, "if your hiring process relies primarily on interviews, reference checks, and personality tests, you are choosing to use a process that is significantly less effective than it could be if more effective measures were incorporated." The most effective measures are cognitive ability tests, integrity tests, and multi-measure tests that combine

cognitive ability tests with personality and interest assessment. **Cathy O'Neil** has an **opinion piece explaining the legal risk** of using any type of personality test that may be correlated with mental illness as using such a test would be a violation of the 1990 **Americans with Disabilities Act**.

Google's Gmail program defaults users into a "tabbed" experience in which many emails are **filtered out of the main inbox** that is automatically displayed into tabs for "social" and "promotional" email. Both **Elizabeth Warren** and **Joe Biden** saw 0% of their campaign mail in the main inbox, 85% went into Promotions tab, and 15% ended up in the dreaded spam basket (which isn't even a tab). Hard to claim Gmail swayed the primary vote.

Google has released a **free technical writing course**.

Google also **announced** it will hire 3500 workers across three new or expanded sites in Canada in **Toronto**, **Kitchener-Waterloo**, and **Montreal**.

The good people at **Mozilla** put together a **student's guide to being an ethical tech worker**. It's lovely. It's maybe not as practical as most of you would prefer. My two cents on how to be an ethical tech worker: don't waste time interviewing with companies that have a bad ethical track record. It's incredibly difficult to change an entrenched culture. Start at a non-toxic company and grow from there.

Government Data Science News

In a case of GDPR vs. facial recognition, **GDPR** carried the day in **France**. **The case involved** two French high schools in Nice and Marseille that had installed cameras at the gate to the high school which would unlock the gates only for registered students. The students and/or their legal guardians signed forms indicating their consent to this arrangement. The courts however, found that the GDPR requires "less intrusive" mechanisms to be used where possible and that the schools could have used badges. On the issue of consent as a legal basis for processing: "the receipt of consent by simple signature in a form, from the student or his legal guardian, is not sufficient to meet the conditions set out in the GDPR. Indeed, the fact that the students are placed under the direct authority of the heads of their schools does not offer sufficient guaranties that the consent was given in a free, specific, informed and univocal way." This is in contrast to the increasing amount of public, state-sanctioned surveillance and facial recognition happening in the US, largely fueled by private companies 'disrupting' criminal justice.

The **U.S. Securities and Exchange Commission** plans to **add "depth of book" data** to its publicly available Securities Information Processor (SIP) feed, the real-time consolidated data feed of the U.S. financial market, which could matter for algorithmic traders and others in the finance community.

The **Federal Communications Commission** (FCC) has proposed fines for wireless carriers that **sold customers' location data**:
+ **T-Mobile** $91 million
+ **AT&T** $57 million

+ **Verizon** $48 million
+ **Sprint** $12 million.
Some, including **Oregon Senator Ron Wyden** (D) are frustrated that the fines are "comically small."

Pope Francis has put out a **statement on AI ethics**. He wants AI to be explainable and to avoid objectionable bias. He didn't directly say anything about facial recognition, but he named dropped it. Honestly, the statement is short and it's doesn't bring anything new to the data ethics conversation. **Microsoft** and **IBM** have signed off; maybe there's some value to having a Pope-approved AI Ethics document for corporate PR reasons?

The **U.S. Census** has faced one challenge after the next. Because the **Trump** administration tried to add a question about citizenship - which was blocked by federal judges - **"nearly 70 percent of adults think the 2020 Census questionnaire will ask which people in their households are citizens."** Spinning that out, it turns out that "32% of adults are extremely or very concerned about how their answers to the 2020 Census questionnaire will be used and with whom they will be shared." If we don't get an accurate count, funding allocations to the neediest tend to shrink. Black and Native American populations are the **most likely to be under-counted, leaving already disadvantaged communities without funds they might otherwise have gotten**. [This is what structural racism looks like.] Poor Census data also presents significant difficulty for data scientists who often use it as ground-truth data for estimating sampling error.

The **Partnership on AI** held a summit on **the state of AI for Good** now that we're ten years in. The event was back in September, but the recap was published recently. Big takeaways:
+ There's a lot of needless duplication in the AI for Good space. Different cities and labs need better ways to learn from each other, share data (where appropriate), and consolidate best practices. It might also help if there were more sustainable funding for this kind of work.
+ The community also addressed the increased importance of data ethics and social impact assessment when working with the public, especially when vulnerable populations are involved.
+ Usability came up, too, which makes sense. Lots of AI for Good is working on shoestring budgets or project-based engagements that don't get all the way through the fit and polish process a for-profit entity might employ.
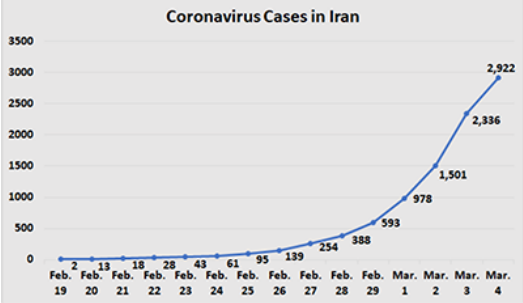
Extra Extra

In my (delicious) experience, I have found physicists to be especially skilled in the kitchen. A new study by physicists **Hungtang Ko** and **David L. Hu** of **Georgia Tech** explains **the kinetic pro-tips for perfect wok-tossed fried rice**.

In case you care, there's a **new world ranking of CS departments** out in **ACM Communications Careers** by education consultants, **QS Quacquarelli Symonds Limited**.

If you're looking for a teachable case study for an intro stats or data science course this summary of the way **the coronavirus caseload is being quantified** is just about perfect.

Data Visualization of the Week

United States Institute of Peace, The Iran Primer from March 05, 2020

Coronavirus Cases in Iran

Tweet of the Week

Twitter, Seattle Coronavirus Updates from March 04, 2020

Seattle Coronavirus Updates
@Seattle2019nCov

HAPPY STORY ALERT: @UW @UWVirology has developed it's own #COVID19 test & expects to be testing 1k–1.5k people per day by the end of the week.
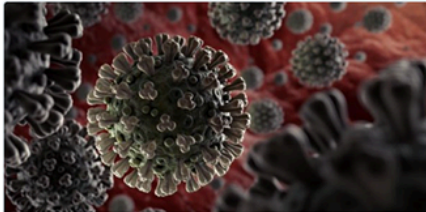
Preparation began @ end of 2019, *just in case* it reached U.S.

That decision will save lives.

Go Dawgs.

q13fox.com/2020/03/04/uw-...

UW Medicine develops its own COVID-19 test
q13fox.com

4:19pm · 4 Mar 2020 · Twitter Web App

EVENTS

Excited to announce our [Academic Data Science Alliance] first Diversity, Equity, and Inclusion SIG call!

Online March 19, starting at 10 a.m. Pacific time. "We'll hear from [**Laura Noren**] on her recent research brief from **Obsidian Security** 'Who's Building Your AI?'" [registration required]

Academic Data Science Alliance, Data Science Education Special Interest Group Call

Online April 2, starting at 10 a.m. Pacific. "Is your institution considering starting a data science degree program? Come hear three ADSA community members discuss how their organization decided to create data science degree programs and how these programs took the shape they did. We'll be hearing from **Sarah Stone** (**University of Washington**), **Ajay Anand** (**University of Rochester**), and **HV Jagadish** (**University of Michigan**)." [registration required]

The Quant Conference

New York, NY April 3. "TQC is a forum that engages the brightest young minds and foremost thought leaders from the industry and academia to dive into the latest innovations in quant finance, foster collaboration and facilitate opportunities." [$$$]

Save the Date! Ethics of Big Data Symposium

Milwaukee, WI May 7 at **Marquette University**. "This year we are bringing together a cross-section of practitioners, researchers, and policymakers to discuss the broad complexities of engaging with "ethics" within data-driven organizations, ranging from big data analytics firms, social media platforms, and developers of artificial intelligence." [save the date]

Pfizer/ASA/Columbia University Symposium on Risks and Opportunities of AI in Clinical Drug Development

New York, NY May 18, starting at 8:30 a.m., **Columbia University**. "This forum is, therefore, expected to serve as a platform for distinguished statisticians, data scientists, regulators, and other professionals to address the challenges and opportunities of AI in pharmaceutical medicine; to foster collaboration among industry, academia, regulatory agencies, and professional associations; and to propose recommendations with policy implications for proper implementation of AI in promoting public health." [$$$]

registration open for NIMBLE short course, June 3-4, 2020 at UC Berkeley

Berkeley, CA June 3-4. "Registration is now open for a two-day training workshop on NIMBLE, June 3-4, 2020 in Berkeley, California. NIMBLE is a system for building and sharing analysis methods for statistical models, especially for hierarchical models and computationally-intensive methods (such as MCMC and SMC)." [$$$]

Embodied AI Workshop at CVPR 2020

Seattle, WA June 14-15. "There is an emerging paradigm shift from 'Internet AI' towards 'Embodied AI' in the computer vision, NLP, and broader AI communities. In contrast to Internet AI's focus on learning from datasets of images, videos, and text curated from the internet, embodied AI enables learning through interaction with the surrounding environment." [registration required]

2020 Causal Inference Summer Institute

Philadelphia, PA June 22-25 at **University of Pennsylvania**. "The **Center for Causal Inference** is proud to announce its fourth annual Causal Inference Summer Institute, a four-day, intensive learning experience that will take place at the Jordan Medical Education Center (3400 Civic Center Blvd, **Perelman School of Medicine**). [$$$]

Next Generation Genomics

New York, NY August 5-6 at **New York University**. "In this conference, we aim to showcase the 'next-generation' of genomic methods and technologies, and in particular to provide a platform for junior faculty to present recent discoveries and innovations." [$$$]

Join us for Data for Black Lives III

Cambridge, MA December 11-13 at **MIT Media Lab**. "Returning to Cambridge will make it possible for us to accomplish all of our goals: expand our programs, continue to build our network, and host the best conference yet." [save the date]

DEADLINES
Contests/Award

ASA Data Visualization Poster Competition for Grades K–12

"The **ASA/NCTM Joint Committee on Curriculum in Statistics and Probability** and the **ASA's education department** encourage students and their advisers to participate in its annual Data Visualization Poster Competition." Deadline for poster submissions is April 1.

Microsoft Research Latin America PhD Award

"**Microsoft** actively seeks to foster greater levels of diversity in our workforce and in our pipeline of future researchers. We are always looking for the best and brightest talent and celebrate individuality. We invite candidates to come as they are and do what they love." Deadline for submissions is April 15.

DCASE2020 Challenge - IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events

"Sounds carry a large amount of information about our everyday environment and physical events that take place in it. We can perceive the sound scene we are within (busy street, office, etc.), and recognize individual sound sources (car passing by, footsteps, etc.). Developing signal processing methods to automatically extract this information has huge potential in several applications." Deadline for challenge submissions is June 15.

Conferences

DEEM Workshop at SIGMOD2020

Portland, OR June 14, in conjunction with **SIGMOD 2020**. "DEEM brings together researchers and practitioners at the intersection of applied machine learning, data management and systems research, with the goal to discuss the arising data management issues in ML application scenarios." Deadline for submissions is March 8.

Call for Papers and Abstracts at SocInfo 2020!

Pisa, Italy October 6-9. "The **International Conference on Social Informatics** (SocInfo2020) is an interdisciplinary venue that brings together researchers from the computational and social sciences to help fill the gap between the two communities." Deadline for submissions is April 15.

Education Opportunities

Barclays Accelerator, powered by Techstars - New York

"The **Barclays Accelerator**, powered by **Techstars**, is a top-tier fintech accelerator for early stage startups. Every year we invest in 10 fintech companies and help them get their businesses to the next level." Deadline for applications is May 10.

RFP

Announcing the 2020 Frictionless Data Tool Fund

"**Open Knowledge Foundation** is launching the second round of the Frictionless Data Tool Fund, a mini-grant scheme offering grants of $5,000 to support individuals or organisations in developing an open tool for reproducible science or research built using the Frictionless Data specifications and software. We welcome submissions of interest until 17th May 2020."

TOOLS & RESOURCES
p5.js 1.0 is Here!

Medium, Processing Foundation, Lauren McCarthy from February 29, 2020

"We are excited to announce the 1.0 Release of p5.js! p5.js is a JavaScript library that aims to make creative expression and coding on the web accessible and inclusive for artists, designers, educators, and beginners. While it's been nearly seven years since p5.js began, we intentionally embarked on the path to reach 1.0 a year ago when **Kate Hollenbach** worked on a first version of a roadmap for this."

How We Improved Data Discovery for Data Scientists at Spotify

Spotify Labs, Andrew Maher from February 27, 2020

"At **Spotify**, we believe strongly in data-informed decision making. Whether we're considering a big shift in our product strategy or we're making a relatively quick decision about which track to add to one of our editorially-programmed playlists, data provides a foundation for sound decision making. An insight is a conclusion drawn from data that can help influence decisions and drive change. To enable Spotifiers to make faster, smarter decisions, we've developed a suite of internal products to accelerate the production and consumption of insights. One of these products is Lexikon, a library of data and insights that help employees find and understand the data and knowledge generated by members of our insights community."

"We've learned a lot since we first launched this product. In this blog post, we want to share the story of how we iterated on Lexikon to better support data discovery."

PDF text extraction

FilingDB from March 02, 2020

"As part of building FilingDB, we've extracted text data from tens of thousands of PDF documents. In the process, we have seen how every single assumption we had about how PDF files are structured was proven incorrect. Our mission was particularly difficult as we had to process PDF documents coming from a variety of sources, with wildly different styling, typesetting and presentation choices."

New Open Source App: Data Science Education

datanami, George Leopold from March 04, 2020

"An open source project shepherded by the **Linux Foundation** aims to accelerate data science curricula while benefitting from the contributions of students and teachers. OpenDS4All is funded by **IBM** and is being developed by the **University of Pennsylvania**. The effort would give educators free access to information needed to develop data science coursework. In return, successful approaches would be folded back into what project promoters call 'constantly evolving and improving' curricula."

"A starter 'curriculum kit' includes a set of open source building blocks that could be used to launch data science programs. Based on the Python programming language, the tools and frameworks include code, documentation and data sets, organizers said."

Yoshua Bengio and Team Introduce GNN Benchmarking Framework

Synced from March 04, 2020

"A new study introduces a reproducible graph neural network (GNN) benchmarking framework to study and quantify the impact of theoretical developments for GNNs. In the field of analyzing and learning from data on graphs, GNNs have become an essential tool. With promising applications in different domains such as chemistry, physics, social sciences, knowledge graphs, recommendation, and neuroscience, how to study and build more powerful GNNs is a hot topic."

Sno — Distributed version-control for geospatial and tabular data

koordinates from March 04, 2020

"Sno stores geospatial and tabular data in **Git**, providing version control at the row and cell level."

CAREERS
Tenured and tenure track faculty positions

Executive Director, Computational Social Systems
University of Pennsylvania, The Wharton School; Philadelphia, PA

Assistant Professor in Interaction Design
KTH Royal Institute of Technology, Media Technology & Interaction Design Division; Stockholm, Sweden

Computing & Data Sciences - All Ranks
Boston University, Faculty of Computing & Data Sciences; Boston, MA

Director of the Ethical Tech Law and Policy Clinic
Duke University, School of Law; Durham, NC

Full-time, non-tenured academic positions

Senior Data Analyst
University of Minnesota, IPUMS; Minneapolis, MN

Data Analyst
University of Minnesota, IPUMS; Minneapolis, MN

Assistant Professor, Library Chief Data Strategist
University of Rhode Island, University Libraries; Kingston, RI

Sea Grant Marine Extension Associate (Climate Community Resilience)
University of Maine Cooperative Extension, Maine Sea Grant; Orono, ME

Sr. Research Data Scientist
University of California-Davis, DataLab: Data Science & Informatics; Davis, CA

Application for Data Scientist - Schmidt DataX Project
Princeton University, Center for Information Technology Policy; Princeton, NJ

Research Assistant in the Computational Approaches to Modeling Language (CAMeL) Lab - Division of Science
New York University, NYU Abu Dhabi; Abu Dhabi, United Arab Emirates

Computing Sciences Researcher I
University of Arizona, Agricultural Data Science group; Tucson, AZ

R&D Software Engineer I
University of Arizona, Agricultural Data Science group; Tucson, AZ

Postdocs

Postdoctoral Appointee - Machine Learning for Multiphase Flow Modeling
Argonne National Laboratory; Lemont, IL

Postdoctoral Associate
Stony Brook University, Department of Political Science; Stony Brook, NY

Postdoctoral Associate/Senior Researcher
Stony Brook University, Department of Political Science; Stony Brook, NY

Postdoctoral Fellow

> Northeastern University, Institute for Experiential AI; Boston, MA

Postdoctoral Research Assistant in High Precision Radial Velocity Searches for Exoplanets (2 posts)

> University of Oxford, Department of Physics; Oxford, England

Postdoc in Music Cultures & AI

> KTH Royal Institute of Technology, Department of Intelligent Systems; Stockholm, Sweden

Postdoctoral Researcher, Machine Learning and Visualization

> Apple, Machine Learning and AI; Pittsburgh, PA

Knight Postdoctoral Research Fellow - "algorithms, misinformation, polarization, propaganda, political institutions, and platform companies"

> University of North Carolina, School of Information and Library Science; Chapel Hill, NC

Hewlett Postdoctoral Research Fellow - "misinformation, partisanship, polarization, propaganda, political institutions, and journalism"

> University of North Carolina, School of Information and Library Science; Chapel Hill, NC

AI/Machine Learning/Encryption Postdoctoral Research Associate (CS Position)

> University of North Carolina, School of Information and Library Science; Chapel Hill, NC

AI/Machine Learning/Encryption Postdoctoral Research Associate (Social Science PhD)

> University of North Carolina, School of Information and Library Science; Chapel Hill, NC

Full-time positions outside academia

Research Fellow - Physiology & Biomedical Engineering

> Mayo Clinic, Multimodal Neuroimaging Lab; Rochester, MN

Human Computation Architect - CTO Data Science

> Bloomberg LLP; New York, NY

Critical Designer, Data Genesis

> New York University, AI Now Institute; New York, NY

Design Director or Senior Digital Designer

> Infographics Group GmbH; Berlin, Germany

Data Science Assistant

> Pew Research Center, Data Labs; Washington, DC

Deputy Chief Technology Officer for Policy & Future Planning

> New York City, Mayor's Office of the Chief Technology Officer; New York, NY

Data Engineer - Informatics Workflows

> Sage Bionetworks; Seattle, WA

Principal Software Automation Engineer, Machine Learning Toolchain

iRobot; Bedford, MA

Research Scientist, HCI Research - AI/ML

> Apple, Machine Learning & AI; Seattle, WA

Sr. Grants Management Specialist (Senior Programs Officer)

> National Foundation on the Arts and the Humanities, Institute of Museum and Library Services; Washington, DC

Natural Resources Specialist (Monitoring Data Coordinator)

Department of the Interior, Bureau of Land Management; Lakewood, CO

Director of Technology Projects

Electronic Frontier Foundation; San Francisco, CA

Diversity and Inclusion Research Fellow

Partnership on AI; San Francisco, CA

Research and Analytics Director, Chatbots

National Domestic Workers Alliance, NDWA Labs; New York, NY, or Remote

Data Science Manager

Center for Policing Equity; Washington, DC

Research Analyst (Health Policy)

The Urban Institute; Washington, DC

Senior Researcher

Girl Scouts; New York, NY

Research Analyst, Africa Growth Initiative

The Brookings Institution, Global Economy and Development program; Washington, DC

Internships and other temporary positions

Seasonal field research assistants, Great Basin

University of California-Davis, Earth Research Institute; Nevada and eastern California

Project Manager for Data Science for Social Good

Alan Turing Institute, University of Warwick; Coventry, England

Microsoft Research Data Science Summer School

Microsoft Research; New York, NY

the Congressional innovation Scholars program

TechCongress; Washington, DC

Basketball Analytics Summer Intern

Los Angeles Lakers; El Segundo, CA

Click here to receive the Data Science Community Newsletter and/or to have us follow your twitter feed so that our data science twitter bot can easily grab links from your tweets.

To send us an announcement for the newsletter, please email laura.noren@nyu.edu and brad.stenger@gmail.com. We retain curatorial discretion.

Data Science Community Newsletter Issue 192.