NYU Data Science Community Newsletter features journalism, research papers, events, tools/software, and jobs for July 22, 2016

Data Science News

Tweet of the Week
Twitter, Pinboard from July 21, 2016

Pinboard
@Pinboard

"We don't accept such incompetence in Silicon Valley". Seriously wondering if there's some second valley that I'm not aware of

9:24pm · 21 Jul 2016 · YoruFukurou

137 RETWEETS   284 LIKES

Can Silicon Valley Really Do Anything to Stop Police Violence?
The New York Times Magazine, Jenna Wortham from July 19, 2016
Ultimately, what the tech industry really cares about is ushering in the future, but it conflates technological progress with societal progress. And perhaps all of us have come to rely too deeply on machinery and software to be our allies without wondering about the cost, the way technology doesn't fix problems without creating new ones.

More data + social movements:
• Can Big Data Help Head Off Police Misconduct? (July 19, NPR, All Tech Considered)
• Here's what 29 million tweets can teach us about Brexit (July 20, The Washington Post, Monkey Cage blog; **Alexandra Siegel** and **Joshua Tucker**)
• Tweeting Turkey, or how social media may have fundamentally changed the future of coups (July 19, The Washington Post, Monkey Cage blog; Joshua Tucker)
• Another tragedy, another #PrayFor, but what does it really say about who cares for whom? (July 18, The Conversation, **Drew Margolin**)

One immigrant's path from cleaning houses to Stanford professor

CNN Money from July 21, 2016
Fei-Fei Li arrived in the U.S. from China at age 16 with many big dreams. And it took many odd jobs to help her achieve them.

Also in gender and class inclusivity:
- Economic Diversity and the Academy: Statistical Scienc (**Sherri Rose**, PhD from July 21, 2016)
- Women are being left behind by the sports data revolution (**Nikita Taparia** from 21 July 2016)

Art and Artificial Intelligence

Medium, MTArt from July 08, 2016
Can intelligent machines really match human creativity? Surely when we think artist, we think human and when we think artificial intelligence, we think machine. At least in the art field, we cannot seem to come to terms with the possibility of a machine truly making art. Why?

Also in data + art:
- There is no difference between computer art and human art (July 20, Aeon Ideas, **Oliver Roeder**)
- SIGGRAPH 2016 Art Gallery (begins July 24, SIGGRAPH)
- Rain Room, by **Random International** (began July 14, **The Los Angeles County Museum of Art**)

Google Maps is turning its over a billion users into editors

TechCrunch, Sarah Perez from July 21, 2016
Google has begun to further tap into the power of the crowd in order to improve its Google Maps application, the company announced this morning. This is being done through the introduction of a number of features that will allow users to more easily share location details, as well as confirm edits suggested by others. Many users had already seen these changes rolling out, but today Google is making them official – an indication that the broader rollout is completing.

More Google:
- Google Sprints Ahead in AI Building Blocks, Leaving Rivals Wary (July 21, Bloomberg, **Jack Clark**)
- Introducing Cloud Natural Language API, Speech API open beta and our West Coast region expansion (July 20, Google Cloud Platform Blog)
- Google Cuts Its Giant Electricity Bill With DeepMind-Powered AI (July 19, Bloomberg, Jack Clark)
- Lessons To Learn From How Google Stores Its Data (July 07, SmartData Collective, **Anand Srinivasa**)

Legendary Hedge Fund Wants to Use Atomic Clocks to Beat High-Speed Traders

Bloomberg, Miles Weiss and Zachary Mider from July 07, 2016

Patent application no. 14/451,356 has one goal: to outrun the speed demons of Wall Street. ... And if it works as advertised, one of the most illustrious names in the hedge-fund business [**Renaissance Technologies**] could gain exclusive U.S. rights to a weapon capable of thwarting even the most predatory of high-speed traders. [video autoplays, 3:21]

SciPy 2016 Retrospective

Camille Scott from July 19, 2016
SciPy, by my accounting, is a curious microcosm of the academic open source community as a whole. It is filled with great people doing amazing work, releasing incredible tools, and pushing the frontiers of features and accessibility in scientific software. It is also is marked by some of the same problems as the larger community.

More SciPy;
- BIDS at SciPy 2016 (July 18, **Berkeley Institute for Data Science**)
- YouTube vids of SciPy 2016 Keynotes and Talks (July 22, 2016; Enthought, YouTube.com)
- **Daniel Chen's** GitHub repo of SciPy 2016 tutorials, notes on talks/keynotes (15 July 2016; Daniel Chen, github.com)

The Gordon and Betty Moore Foundation Grant for Numba and Dask
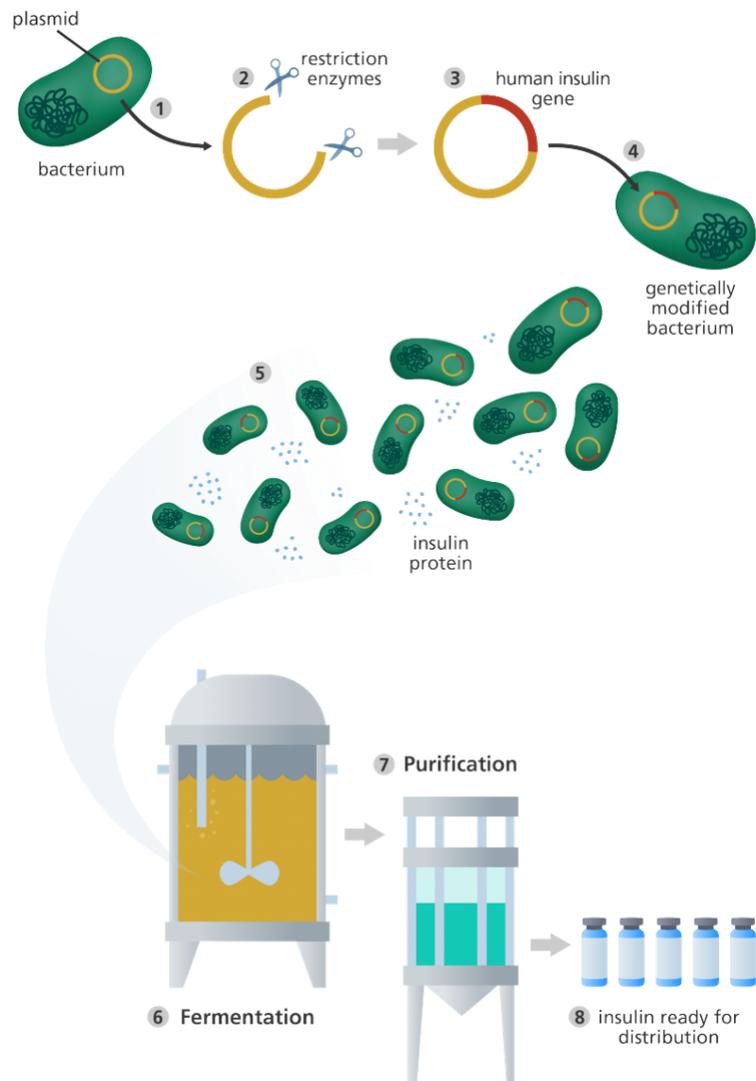
Continuum Analytics from July 14, 2016
I am thrilled to announce that the **Gordon and Betty Moore Foundation** has provided a significant grant in order to help move Numba and Dask to version 1.0 and graduate them into robust community-supported projects.

Additional funding announcements:
- University library Data Curation Network supported by the **Sloan Foundation** (16 May 2016)

Data Science at Zymergen

Medium, @squarecog, Dmitriy Ryaboy from July 19, 2016

Zymergen radically speeds up the process of creating and improving specialized strains. We massively parallelize introduction of small genetic changes into microbes and evaluation of results of those changes. More trials, more errors, more successes, less human effort. Lots of data.

[1607.06450] Layer Normalization

arXiv, Statistics > Machine Learning; Jimmy Lei Ba, Jamie Ryan Kiros, Geoffrey E. Hinton from July 21, 2016

A recently introduced technique called batch normalization uses the distribution of the summed input to a neuron over a mini-batch of training cases to compute a mean and variance which are then used to normalize the summed input to that neuron on each training case. This significantly reduces the training time in feed-forward neural networks. However, the effect of batch normalization is dependent on the mini-batch size and it is not obvious how to apply it to recurrent neural networks. In this paper, we transpose batch normalization into layer normalization by computing the mean and variance used for normalization from all of the summed inputs to the neurons in a

layer on a single training case.

Events

Who's speaking at Data Day Seattle 2016?

Data Day Seattle will be held at the newly-renovated Westin Seattle. This year we're taking the entire building! Enough room for 50+ talks, six tracks, workshops, office hours, booksignings and a post-conference happy hour / job fair.

Seattle, WA Saturday, July 23. [$$$]

Discover R and RStudio at JSM 2016 Chicago!

The JSM conference is one of the largest to be found on statistics, with many terrific talks for R users. We've listed some of the sessions that we're particularly excited about below. These include talks from RStudio employees, like Hadley Wickham, Yihui Xie, Mine Cetinkaya-Rundel, Garrett Grolemund, and Joe Cheng, but also include a bunch of other talks about R that we think look interesting.

Chicago, IL Sunday-Thursday, July 31-August 4. [$$$]

Data Science for Social Good Conference

This conference will highlight the successes, opportunities, and challenges faced by the growing Data Science for Social Good community by bringing key members from each community (academia, governments, non-profits, foundations, social enterprises, and corporations) together to share best practices, learn from each other, and generate new collaboration opportunities.

Chicago, IL Wednesday-Thursday, August 24-25, at the University of Chicago Gleacher Center. [$$]

CCS Warm-Up: a school on complex systems to prepare for CCS2016 conference

Coinciding with the Conference on Complex Systems and profiting the opportunity offered by the presence of a wide variety of experts in different topics in Amsterdam, we are organising a school for PhD students and young postdocs to prepare for such a big event. The school aims to offer young researchers the opportunity to learn new methods, present their work and meet fellow researchers.

Amsterdam, The Netherlands Friday-Sunday, September 16-18. [$$]

Introducing the Microsoft Data Science Summit, Sep 26-27

Join us to hear from thought leaders and Microsoft engineers on the latest Big Data, Machine Learning, Artificial Intelligence, and Open Source techniques and technologies.

Atlanta, GA Monday-Tuesday, September 26-27. [$$$]

Deadlines

2016 Financial Stability Conference - Innovation, Market Structure, and Financial Stability

deadline: Washington DC

Washington, DC The **Federal Reserve Bank of Cleveland** and the **Office of Financial Research** invite the submission of research and policy-oriented papers for the 2016 Financial Stability Conference to be held December 1-2, 2016, in Washington, D.C. The objectives of this conference are to highlight research and advance the dialogue on financial market dynamics that affect financial stability, and to disseminate recent advances in systemic risk measurement and forecasting tools that assist in macroprudential policy development and implementation.

The deadline for submissions is July 31.

NYC Digital Humanities - Third Annual Graduate Student Project Award

deadline: Education Opportunity

We are pleased to announce our third annual cross-institutional NYCDH digital humanities graduate student project award. We invite all graduate students attending an institution in New York City and the metropolitan area to apply.

Deadline to apply is Monday, August 15.

UX, IoT & interaction conference: O'Reilly Design, March 19 - 22, 2017, San Francisco, CA

deadline: Conference

Designers are constantly asked to design for the future—a future that by its nature is a moving target. The O'Reilly Design Conference will provide designers with the skills, connections, and inspiration they need to shape the products and services of today and tomorrow.

San Francisco, CA Sunday-Wednesday, March 19-22.

Deadline for speaking proposals is Wednesday, September 7.

MLconf Industry Impact Student Research Award, Sponsored by Google

deadline: Contest/Award

The MLconf Industry Impact Student Research Award sponsored by **Google** identifies researchers with the potential to disrupt industry. Our committee of distinguished ML professionals will review nominations for 2016.

Deadline for submissions is Friday, October 28.

CDS News

JuliaCon 2016 | Julia 1.0 | Stefan Karpinski

YouTube, JuliaLanguage from July 11, 2016
Karpinski discusses the Julia 1.0 roadmap.

Automatic Machine Learning? | SciPy 2016 | Andreas Mueller

YouTube, Enthought from July 15, 2016
Recent years have seen a widespread adoption of machine learning in industry and academia, impacting diverse areas from advertisement to personal medicine.

As more and more areas adopt machine learning and data science techniques, the question arises on how much expertise is needed to successfully apply machine learning, data science and statistics. [video autoplays, 29:38]

Tools & Resources

Odo: Shapeshifting for your data — odo 0.4.0+38.g00361e1 documentation

Continuum Analytics from September 05, 2015
odo takes two arguments, a source and a target for a data transfer. ... It efficiently migrates data from the source to the target through a network of conversions.

Join d3 on Slack!

Slack from July 19, 2016
840 registered participants as of July 22.

Careers

Postdoctoral Fellowships in Humanities and Social Sciences 2017-2020

HASTAC from July 15, 2016
Princeton, NJ The Princeton Society of Fellows, an interdisciplinary group of scholars in the humanities, social sciences, and selected natural sciences, invites applications for the 2017-2020 Fellowship competition.

Postdoctoral Researcher x 2 (14 month contract) - Programmable City Project, National Institute for Regional and Spatial Analysis (NIRSA)

Maynooth University from July 18, 2016
Maynooth, Co Kildare, Ireland Funded by a **European Research Council** Advanced Investigator Award to **Prof. Rob Kitchin**, we are seeking two postdoctoral researchers to work on The Programmable City project.

18F — Technology Transformation Service looking for new commissioner

18F, U.S. General Services Administration from July 13, 2016
Washington, DC Now's the time to get involved in transforming how the government builds, buys, and shares technology. Now's the time to lead a growing office of

talented and motivated people looking to help federal agencies deliver better services to the public at a lower cost. Sound interesting? Then we've got a job for you: Commissioner of the Technology Transformation Service.

Careers - The Institute for the Interdisciplinary Study of Decision Making

NYU Institute for the Interdisciplinary Study of Decision Making (IISDM) from July 13, 2016

New York, NY IISDM is launching the most audacious study of human beings ever conducted, involving some 10,000 New York City residents and lasting 10-20 years. We need dynamic team members with a strong sense of intellectual curiosity and a desire to affect positive change in the world who will thrive in a start-up like environment while benefiting from the stability of working for one of the premiere universities in the world.

Program Administrator – apply by July 30

Data & Society Research Institute from July 22, 2016

New York, NY You are energetic, friendly, and happy to jump into a range of tasks to keep the programs and events at Data & Society running smoothly. You will be supporting the Director of Programs in administering the D&S residency program, and supporting D&S staff from a variety of teams in small and large events management.