NYU Data Science Community Newsletter features journalism, research papers, events, tools/software, and jobs for July 1, 2016

Data Science News

March of the machines - What history tells us about the future of artificial intelligence—and how society should respond

The Economist from June 25, 2016
Experts warn that "the substitution of machinery for human labour" may "render the population redundant". They worry that "the discovery of this mighty power" has come "before we knew how to employ it rightly". Such fears are expressed today by those who worry that advances in artificial intelligence (AI) could destroy millions of jobs and pose a "Terminator"-style threat to humanity. But these are in fact the words of commentators discussing mechanisation and steam power two centuries ago. Back then the controversy over the dangers posed by machines was known as the "machinery question". Now a very similar debate is under way.

More perspectives on AI:
• From not working to neural networking - The artificial-intelligence boom is based on an old idea, but with a modern twist (June 25, The Economist)
• Enter the Matrix: Developing a Big Red Button for AI and Robots (June 26, Medium, Mark Riedl)
• Microsoft CEO Satya Nadella: Humans and A.I. can work together to solve society's challenges (June 28, Slate)

Request for Information on Artificial Intelligence

Federal Register from June 27, 2016
The **Office of Science and Technology Policy** is interested in developing a view of AI across all sectors for the purpose of recommending directions for research and determining challenges and opportunities in this field. The views of the American people, including stakeholders such as consumers, academic and industry researchers, private companies, and charitable foundations, are important to inform an understanding of current and future needs for AI in diverse fields. The purpose of this RFI is to solicit feedback on overarching questions in AI, including AI research and the tools, technologies, and training that are needed to answer these questions.

Also in government:
• How to Prepare for the Future of Artificial Intelligence (June 27, whitehouse.gov, Ed Felten)
• Researchers Sue the Government Over Computer Hacking Law (June 29, WIRED, Security)
• Biden threatens funding cuts for researchers who fail to report clinical trial results (June 29, STAT)
• USGS finds data fraud, closes chemistry lab (June 24, Chemical & Engineering News)

Why Scientists Are So Worried about Brexit

MIT Technology Review from June 20, 2016
A report by the **House of Lords** reported in April that "the overwhelming balance of opinion from the UK science community" opposed Brexit. ... Why? Partly because the EU funds a lot of science and technology research for its member countries, with 74.8 billion euros budgeted from 2014 to 2020.

More Brexit:
- Brexit polling: What went wrong? (June 24, Andrew Gelman, Statistical Modeling, Causal Inference, and Social Science blog)
- Brexit: voter turnout by age (June 24, Financial Times, John Burn-Murdoch)
- After Brexit, the race is on to replace London as Europe's startup capital (June 27, Quartz, Joon Ian Wong)
- Channelling Brexit anger (June 30, Times Higher Education)

How Manufacturing Giants Are Building The Massive Industrial Internet Of Things

ARC from June 24, 2016
Over the next 10 years, investment in the Industrial Internet of Things is predicted to be hundreds of billions of dollars. Companies such as **General Electric** and **Hitachi** have already started pouring significant amounts of money into dedicated research and software platforms to make the most of the opportunities available.

NYU Stern Just Launched A FinTech MBA Program — It Will Explore Blockchains And Big Data

BusinessBecause from June 22, 2016
NYU Stern School of Business is expanding its MBA program to prepare students for future jobs in the fintech, as the sector faces a tsunami of innovation. Starting Fall 2016, NYU Stern MBAs will be able to choose a dedicated fintech specialization, which NYU Stern says is a first among top business schools, plus eight new electives.

More FinTech:
- The new fintech generation: don't fight the banks, embrace them (June 21, Lending Times)
- Rogue Machine Intelligence and A New Kind of Hedge Fund (June 21, Medium, Numerai)

[1606.08813] EU regulations on algorithmic decision-making and a "right to explanation"
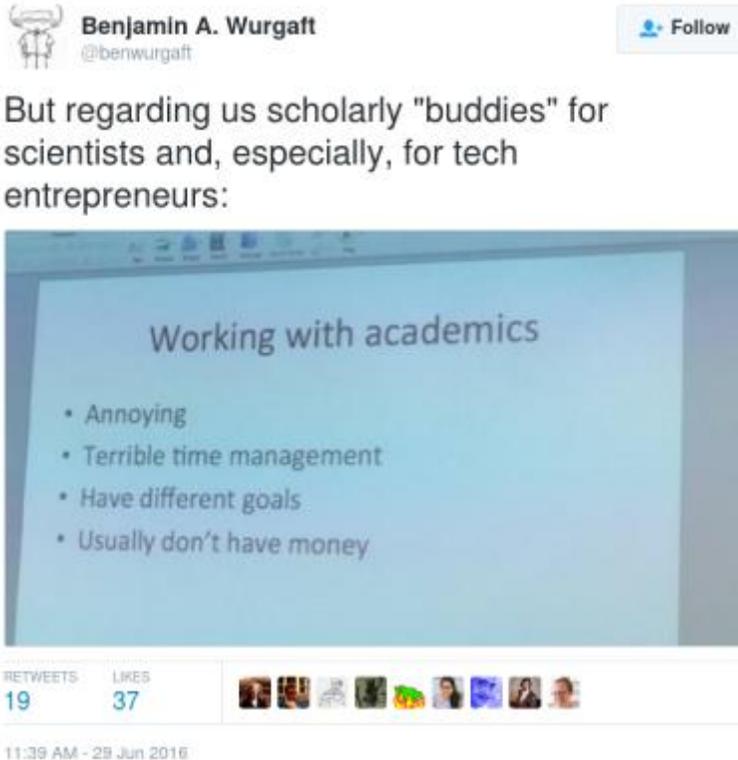
arXiv, Statistics > Machine Learning; Bryce Goodman, Seth Flaxman from June 28, 2016
We summarize the potential impact that the **European Union's** new General Data Protection Regulation will have on the routine use of machine learning algorithms. Slated to take effect as law across the EU in 2018, it will restrict automated individual decision-making (that is, algorithms that make decisions based on user-level predictors) which "significantly affect" users. The law will also create a "right to

explanation," whereby a user can ask for an explanation of an algorithmic decision that was made about them. We argue that while this law will pose large challenges for industry, it highlights opportunities for machine learning researchers to take the lead in designing algorithms and evaluation frameworks which avoid discrimination.

Tweet of the Week

Twitter, Benjamin A. Wurgaft from June 29, 2016

Events

AI Now

The **White House** and **New York University's Information Law Institute**, with support from **Google Open Research**, **Microsoft Research** and the **MacArthur Foundation** will host a major public symposium to address the near-term impacts of AI technologies across social and economic systems.

The focus will be the challenges of the next 5-10 years, specifically addressing four themes: social inequality, labor, healthcare, and ethics.

New York, NY Thursday, July 7, at NYU Skirball Center.

12th International Workshop on Mining and Learning with Graphs (MLG 2016)

San Francisco, CA Held in conjunction with KDD'16 on Sunday, August 14.

O'Reilly Artificial Intelligence Conference

Discover the real-world opportunities of applied artificial intelligence

New York, NY Monday-Tuesday, September 26-27.

Peter Hall Memorial Conference

Please join the UC Davis Department of Statistics for the Peter Hall Memorial Conference in honor of Distinguished Professor **Peter Hall** who sadly passed away in January 2016.

Davis, CA Friday-Saturday, September 30-October 1 at the UC-Davis Conference Center.

Women in Statistics and Data Science Conference

The **American Statistical Association** invites you to join us for the 2016 Women in Statistics and Data Science Conference—the only one for the field tailored specifically for women!

Charlotte, NC

1-day Reproducibility Conference Coming to Columbia University December 2016!

Columbia University and other New York City research institutions are hosting a one-day symposium to showcase a robust discussion of reproducibility and research integrity among leading experts, high-profile journal editors, funders and researchers.

New York, NY Friday, December 9, at Columbia University. Registration coming soon.

Deadlines

AMTA 2016 | Call for MT Research Papers

deadline: Conference

AMTA-2016 solicits original research papers that will advance the field of Machine Translation. We seek submissions across the entire spectrum of MT-related research activity. ... This year, we are particularly interested in submissions discussing research on translation system design, implementation, and scaling.

Austin, TX Saturday-Wednesday, October 29-November 2.

Deadline for submissions is Monday, July 11.

COLING 2016

deadline: Conference

Osaka, Japan COLING 2016, the 26th International Conference on Computational Linguistics, will be organized by the Association for Natural Language Processing (ANLP) from Sunday-Friday, December 11-16.

Deadline for submissions is Friday, July 15.

Y Combinator - New Cities
deadline: Contest/Award

Some existing cities will get bigger and there's important work being done by smart people to improve them. We also think it's possible to do amazing things given a blank slate. Our goal is to design the best possible city given the constraints of existing laws. ... We've begun research and are now forming a team to work on it full-time. We need people with strong interests and bold ideas in architecture, ecology, economics, politics, technology, urban planning, and much more.

Deadline to apply is Saturday, July 30.

2016 Workshop on Visualization for the Digital Humanities
deadline: Conference

Baltimore, MD We invite contributions for the 2016 Workshop on Visualization for the Digital Humanities. This will be a one day workshop taking place as part of IEEE VIS 2016.

Deadline for submissions is Saturday, July 30.

NYC Digital Humanities - Third Annual Graduate Student Project Award
deadline: Contest/Award

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.

Leamer-Rosenthal Prizes for Open Social Science
deadline: Contest/Award

In order to promote transparent research, and to offer recognition and visibility to scholars practicing open social science, the **John Templeton Foundation** is generously supporting the **Berkeley Initiative for Transparency in the Social Sciences** to launch prizes named for pioneers who helped lay the foundations of research transparency: economist **Edward E. Leamer** and psychologist **Robert Rosenthal**.

Deadline for nominations is Friday, September 16.

Call for Computer Vision Research Proposals with New Amazon Bin Image Data Set

deadline: RFP

The Amazon Academic Research Awards (AARA) program is soliciting computer vision research proposals for the first time. The AARA program funds academic research and related contributions to open source projects by top academic researchers throughout the world.

Deadline for application submissions is Saturday, October 1.

Tools & Resources

---

rltorch: A RL package for Torch that can also be used with openai gym

GitHub - ludc from June 26, 2016
This package is a basic Reinforcement Learning package written in LUA for Torch. It implements some simple environments and learning policies (Policy Gradient and Deep Q Learning). It also can be easily used with the OpenAI Gym package by using lutorpy (example given in the opeanaigym directory).

Juno for Jupyter: Taking Your Analysis to the Data

Medium, Timbr.io from June 23, 2016
As we've worked with Jupyter notebooks one of our frustrations has been moving around large data sets to a local or hosted notebook environment. One of the most challenging datasets to move around and analyze is satellite imagery. Each individual image can be gigabytes in size. This can make downloading and analyzing imagery slow and painful.

To make life a bit easier the **Timbr.io** team has created a new project called Juno. Using Juno you can connect to, execute commands on and get output from remote Jupyter kernels from within a local or hosted Jupyter Notebook, while also sending the outputs to the app for sharing with others. The remote kernels you spin up in Juno can be accessed from your notebook via our new, open source Jupyter Notebook extension called juno-magic.

Release v4.0.0 · d3

GitHub - d3 from June 28, 2016
D3 is now modular, composed of many small libraries that you can also use independently. Each library has its own repo and release cycle for faster development. The modular approach also improves the process for custom bundles and plugins.

There are a lot of improvements in 4.0: there were about as many commits in 4.0 as in all prior versions of D3. Some changes make D3 easier to learn and use, such as immutable selections. But there are lots of new features, too!

working with spatial data - workshop materials

GitHub - enjalot from June 25, 2016
This workshop is designed to be very hands-on, with many examples that can be

extended as exercises. It would be impossible to touch everything that we could find interesting in web mapping, so the hope is that after going through these three acts you will feel empowered to swap in your own data and leverage hundreds of examples in your own data visualization projects!

Going beyond full utilization: The inside scoop on Nervana's Winograd kernels

Nervana Systems, Urs Köster and Scott Gray from June 28, 2016
This is part 2 of a series of posts on how Nervana uses the Winograd algorithm to make convolutional networks faster than ever before. In the first part we focused on benchmarks demonstrating a 2-3x algorithmic speedup. This part will get a bit more technical and dive into the guts of how the Winograd algorithm works, and how we optimized it for GPUs.

Careers

Join Our Team || Databrary: An Open Data Library for Developmental Science

New York University from June 21, 2016
New York, NY Databrary is a web-based video data library where scientists who collect and analyze research video can store and share their data openly with other researchers. ... We are growing a team to design and build new open source solutions for science using the latest technologies.

Supervisory Health Science Policy Analyst - OD - DE, National Institutes of Health

USAJOBS from June 27, 2016
Montgomery County, MD The Precision Medicine Initiative (PMI) Cohort Program is a newly established office within the **NIH Office of the Director**. The highly visible initiative will result in the establishment of the nation's premier research cohort.

Are you our new Managing Director?

University of Virginia Library from June 29, 2016
Charlottesville, VA The **University of Virginia Library** seeks a Managing Director to help shape and steward our internationally respected Scholars' Lab. This individual will work closely with the Academic Director to develop and support the deep resources for digital scholarship within the vibrant intellectual community at UVA.

Big Data-Scientist Training Enhancement Program (BD-STEP)

U.S. Department of Veterans Affairs, Office of Academic Affiliations from June 29, 2016
VA Hospital sites across the U.S. Junior-level physical scientist trainees from accredited programs in relevant disciplines (including engineering, chemistry, computer science, mathematics, and physics) will be identified by the **National Cancer Institute's Center for Strategic Initiatives**. BD-STEP will draw on the expertise of these highly skilled data scientists-in-training to facilitate the execution of large-scale system changes in clinical care to expeditiously improve patient outcomes. BD-STEP matches trainees with VA medical centers across the country to leverage

VA data systems, supporting clinically-relevant, year-long training and research opportunities in collaboration with VA clinician scientists.

Deadline to apply is Tuesday, July 5.