PROLOG

DECEMBER 2020

Hello, and welcome to the first edition of Prolog, the GRAIL Network newsletter! Here, we'll provide regular updates on AI-related happenings within the U.S. federal government and European Union, topical writing from the Center for Democracy & Technology (CDT) and R Street Institute, and work and perspectives from within the Network. More about the Governance Research in Artificial Intelligence Leadership Network is available on our website.

While this newsletter is created primarily for GRAIL Network members, we hope to also create a useful and relevant resource for Congressional staff, folks at federal agency offices, and other policymakers.

To help us off to a good start, we hope you'll take a short anonymous survey that will inform what we feature. You can also send questions, comments, news items, and recent work to info@grailnetwork.org.

TAKE SURVEY

Spread the Word: The CDT research team is hiring a Postdoctoral Fellow to work on developing models and tools to improve the explainability and auditability of AI-based content moderation systems. More on the position is available here, and anyone interested is welcome to contact CDT with questions. We hope you'll share the position with colleagues and students who might qualify.

Named-Entity Recognition: New Work from Network Members



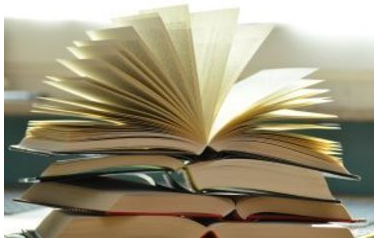
- Shira Mitchell and Kristian Lum examine the assumptions and choices made to justify the use of prediction-based decision making, discuss how those choices and assumptions raise fairness concerns, and offer a more

consistent catalog of fairness definitions in "Algorithmic Fairness: Choices, Assumptions, and Definitions."

- Jonathan Frankle examines the tradeoffs associated with distributed training methods for neural networks, finding that using local stochastic gradient descent (SGD) results in faster training times, but lower accuracy, in "Trade-offs of Local SGD at Scale: An Empirical Study."

- Avi Goldfarb examines how an increased reliance on prediction-based systems in warfare will also increase the need for, and value of, human judgement, in "Artificial intelligence in war: Human judgment as an organizational strength and a strategic liability."

Data Mining: Contributions from CDT and R Street



- Challenging the Use of Algorithm-driven Decision-making in Benefits Determinations Affecting People with Disabilities: Algorithm-driven decision-making tools are often implemented to assess people's eligibility for, or the distribution of, public benefits. They frequently reduce and deny benefits, though, often with unfair and inhumane results and disproportionate and particular harm to people with disabilities. This CDT report analyzes legal challenges to these tools that have been filed within the past 10 years, identifies key insights into what went wrong, and analyzes the legal arguments that plaintiffs have used to challenge those systems in court.

- Algorithm-driven Hiring Tools: Innovative Recruitment or Expedited Disability Discrimination?: The Americans with Disabilities Act has explicit prohibitions against the use of hiring processes that discriminate on the basis of disability, but many algorithm-driven hiring tools fall far short of these standards. This CDT report seeks to highlight how hiring tools may profoundly affect people with disabilities, the legal liability employers may face for using such tools, and concrete steps for employers and vendors to mitigate some of the most significant areas of concern.

- CDT Comments to NIST on its Four Principles of Explainable Artificial Intelligence: CDT commented in support of NIST's principles, which establish that explanations for how AI systems work must provide accurate information about the reasoning for an output from an AI system, and give the audience the right level of knowledge and understanding for their circumstances. CDT also suggested several ideas to improve the utility of NIST's principles.

- Civil Rights Principles on the Use of AI in Hiring: Employers are increasingly relying on AI-driven tools to process and select candidates, which risk perpetuating existing patterns of inequality and foreclosing economic opportunity. To ensure that algorithmic hiring tools do not erect artificial barriers to employment, CDT was proud to partner with the Leadership Conference on Civil and Human Rights, Upturn, and numerous other civil rights groups in releasing principles on the use of AI in hiring and working. The principles underscore particular ways in which AI-driven hiring tools may impact disabled workers, and highlight responsibilities for employers and vendors to address these risks.

Intelligence Explosion: Applications in AI



- AI Solves 50-Year Old Science Problem In 'Stunning Advance' That Could Dramatically Change How We Fight Diseases, Researchers Say
- Spotify Just Invented AI Technology That Will Police Songwriter Plagiarism



Committee Machine: What We're Tracking

Legislation

The 2021 National Defense Authorization Act (NDAA), if signed into law, would enact several military policy changes regarding military cybersecurity and artificial intelligence, and elevate the Department of Defense's Joint AI Center (JAIC). The legislative package also included a handful of AI-related bills not connected to the DoD or JAIC:

- **S. 1558, Artificial Intelligence Initiative Act** (last action May 21, 2019): Organizes a coordinated national strategy for developing AI, and authorizes a $2.2B federal investment over five years to build an AI-ready workforce, accelerating the responsible delivery of AI applications from government agencies, academia, and the private sector over the next 10 years. It would create a new National AI Initiative Office in the White House's Office of Science and Technology Policy and another federal advisory committee on AI. The NDAA increases the legislation's proposed authorization of funds to more than $6 billion.

- **S. 2065, Deepfake Report Act** (last action October 28, 2019): Directs the Department of Homeland Security (DHS) to conduct an annual study of deepfakes and other types of similar content. It requires DHS to assess the technology used to generate deepfakes, the uses of deepfakes by foreign and domestic entities, and available countermeasures to deepfakes.

- **S. 3890, National AI Research Resource Task Force Act of 2020** (last action June 4, 2020): Would direct the National Science Foundation to create a task force to investigate methods for AI research funding. The task force would develop a detailed roadmap for a national cloud computer for AI research. The legislation would also convene a group of technical experts across academia, government, and industry to develop a detailed roadmap for how the United States can build, deploy, govern, and sustain a national research cloud.

- **S. 4082, Artificial Intelligence Standards and National Security Act** (last action June 25, 2020): Requires DoD to report on its role in the development of AI standards. The DoD must specifically assess the ways in which an AI standards strategy will improve the national security; and the feasibility and current status of assigning members of the Armed Forces on active duty to the Joint AI Center of DoD.

- **S. 4901, Ensuring American Leadership over International Standards Act of 2020** (last action November 16, 2020): Requires the Director of NIST to commission a study on the impact of the Chinese government's influence in setting global standards for emerging technologies and would provide feedback on how the United

States and our global allies can continue to ensure that international standards setting continues in a transparent and democratic manner.

Executive and Federal Agency Actions

- **Presidential Executive Order on Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government** (issued December 3, 2020): This order is a follow-up to the Executive Order issued earlier this year. It lays out some more high-level guidelines for federal agencies, but takes a largely non-regulatory approach.

- **NIST Explainability Workshop**: As a follow up to its paper, "Four Principles of Explainable Artificial Intelligence," NIST will hold a three-day workshop (Jan. 26-28) that will delve further into developing an understanding of explainable AI.

Court Cases

- ***Chavis v. Delaware***: This appeal to the Supreme Court asks whether DNA evidence derived from a multi-analyst testing process can be introduced through only one of the testing analysts. This isn't exactly an algorithmic decision issue, but it does raise interesting questions about the explainability and verification of complex systems in court, generally.

NGOs

- The fourth annual Association for Computing Machinery's Fairness Accountability, and Transparency (ACM FAccT) conference (Mar. 3-10, 2021) has an open call for workshops, panels and other presentation proposals with a deadline of January 4, 2021.

Please send opportunities you'd like us to promote, questions, tips, comments, news items, and any recent work we should share to info@grailnetwork.org.

Share this email:

emma

Manage your preferences | Opt out using TrueRemove®
Got this as a forward? Sign up to receive our future emails.
View this email online.

1401 K St NW Suite 200
Washington, DC | 20005 United States

This email was sent to .
To continue receiving our emails, add us to your address book.

Subscribe to our email list.