March 4, 2022

COMMENTS IN RESPONSE TO RFI TO THE UPDATE OF THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN (DOCUMENT NUMBER 2022-02161)

The non-profit Association for Computing Machinery (ACM), with more than 50,000 U.S. members and approximately 100,000 worldwide, is the world's largest educational and scientific computing society. ACM's US Technology Policy Committee (USTPC), currently comprising more than 160 members, serves as the focal point for ACM's interaction with all branches of the US government, the computing community, and the public on policy matters related to information technology. It is charged with providing policy and law makers throughout government with timely, substantive and apolitical input on computing technology and the legal and social issues to which it gives rise.[1]

In response to the Office of Science and Technology Policy's Request for Information to the Update of the National Artificial Intelligence Research and Development Strategic Plan of February 1, 2022 (RFI),[2] USTPC is pleased to submit the following comments:[3]

First, while we support all eight of the strategies outlined in the "The National Artificial Intelligence R&D Strategic Plan: 2019 Update," we are pleased that the overall plan is being reviewed and updated. We especially encourage a focus on strategies 3 and 4: "understand and address the ethical, legal, and societal implications of AI" and "ensure the safety and security of AI systems," respectively. The Committee notes that building systems that achieve these aims is difficult. We believe, therefore, that emphasizing and enabling research to advance the field of accountable AI system design is especially important.

[1] To arrange for a technical briefing from USTPC and other ACM expert members, please contact Adam Eisgrau, ACM Director of Global Policy & Public Affairs, at acmpo@acm.org or 202-580-6555.

[2] *See* 87 FR 5876 (February 2, 2022) at https://www.federalregister.gov/documents/2022/02/02/2022-02161/request-for-information-to-the-update-of-the-national-artificial-intelligence-research-and.

[3] The lead author of these Comments for USTPC was its Artificial Intelligence & Algorithms Subcommittee Chair Prof. Jeanna Matthews of Clarkson University. Also contributing were USTPC members L. Jean Camp, Charalampos Chelmis, Thomas Chen, Carlos Jiménez, Arnon Rosenthal, Ben Schneiderman, and Kenneth Zhang.

ACM U.S. Technology Policy Committee
1701 Pennsylvania Ave NW, Suite 200
Washington, DC 20006

+1 202.580.6555
acmpo@acm.org
www.acm.org/public-policy/ustpc

Second, we recommend that the Strategic Plan adopt the broadest possible definition of artificial intelligence to include, specifically, automated or algorithmic decision-making systems more broadly. This is appropriate and necessary because, when automated systems are used to make critical decisions impacting society and the lives of individuals, the ethical, legal, societal, safety and security issues are similar regardless of the complexity or interpretability of the algorithms. Analysis of safety and security should include comprehensive evaluation of data compilations used for training, the accuracy of decision-making systems, and the potential for the abusive use of platforms.

Third, we encourage revision of the current plan to rank tiers of systems based on the critical nature of their impact on individuals and society and to hold systems classified in higher tiers to proportionately higher standards of verification and validation, testing, documentation and explanation. The criteria for determining the level of rigor applied to a system should be dependent on its impact on individuals and society, rather than the complexity of its algorithms, or of the size or nature of the company producing it. Automated decision-making systems impacting human life and liberty should be held to the highest standards including independent verification and validation, audit trails, and retrospective analyses of failures. The same should be true for systems deployed in high impact or highly regulated areas such as hiring, housing, credit, and the allocation of public resources, and others.

Finally, in updating the National Artificial Intelligence Research and Development Strategic Plan, we respectfully commend the agency's attention to the attached Statement on Algorithmic Transparency and Accountability[4] and its seven associated principles: 1) awareness; 2) access and redress; 3) accountability; 4) explanation; 5) data provenance; 6) auditability; and 7) validation and testing. The Statement is a joint product of ACM's Europe and US policy committees.

ACM's US Technology Policy Committee looks forward to assisting OSTP, NSF and other agencies throughout the process of reconsideration and revision of the 2019 Strategic Plan and welcomes all inquiries to that end. For further information, or should you have any other questions, please contact ACM's Director of Global Public Policy, Adam Eisgrau, at 202-580-6555 or eisgrau@acm.org.

[4] https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf

Association for Computing Machinery
US Public Policy Council (USACM)

usacm.acm.org
facebook.com/usacm
twitter.com/usacm

January 12, 2017

Statement on Algorithmic Transparency and Accountability

Computer algorithms are widely employed throughout our economy and society to make decisions that have far-reaching impacts, including their applications for education, access to credit, healthcare, and employment.[1] The ubiquity of algorithms in our everyday lives is an important reason to focus on addressing challenges associated with the design and technical aspects of algorithms and preventing bias from the onset.

An algorithm is a self-contained step-by-step set of operations that computers and other 'smart' devices carry out to perform calculation, data processing, and automated reasoning tasks. Increasingly, algorithms implement institutional decision-making based on analytics, which involves the discovery, interpretation, and communication of meaningful patterns in data. Especially valuable in areas rich with recorded information, analytics relies on the simultaneous application of statistics, computer programming, and operations research to quantify performance.

There is also growing evidence that some algorithms and analytics can be opaque, making it impossible to determine when their outputs may be biased or erroneous.

Computational models can be distorted as a result of biases contained in their input data and/or their algorithms. Decisions made by predictive algorithms can be opaque because of many factors, including technical (the algorithm may not lend itself to easy explanation), economic (the cost of providing transparency may be excessive, including the compromise of trade secrets), and social (revealing input may violate privacy expectations). Even well-engineered computer systems can result in unexplained outcomes or errors, either because they contain bugs or because the conditions of their use changes, invalidating assumptions on which the original analytics were based.

The use of algorithms for automated decision-making about individuals can result in harmful discrimination. Policymakers should hold institutions using analytics to the same standards as institutions where humans have traditionally made decisions and developers should plan and architect analytical systems to adhere to those standards when algorithms are used to make automated decisions or as input to decisions made by people.

This set of principles, consistent with the ACM Code of Ethics, is intended to support the benefits of algorithmic decision-making while addressing these concerns. These principles should be addressed during every phase of system development and deployment to the extent necessary to minimize potential harms while realizing the benefits of algorithmic decision-making.

[1] Federal Trade Commission. “Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues.” January 2016. https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report.

Principles for Algorithmic Transparency and Accountability

1. Awareness: Owners, designers, builders, users, and other stakeholders of analytic systems should be aware of the possible biases involved in their design, implementation, and use and the potential harm that biases can cause to individuals and society.

2. Access and redress: Regulators should encourage the adoption of mechanisms that enable questioning and redress for individuals and groups that are adversely affected by algorithmically informed decisions.

3. Accountability: Institutions should be held responsible for decisions made by the algorithms that they use, even if it is not feasible to explain in detail how the algorithms produce their results.

4. Explanation: Systems and institutions that use algorithmic decision-making are encouraged to produce explanations regarding both the procedures followed by the algorithm and the specific decisions that are made. This is particularly important in public policy contexts.

5. Data Provenance: A description of the way in which the training data was collected should be maintained by the builders of the algorithms, accompanied by an exploration of the potential biases induced by the human or algorithmic data-gathering process. Public scrutiny of the data provides maximum opportunity for corrections. However, concerns over privacy, protecting trade secrets, or revelation of analytics that might allow malicious actors to game the system can justify restricting access to qualified and authorized individuals.

6. Auditability: Models, algorithms, data, and decisions should be recorded so that they can be audited in cases where harm is suspected.

7. Validation and Testing: Institutions should use rigorous methods to validate their models and document those methods and results. In particular, they should routinely perform tests to assess and determine whether the model generates discriminatory harm. Institutions are encouraged to make the results of such tests public.