The Value Proposition for Statistics, as a Discipline and Practice, in AI Research and Development[1]

ASA's comments to the Request for Information: Identifying Priority Access or Quality Improvements for Federal Data and Models for Artificial Intelligence Research and Development, and Testing; Extension of Comment Period

August 9, 2019

As the National Research & Development Strategic Plan articulates, *artificial intelligence (AI) holds tremendous promise to benefit nearly all aspects of society, including the economy, healthcare, security, the law, transportation, and even technology itself. With this tremendous potential comes great responsibility to minimize harm to society and ensure the US's leadership in cutting-edge advances in AI.*

As clearly stated in the plan, machine learning is a cornerstone for advancements of AI in the data science age. Machine learning is a frontier field at the interface of statistics and computer science. It is an important part of modern statistics that has been built on statistical ideas while taking computational considerations up-front. With that the American Statistical Association (ASA) recommends highlighting the role of statisticians in the development of AI in the US's Research and Development Strategic Plan. Promoting statisticians in this way will help to ensure that the *United States continues to lead the world in cutting-edge advances in AI that will grow our economy, increase our national security, and improve quality of life.*

Statistics brings a rich and deep history of data methodology, practice, collaboration, & ethics, where statistics is the underlying data methodology that strengthens the reliability and stability of AI. Statisticians bring our rich expertise and contributions in the following crucial areas of AI to implement the National AI Research & Development Strategic Plan:

- ***Statistics is a key part of AI:*** *All data methodologies that support AI are rooted in statistical concepts, principles, ideas, best practices and theory.* ***Statisticians develop***

[1] The ASA thanks the Statistical Learning and Data Science Section for its leadership on these comments. In particular we are grateful to Sarah Kalicin, Bin Yu, Joanne Wendelberger, Edgar Dobriban, and Nick Horton.

and validate many data science (or AI) methods and algorithms that address concerns in the Strategic Plan. For instance, predictability, explainability, transparency, and robustness can all be applied to ensure maximum societal benefit (minimize do-no-harm) and enhance US's AI Leadership.

- **Statistics leads and refines data problem-solving frameworks for knowledge discovery:** While data (statistical) methodologies extract knowledge from data, problem-solving frameworks - based on the scientific method - provide problem, data, assumption, and requirement identification focus that are crucial for developing sustainable and stable AI. Many problems in society that will be identified by AI knowledge discovery will require bringing together diverse team members to use comprehensive, efficient approaches to ensure that the right solutions are developed. ***Statisticians lead in the refinement of problem-solving frameworks - based on the scientific method - to foster better data problem identification techniques and data (statistical) methodologies.***

- **Statistical literacy fosters better AI adoption and effectiveness**: For AI to be effective, it needs to be trusted and adopted by society. This means that members of society need to know how to question, understand, embrace (when it is a benefit), and reject (when it is not a benefit) the underlying data (Statistical) methodology that AI has developed. Statistical literacy -- the ability to understand data use, assumptions, interpretability, communication, associated uncertainty, and results -- is not just crucial for society to understand, but for AI researchers to use in AI development to articulate and quantify benefits and risks. ***Statisticians, from academia and industry, have been teaching statistical literacy to foster better collaboration, alignment, and buy-in among key project stakeholders to support faster adoption and integration of data (Statistical) methodologies that support AI.***

Moreover, statistics contributes to the efficacy of the AI research infrastructure because it brings a long tradition of scholarship and language to the table. For many problems of recent interest in AI (such as sample and data efficiency, confidence, modeling, causality, reproducibility, robustness), valuable solutions have been proposed in the statistics research community over decades. This can help the entire AI research community and society save valuable resources as we build on existing tools without needing to reinvent them.

The ASA supports the National AI Research & Development Strategic Plan in furthering benefit to our society and ensuring the US's competitive lead in AI. We encourage the committee to recognize the indispensable role of statistics in AI development, by explicitly stating our role in the National AI Research and Development Strategic Plan, by including statisticians in the relevant decision-making committees, and by directing appropriate federal funding to support statistics' contributions in AI Research & Development Strategic investments. Please feel free to reach out to the ASA for more details on furthering our engagement in AI.