ASA Response to Department of Education's Proposed Priority and Definitions—Secretary's Supplemental Priority and Definitions on Advancing Artificial Intelligence in Education

August 20, 2025

Artificial intelligence (AI) is built on data, and understanding data requires statistics. Without statistical and data literacy, AI education risks producing graduates who can operate AI tools but cannot judge their accuracy, detect statistical bias, or improve their performance. To ensure that U.S. students become informed creators of AI technologies, not just passive consumers, the Department of Education's priorities must explicitly include statistics as a core pillar of AI literacy, in addition to computer science and domain expertise.

The American Statistical Association appreciates the opportunity to comment on the proposed rule by the Department of Education, Proposed Priority and Definitions—Secretary's Supplemental Priority and Definitions on Advancing Artificial Intelligence in Education.

We strongly agree with the following statements in the July 21 Notice of Proposed Rulemaking:

> … it is increasingly important for students to develop AI literacy. A strong foundation in AI literacy will help ensure students are prepared to navigate and contribute to a society where these technologies play a growing role in decision-making, communication, innovation, and career readiness.

However, we believe the rule could be substantially stronger and more successful if students understood statistics as well as computer science, both of which are foundational to AI. As data science comprises statistical, computer science and subject matter expertise, we understand that AI sciences are an evolution of that field, where the same three pillars will be essential. As the science of learning from data, statistics permeates nearly all aspects of AI, including the following four ways: (i) ensuring data

quality for reliable AI performance; (ii) accurate modeling; (iii) enhancing AI efficiency; and (iv) addressing AI accuracy and generalizability. Statistical and data literacy is therefore foundational to AI literacy and education. Every AI model has limitations and makes assumptions. Statistical and data literacy helps students to understand the limitations of AI and therefore critically interpret results and adapt accordingly. AI literacy efforts that fail to include statistical and data literacy risk producing students who can use AI but cannot evaluate or improve it. As importantly, enhancing statistical and data literacy skills for all US students will generally prepare them for 21st century jobs.

We therefore urge the following changes in the proposed rule:

1. In the background, amend the third paragraph with the underlined additions and deletions denoted by strike-throughs:

   > To move from passive users of AI to active creators and innovators, students must also understand foundational computer science and statistics. Exposure to core concepts such as algorithms, data processing, data acquisition, data analysis, statistical modeling, model evaluation, and computational thinking can deepen their understanding of how AI systems function. Incorporating computer science and statistics where appropriate can reinforce efforts to build meaningful AI literacy in educational settings.

2. Edit the following "*Proposed Priority"* bullets and sub-bullets as follows:

   > a(i) Support the integration of AI and statistical literacy skills and concepts into teaching and learning practices to improve educational outcomes for students, including how to detect AI generated disinformation or misinformation online;
   >
   > a(ii) Expand offerings of AI, statistics, and computer science education in K-12 education;
   >
   > a(iii) Expand offerings of AI, statistics, and computer science courses as part of an institution of higher education's general education and/or core curriculum;
   >
   > a(iv) Embed AI, statistics, and computer science into an institution of higher education's general preservice or in-service teacher professional development or teacher preparation programs;
   >
   > a(vi) Provide professional development in foundational computer science, statistics, and AI, preparing educators to effectively teach AI in stand-alone computer science, statistics, and other relevant courses, including instruction about how to use AI responsibly;

3. Under *Proposed Definitions*, we strongly recommend adding a definition of statistics:

   > Statistics.—"Statistics" is the science of learning from data, and of measuring and communicating uncertainty, often for the purpose of making better decisions.

Thank you for your consideration. Please see below for more on the foundational role of statistics in AI.

The Foundational Role of Statistical Literacy in AI

AI and data science, a field that includes statistics, are rapidly transforming nearly every sector, profoundly impacting scientific innovation and daily life, as the Priority clearly implies. This transformation, evident in personalized learning instruction, increased learning engagement, advanced speech recognition systems, and predicting the probability of diseases such as cancer, is largely a testament to the innovative and entrepreneurial spirit characterizing these burgeoning areas.[1]

AI, broadly conceptualized as "Software with the Ability of Performing Human Mental Work," encompasses machine learning, data mining, and statistical modeling.[2] The inherently interdisciplinary nature of statistics and AI necessitates substantial collaborative efforts across various fields. This necessitates that statisticians and data scientists be extensively involved in all AI initiatives to fully realize their potential for productivity, innovation, and problem-solving.[3] The ASA unequivocally states that statistics plays a "central role" in AI, particularly within the domains of machine learning and deep learning.[4]

A strong statistical foundation is crucial for AI system developers and practitioners. It directly improves model reliability and helps prevent common pitfalls, biased decision-making, and unreliable predictions. Statistical rigor is also essential for ensuring the trustworthiness of AI systems, encompassing accountability, transparency, and fairness. For instance, statistical practitioners are responsible for disclosing known biases in data collection and the limitations of AI systems, and for applying methods to assess and mitigate risks of unfair outcomes. Without this literacy, it becomes challenging to evaluate AI decisions, ensure computational reproducibility, or maintain statistical validity.

The demand for AI talent has been growing at an even faster rate than the increasing supply of AI professionals trained in U.S. colleges and universities, as reported by the Council of Economic Advisers in January 2025.[5] This highlights a critical and widening gap in the workforce that foundational education can address. Because AI's core functionality is intrinsically built upon statistical principles, then education in AI must necessarily commence with, or deeply integrate, education in statistics. A curriculum that focuses solely on the mechanical explanation of algorithms without grounding in the underlying statistical rationale would produce practitioners ill-equipped to understand model limitations, diagnose failures, or innovate responsibly. This points to a fundamental need for curriculum reform that prioritizes

---

[1] American Statistical Association. (2023, August 4). *ASA Statement on The Role of Statistics in Data Science and Artificial Intelligence*.
[2] American Statistical Association. (2025, July 10). ASA Comments on the Future of the National Library of Medicine Biomedical and Data Science Extramural Research Programs.
[3] American Statistical Association. (2023, August 4). *ASA Statement on The Role of Statistics in Data Science and Artificial Intelligence*.
[4] American Statistical Association. (2023, August 4). *ASA Statement on The Role of Statistics in Data Science and Artificial Intelligence*.
[5] American Statistical Association. (2025, May 1). Policy Actions the White House Should Take on AI. Amstat Magazine Blog.

statistical and mathematical foundations as prerequisites for advanced AI studies, rather than as separate or elective subjects. Therefore, for any proposed rule concerning AI in education, integrating data and statistical literacy is not merely an enhancement but essential for cultivating a generation of AI-literate, workforce-ready professionals capable of building and managing AI systems responsibly and effectively. This ensures that AI education moves beyond a "mechanical explanation of algorithms" to encompass the entire data science process and its supporting statistical concepts, from data acquisition to model interpretation and ethical deployment.

Statistical Literacy as a "Statistical Bridge" for Effective Collaboration

One commonly known issue is "...the interdisciplinary nature of data science and AI means a substantial collaborative effort is needed, and that statisticians—who themselves are data scientists—should be extensively involved".[6] It is also similarly suggested that "it is unrealistic to expect one person working in data science or AI to possibly be an expert in all relevant areas".[7] This concept is reinforced by the emphasis on "Collaboration between methodological experts and domain scientists" being "central to data science".[8] The recognition that no single individual can master all facets of AI underscores the imperative for effective teamwork. In this context, statistical literacy can serve as a crucial common language and conceptual framework, enabling diverse experts—from computer scientists and engineers to domain specialists and ethicists—to communicate effectively, understand the inherent uncertainties and limitations of models, and collectively ensure the integrity, reliability, and ethical deployment of AI systems. Without this shared foundational literacy, interdisciplinary collaboration risks becoming fragmented, leading to misinterpretations, inefficient development, and potentially suboptimal or even harmful AI outcomes. This highlights the need for statistical education that emphasizes not just technical proficiency but also durable skills , such as communication, critical thinking, and collaborative problem-solving within complex interdisciplinary teams. Durable skills are a combination of how you use what you know, including skills like critical thinking, communication, collaboration, and creativity, as well as character skills like fortitude, growth mindset, and leadership.[9]

Knowledge Extraction from Data: The Statistical Bedrock of AI

AI and data science fundamentally depend on statistics, mathematics, and computer science to derive insights and knowledge from data.[10] These fields collectively provide the necessary tools to interact with data, offer effective and meaningful summaries or inferences, and, in the case of AI, develop systems

---

[6] American Statistical Association. (2023, August 4). *ASA Statement on The Role of Statistics in Data Science and Artificial Intelligence*.
[7] American Statistical Association. (2023, August 4). *ASA Statement on The Role of Statistics in Data Science and Artificial Intelligence*.
[8] American Statistical Association. (2025, March 14). *ASA Response to the OSTP Request for Information on the Development of an Artificial Intelligence Action Plan*.
[9] https://americasucceeds.org/wp-content/uploads/2025/08/Durable-by-Design_July-2025-.pdf
[10] American Statistical Association. (2023, August 4). *ASA Statement on The Role of Statistics in Data Science and Artificial Intelligence*.

capable of performing tasks typically requiring human intellectual processes.[11] Statistics holds a "central role" in data science and AI, particularly within the domains of machine learning and deep learning.[12]

Key Table 1: Core Statistical Principles Underpinning AI Functionality

Statistical Principle	AI Application/Methodology	Benefit to AI System Quality/Reliability
Statistical Inference	Machine Learning Model Development, Deep Learning	Separating signal from noise, Formulating questions in terms of underlying processes
Uncertainty Quantification	Predictive Modeling, AI Decision-Making	Quantifying confidence in predictions, Addressing unmet needs in complex data environments
Causal Inference	Policy Interventions, Actionable Insights	Distinguishing causation from correlation, Identifying effective interventions that cause desired changes
Reproducibility	Algorithm Development, Model Deployment	Ensuring predictable and verifiable outcomes
Model Validation	AI System Evaluation, Quality Assurance	Preventing overfitting, Ensuring models perform well on real-world data

[11] American Statistical Association. (2023, August 4). *ASA Statement on The Role of Statistics in Data Science and Artificial Intelligence*.
[12] American Statistical Association. (2023, August 4). *ASA Statement on The Role of Statistics in Data Science and Artificial Intelligence*.

Bias Detection & Mitigation	Data Collection & Preprocessing, Ethical AI Development	Identifying and addressing statistical biases in data and algorithms, Promoting fairness, Ensuring accurate and unbiased information

The Role of Human Oversight and Statistical Competence in AI Decision-Making

The principle of "human-in-the-loop" is a defining characteristic of trustworthy AI, ensuring that human oversight is systematically integrated into AI decision-making processes.[13] The entire ethical AI framework, as articulated by the ASA, is built upon the guiding principles of accountability, transparency, and fairness.[14] Each of these principles directly maps to and requires a deep understanding of core statistical concepts. For example, "disclos[ing] known biases inherent in the data collection process" and understanding the "limitations and implications of the use of AI"[15] demands sophisticated statistical knowledge of sampling, measurement error, confounding, and model assumptions.[16] The statement that "Fairness is a fundamental value in statistical practice, as it aims to provide accurate and unbiased information"[17] explicitly links ethical goals to statistical methodology. Furthermore, the inherent challenges of AI—such as "enormous data sets, often collected in ways that may compromise their integrity," "uncertainties about the accuracy and reliability of the data," and a "high risk of overfitting"[18]—are fundamentally statistical problems that necessitate statistical solutions for ethical resolution. Therefore, ethical AI is not merely a philosophical or legal aspiration; it is, at its core, a statistically informed endeavor.

Without a robust foundation in statistical literacy, eventual practitioners cannot effectively identify, quantify, measure, or mitigate the biases, uncertainties, and risks inherent in AI systems, which can lead to discriminatory, unfair, or otherwise harmful outcomes. Consequently, statistical education must deeply integrate ethical considerations, demonstrating how statistical methods are not just analytical tools but also indispensable instruments for achieving and upholding ethical AI principles in practice.

---

[13] Johns Hopkins Center for Health Security. (2025, May 29). Response to Request for Information on the Development of a 2025 National Artificial Intelligence (AI) Research and Development.
[14] Johns Hopkins Center for Health Security. (2025, May 29). Response to Request for Information on the Development of a 2025 National Artificial Intelligence (AI) Research and Development.
[15] Johns Hopkins Center for Health Security. (2025, May 29). Response to Request for Information on the Development of a 2025 National Artificial Intelligence (AI) Research and Development.
[16] Generally, bias, in everyday life, is a preconceived idea or feeling that affects how you think or make decisions. In statistics, statistical bias is a consistent, built-in error in your data or method that causes your results to be wrong in the same direction every time. It's like having a scale that always reads two pounds heavier—it's consistently off from the real weight.
[17] Johns Hopkins Center for Health Security. (2025, May 29). Response to Request for Information on the Development of a 2025 National Artificial Intelligence (AI) Research and Development.
[18] Johns Hopkins Center for Health Security. (2025, May 29). Response to Request for Information on the Development of a 2025 National Artificial Intelligence (AI) Research and Development.

The ASA notes that "AI applications present issues distinct from standard statistical practice for statistical practitioners".[19] This critical observation implies that traditional statistical training, while foundational and necessary, may not be entirely sufficient on its own to address the unique complexities posed by modern AI. Statistical education for AI professionals must dynamically adapt to address these emerging complexities and distinct challenges. This means incorporating advanced topics related to large-scale data integrity, novel model assessment methods specifically designed for high-dimensional and complex models, and a deeper understanding of the direct societal impact of AI decisions. It suggests a need for a living, adaptive curriculum that responds to the rapid pace of AI innovation, ensuring that statisticians remain at the forefront of developing and deploying ethical, robust, and responsible AI systems. This evolution is crucial for maintaining the relevance and impact of statistical expertise in the AI era.

Promoting Statistical Literacy: A National Strategic Asset

Explicitly promoting statistical literacy among AI system developers and practitioners is identified as a fundamental policy action to directly improve model reliability and prevent common pitfalls.[20]

The consistent and explicit calls from the ASA, in its responses to various governmental RFIs (NIH, NSF, Office of Science and Technology Policy), for the integration of statistical rigor and the promotion of statistical literacy into national AI strategies signifies more than just an academic suggestion. This represents a strategic policy recommendation from a leading professional body. The specific advocacy for "standardized statistical benchmarks" and "incentiviz[ing] data integrity and model validation procedures"[21] demonstrates a clear understanding that statistical principles are not merely theoretical concepts but practical, actionable tools for effective governance, quality assurance, and responsible innovation in AI. This indicates that policy recommendations reflect a recognition of statistical literacy as a national strategic asset. For policymakers, investing in statistical literacy within AI education and workforce development is not merely an educational expenditure but a strategic investment in national AI competitiveness, security, and ethical leadership on the global stage. This implies that government funding, policy frameworks, and inter-agency collaborations should actively support programs, research initiatives, and infrastructure that strengthen the statistical foundations of AI education and practice, recognizing it as a critical component of the national infrastructure for a thriving and trustworthy AI ecosystem.

The ASA highlights the importance of "harness[ing] government and other data for AI methodological and substantive research" through initiatives like NAIRR and NSDS.[22] Crucially, this is coupled with the necessity for "robust statistical system data governance with strong ethical, methodological and legal frameworks".[23] The NIH response further reinforces this by linking "investments in data-sharing

---

[19] Johns Hopkins Center for Health Security. (2025, May 29). Response to Request for Information on the Development of a 2025 National Artificial Intelligence (AI) Research and Development.
[20] European Commission. (2019, April 8). *Ethics Guidelines for Trustworthy AI*
[21] American Public Human Services Association (APHSA). (2025, May 29). *Response to the 2025 National AI R&D Strategic Plan RFI*
[22] European Commission. (2019, April 8). *Ethics Guidelines for Trustworthy AI*
[23] European Commission. (2019, April 8). *Ethics Guidelines for Trustworthy AI*

infrastructure" directly with "investments in statistical methods for harmonization, uncertainty quantification, bias correction, and missing data handling".[24] This demonstrates a holistic view where data availability alone is insufficient without the statistical expertise to manage and interpret it. This illustrates the interconnectedness of data access, statistical rigor, and ethical AI development. Simply having vast datasets is insufficient; the data must be treated with statistical rigor to ensure its integrity, address inherent biases, quantify uncertainty, and enable valid and reliable inferences. This implies that AI education must extend beyond mere algorithmic training to encompass comprehensive topics in data governance, privacy, data quality assessment, and the broader ethical implications of data collection and use, all of which are fundamentally underpinned by statistical principles.

Key Table 2: ASA Recommendations for Integrating Statistical Literacy into AI Policy and Practice

Policy Area/Challenge	Key Recommendation	Specific Action/Rationale
AI R&D Strategy & Funding	Embed Statistical Rigor	Integrate into all phases of the data science pipeline; Emphasize biostatistics/statistics as crucial developers of ethical data science methods.
AI Model Performance & Evaluation	Standardize Statistical Benchmarks	Define metrics for fairness, accuracy, and uncertainty quantification to evaluate AI models reliably.
Data Governance & Integrity	Incentivize Data Integrity & Model Validation Documentation	Encourage thorough documentation and sharing of procedures, including evidence of privacy and confidentiality maintenance.

[24] American Statistical Association. (2025, March 14). *ASA Response to the OSTP Request for Information on the Development of an Artificial Intelligence Action Plan*. Retrieved from https://www.amstat.org/docs/default-source/amstat-documents/pol-ostp-ai-action-plan-comment.pdf

Ethical AI Development	Foster Interdisciplinary Partnerships	Establish collaborations among AI developers, statisticians, and policymakers to accelerate advancements in explainability, fairness, and risk assessment.
Workforce Development & Education	Promote Statistical Literacy	Ensure AI system developers and practitioners have fundamental statistical understanding to improve model reliability and prevent common pitfalls like overfitting and bias.

The ASA urges the Department of Education to consider these points and explicitly incorporate the fundamental role of statistical literacy into its proposed priority and definitions for Advancing AI in Education.