18 November 2021

National Institute of Standards and Technology (NIST)
U.S. Department of Commerce
100 Bureau Drive
Gaithersburg, MD 20899
Via Email: scientificfoundationreviews@nist.gov

Re: RFC Response: NIST Internal Report 8351-DRAFT *DNA Mixture Interpretation: A NIST Scientific Foundation Review*

IEEE-USA is pleased to submit these comments on the above-captioned, Request for Comment on NIST's DNA Mixture Interpretation: A NIST Scientific Foundation Review (8351-DRAFT, "the Review").

IEEE-USA represents approximately 150,000 engineers, scientists, and allied professionals in United States, many of whom are actively conducting research and development into artificial intelligence, software engineering, cybersecurity, and advanced computing, as well as other foundational and emerging technologies. We are the American component of the IEEE – the largest organization of technology professionals in the world, representing more than 400,000 engineers, scientists, and allied professionals worldwide.

The IEEE Standards Association (IEEE-SA), the leading developer of global technical standards used in power and energy, telecommunications, biomedical and healthcare, information technology, transportation, and information assurance products and services, is developing technical standards and frameworks that show how professionals can and should prioritize ethical considerations in the design, development, and deployment of artificially intelligent and autonomous systems (hereinafter referred to collectively as AI systems).[1] Of note, IEEE is developing IEEE P3119 Standard for the Procurement of Artificial Intelligence and Automated Decision Systems which aims to address the needs of government workers, policymakers, and technologists to make meaningful, accountable, and transparent choices regarding the socio-technical considerations and impact of AI products, services, and/or systems encountered by the public.[2]

[1] See, e.g., The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition. IEEE, 2019. https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/Autonomous-systems.html (IEEE Ethically Aligned Design); IEEE P7000 Series Standards and Projects addressing topics including transparency, data privacy, and algorithmic bias https://ethicsinaction.ieee.org/p7000/; IEEE Model Process for Addressing Ethical Concerns During System Design, IEEE Standard IEEE 7000-2021; IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being, IEEE Standard 7010-2020.

[2] IEEE SA Working Group, "Process Model and Requirements Aimed at AI Procurement in a New IEEE Standard," (20 October 2021) https://beyondstandards.ieee.org/process-model-and-requirements-aimed-at-ai-procurement-in-a-new-ieee-standard/.

IEEE-USA believes that we stand at an important juncture that pertains less to what new levels of efficiency AI systems can enable, and more to whether these technologies can become a force for good in ways that go beyond efficiency. We have a critical opportunity to use AI systems to help make society more equitable, inclusive, and just; make government operations more transparent and accountable; and encourage public participation and increase the public's trust in government. When used according to these objectives, AI systems can help reaffirm and protect our democratic values.

If, instead, we miss the opportunity to use these technologies to ensure protection of human values and trustworthiness, we risk reinforcing disparities in access to goods and services, discouraging public participation in civic life, and eroding the public's trust in government. Put another way: responsible development and use of AI systems to further safeguard human values and ensure trustworthiness is an approach that leads to a sustainable ecosystem of innovation. It is this type of approach that our society will trust and accept.

IEEE-USA believes that the software and hardware used to perform DNA mixture interpretation, including probabilistic genotyping systems (PGS) (hereinafter referred to collectively as DNA software), are automated decision-making systems that impact the life and liberty of individuals, and should be governed by the same rigorous standards and requirements as other automated decision-making systems such as AI systems.[3]

While there is no single test for determining whether a software system is an 'artificially intelligent' system, modern DNA software, including PGS, is inarguably complex scientific software that leverages the power of computing to automate portions of forensic DNA analytical and decision-making processes. The correct development, verification, validation, and use of DNA software requires specialized technical understanding of complex mathematical, statistical, and computing methods. Under the federally codified definition of AI, DNA software undoubtedly meets the definition of AI.[4] Additionally, DNA software, like many forensic technologies, is an engineered product incorporating scientific models and numeric methods. Too often, DNA software is narrowly seen only as a forensic technology governed by forensic analysts, when it is also software and hardware to be governed by software engineers and other related professionals. Many of those professionals are represented by IEEE-USA. In sum, IEEE-USA believes that general concerns, requirements, standards, and policies regarding AI systems should and do apply to modern DNA software.[5]

[3] M. Canellas, "Defending IEEE Software Standards in Federal Criminal Court." Computer, vol. 63, no. 6, pp. 14-23, 2021. doi: 10.1109/MC.2020.3038630.

[4] Section 238(g) of the John S. McCain National Defense Authorization Act for Fiscal Year 2019, Pub. L. No. 115-232, 132 Stat. 1636, 1695 (Aug. 13, 2018) (codified at 10 U.S.C. § 2358, note), defined AI to include the following:

1) Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.
2) An artificial system developed in computer software, physical hardware, or another context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.
3) An artificial system designed to think or act like a human, including cognitive architectures and neural networks.
4) A set of techniques, including machine learning, that is designed to approximate a cognitive task.
5) An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision-making, and acting.

[5] While this Comment focuses particularly on DNA software, the same general concerns, requirements, standards, and policies apply to other forensic technologies that serve as automated decision systems such as face recognition, image recognition, fingerprint identification, and predictive policing.

The most important line in NIST's current draft is KEY TAKEAWAY #4.3 which concludes that "Currently, there is not enough publicly available data to enable an external and independent assessment of the degree of reliability of DNA mixture interpretation practices, including the use of probabilistic genotyping software (PGS) systems." Because this is the reality for a process and for software used in decisions of whether to deprive people of their rights and liberties, is an indictment of the lack of trustworthiness for these systems and software. With this lack of ability to determine the reliability of DNA mixture interpretation practices including PGS, neither these systems, nor their results can be considered reliable or trustworthy. To remedy this issue, we draw upon our collective expertise in line with the goals of the RFC to provide the following recommendations:

1. **DNA mixture interpretation using DNA software should only be deemed reliable based on objective information gathered through independent verification and validation as determined by IEEE Standard 1012.**

The use of DNA software in criminal court can result in catastrophic failures through false imprisonment and the deprivation of people's rights. Scientists and engineers have long demanded that safety-critical software and hardware be the right systems built the right way. Therefore, DNA software should be independently verified and validated (IV&V) prior to deployment, or prior to informing decisions in the legal system, law enforcement, governance, and related compliance. Specifically, DNA software ought to be independently verified and validated in accordance with IEEE Standard 1012, IEEE Standard for System, Software, and Hardware Verification and Validation,[6] and be subject to recurring post-deployment audit, including with respect to their operators. IEEE-USA encourages NIST uphold these same requirements.

Sponsored by the IEEE Computer Society, IEEE Standard 1012 is a universally applicable and broadly accepted process for ensuring that the right product is correctly built for its intended use. It is used to verify and validate Department of Defense nuclear weapons systems and NASA manned space systems and critical space exploration probes, among many others.

IV&V are interrelated and complementary processes that build quality into any system. Verification is focused on a product, providing objective evidence for whether the product conforms to requirements, standards, and practices. Validation is focused on customers and stakeholders, providing evidence for whether a product is accurate and effective, solves the right problem, and satisfies the intended use and user needs in the operational environment. In short, verification ensures that a product is correctly built, while validation ensures that the right product is built.

In the context of DNA software, IV&V answers the following types of questions:[7] Is the model of DNA analysis used by the software the best available, coded as designed, and appropriate for the problem? Does DNA software systematically favor including defendants? How likely are false negatives and false positives? Would outside experts agree with the software's results at each stage of analysis?

[6] IEEE Standard for System, Software, and Hardware Verification and Validation, IEEE Standard 1012-2016, Sept. 2017 (hereinafter referred to as IEEE Standard 1012).

[7] N. Adams, R. Koppl, D. Krane, W. Thompson, and S. Zabell, "Letter to the editor — appropriate standards for verification and validation of probabilistic genotyping systems," J. Forensic Sci., vol. 63, no. 1, pp. 339–340, 2018. doi: 10.1111/1556-4029.13687.

To appropriately perform IV&V, IEEE Standard 1012 requires that each software and hardware component be assigned an integrity level that increases depending on the likelihood and consequences of a failure: negligible, marginal, critical (causing "major and permanent injury, partial loss of mission, major system damage, or major financial or social loss"), and catastrophic (causing "loss of human life, complete mission failure, loss of system security and safety, or extensive financial or social loss").[8] As the integrity level increases, so too does the intensity and rigor of the required IV&V tasks.

DNA software analysis tools, like all software, should undergo IV&V according to its integrity level as defined by IEEE Standard 1012. Because a thorough and public conversation is yet to take place, there is presently no consensus on such an integrity level. However, the likelihood of DNA software to cause wrongful convictions in the criminal justice system clearly constitutes catastrophic failure, and therefore should be held to the highest integrity level, the level where IV&V should be performed independently.

The IV&V process must be independent to avoid conflicts of interest that could lead to catastrophic failure. To this end, IEEE Standard 1012 requires technical, managerial, and financial IV&V when testing software and hardware where catastrophic consequences could occasionally occur and where critical consequences will probably occur. Moreover, letting developers certify their own software is a clear conflict of interest, and the IEEE/Association for Computing Machinery Code of Ethics for Software Engineers is clear about the obligation of developers to manage competing aims.[9] Full definitions of technical, managerial, and financial independence from IEEE Standard 1012 are below, but, in brief, the following must all be independent from the group that oversees the design and building of software: personnel, problem formulation, test and analysis tools for IV&V (technical), responsibility for IV&V (managerial), and control of the budget for IV&V (financial).[10]

Specifically, technical independence "[r]equires the IV&V effort to use personnel who are not involved in the development of the system or its elements. The IV&V effort should formulate its own understanding of the problem and how the proposed system is solving the problem."[11] "Technical independence means that the IV&V effort uses or develops its own set of test and analysis tools separate from the developer's tools."[12] And if sharing tools is necessary, "IV&V conducts qualification tests on tools to assure that the common tools do not contain errors that may mask errors in the system being analyzed and tested."[13] This independence requires the exclusion of parties with a stake in the outcome, which for forensic technologies includes forensic labs that, while not financially dependent on developers, have a shared interest in software's acceptance.

Managerial independence "[r]equires that the responsibility for the IV&V effort be vested in an organization separate from the development and program management organizations. Managerial independence also means that the IV&V effort independently selects the segments of the software, hardware, and system to analyze and test, chooses the IV&V techniques, defines the schedule of IV&V activities, and selects the specific technical issues and problems to act on."[14] The IV&V effort must be

[8] IEEE Standard 1012, p. 196.
[9] D. Gotterbarn, K. Miller, and S. Rogerson, "Computer society and ACM approve software engineering code of ethics," Computer, vol. 32, no. 10, pp. 84–88, 1999. doi: 10.1109/MC.1999.796142.
[10] IEEE Standard 1012, p. 198.
[11] IEEE Standard 1012, p. 198.
[12] IEEE Standard 1012, p. 198.
[13] IEEE Standard 1012, p. 198.
[14] IEEE Standard 1012, p. 198.

"allowed to submit to program management the IV&V results, anomalies, and findings without any restrictions (e.g., without requiring prior approval from the development group) or adverse pressures, direct or indirect, from the development group."[15]

Financial independence "[r]equires that control of the IV&V budget be vested in an organization independent of the development organization. This independence prevents situations where the IV&V effort cannot complete its analysis or test or deliver timely results because funds have been diverted or adverse financial pressures or influences have been exerted."[16]

It is clear from these definitions that peer-reviewed publications, while a priceless tool for scientific inquiry, are not a substitute, nor a valid approximation of IV&V when determining reliability or trustworthiness of a deployed system. Peer reviewed publications form the foundation of scientific advancement, but peer reviewers of scientific publications are not tasked with answering questions like "Should the DNA software be admissible in court? Is the DNA software fit for the evidence in this legal case?" Peer reviewers do not have access to the system itself and are not tasked with assessing its reliability. Peer reviewers are assessing whether a publication deserves the attention of the scientific community, whether the results described deserve the attention of other scientists. With respect to specific legal cases, any individual case could go well beyond the bounds of the published studies. For example, a case could involve more contributors, a smaller evidence sample, or a different version of the software than was examined in the peer-reviewed studies.

Moreover, as exemplified by the Review's Table 4.3, most peer-reviewed studies of probabilistic genotyping software are not independent, violating a fundamental tenet of IV&V, and making them insufficient to determine reliability. The peer-reviewed studies of TrueAllele or STRmix were almost exclusively authored by employees of the companies who developed the software, or laboratories who have at least an implied conflict of interest since the software they use needs to be viewed as reliable in order for it to be admissible under law. The lack of independent review raises serious concerns about the reliability of the studies themselves and was the chief criticism of DNA software in a report by the President's Council of Advisors on Science and Technology (hereinafter the "PCAST report").[17] The PCAST Report called for more testing that "should be performed by or should include independent research groups not connected with the developers of the methods and with no stake in the outcome."

To address the fact that peer-reviewed studies are no substitute for IV&V and the lack of independence in those peer-reviewed studies, IEEE-USA specifically recommends that:

- TAKEAWAY #4.1 be rewritten to state that the "The degree of reliability of a component or a system can be assessed using empirical data obtained through *technically, managerially, and financially independent verification and* validation studies, *and post-deployment audits*" (changes emphasized).
- The analysis of validation experiments in Chapter 4 includes managerial, technical, and financial independence as "influencing factors" in their analysis.

[15] IEEE Standard 1012, p. 198.

[16] IEEE Standard 1012, p. 198.

[17] President's Council of Advisors on Science and Technology, Report to the president Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, Washington DC, 2016. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf.

- NIST state that IEEE Standard 1012 is applicable to DNA mixture interpretation, especially when DNA software or hardware are used.
- NIST state that DNA software should undergo IV&V according to its integrity level as defined by IEEE Standard 1012.

2. **DNA software should be tested against and governed by standards adhering to principles of due process, openness, consensus, balance, and right of appeal.**

By definition, standards are "published documents that establish specifications and procedures designed to maximize the reliability of the materials, products, methods, and/or services people use every day."[18] They are the basis on which the safety and credibility of new products and markets are verified, making them fundamental to the modern economy.[19] Because standards have such a profound effect, standards-setting organizations (SSOs), like IEEE SA, have significant legal obligations regarding the standards they develop and the processes by which they craft those guidelines, including contract, intellectual property, and antitrust law.[20] Among the many U.S. Supreme Court opinions dealing with SSOs, there are two particularly relevant rules the organizations must abide by to avoid liability: fair processes and independence.[21]

As a result, the IEEE SA standards-development process follows a well-defined and documented path, from concept to completion, guided by a set of five basic principles and imperatives that ensure fairness and good practices during the development cycle.[22]

- Due process: having highly visible procedures for standards creation and following them.
- Openness: ensuring that all interested parties can participate and are not restricted to a particular type or category.
- Consensus: requiring a supermajority of a group to approve a draft standard (75% of the ballots must be returned, with 75% of them voting yes).
- Balance: ensuring that voting groups include all interested participants and avoid an overwhelming influence by any one party.
- Right of appeal: allowing anyone to appeal a standards development decision at any point, before or after approval.

IEEE-USA recommends that NIST distinguish between guidelines and standards, and evaluate whether the guidelines or standards from the Scientific Working Group on DNA Analysis Methods (SWGDAM), the International Society for Forensic Genetics (ISFG), the European Network of Forensic Science Institutes (ENFSI), the UK Forensic Science Regulator, the American National Standards Institute and AAFS Standards Board (ANSI/ASB), and the IEEE Standards Association adhere to the five principles of standard-development process and address the potential for conflicts of interest.

[18] "Developing standards." IEEE Standards Association. https://standards.ieee.org/develop/index.html (accessed Nov. 14, 2021).
[19] "Developing standards." IEEE Standards Association. https://standards.ieee.org/develop/index.html (accessed Nov. 14, 2021).
[20] A. Updegrove. "Laws, cases and regulations in the essential guide to standards." ConsortiumInfo, 2013. https://www.consortiuminfo.org/essentialguide/laws.php (accessed Nov. 14, 2021).
[21] A. Updegrove. "Laws, cases and regulations in the essential guide to standards." ConsortiumInfo, 2013. https://www.consortiuminfo.org/essentialguide/laws.php (accessed Nov. 14, 2021).
[22] "Developing standards." IEEE Standards Association. https://standards.ieee.org/develop/index.html (accessed Nov. 14, 2021).

3. **There must be standards and certifications for DNA software and their operators, and recurring benchmarking exercises and independent studies to ensure DNA software's effectiveness, competence, inclusiveness, accountability, and transparency in operation.**

The Review highlights a significant lack of guidance for testing and evaluating DNA software and its protocols[23] and the lack of publicly available data.[24] To address these deficiencies, IEEE-USA believes, and NIST should recommend, that governments should make the reports documenting the required IV&V and audits of their DNA software public. Furthermore, we believe that governments, including NIST, should encourage, develop, and update standards and certifications for DNA software and their operators, and fund recurring benchmarking exercises and independent studies to ensure their effectiveness, competence, inclusiveness, accountability, and transparency in operation. Specifically, we believe these standards, certifications, exercises, and studies should address:

- The requirements for informed trust by the general public in DNA software (see Recommendation #6 below) and the development of metrics that are immediately and easily accessible by experts and non-experts alike;
- The existence or absence of reliable and unbiased underlying scientific principles and methods in DNA software;
- The requirements for recurring testing and auditing of the operation of DNA software, including the operators, field conditions, testing data, environments, methodologies, and performance metrics;
- The requirements for publicly available documentation by developers and testers of DNA software, and of the use of DNA software in individual and aggregate cases and decisions;
- The requirements for certification or loss of certification of operators and DNA software, and for their validation for DNA software already in use;
- The requirements for individuals to be able to access, review, contest, and correct the data about them, to review and contest the decisions that affect them, and to request human review of such data and decisions;
- The requirements for operation in an ethical manner; and,
- The requirements for identifying and addressing vulnerabilities and threats to security, safety, and privacy such as spoofing, evasion attacks, transfer learning attacks, and data poisoning.

4. **Determining the reliability and trustworthiness of forensic technologies like DNA software requires evaluating them in their operational environments, their use in legal proceedings and how fit the technology was for those uses.**

IV&V is predicated on the value of testing technology in operational environments. No software or hardware is "generally" reliable -- any technology is only fit for certain purposes. Even technologies that are widely considered to be reliable have known failure modes. For example, cellular telephones are widely considered to be reliable but are not classified as "generally" reliable because they do not work effectively in tunnels or underground. Further, the desire for a technology to be classified as "generally" reliable rather than to consider its reliability in a particular case is misguided. A core premise of labeling

[23] KEY TAKEAWAY #A1.3 "Limited information has been provided in guidance documents, such as the FBI Quality Assurance Standards or the SWGDAM guidelines, regarding suggested or required studies to inform mixture interpretation protocols."

[24] KEY TAKEAWAY #4.3: "Currently, there is not enough publicly available data to enable an external and independent assessment of the degree of reliability of DNA mixture interpretation practices, including the use of probabilistic genotyping software (PGS) systems."

a product or process as “well-engineered” is that these operating conditions are specifically defined, tested against pre-defined standards, and accompanied with estimated rates of failure. Systems like DNA software are engineered products incorporating scientific models and therefore require not only the perspective of researchers who have published proofs-of-concept but also engineers who have used product trials and operational testing and evaluation to demonstrate system performance in operating conditions, against predefined standards, and estimated rates of failure.

Therefore, a scientific foundation review of the reliability and trustworthiness of forensic technologies cannot be effective if detached from an analysis of how the technology is used in legal proceedings, in the forensic technology’s operational environment -- yet that is exactly what this Review is purporting to do. The Review examines the peer-reviewed and forensic laboratory studies but does not compare that to any of the thousands of criminal cases where DNA software has been used. Notwithstanding the concerns over peer-reviewed studies discussed above, if the DNA profiles and contribution proportions analyzed in legal proceedings are not similar to the samples used in the peer-reviewed studies, the studies have little value. This is particularly true for many types of DNA software, especially PGS, because of its non-continuous nature, meaning that a small set of inputs cannot be reliably interpolated into cases involving different sets of inputs.

To determine the reliability of DNA software systems, IEEE-USA recommends NIST catalog and evaluate how DNA software is being used in legal proceedings and how fit the technology is for those uses.

5. **Users of DNA mixture interpretation include far more than forensic scientists, attorneys, judges, and juries. They include the public upon whom these systems are used, litigants, academics, journalists, and other researchers. For those users to assess the degree of reliability, validity, and whether that information is fit-for-purpose, they need appropriate access to the software.**

IEEE-USA believes that users are too often inappropriately denied access or forced to overcome improper and unnecessary barriers to access DNA software in order to determine the degree of reliability, validity, and whether that information is fit-for purpose. It is true that providers and users have responsibilities as described in KEY TAKEAWAY #4.2 and Section 4.1.5 but there are many more users of DNA mixture interpretation than merely forensic scientists, judges, or juries. Independent testing of proprietary or government DNA software by litigants, academics, journalists, and other researchers is needed to ensure that DNA software are properly vetted and held accountable. NIST should recommend governments clarify whether and how proprietary DNA software may be reverse engineered, modified, and evaluated under laws such as the Computer Fraud and Abuse Act and the anti-circumvention provision of the Digital Millennium Copyright Act, and rules of procedure and evidence. More broadly, NIST should recommend governments take steps to affirmatively promote awareness, access, research, and testing including:

- Ensuring accountability and transparency in government procurement and contracting for DNA software;
- Identifying and disclosing the DNA software used by the government;
- Adopting clear procedures relating to collection, usage, storage and sharing of personal information used by DNA software;

- Providing constituents notice about DNA software decisions, explanations for those decisions, and processes for challenging decisions or data; and,
- Specifically, in legal disputes, tribunals should permit disclosure under appropriate protective orders of intellectual property related to DNA software when necessary to obtain evidence in compliance with other judicial requirements, including constitutional requirements, discovery laws, or subpoenas.

6. **Trustworthiness is determined by more than reliability, and therefore, to determine trustworthiness, one must assess the processes and procedures where these systems are deployed.**

Technical assessments of reliability as surveyed in the Review are not the sole determination of trustworthiness. There are eight principles for creating and operating AI systems that further human values and ensure trustworthiness:[25] (i) human rights: AI systems shall be created and operated to respect, promote, and protect internationally recognized human rights; (ii) well-being: AI system creators shall adopt increased human well-being as a primary success criterion for development; (iii) data agency: AI system creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity; (iv) effectiveness: AI system creators and operators shall provide evidence of the effectiveness and fitness for the purpose of AI systems; (v) transparency: the basis of a particular AI system decision should always be discoverable; (vi) accountability: AI systems shall be created and operated to provide an unambiguous rationale for all decisions made; (vii) awareness of misuse: AI system creators shall guard against all potential misuses and risks of AI systems in operation; and, (viii) competence: AI system creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.

Therefore, IEEE-USA recommends that NIST evaluate more than the technical assessments of reliability to determine trustworthiness. Below we list additional requirements for ensuring the trustworthiness of AI systems in general which includes the automated decision systems such as DNA software and many of the forensic technologies used today. These requirements should be used as factors of analysis in Chapter 4 when evaluating the various forensic laboratories and when developing an overall assessment for forensic technologies as described in Chapter 6. While Chapter 6 recommends assessing (i) how a new technology works, (ii) what its limitations are, and (iii) how it might specifically address the problem to be solved, these three factors are insufficient to ensure trustworthy use of AI systems, DNA software or any forensic technology. If those providing or using DNA software or any other forensic technologies do not adhere to these requirements, then they should not be deemed trustworthy or fit for their use in determining or affecting people's rights and liberties.

To ensure trustworthiness of AI systems, DNA software, and other forensic technologies, IEEE-USA believes that governments and forensic laboratories should be required to:

a. Ensure awareness, access, and research on the existence, fairness, safety, security, privacy, and ethical and societal impacts of DNA software.

[25] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition. IEEE, 2019. https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/Autonomous-systems.html

Governments should: (i) publicly identify and disclose the DNA software used by the government; (ii) conduct and publicly disclose a methodological validation study that establishes the value of using new DNA software in place of existing practices prior to deploying DNA software; (iii) adopt clear procedures relating to the collection, usage, storage, and sharing of personal information in the context of developing, using, and validating a given DNA software in a privacy-preserving manner; and (iv) prevent intellectual property, confidentiality claims, lack of funding, or lack of an designated independent body within government to monitor compliance from impeding duly limited independent validation and verification and publicly disclosed review of the fairness, safety, security, privacy, and ethical and societal impacts of DNA software. DNA software ought to be submitted voluntarily, to the agency performing validation and verification thereof, and the agency using related private intellectual property or proprietary data in its evaluation must adopt rules to protect such private rights from misappropriation.

Users and the public should be allowed to (i) request and receive an explanation of how a government determination using DNA software was reached; (ii) determine whether the DNA software used in government decision-making disproportionately impacts a protected class; and (iii) rectify, challenge, or complete inaccurate or incomplete personal data that is part of the DNA software system or decision.

b. Commit to removing barriers to parties' access to information needed to ascertain relevant evidence about and from DNA software in legal disputes.

Specifically, in legal disputes where judges, juries, and lawyers are the users of DNA software results, barriers to parties' access to information needed to ascertain relevant evidence about and from DNA software should be eliminated.[26] Intellectual property protections should not be used as a shield to prevent duly limited disclosure of information needed to ascertain whether DNA software meets acceptable standards of effectiveness, fairness, and safety. Specifically, in legal disputes, tribunals should permit disclosure under appropriate protective orders of intellectual property related to DNA software necessary to obtain evidence in compliance with other judicial requirements, including constitutional requirements, discovery laws, or subpoenas. Furthermore, laws, procedures, and public funding should not make it more difficult for non-government parties in legal disputes to develop, obtain expertise regarding, or gain access to evidence from DNA software than for government parties to do so.

c. Ensure accountability and transparency in procurement and contracting for DNA software.

To support awareness, access, and research on the existence, fairness, safety, security, privacy, and ethical and societal impacts of DNA software, there must be accountability and transparency in

[26] For example, when source code is ordered to be provided, "information needed" requires providing sufficient information for the recipient to build, run, and test the software themselves including, at minimum:

- All software dependencies including third-party code libraries, toolboxes, plugins, frameworks, and databases;
- Software engineering and development materials describing the development, deployment, and maintenance of the version(s) of the software system used in the instant case, including software engineering documents and build instructions;
- All records of software glitches, crashes, bugs, or errors encountered during the developmental validation study;
- Software version numbers of the components of the system used for the developmental validation study; and,
- All records of unexpected results, including false inclusions, false exclusions and the conditions under which the unexpected results were achieved.

When source code is ordered to be provided, "access" requires, at minimum, that the source code be made available for inspection, in a format allowing it to be reasonably reviewed, searched, and tested, during normal business hours or at other mutually agreeable and reasonable times, and at mutually agreeable and reasonable locations.

government procurement and contracting for DNA software. The government should not procure DNA software that (i) require the governmental entity to indemnify vendors for any and all negative outcomes; (ii) do not adhere to the eight principles in IEEE's Ethically Aligned Design for creating and operating DNA software that further human values and ensure trustworthiness (as may be reflected in articulated guidelines, standards, certifications, audits, and other sound documentation);[27] (iii) do not comply with federal, state, and local anti-discrimination laws; or, (iv) are shielded from independent validation and verification, and public review.

7. **To ensure DNA software is reliable and trustworthy, governments should provide sufficient funding for testing, evaluation, certification, and investigation of DNA software.**

Throughout the document, the Review highlights the value of government funding in the development of research on DNA software (e.g., Sections 3.1.5, A2.3, and A2.3.2). IEEE-USA believes NIST should go further by including a KEY TAKEAWAY recommending that governments provide sufficient funding for testing, evaluation, certification, and investigation of DNA software. The adoption and acceptance of DNA software requires developing and sustaining public confidence in their quality, reliability, and compliance with regulations and social norms. Increased government funding for government and independent third-party evaluation and certification of DNA software is essential to ensure efficacy, transparency, traceability, accountability, and competency. Development of design requirements, methods, metrics, and environments so that DNA software can be tested and evaluated for interactions with different autonomous agents, including humans, and adversarial exploitation is critical in the adoption and acceptance of DNA software. To this end, mechanisms must be developed for identifying and accounting for the features of DNA software that could cause current testing, evaluation, certification, and investigation methods to misinform decision makers or the public about the risk of system deployment or the causes of system malfunction.

IEEE-USA thanks NIST for considering these comments in the agency's revisions to the Request for Comment on NIST's DNA Mixture Interpretation: A NIST Scientific Foundation Review. We would welcome any further discussions with the agency on these matters. If you have questions, please do not hesitate to contact Erica Wissolik at (202) 530-8347 or e.wissolik@ieee.org.

Sincerely,

Will Robinson

William Robinson
IEEE-USA Vice President, Government Relations

[27] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition. IEEE, 2019. https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/Autonomous-systems.html