Alliance for AI in Healthcare (AAIH)
1340 Smith Ave, Suite 400
Baltimore, MD 21209
Oct 21st 2019

Subject: AAIH Comments on "Request for Information on the Bioeconomy"

Dear Office of Science and Technology Policy (OSTP)

The **Alliance for Artificial Intelligence in Healthcare (AAIH)** welcomes the opportunity to comment on OSTP's RFI: "**Request for Information on the Bioeconomy" (Document Citation: 84 FR 47561, Document No: 2019-19470).** Below we provide a brief introduction of who we are followed by our responses.

The AAIH is an **international 501(c)4 not-for-profit multi-stakeholder advocacy organization,** launched in 2019 to promote **scientific, legislative, and regulatory initiatives** necessary for the development, accessibility, and implementation of Artificial Intelligence (AI) powered healthcare solutions. The AAIH mission is to enable the **advancement and responsible use of AI** in healthcare to improve patients' lives by focusing on resolution of the sector's most **pressing and significant issues** for patients and developers. We achieve this objective by working collectively to identify and address **industry wide concerns.** AAIH is comprised of **over 25 organizations**[1] that develop or utilize **AI in biomedical R&D and clinical applications** who have recognized the need for **collaborative engagement** of the industry (and government) to realize common goals for the growth and adoption of the AI health sector and to significantly **improve quality of care**. Our members include leading-edge start-ups, biopharma, diagnostics and device manufacturers, academia, and technology and infrastructure developers.

The RFI encourages alliances to respond. We echo and fulfill this requirement for diverse input through our primary constituency of **industry and academia** member organizations. As an organization, the AAIH is a staunch supporter of **Public Private Partnership** frameworks. We cooperate with stakeholders from government, NGOs, key opinion leaders, and other international organizations to support **innovation and thoughtful application of AI.** By convening diverse stakeholders with a wide range of relevant expertise and viewpoints, we are working to **establish responsible, ethical, and reasonable principles** for the **development and implementation of AI** in healthcare. As an example, we are convening an open workshop on November 19, 2019, in Washington, DC to explore how AI will impact **patient-centered end-to-end drug development**, incorporating **regulatory, economic, ethical**, and **infrastructure considerations.** We look forward to working with our partners on such efforts and are eager to leverage our expertise to enable the **mission of fellow public and private stakeholders**. It is in this spirt and with this commitment to **collaborative innovation** that we submit our comments today starting with our overarching position followed by specific responses to your questions.

Given that our organizational position is that **AI will play a key role in advancing the bioeconomy**, the questions you posed represent issues of keen interest to us. In fact, the thematic areas you cover are echoed in the mandates of our main standing committees: **Education and Accreditation, Federal Engagement and Regulatory Affairs, and Technology & Standards Development**. We believe that especially in the **highly regulated arena of healthcare**, the establishment, utilization, and concerted improvement of standards, scientific discovery and technological advances, and a well-versed viable and sustainable workforce in the AI space requires a multipronged and coordinated approach across the healthcare continuum. This tenet is the basis of our philosophy around a **member-driven approach** to **Public Private Partnership**, which will be critical to informing governmental efforts on bolstering the U.S. Bioeconomy.

[1] https://www.theaaih.org/members

We also realize the linkages and implications of **international policy** in **normalizing AI standards** among emerging governmental strategies. To that end, led by our **Federal Engagement and Regulatory Affairs Committee**, we recently hosted an interactive workshop with industry, academia, and Federal Agency perspectives (NIH, DOE, NCI, FDA, etc.) around the need and potential areas for **Public Private Partnership**, and are continuing to engage our member organizations, collaborators, and stakeholders in other **jurisdictions to foster cross-border harmonization**.

Please see our comments on the RFI around governmental efforts and priorities in advancing the U.S. Bioeconomy in the following pages with a **focus on AI in Healthcare.** We look forward to continued engagement and collaboration with the Office of Science and Technology Policy at this critical juncture in the development of American leadership in AI.

Respectfully submitted *on behalf of AAIH members*,
Annastasiah Mudiwa Mhaka, PhD (Co-Founder and Convenor, AAIH)
Aaron Chang (Strategic and Technical Advisor, AAIH)

RESPONSES TO SPECIFIC QUESTIONS/COMMENTS

These comments represent focus areas led by 3 of our standing committees Education and Accreditation Committee, Federal Engagement and Regulatory Affairs Committee, Technology & Standards Development Committee

1. What specific actions could the U.S. Government take to reinforce a values-based ecosystem that will guide the transformation and expansion of the U.S. Bioeconomy, in both the short- and long-term?

There are several specific actions the US Government could take to foster a values-based environment to support the U.S. Bioeconomy. Several suggestions, led by our **Federal Engagement and Regulatory Affairs Committee** follow:

a. Policy or regulatory opportunities and gaps throughout the continuum of basic science translation, product development and commercialization;

Several upcoming regulatory and policy opportunities exist across the continuum of healthcare AI. From the FDA perspective, developing regulations around Software as a Medical Device and how they relate to various clinical processes, including Clinical Decision Support guidelines, remain an area that will affect AI powered systems particularly at the **point of clinical care**. The FDA has continued to request and incorporate public input into their guidance documents.

AAIH member organizations see an opportunity to build on the momentum of the FDA's current device related advances by expanding scope to also include **AI enabled Investigational New Drug (IND) and New Drug Application (NDA) applications**. To this end, on November 19, 2019 AAIH is convening a workshop open to public and private sector stakeholders to explore how AI will impact **patient-centered end-to-end drug development**, incorporating **regulatory, economic, ethical**, and **infrastructure considerations.** We envision this initiative leading to the establishment of an **engaged, transparent, and forward-looking** body that continues to advance our collective thinking on how AI may meaningfully impact the continuum of healthcare from biomedical discovery to commercialization -- to lower costs, accelerate innovations and to ultimately benefit patients and other key stakeholders alike.

Concerning **data and intelligence collaborations**, there is a need to establish how we standardize research, clinical and social data in order to make it **AI-ready** and to assist in multi-modal based decision-making in R&D as well as disease prevention, treatment, and management. We must have common and basic metrics ("standards") for generating and sharing data along with formalized outcome measures against which the community can test and validate their algorithms and tools. Further, what one might consider general terms with universal meaning might connote divergent meanings to different stakeholders. A good example is in how "privacy" and "ownership" are defined by patients, healthcare providers, and other parties, thus requiring us to develop and drive common contextual frameworks. AAIH is starting to socialize its preliminary road-map around vocabulary and standards development in order to establish rich collaborations while avoiding duplication of work. We have just released an ***AI in Health Primer*** to begin to set a common grounding around our future work and communications.

With the increase in use of AI driven systems, **balancing appropriate precautions** against **misuse of individuals' confidential information** and **data sharing measures** through **interoperability** remains a crucial area in which broad implementation of optimal solutions have not taken hold. The U.S. government can promote initiatives which **expand access** to the records of federally supported health care programs in ways which still **comply with patient privacy regulations**. The resultant effect would lie in better, more accurate, and more complete identification of interventions that are both clinically superior and cost-effective, concomitant with the ability to distinguish them from other less effective or unacceptably risky

interventions. Cross referencing these data with health outcomes and claims data provides potential for identifying areas for healthcare cost savings in a value-based care framework.

b. Scientific areas where research funding could be strategically targeted to stimulate discovery;

Several scientific areas in healthcare AI and their associated data models can be used in tandem to strategically **target research funding** in order to stimulate discovery that fits well with value-based care frameworks.

Areas in Healthcare:

Investment in AI can lower drug development costs and reduce failure rates by increasing the number and quality of available targets, designing and testing fewer molecules that are more effective and less toxic, and selecting the right patients at the right time for the right treatment in clinical trials.

In clinical care, strategic investment in AI can aid in health care decision-making that augments disease prevention, diagnosis, and treatment while improving patient follow-up. AI powered approaches can permit early, accurate diagnoses by aggregating disparate pieces of information, extracting key data patterns, and identifying effective interventions early.

Relevant Data Models to Inform Investment:

These models include R&D, public health models, root cause analysis models, and high-quality models to be utilized in NIH and FDA's initiatives.

Models for Research and Drug Development: Models that use AI to replace or advance some aspects of R&D processes in a regulatory compliant manner would most quickly facilitate drug development. Examples, among others, include therapeutic window algorithms, augmented translational animal models during the preclinical phase, and human adverse event toxicology prediction models.

Public Health Models: The US spent $3.5T last year on healthcare and we are projected to reach $5T by 2025. AI based approaches to healthcare data can support and enhance medical discovery and optimize clinical workflows to prevent our healthcare system from financially collapsing.

Root Cause Analysis Models: Models that can uncover root causes of disease are very useful. Probabilistic programming and explainable models could be of most benefit to this domain. Federal Agencies typically are not required to employ these types of models that utilize explainable AI, but the public and wider research community would benefit if they did.

Quality-Assured Models in Research: Flawed models are prolific in academic research as there is often a rush to publish. The FDA and the NIH should remain vigilant of these issues when respectively approving devices and drugs and awarding grants to academic researchers in the field. Improved approval and grant awarding processes could help reward fundamental best-practices and better training for subsequent generations of trainees who will be developing commercial AI products in the near future.

c. Novel public-private partnership mechanisms;

The U.S. Government has initiated several agency led public private mechanisms that support the U.S. Bioeconomy through AI in healthcare, such as the new CMS Artificial Intelligence Outcomes Challenge.[2] In addition, to the Department of Health and Human Services (HHS) agencies, other federal departments such as the Department of Energy and the Department of Veterans Affairs have entered into innovative public private partnerships utilizing AI in health. For example, some of AAIH's board members were founding members of the Accelerating Therapeutics for Opportunities in Medicine (ATOM) consortium, a public-private partnership with the mission of transforming drug discovery by accelerating the development of more effective therapies for patients. The ATOM partnership was founded in 2017 by GlaxoSmithKline, Lawrence Livermore National Laboratory, Frederick National Laboratory for Cancer Research, and the University of California, San Francisco.

Tasking and funding more federal agencies and centers with clear areas of responsibility to run open innovation challenges that supply AI-ready federal data for creative use can lead to novel, cost effective models in advancing the Bioeconomy. Furthermore, other incentives such as offering tax breaks or extension of exclusivity terms through Intellectual Property Rights to winners of such challenges, in addition to prize money, may encourage wider participation in such efforts from the private sector.

d. International opportunities;

Many opportunities exist for international collaboration in order to bolster the global bioeconomy. In order for the U.S. to retain the global leadership role in this effort, **elevation of standards for practice and standards for collaboration** stand as key opportunities for distinction.

Key leadership personnel in both domestic and international regulatory frameworks is key. One example of such leadership is Bakul Patel, the FDA Director of the Division of Digital Health, who also serves as the committee chair on Software as a Medical Device for the International Medical Device Regulators Forum. Increasing the number of regulatory officials who lead on international bodies stands as one strategic area from which to influence and lead global bioeconomy efforts with American values.

Increasing shared clinical data with selected countries that truly meet Good Clinical Practice Guidelines will both encourage elevation of global standards in ethical clinical practice and data use while also enabling international efforts in **reducing the bias** that occurs when clinical interventions are only developed using certain global populations and demographics.

We believe that a global perspective is required in order to ensure that the wave of AI that is spreading across healthcare leads to **equitable and sustainable development of technologies** that advance the global bioeconomy. We are keen to promote **cross-border harmonization** with like-minded jurisdictions in order to drive solutions for global benefit. The AAIH is in the process of engaging U.S., EU, and Canada to better understand key needs and opportunities for advancement in the AI Health sector.

e. Challenges to taking identified actions or implementing change.

Several challenges remain in light of the above-mentioned potential opportunities and positive initiatives.

[2] CMS AI Health Outcomes Challenge. https://innovation.cms.gov/initiatives/artificial-intelligence-health-outcomes-challenge/

The overarching challenge is that industry, academia and government may be differently incentivized thereby impacting much needed fruitful collaboration around gaps and how best to address them. We envision that **bringing together our diverse capabilities** across industry and academia, and working together with government, we are better positioned to surmount this challenge by creating a **jointly informed roadmap for the role of AI in healthcare.**

Domestically, interoperability standards still have quite a way to go. The ONC's collaboration with HIMSS on FHIR based interoperability seeks to address initially identified issues as a result of several RFI's. However, challenges in widespread health system adoption from both an implementation and incentivization perspective remain.

Internationally, countries such as China have the population and industries in need of clinical studies, but their infrastructure still lacks transparency and protections sufficient for trial subjects. Furthermore, the intellectual property protection measures necessary to encourage broader U.S. and international bioeconomy investments still need further refinement.

2. In what ways can the U.S. Government partner with the private sector, industry, professional organizations, and academia to ensure the training and continued development of a skilled workforce to support the growth of the Bioeconomy? Please consider:

a. Potential needs and solutions at the skilled technical, undergraduate, professional master's program or graduate level;

The traditional challenge of integrating the role of industry in a traditional STEM education is defining the specific skill set that students need to fill jobs in the modern economy. Universities are often primarily equipped to prepare students for a career in research. Some of the skills students acquire are applicable to both industrial and academic careers, but there may be some mismatch.

To better prepare students for an AI health workforce, it is crucial that **industry and government intervene early on in the education process** (preferably K-12) with **continued reinforcement** through skilled technical, undergraduate, graduate and professional levels. Given Health AI's nature as a convergence of computer science and biomedicine, whose general concepts and implications are increasingly taught at an earlier level, we believe that it is possible to design educational initiatives that start at lower educational levels. The more elaborate components starting at the undergraduate and perhaps vocational level require that academic programs proactively interface with industry in order to assure that students are 1) armed with competencies matching the needs of employers in the rapidly accelerating healthcare AI space and 2) equipped with the knowledge to produce technologies which effectively address underserved needs and carry higher chances of commercialization.

Following is an example list of some of the **key attributes that students should acquire** for successful industrial careers and are increasingly critical for emerging AI Health sector:

- The ability to work in interdisciplinary teams
- The ability to quickly tackle new problems and develop plans to solve them
- The ability to communicate with different levels of management and technical personnel
- "Self-starting" ability; that is, the ability to take independent initiative
- The ability to stick to a defined timeline and budget
- The ability to stay current with the latest techniques, tools, and research in the field
- Literacy in scientific computing and software
- Literacy in policy and regulatory frameworks and how these impact AI health advances

- Strong oral presentation and writing skills
- The vision to balance between long-term strategic thinking and short-term goals
- Appreciation for general business and commercialization concepts

Approaches that integrate internship or co-op programs, often involving university alumni, into degree programs are one potential solution to this challenge. However, these traditional internship programs which match students with a single employer are limited in scale and only expose students to one experience from which they may base their future impression of career paths in a particular industry. To avoid the chance that one bad experience can turn a student off to the field, graduate education frameworks incorporating industry consortiums and alliances stand in a unique position to create workforce development programs with high industry retention rates. Furthermore, while scalable, industry consortiums can also play a role in elevating the translational caliber of academic centers by fostering more interactions between academic faculty and industry representatives. In fact, findings from industry consortium group AdvaMed found that one of the keys to academic environments producing a high level of translational projects occurred in the context of increased informal industry academia relationships and points of contact.[3]

On these bases, AAIH has formulated and submitted a framework for grant review to conduct a pilot program to support a long-term, scalable program around Innovative Graduate Education (IGE) that incorporates critical components of an effective education program. Member and partner universities will work with AAIH member companies to place graduate students for a 3-6 month period to better prepare students for the workforce. Our framework seeks to address the dual-training needs outlined above as well as reinforce our mission elements around educating students on policy, regulatory, business and financing frameworks that impact AI Health R&D and Commercialization.

b. Specific needs within basic science, translational research, product development, and commercialization;

Please ***see 1c*** above as it relates to US government opportunities to address key needs and ***see 2a*** above as it relates to needs around translational education.

c. Approaches for the development of non-traditional, multi-disciplinary educational backgrounds that address the convergent nature of emerging technologies and integrate core values including safety and security;

We recommend evaluating curricula such as the physician-scientist-in-residence program at the Virginia Commonwealth University School of the Arts,[4] the Medical Innovators Development Program at Vanderbilt University,[5] and the Bioengineering Innovation and Design program from the Johns Hopkins Department of Biomedical Engineering,[6] for concepts and approaches to nurture cross-disciplinary approaches to innovation. Another example of a program to evaluate is the Stanford Biodesign Innovation Fellowship, which invites applicants with advanced degrees and/or substantial work experience in the engineering, science, computer science, business, product design, law, medical, or nursing fields.[7]

Lastly, the components described in ***section 2a*** relating to the AAIH Graduate Fellow program will be scaled if piloted successfully.

[3] University Technology Transfer Best Practices Guide. Advamed Accel and WSGR. 2018.
[4] https://news.vcu.edu/article/VCU_schools_of_the_Arts_and_Medicine_launch_physicianscientistinresidence
[5] https://medschool.vanderbilt.edu/midp/midp-curriculum/
[6] https://cbid.bme.jhu.edu/
[7] http://biodesign.stanford.edu/programs/fellowships/innovation-fellowships.html

d. Effective geographic distribution of workforce and talent development across the United States;

Increasing emphasis and investment on the 'flyover states' and regions of socioeconomic distress will ensure that the U.S. leads in finding untapped talent within its own borders. Initiatives similar to those pioneered by NEOMED in expanding access to medical education provide case studies in ways to nurture talent in disadvantaged communities.[8]

e. The development of a public and private ecosystem that will attract and retain domestic and foreign talent within the United States at all skill levels.

To encourage retention of talent, the U.S. Government should consider revising the current system of regulating immigration to enable substantially larger numbers of highly educated, highly skilled foreigners to immigrate to the U.S. and become citizens with only the degree of difficulty required by national security considerations. Given the legacy of immigrant inventors and scientific pioneers in this nation, it behooves the government to recognize that individuals such as these will enhance, not retard, American leadership in technology and contribute to a prosperous Bioeconomy.

3. In what ways can the U.S. Government partner with the private sector, industry, professional organizations, and academia to establish a more robust and efficient Bioeconomy infrastructure? Please consider:

a. Current infrastructure—from databases to world-class technology and manufacturing capabilities;

Data is the lifeblood for AI applications, and the old adage of garbage in garbage out lies at the core of why **well-curated, annotated, and featurized data is essential for quality AI R&D**. Several aspects of ideal datasets and models identified by the private sector, which the U.S. Government should support in the healthcare AI space, include the following:

- Large datasets
  - The larger, more inclusive, and evenly sampled the training set, the higher quality the model. A better training set will ensure that the model, when evaluated on the test set is interpolating and not extrapolating (i.e. the training set should contain examples similar in feature space to the test set).
- Multiple dimensions or measurements
  - Test set data selection methodology and data redundancy. Randomly generated test sets lead to overestimated practical utility when training data has many related entities. For example, two medical images generated for the same patient could be split between the training and testing sets in a manner that provides a prediction in the evaluation criteria that is not representative of real-world data.
- Large overlap of measurements between samples
- Well populated metadata, even in free text
  - The data need to be released with extensive metadata, such as; a data format with a clear explanation of the data fields and clear explanations of how the data were generated and why.
- Clear data access rules and restrictions
- Clear and appropriate consents for use

[8] https://www.cbsnews.com/news/auto-doctor-fulfills-childhood-dream-and-becomes-a-medical-doctor-in-his-40s/

- Proper data blinding for privacy or other concerns
- Model Evaluation Metrics
    - A critical review of the experimentation process that was put in place to evaluate the models, including feature space analysis would help attenuate these issues. Datasets with low clustering of feature space are ideal in this practice.
    - Agencies should consider developing standardized test sets, e.g., molecules that are known to FDA to have a given effect on humans.

b. Geographic distribution of manufacturing capabilities compared to future manufacturing needs;

We are not in a position to address this.

c. Leveraging existing public-private partnerships and identifying trusted information sharing mechanisms to accelerate innovation and facilitate fruitful, equitable domestic and international collaborations;

Please ***see 1c.***

d. Institutional models for achieving translation of basic science discoveries to application and/or entry into the marketplace.

Please ***see 2a***. Many basic science discoveries that fail to commercialize originate in the academic sector. Translational educational programs play a role in shifting this institutional culture to encourage more technology transfer to the marketplace.

4. Across the spectrum, from basic discovery to practical application, what data policies, information-sharing mechanisms, and safeguards will be necessary for a prosperous U.S. Bioeconomy? Please consider:

a. Scientific, regulatory, manufacturing standards and/or benchmarks and/or best practices around data that should be developed to best accelerate Bioeconomy growth;

In general, the U.S. should follow **open and existing standards for data formats, metadata, and content**. This way, existing methods can be **reused** and be **quickly integrated** into existing business processes. By combining this with baseline benchmarks you immediately make available an example implementation and a way to compare various datasets, standards, and methodologies. Without standardized benchmarks you will not be able to properly compare various methods and datasets. **These standardized frameworks will allow the Federal government to more readily collaborate with the private industry - a key aspect of stimulating informed growth in the U.S. Bioeconomy.**

More specifically, AAIH members see use in several types of federal data including certain Center for Medicare & Medicaid Services (CMS) claims data, FDA data, NIH data, multimodal data, and biological datasets. Without **available and curated proprietary and public datasets**, the efficiency of AI applications in areas including drug design and clinical care including telehealth is limited.

Greater **accessibility and interoperability** of currently disparate patient trajectory data (ICD codes, CPT codes, notes, lab results, costs, etc.) would accelerate the development of predictive AI models for patient outcomes, support AI-based decision systems, manage costs, help triage patients, etc. CMS is the agency that has the data, and several of the FHIR-based interoperability initiatives from ONC and HHS are a step in the right direction. Separately, claims history at CMS for Medicare reimbursement for telehealth

services would be invaluable in assessing whether, as the Congressional Budget Office claims, expanded reimbursement for telemedicine will break the bank or whether, instead, reimbursement will be cost-neutral or even cost-saving.

The FDA has data in the FDA Adverse Event Reporting System (FAERS) database, but it is not optimally formatted and curated to the point of being barely usable. This database needs to be improved and the usefulness of this data for modeling and other data science uses needs to be emphasized during the reformatting process. The FDA has also gathered an enormous amount of data from numerous clinical trials, including why they were terminated, and observed adverse drug reactions (ADRs). This should be made readily available.

In general, the Gene Expression Omnibus (GEO) database serves as a good repository for RNA sequencing and microarray datasets post publication, but GEO is not widely used across NIH. Furthermore, validation datasets (e.g., Western Blots that target specific genes and proteins) are not publicly available. Validation datasets are often reported through tables and graphs in papers, but those datasets are also not publicly available. This practice undermines the ability of outside researchers to validate this research. Access to the raw data enables reproducibility.

b. Possible safeguards for technology, data, and emergent products, such as patent/intellectual property protection, data quality and provenance validation, and privacy and security assurances.

Some possible safeguards and approaches include 1) automatic checks on personally identifiable information to ensure patient privacy when creating and sharing datasets and 2) creation of sensitive datasets that can incorporate tight access controls to ensure data cannot be shared widely. These datasets should follow the same standards as public datasets but can be placed behind secure APIs to prevent them from being distributed too widely. This same system could possibly be used to ensure that companies receive compensation when their datasets are being used, which could incentive private sector participation in data sharing and knowledge increase.