April 2, 2018

National Institutes of Health

9000 Rockville Pike

Bethesda, Maryland 20892

RE: Response submitted electronically at http://grants.nih.gov/grants/rfi/rfi.cfm?ID=73

Dear Dr. Francis S. Collins,

The Alliance for Nursing Informatics (ANI) advances nursing informatics leadership, practice, education, policy and research through a unified voice of nursing informatics organizations. ANI has reviewed the Request for Information (RFI) for the National Institutes of Health (NIH) *Strategic Plan for Data Science*. ANI supports the NIH's overarching goals, strategic objectives, and implementation tactics for promoting the modernization of the NIH-funded biomedical data science ecosystem. ANI commends NIH's planned engagement of industry partners who have expertise in key areas related to information technology to complement the research strengths of NIH and the academic community and facilitate efficiencies so that investigators time can focus on data driven knowledge development, rather than primarily on data management.

We offer our comments as nursing stakeholders aligned with the information requested for the following topics: The appropriateness of the goals; Opportunities for partnership; Additional concepts that should be included; Performance measures and milestones.

Topic 1: The appropriateness of the goals of the plan and of the strategies and implementation tactics proposed to achieve them

We find the goals of the Strategic Plan for Data Science to be highly appropriate to support the continuing generation and analysis of biomedical research data and translation into practice. We applaud the goal to support an efficient and effective research data infrastructure and agree with the tactic of leveraging and strengthening existing systems, as well as continued emphasis on data security. Similarly, the goal to promote "Modernization of the Data-Resources Ecosystem" is an important mechanism to support investigators focus on data driven knowledge development, rather than data management. In terms of implementation tactics, we support the creation of efficient linkages among NIH data resources that contain clinical and observational information, and encourage the inclusion of diverse types of data, including nursing-and patient generated data.

We fully endorse the goal to support the Development and Dissemination of Advanced Data Management, Analytics, and Visualization Tools. As part of implementation tactics, we encourage the inclusion of other nursing databases, (e.g. CINAHL, Scopus) and the

establishment of a curated best practices repository related to evidence-based practices, workflows and data governance, storage and use. In addition, we note that the National Library of Medicine (NLM) lacks vendor-neutral, evidence-based best practices for nursing (and interprofessional) informatics. The Nursing Management Minimum Data (NMMDS), recognized by the American Nurses Association (ANA) and incorporated into the national LOINC standard supports measures to address the Quadruple Aim. Having a national repository for NMMDS and related data would empower big data science preparation in this expanded area of health of the providers[1] and we recommend that NIH consider this effort.

We are pleased with the emphasis on enhancing the nation's biomedical data science research workforce through improved training programs and novel partnerships as this will be critical for sustainability and continued growth in this area. However, we believe this needs to be expanded to include continuing professional education for the healthcare workforce to achieve role-based informatics and data science competencies, including nurses in all professional roles. This is particularly important to ensure the quality and accuracy of clinical data and its efficient use at the bedside. We do applaud the emphasis on engaging with the broader community, including the use of citizen scientists. As nurses, we stress the importance of nurse-patient collaboration and shared decision-making, and would like the engagement of patient and family stakeholders to be highlighted more in implementation tactics for this goal.

Topic 2: Opportunities for NIH to partner in achieving these goals

In order to achieve these laudable goals, partnering with ANI as a group of nursing scientists and informatics experts is critical to provide nursing domain knowledge as needed. ANI members can be used as a resource to support the development of use cases and testing of accessible tools and workflows. Specifically, the Nursing Knowledge Big Data Science Initiative, a collaborative, composed of experts in the field of nursing and data science, is also well positioned to partner on achieving these goals.

Additionally, AMIA is well positioned as a partner for evaluation efforts to conduct rigorous studies to measure impact. Nursing policy experts within ANI work closely with AMIA and can provide insights to support creation of the FAIR (Findable, Accessible, Interoperable, and Reusable) data ecosystem. We also support the alignment of the National Library of Medicine (NLM) Strategic Plan and the Office of the National Coordinator (ONC) Health Information Technology Federal Advisory Committee (HIT-FACA) workgroups, specifically with regards to the domain of Social and Behavioral Determinants of Health (SBDOH).

Topic 3: Additional concepts that should be included in the plan

We fully support the Five *Overarching Goals* and the corresponding *Strategic Objectives*. However, we note that most of the data science initiatives to date focus on human biologic (including neurologic and genomic) variables, as well as increasing progress with clinical data contained in the electronic health record. Missing, but achievable, is the inclusion of other

---

[1] Pruinelli L, Delaney CW, Garcia A, Caspers B, Westra B. (2016). Nursing Management Data Set (NMMDS): Cost Effective Tool to Demonstrate the Value of Nursing Staffing in the Big Data Science Era. Nursing Economic$, 34(2), 66-71.

critically important data that together impact health outcomes, community-based, and population health. This list includes: social behavioral determinants of health, exterior environmental data[2], data from the built environment[3], patient generated data[4] and healthcare delivery data[5,6]. We strongly encourage funding for the advancement of this broader array of data that impact human health and are useful for longitudinal care plans and new funding models, such as value-based care. We recommend that the integration among these disparate sources be a major objective for NIH.

We were pleased to see that the strategic plan includes "*making the data useable for as many people as possible (including researchers, institutions, and the public*)" (p.4). With the large investment of public spending on health data science, usefulness to a broad audience is imperative. While we agree with the FAIR principles[7], we think it's important to acknowledge that these principles alone will not make this complex array of data useable to non-data scientists. Therefore, usability research must be an important focus for achieving the potential scientific discoveries.

We are also concerned that the varying health literacy levels in the lay public can lead to erroneous conclusions if the data tools available to the public are not understandable and highly useable[8,9, 10]. Therefore, funding must be available for rigorous usability testing across health

---

[2] Zenk, S.N., Matthews, S.A., Kraft, A.N. and Jones, K.K., 2018. How many days of global positioning system (GPS) monitoring do you need to measure activity space environments in health research?. *Health & place*, *51*, pp.52-60.

[3] National Academies of Sciences, Engineering, and Medicine, 2017. *Microbiomes of the built environment: a research agenda for indoor microbiology, human health, and buildings*. National Academies Press.

[4] Julia Adler-Milstein, Peter J Embi, Blackford Middleton, Indra Neil Sarkar, Jeff Smith; Crossing the health IT chasm: considerations and policy recommendations to overcome current challenges and enable value-based care, *Journal of the American Medical Informatics Association*, Volume 24, Issue 5, 1 September 2017, Pages 1036–1043, https://doi.org/10.1093/jamia/ocx017

[5] Qato, D.M., Zenk, S., Wilder, J., Harrington, R., Gaskin, D. and Alexander, G.C., 2017. The availability of pharmacies in the United States: 2007–2015. *PloS one*, *12*(8), p.e0183172.

[6] Aiken, L.H., Sloane, D., Griffiths, P., Rafferty, A.M., Bruyneel, L., McHugh, M., Maier, C.B., Moreno-Casbas, T., Ball, J.E., Ausserhofer, D. and Sermeus, W., 2017. Nursing skill mix in European hospitals: cross-sectional study of the association with mortality, patient ratings, and quality of care. *BMJ Qual Saf*, *26*(7), pp.559-568.

[7] Wilkinson MD et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018.

[8] Van Der Heide, I., Heijmans, M., Schuit, A.J., Uiters, E. and Rademakers, J., 2015. Functional, interactive and critical health literacy: Varying relationships with control over care and number of GP visits. *Patient education and counseling*, *98*(8), pp.998-1004.

[9] Kaphingst, K.A., Stafford, J.D., McGowan, L.D.A., Seo, J., Lachance, C.R. and Goodman, M.S., 2015. Effects of racial and ethnic group and health literacy on responses to genomic risk information in a medically underserved population. *Health Psychology*, *34*(2), p.101.

[10] Diviani, N., van den Putte, B., Giani, S. and van Weert, J.C., 2015. Low health literacy and evaluation of online health information: a systematic review of the literature. *Journal of medical Internet research*, *17*(5).

literacy levels to determine both usability, utility and usefulness of the data tools available[11,12,13]. Established usability methods can begin this process, but there is a dearth of research that focuses on data science methods and the potential for unique usability concerns.

Topic 4: Performance measures and milestones that could be used to gauge the success of elements of the plan and inform course corrections

Efforts to enhance the NIH data science workforce through recruiting a cohort of data scientists and others with expertise in areas such as project management, systems engineering, and computer science to provide innovation and expertise not readily available within the federal government are highly supported by ANI. We encourage NIH to be highly strategic in its approach to workforce development to avoid pockets of expertise that are not geographically or experientially diverse. We support the development of materials to train healthcare providers in data science-related clinical applications as a path to engage the broader community. In addition, we point to the need for data scientists to learn from and complement their domain areas with diverse expertise for accurate and deep understanding of other domain-specific data, such as observational patient data recorded in EHRs as a byproduct of nursing practice. In this context, we recommend a performance measure for NIH to track that adequate data science-clinical domain partnerships are established and grown within efforts to enhance the NIH data science workforce and through the launch of the NIH Data Fellows program. Further, we encourage innovative programming, such as a Big Data to Knowledge (BD2K) program that targets nurses, to educate the nurse workforce on data science principles given the domain dependencies in identifying many important data science questions. We also encourage that these approaches include other health professions and the breadth and diversity of these programs are measured to identify gaps overtime.

As noted within NIH's draft strategic plan for data science, "scientists across a wide array of fields said they spend most of their work time (about 80 percent) doing what they least like to do: collecting existing data sets and organizing data...that leaves less than 20 percent of their time for creative tasks like mining data for patterns that lead to new research discoveries."   ANI recommends that NIH establish measures to track researcher's efforts for data collection, curation, and management versus deep learning, knowledge development and sharing. These measures should be used to benchmark and trend the impact of the NIH-funded biomedical data science ecosystem across each institute within NIH to identify areas in need of more robust infrastructure and stronger partnerships between industry, data scientists, and health domains.

We support NIH's encouragement of researchers to use common data elements to help improve accuracy, consistency, and interoperability among data sets within various areas of health and disease research; however, we point to the dependencies on HIT vendors and healthcare

[11] Jake-Schoffman, D.E., Silfee, V.J., Waring, M.E., Boudreaux, E.D., Sadasivam, R.S., Mullen, S.P., Carey, J.L., Hayes, R.B., Ding, E.Y., Bennett, G.G. and Pagoto, S.L., 2017. Methods for Evaluating the Content, Usability, and Efficacy of Commercial Mobile Health Apps. *JMIR mHealth and uHealth*, *5*(12).
[12] Schumacher, R.M. and Lowry, S.Z., 2010. NIST guide to the processes approach for improving the usability of electronic health records. *National Institute of Standards and Technology*.
[13] Zhang, J. and Walji, M.F., 2011. TURF: Toward a unified framework of EHR usability. *Journal of biomedical informatics*, *44*(6), pp.1056-1067.

delivery organizations to adopt common data elements upstream in systems used at the point of care. In response, ANI recommends that the distribution of researcher's efforts continue to be quantified with a specific focus on electronic health records (EHRs) and other clinical data for health research. We suggest that these efforts be in collaboration with NLM and ONC, to understand how to better support researchers in these domains with a more efficient HIT data science pipeline and the need for additional funding mechanisms.

The measures proposed to "Support the Development and Dissemination of Advanced Data Management, Analytics, and Visualization Tools," are important but are largely focused on the quantity of new tools. ANI recommends that additional measures should focus on the *quality* of new tools developed for their capability to promote community development and adoption of uniform standards for data indexing, citation, and modification-tracking (provenance).

Finally, ANI strongly supports NIH's strategic plan related to FAIR-compliant data and ensuring open-source data-analysis tools are available. To expedite the path forward for NIH funded efforts to include FAIR-compliant data, NIH should consider partnering with academic institutions to align NIH's FAIR-compliant data reward and expectation system for investigators with academic criteria for promotion and tenure.

ANI appreciates the opportunity to contribute to the conversation on NIH's strategic plan for data science to promote the modernization of the NIH-funded biomedical data science ecosystem. We believe that the priorities ANI has endorsed in this letter can have a significant impact on improving data and information access to enhancing health and wellness.

Sincerely,

Charlotte Weaver, PhD, RN, MSPH, FHIMSS, FAAN
ANI Co-chair
E-mail: caweaver2011@gmail.com

Mary Beth Mitchell, MSN, RN, BC, CPHIMS
ANI Co-chair
Email: marybethmitchell@texashealth.org