

Reference ID: 11226462762\_YANG

---

**Reference ID:** 11226462762\_YANG

**Submission Date and Time:** 12/16/2019 1:06:21 PM

This contribution was submitted to the National Science Foundation in response to a Request for Information, <https://www.nsf.gov/pubs/2020/nsf20015/nsf20015.jsp>. Consideration of this contribution in NSF's planning process and any NSF-provided public accessibility of this document does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government. The content of this submission is protected by the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>).

*Consent Statement:* "I hereby agree to give the National Science Foundation (NSF) the right to use this information for the purposes stated above and to display it on a publicly available website, consistent with the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>)."

**Consent answer:** I consent to NSF's use and display of the submitted information.

#### **Author Names & Affiliations**

Submitting author: Kesong YANG - University of California San Diego

**Additional authors:** None

**Contact Email Address** (for NSF use only): (hidden)

#### **Research domain(s), discipline(s)/sub-discipline(s)**

computational materials science

#### **Title of Response**

Data-driven Research Challenge for Hybrid Materials

#### **Abstract**

The hybrid organic-inorganic halide perovskites have emerged as one class of most promising optoelectronic materials for next-generation technological applications ranging from photovoltaics to quantum information technology. Here we describe current research challenges for large-scale production of accurate quantum materials data and data analysis for hybrid semiconductor materials,

and the emerging research need to develop related software frameworks for overcoming these challenges. It is expected that the investment of NSF in the development of related software framework packages will significantly promote the computational simulation and discovery of novel hybrid materials, further promoting the application of artificial intelligence to the community of materials research. The training of next-generation workforce in this field is also in a great support of the national Materials Genome Initiative.

**Question 1 (maximum 400 words) – Data-Intensive Research Question(s) and Challenge(s).** Describe current or emerging data-intensive/data-driven S&E research challenge(s), providing context in terms of recent research activities and standing questions in the field. NSF is particularly interested in cross-disciplinary challenges that will drive requirements for cross-disciplinary and disciplinary-agnostic data-related CI.

The hybrid organic-inorganic halide perovskites have emerged as one class of most promising optoelectronic materials for next-generation technological applications including photovoltaics and light-emitting diodes, photodetectors, memory devices, and even quantum information technologies. This is mainly attributed to their exceptional optoelectronic properties such as tunable bandgaps, high absorption coefficient, and excellent charge carrier mobility and the ease of solution-based processability that allows for large-scale fabrication. The interdisciplinary nature and multiple functions of hybrid halide perovskites have attracted increasing attention from a large community of researchers around the world, particularly for the research of perovskite solar cells. On the basis of information from Clarivate Analytics, researchers at more than 3700 organizations worldwide are publishing their research results on the halide perovskite-related photovoltaics and optoelectronics, and this has produced more than 1,4000 scientific papers in the field with a speed of more than 3500 per year now. Hence, these promising applications of hybrid perovskites are expected to continue promoting the explosive growth of the hybrid materials research field in the next following decade. To accelerate the discovery and development of novel hybrid semiconductor materials for various types of optoelectronic applications, it is very necessary to produce accurate quantum materials data of this class of materials from large-scale quantum-mechanics calculations, i.e., ab-initio calculations. However, the hybrid semiconductor materials consist of both organic cations and inorganic framework, and the complex geometrical structures hold a great challenge in the ab-initio calculations mainly because of the great diversity of organic cations. Therefore, an easy-to-use tool to build hybrid material structures for large-scale ab-initio calculations and analysis of materials properties is highly desired for the accelerated discovery and design of hybrid semiconductor materials.

**Question 2 (maximum 600 words) – Data-Oriented CI Needed to Address the Research Question(s) and Challenge(s).** Considering the end-to-end scientific data-to-discovery (workflow) challenges, describe any limitations or absence of existing data-related CI capabilities and services, and/or specific technical and capacity advancements needed in data-related and other CI (e.g., advanced computing, data services, software infrastructure, applications, networking, cybersecurity) that must be addressed

to accomplish the research question(s) and challenge(s) identified in Question 1. If possible, please also consider the required end-to-end structural, functional and performance characteristics for such CI services and capabilities. For instance, how can they respond to high levels of data heterogeneity, data integration and interoperability? To what degree can/should they be cross-disciplinary and domain-agnostic? What is required to promote ease of data discovery, publishing and access and delivery?

Today, large-scale ab-initio electronic structure calculations are being actively used for producing quantum materials databases such as Materials Project, AFLOWLIB, and Open Quantum Materials Database (OQMD), with a purpose of accelerating materials discovery. However, a universal software framework that supports large-scale ab-initio calculations for organic-inorganic hybrid materials (beyond perovskite materials) have, thus far, been lacking. The key reasons for this are three-fold. First and also most important, there has been no open and easy-to-use software to generate hybrid materials in a universal fashion so far, mainly because of the structural complexity of hybrid materials. Second, the hybrid materials have more atoms and more complex symmetry than their corresponding inorganic analogs, resulting in computational difficulties for large-scale ab-initio calculations. Third, the long-range dispersion interaction between organic molecules in the hybrid materials often plays important roles, and thus the van der Waals interaction needs to be included in the calculations. To implement the computational setup of van der Waals interaction into high-throughput ab-initio calculations, further development of interface workflow is also needed. In addition, it is worth mentioning that although hybrid materials have significant advantages than their corresponding inorganic analogs, one major challenge of the hybrid semiconductor materials is their relatively poor stability, and thus a series of comprehensive energetic calculations and analysis are necessary to design novel hybrid semiconductor materials with robust materials stability and desired optoelectronic materials. One key energetic parameter closely related to the thermal stability of hybrid materials is the decomposition enthalpy ( $\Delta H_d$ ), which measures the stability of materials with respect to decomposition. To ensure the stability of selected compounds with respect to decomposition,  $\Delta H_d$  should be positive for all possible decomposition pathways. It is noted that such calculations are not trivial at all, which requires a large number of thermodynamic data for all the possible decomposition compounds and most of them are hybrid materials that are not available in any open quantum materials databases. Moreover, considering the important role of van der Waals interaction in the hybrid materials, to get accurate decomposition enthalpies for van der Waals calculations, all the thermodynamic data for the decomposition compounds require van der Waals calculations, which will double the research efforts and computational time with respect to the standard ab-initio calculations. Therefore, the development of such algorithms to automatically calculate the decomposition enthalpies of hybrid materials with respect to various decomposition compounds is extremely challenging but also highly useful for designing novel hybrid materials in a broad spectrum of technological applications. To the best of our knowledge, there are no open-source software or toolkits that are available for hybrid materials to perform above mentioned energetic analysis.

**Question 3 (maximum 300 words) – Other considerations.** Please discuss any other relevant aspects, such as organization, processes, learning and workforce development, access and sustainability, that need to be addressed; or any other issues more generally that NSF should consider.

The hybrid semiconductor materials are one of the fastest-growing research areas in materials science and engineering, are showing an increasing relevance in many areas of materials applications from solar cells to light emitters and quantum information technologies. The investment of NSF in the development of related software framework and package will significantly promote the computational simulation and discovery of novel hybrid materials. This will enable great expansion of materials classes with respect to the pure inorganic materials and organic molecules, and extend the “genome” of materials into hybrid materials, in support of the national Materials Genome Initiative, further promoting the application of artificial intelligence to the community of hybrid materials research. In addition, it is also very necessary to train next-generation workforce in this field.

-- End Submission --