Opportunity Title: Machine-led Discovery of Novel Materials for Automated Chemical Synthesis
Opportunity Reference Code: ICPD-2021-55

Organization Office of the Director of National Intelligence (ODNI)

Reference Code ICPD-2021-55

How to Apply **Create and release your Profile on Zintellect** – Postdoctoral applicants must create an account and complete a profile in the on-line application system. **Please note: your resume/CV may not exceed 2 pages.**

Complete your application – Enter the rest of the information required for the IC Postdoc Program Research Opportunity. The application itself contains detailed instructions for each one of these components: availability, citizenship, transcripts, dissertation abstract, publication and presentation plan, and information about your Research Advisor co-applicant.

Additional information about the IC Postdoctoral Research Fellowship Program is available on the program website located at: https://orise.orau.gov/icpostdoc/index.html.

If you have questions, send an email to ICPostdoc@orau.org. Please include the reference code for this opportunity in your email.

Application Deadline 2/26/2021 6:00:00 PM Eastern Time Zone

Description **Research Topic Description, including Problem Statement:**

The interest in utilizing theoretical science to predict new materials and synthesis pathways has been longstanding. Computing power has accelerated in the 21st century and corresponding advancements in artificial intelligence (AI) and machine learning (ML) have increased. Recent research is applying AI, ML, and autonomous systems to the field of chemical synthesis. At present, human-led discovery of new materials through manual practices can take decades of research, significant continuous funding, and can result in a high degree of risk. This fellowship would focus on applying ML for the discovery of novel materials to increase efficiency and reduce cost and risk.

Recent research across academia has led to the creation of software to translate bulk text into low-level instructions using natural language processing1. Development in this space allows for the optimization of experiments based on prior experiences and the progress of data-driven materials discovery2. This problem statement is specifically looking at the next stage of the pipeline to identify, implement, and validate a method for machine-led discovery of materials. This may include the discovery of novel materials or the discovery of novel reaction pathways for conventional or traditional materials. Research in this area would also enable the assurance of material supply and identification of potential new material threats.

1 ChemData Extractor: A Toolkit for Automated Extraction of Chemical Information from Scientific Literature, M. Swaine, J. Cole, https://pubs.acs.org/doi/abs/10.1021/acs.jcim.6b00207

2 A Design-to-Device Pipeline for Data-Driven Materials Discovery, *J.Cole,* https://pubs.acs.org/doi/10.1021/acs.accounts.9b00470

Opportunity Title: Machine-led Discovery of Novel Materials for Automated Chemical Synthesis
Opportunity Reference Code: ICPD-2021-55

Example Approaches:

Proposals should include the assessment and development of available algorithms for discovery of new materials using ML. These should be assessed against the following criteria:

- Applicability of algorithms to different material types, such as: synthetic biology, energetics and other hazardous materials, advanced materials (e.g. smart materials/nanomaterials).
- Computational requirements.
- Requirement for transparent methods and explainable results.
- Mitigation of risk associated with false alarms.
- Potential for validation and optimization of the trained model.

The most relevant and applicable method(s) should be implemented and validated against applicable data.

Open-source datasets for model training are available. These include but are not limited to:

- PubChem NCBI: https://pubchem.ncbi.nlm.nih.gov/.
- ChemSpider (RSC): http://www.chemspider.com/.
- NIST Webbook: https://webbook.nist.gov/.
- Crystallography Open Database: http://www.crystallography.net/cod/result.php.

The Cambridge Structural Database (https://www.ccdc.cam.ac.uk/structures/) may also be used; however, this is licensed. It is expected that datasets from different sources may be required to best train a model for a variety of scenarios.

Proposals are also invited to consider future developments in the pipeline such as in-line characterization (e.g. in chemical synthesis) and analysis of predictions.

Technical collaborating from the ML and chemical synthesis aspect will be offered. In addition, classified data will be used for testing of models, and validation and feedback will be provided for further development.

Relevance to the Intelligence Community:

Machine led discovery has applications across the Intelligence Community, including:

- Discovery of novel materials and novel reaction pathways for known materials in a cost-efficient manner that reduces the human risk.
- Efficient derivation of synthetic routes for new and emerging threats and understanding of techniques that could be employed by adversaries.
- Advancement of machine-led decision making algorithms that can be used in remote robotic operations.

This is a vital stage in the discovery, manufacture, and validation pipeline, alongside enhancing materials assurance. Ultimately, this research will enable the discovery of materials beyond human capability. In the future, this capability would also enable the screening of materials that do not obey conventionally understood chemistry and classification. Successful

Opportunity Title: Machine-led Discovery of Novel Materials for Automated Chemical Synthesis
Opportunity Reference Code: ICPD-2021-55

developments would lead to integration with robotics and automated manufacture.

This research has the potential to be used in the domains of pharmaceuticals, synthetic biology, advanced materials, and more.

Key Words: Machine Learning, ML, Artificial Intelligence, AI, Neural Networks, Automated Discovery, Novel Materials, Automated Chemical

Qualifications

Postdoc Eligibility

- U.S. citizens only
- Ph.D. in a relevant field must be completed before beginning the appointment and within five years of the application deadline
- Proposal must be associated with an accredited U.S. university, college, or U.S. government laboratory
- Eligible candidates may only receive one award from the IC Postdoctoral Research Fellowship Program

Research Advisor Eligibility

- Must be an employee of an accredited U.S. university, college or U.S. government laboratory
- Are not required to be U.S. citizens

Eligibility Requirements

- **Citizenship:** U.S. Citizen Only
- **Degree:** Doctoral Degree.
- **Discipline(s):**
  - **Chemistry and Materials Sciences** (12)
  - **Communications and Graphics Design** (2)
  - **Computer, Information, and Data Sciences** (17)
  - **Earth and Geosciences** (21)
  - **Engineering** (27)
  - **Environmental and Marine Sciences** (14)
  - **Life Health and Medical Sciences** (45)
  - **Mathematics and Statistics** (10)
  - **Other Non-Science & Engineering** (2)
  - **Physics** (16)
  - **Science & Engineering-related** (1)
  - **Social and Behavioral Sciences** (27)