

Website: <https://www.arl.army.mil/HTMDEC>

High-Throughput Materials Discovery for Extreme Conditions

(HTMDEC)

Program Description

Purpose: Within the Army science and technology enterprise, DEVCOM-ARL is chartered to conduct disruptive foundational research, engage as the Army's primary collaborative link to the scientific community, and interface to shape future fighting concepts. We crystalize these ideas and the impetus to perform these functions at the pace of innovation as 'Operationalize Science for Transformational Overmatch'. Simply put, we seek to accelerate discovery and transition breakthroughs to the Warfighter.

Rule-based artificial intelligence (AI) and machine learning (ML) tools present powerful avenues for exploring an information landscape in discovering novel materials for applications in extreme conditions (e.g. high-strain rate, high-g loading, high temperature). Such approaches present considerable opportunity in exploring new frontiers for materials used in protection and lethality applications, especially when coupled with new approaches that allow larger and richer datasets, computational tools, and data infrastructure for collaboration. Broadly, AI/ML can be used to augment individual steps in the synthesis-processing-characterization pipeline, be used for scale-bridging to draw greater information from more tractable experimental approaches, and be used to guide a broader research loop.

Advances in synthesis, modeling, and characterization will greatly advance our ability to exploit monolithic materials in extreme conditions. However, there is a need to contemplate how the capabilities of additive manufacturing and other processing techniques can be used to evaluate materials that exhibit spatial variations in composition, anisotropic characteristics, and contain interfaces between multiple materials. The parameter space expands exponentially as these variables compound the system inputs, but truly advanced materials performance will likely be dependent on an integrated systems-level approach to materials design.

ML toolsets coupled with advanced manufacturing and characterization is necessary to achieve accelerated discovery of new materials for application in extreme dynamic (impact, thermal, ablative) conditions. ML toolsets and software exist but may need to be adapted for the

specific requirements of materials discovery and design. Full exploitation of the ML approach will certainly require extension and further development to focus on proof-of-concept for material classes of interest in Army applications. This could be achieved within a generalized and scalable framework that supports rapid, robust and trusted data exchange. New tools that consolidate/organize data and increase throughput throughout the workflow will require a specialized approach to be applied to ephemeral phenomena e.g. shocks, heating, localized deformation, and failure. ML models that incorporate these phenomena will critically rely on physics-based models that adequately capture the underlying driving mechanisms. Critical (targeted by ML approaches) physics models may require further development; ML offers opportunity to consolidate much of these physics into fast-running analytic frameworks compatible with the high-throughput approach and may be used to guide autonomous systems for high-throughput characterization of transient phenomena.

To accelerate improvements in Army armor and weapon system performance, DEVCOM-ARL wants to leverage high-throughput methods in synthesis, processing, characterization, and modeling for materials used in these applications. Machine-learning techniques are in the nascent stage of integration with materials science but may present a path towards accelerated discovery, as these tools may uncover novel links between system performance and material science that have been previously underdeveloped or overlooked. DEVCOM-ARL seeks collaboration with external investigators to leverage (and train experts on) machine-learning techniques in the discovery of materials that perform in extreme environments, but machine-learning techniques require large volumes of quantifiable data in order to best reveal links between the materials science and system performance. High throughput characterization and manufacturing techniques may present a viable approach to satisfy the data volume requirements to bring machine-learning to bear.

In summary, the US Army Modernization Priorities require materials that survive and perform in extreme environments; harsh military environments of high-acceleration (e.g. projectile launch and flight), high-temperature and rapid ablation (e.g. hypersonic flight), and impacts at very high velocity (terminal ballistics). The totality of these environments and accumulating requirements on future materials drives the imperative to consider an increasingly large number of constituent elements, structure and properties. Discovery must now parse through billions of candidate materials to achieve highly specialized and transformational functions. This drives a data-driven approach; one that fuses high-throughput materials synthesis and characterization with machine learning algorithms and close-loop discovery automation.

The overarching goal of this program is to couple automation and machine learning techniques to material manufacturing and characterization to withstand and perform under extreme

conditions. The program will develop the necessary methodologies, models, algorithms, synthesis & processing techniques, and requisite characterization and testing to rapidly accelerate the discovery of novel materials through data-driven approaches. As such, it is expected the results of this program will be the above techniques as well as novel materials exhibiting unprecedented properties at the appropriate scales that have been developed utilizing all of the aforementioned tools which will be provided to DEVCOM-ARL for further analysis and testing.

Proposals may draw from any number of these thrusts but must focus on research that employs high-throughput data-driven techniques to close material design loops connecting material selection, synthesis, and processing to system performance.

In order to achieve this paradigm shift in materials discovery, significant advances are needed in the following general thrust areas:

- Data-driven Material Design - meant to be a comprehensive term for all aspects of the material design phase of the material development cycle which are accelerated through the integration of data-driven methods.
- High-Throughput Synthesis & Processing – to include both modifying existing synthesis & processing methods to accommodate for high-throughput, as well as developing novel techniques.
- High-Throughput Characterization – to include implementation of automation for conventional techniques, and the development of surrogate tests to mimic techniques which are not amenable to automation, especially for experiments in extreme conditions (e.g. high strain rate, high temperature).
- ML-augmented Physics-Based Models – the use of ML tools to identify the most crucial parameters and parametrization experiments for physics-based models is poised to be a tipping point in materials science. To date, nearly all ML algorithms have been developed for big data (e.g. image recognition). It is critical that we discontinue ‘repurposing’ these types of algorithms and begin developing ML algorithms specifically designed for materials discovery, and informed by physics.

The general thrust areas will be complemented by the following targeted thrust areas:

- Program and Workflow Development – optimizing workflow to achieve the best outcomes for the general thrust areas through improved teaming and program planning.
- Data Handling and Management – development and upkeep of a data platform used by all funding recipients and government collaborators.

HTMDEC has been developed in coordination with other related ARL-funded collaborative efforts (see descriptions of ARL collaborative alliances at <https://www.arl.army.mil/www/default.cfm?page=93>) and shares a common vision of highly collaborative academia-industry-government partnerships. However, HTMDEC will be executed with a program model different than previous ARL Collaborative Research/Technology Alliances. Specific components of the program are highlighted below:

- HTMDEC will be a two-step application process, consisting of a White paper stage and a Proposal stage.
- HTMDEC will be executed through an initial funding period (“Seedling” only), followed by subsequent funding periods involving both Seedling and “Center” awards. A Center will be an option period exercised from a seedling award. The only exception to this will be for the Targeted Thrust Area- Data Handling & Management. The Seedling selected for continuation for this particular thrust area will do so as a recurring Seedling for the duration of the HTMDEC program, and will support all of the Centers as the approved data platform for HTMDEC.
- A FOA Opportunity workshop will be held to brief interested Applicants on the long term program goals of this FOA. In FY2022, only Seedling efforts will be awarded. The focus of these Seedlings will be to address either one or more of the general or targeted thrust areas.
- White papers will address one or more of the thrust areas; thrust areas may change on an annual basis in order to reflect current interests. White papers will be evaluated. Applicants with only the most highly rated white papers will receive an invitation from the Government to submit a Proposal.
- Seedling Proposals will address one or more of the thrust areas. Proposals will be evaluated and funding will be provided to those Recipients selected for award of a cooperative agreement (CA) described as the Seedling award.
- Prior to the close of the Seedling award, the Recipients of a Seedling CA are then eligible to submit a proposal for consideration of an option period under the Seedling CA of up to 4 years. This option period will be a Center. Since the option proposal will need to address all four general thrust areas, Seedling Recipients will be encouraged to collaborate and combine during the option period of performance to propose the strongest possible Center option for

selection. Option proposals will be evaluated and funding will be provided to those Recipients selected for the exercise of their option.

- In addition to the Center option proposal addressing all four general thrust areas, Recipients will also have the opportunity to provide for a Graduate Student Fellowship effort for US citizens working in one of the thrust areas for both Center and Seedling proposals?.
- Project review workshops will be held annually, with the intention of allowing all CA Recipients to present the results of their research, as well as interacting with the other Recipients.

White papers and Proposals that are in compliance with the requirements of the FOA will be evaluated in accordance with merit-based, competitive procedures. These procedures will include evaluation factors and an adjectival and color rating system. A review team, consisting of a qualified group of scientists and managers will evaluate the compliant proposals and provide the results of that evaluation to the decision-maker for the Government.

Technical Thrusts

FY2022 Proposals will address a single (targeted or general) thrust as a seedling. In following years, seedlings may address any active thrust area, while centers must address at least the four general thrusts to converge high-throughput and data-driven techniques to demonstrate material selection, synthesis, processing, and performance in a design-cycle manner. FY2022 seedlings for targeted thrusts will have the opportunity to develop a robust framework for the program in these areas: 1. Program & Workflow Development and 2. Data Handling & Management. General thrust areas that focus on research employing high-throughput data-driven techniques to close material design loops connecting material selection, synthesis, and processing to system performance are: 1) Data-driven Material Design, 2) High-Throughput Synthesis & Processing, 3) High-Throughput Characterization, and, 4) ML-augmented Physics-Based Models.

Targeted Thrust Areas:

1. Program & Workflow Development

In FY2022 only, funding will be made available to allow for development of a comprehensive program plan, project workflow, and facilitate teaming among awardees and with government in anticipation of continuing on to a Center in FY2023. As a funded 'seedling', applicants are encouraged to exercise proof-of-principle features of a rudimentary design cycle in order to expose shortcomings of existing methods and suggest research to close existing gaps. Project workflow should include determination of all of the required resources, including but not limited to: processes, facilities, personnel, etc.. As Centers are expected to address all four of

the primary thrust areas, proposals for this particular targeted thrust area should likewise address all four thrust areas.

2. Data Handling & Management

In FY2022 only, proposals may directly address data handling and management within the program. A common data handling and management schema must be established from the onset of the program that will be utilized by program participants throughout the life of the program. This schema should consider all aspects of data handling and management, including, but not limited to, data acquisition, data transfer and sharing protocols, tagging, data analysis, version control and safe/redundant/scalable data storage. The FAIR guiding principles for scientific data management and stewardship defined in 2016 in Scientific Data shall be followed in order to improve the Findability, Accessibility, Interoperability, and Reuse of digital assets. Proposals must specifically address data sharing between program awardees and government collaborators that have strictly regulated information systems (e.g. the DOD high performance computing system). Proposals for this particular task will only be accepted in FY2022, with the intent to down select the most promising proposals (selected on the basis of scientific merit and collaboration within the effort) for award throughout the duration of the program as the approved data handling and management platform for all of the award recipients moving forward.

General Thrust Areas:

1. *Data-driven Material Design*

For the purposes of this FOA, data-driven material design is meant to be a comprehensive term for all aspects of the material design phase of the material development cycle which are accelerated through the integration of data-driven methods. This includes, but is not limited to: data mining through literature sources and existing databases (from data sources such as density functional theory, thermodynamics databases such as Calphad, literature sources, etc.), use of adaptive learning or experimental optimization to predict or guide future experiments, and use of machine learning and artificial intelligence to parameterize the material property space for extreme conditions. Approaches that accelerate material design in multi-objective systems are of particular interest. A trained model is not a sufficient deliverable; a deliverable must include a toolset that can be used to design or discover new materials that offer an improvement in performance in Army application environments. Sub-areas within this thrust could include: data mining, adaptive learning approaches and ML/AI driven design, and uncertainty quantification:

- i. Data mining includes the use of machine learning algorithms for mining of material properties from various databases, repositories, and the open literature. This includes using conventional data science approaches on the larger data sets that currently exist, as well as using approaches such as natural language processing (NLP) to consolidate information that currently exists in an unorganized fashion in literature. While this type of data is sparse as compared to typical 'big data' sets, there is an opportunity to rapidly explore the compositional space of candidate materials as a first-order screening tool. This information can also be utilized to construct quantitative structure-property relationship (QSPR) models for predicting the properties of novel materials.
- ii. Adaptive learning is an AI approach to guide experimental steps for optimal value. Using various approaches, an active learning scheme can be used to target specific experimental parameters (e.g. composition and processing conditions) based on user-defined rules, such as targeting areas with maximum uncertainty in learned QSPR or targeting extrema in predicted materials performance. Such approaches can be critical in materials design and discovery of new candidate materials with strong application-specific performance. Integration of active learning schemes with experimental hardware offers the opportunity to automate or autonomize experimental steps and vastly accelerate the materials design process.
- iii. ML/AI driven design involves use of ML/AI approaches to accelerate steps in the materials design and discovery process. This can include improvement of individual steps in the research pipeline (e.g. use of image recognition approaches in microscopy), connecting approaches together for improved connectivity (e.g. use of ML approaches to address scale bridging between computational models and tools), and use of these tools to learn a broader region of materials space and identify key regions of high performance. In all cases, emphasis of using these tools to design new materials at an accelerated rate is critical.
- iv. Uncertainty quantification will play a prominent role in synthesis, characterization, and modeling for the materials design process. Understanding the scope of uncertainty during synthesis ensures that prototype materials contain desired structure. Analyzing uncertainty propagation in characterization experiments, based on uncertainties during synthesis and within the experiment itself, enables quantifiable estimates of confidence in collected data and enables informed design-of-experiments. Finally, modeling efforts stand to benefit from careful uncertainty quantification for model inputs to better identify the role of the most crucial parameters in the materials design space.

2. High-Throughput Synthesis & Processing

Emerging methodologies such as ICME and 'Materials by Design' have allowed scientists to rethink the early stages of a material's development life cycle by integrating computational models and simulation into the material design phase. This allows researchers to very rapidly hone in on the most promising composition space, but is just the beginning of the material development cycle. In order to ensure high-throughput materials discovery, it is imperative that the accelerated pace continues through the synthesis, processing, and characterization phases. This poses significant challenges, as most existing R&D infrastructure is not designed for high-throughput and/or automated processes.

- i. High-throughput synthesis includes both modifying existing synthesis methods to accommodate for high-throughput, as well as developing novel techniques. Any technique developed should consider sample homogeneity and compositional control, as fluctuations could greatly impact uncertainty downstream. Characterization and testing of these samples will occur over many length scales, so it is imperative that the technologies proposed are amenable to the synthesis of bulk materials, and not solely applicable to thin films. Of particular interest is the integration of in situ diagnostics into the synthesis technique such that real time quality assurance (QA) could be achieved. Challenges to synthesis include a lack of available technologies demonstrated in a high throughput environment, e.g. combinatorial chemistry for metals and ceramics. Challenges also include a disparity across material classes, with polymers and metals being more amenable to high-throughput than ceramics and composites.
- ii. High-throughput processing involves all of the post-synthesis processing required before a material is ready for characterization, testing, or use in a relative environment. This includes, but is not limited to consolidation, heat-treatments, cold and hot working, machining, etc.. Processing maps are often ill-explored and process models tend to be basic in nature. However, processing dictates microstructure evolution and will be paramount to assuring consistency in material properties. Challenges in processing include the fact that there is no straightforward or unified way to represent/encode a material's processing history. In addition, processing techniques vary widely, and are sometimes proprietary and not based on industrial standards. Challenges also include the fact that many processing steps (e.g. sintering) are not easily adapted to high-throughput processes and could pose a rate-limiting bottleneck in the process.

- iii. In situ diagnostic approaches and analysis will play a critical role in any high-throughput process, including both synthesis and processing. Techniques are required that can monitor the quality of a material in real time such that uncertainty can be minimized. Items to be considered include, but are not limited to: homogeneity, presence of defects, texturing, crystallographic phase, etc.

3. *High-Throughput Characterization*

Characterization of materials suffers from being inherently slow. Equipment has been designed for conventional sample-by-sample handling, and very little has been done in the way of high-throughput and/or automation, especially in characterization experiments for extreme conditions or with higher fidelity technologies. Conducting high strain-rate or other extreme mechanical property testing with a high-throughput approach will always hold further challenges than in quasi-static testing. Methodologies to improve the throughput and/or automation of high-strain rate and high-temperature tests, which accurately mimic the extreme environments of the application space, are of significant interest. Modifications to techniques such as laser induced particle impact testing (LIPIT) and laser driven flyer plate experiments are examples of areas ripe for further exploration. High-fidelity characterization techniques (e.g. nano-indentation, scanning electron microscopy, etc.) have seen some limited levels of automation, but the majority of high-fidelity characterization techniques stand to benefit from further adaptations to achieve higher experimental throughput.

- i. High-strain rate surrogate tests are required to enable testing in a relevant environment, but in a high-throughput or automated mode. The application space for the HTMDEC program places materials in extreme conditions such as at temperatures in excess of 1000s of degrees Celsius or at strain rates exceeding 10⁴ per second with pressures in the tens of Gigapascals and rise-times in the nanoseconds. Achieving these conditions is nontrivial in conventional experiments, but these experiments tend to provide the most meaningful screening method for down-selection of the most promising candidate materials and systems. Surrogate experiments may need to focus on a single material property or mechanism that is the primary driver of the material or system response to extreme conditions in order to improve experimental throughput. These accelerated and/or automated surrogate experiments may also need to trade experimental fidelity for throughput, but the resulting larger datasets might guide further testing with slower, conventional higher fidelity traditional testing for extreme

conditions (e.g. Kolsky bar, ballistic testing). Challenges include examining the trade-off space between experimental throughput and fidelity, as well as automating tests of considerable complexity. Ideally the surrogate experimental approach should interrogate system level performance in as realistic of conditions as possible (i.e. experiments should reveal the role of multiple materials, their interfaces, and their geometries at the system level).

- ii. Automated characterization techniques are needed to rapidly accelerate what is often one of the slowest phases of material development. In addition to automating the actual techniques, data acquisition and analysis should also be automated and seamlessly fed into the appropriate ML data sets, it's anticipated that the greatest advances can be realized in lower fidelity, quasi-static techniques than in higher-fidelity, high-strain rate testing, especially in the near term. Challenges include the lack of automation in conventional characterization techniques, which were designed for sample by sample analysis. Challenges also include the fact that with high-throughput there will be increased uncertainty due to trade-offs between speed and accuracy.
- iii. Novel diagnostic approaches and analysis are required to maximize measurement resolution from single high throughput characterization experiments to increase data fidelity and reduce error. Automated instrumentation methods (e.g. machine vision) may present an avenue to obtain more granular, full-field data from a single experiment than conventionally possible, both increasing the throughput of data and reducing experimental burden for individual characterization experiments. Coupling greater diagnostic depth to high-throughput experiments could then provide an even greater amount of experimental data than just the high-throughput experiment itself. ML trained analysis of characterization experiments may reduce experimental uncertainty requiring fewer experiments to make connections between processing and system performance.

4. *ML-augmented Physics-Based Models*

The ultimate goal of this research program is to link material processing and properties to system level performance. Numerical models offer an opportunity to design at the system level in silico but are completely reliant upon correctly identifying these links and developing models that faithfully and completely capture these links. Structural analysis of heterogeneous materials and systems using phenomenological constitutive models is often inaccurate without a connection to the material microstructure and underlying physics. Physics-based models that

can accurately simulate the material response in the application environment are one way to attain this goal, but understanding the role of model inputs remains a challenge, and developing these materials models in realistic materials systems with fidelity to all behaviors across many length scales simultaneously remains in many cases difficult or intractable. However, for many systems ML-augmented physics-based models represent a nascent catalogue of techniques that seek to improve upon the conventional models used with respect to speed or fidelity, and overcome the limitations found with a lack of data necessary to use more conventional data-driven learning. Approaches range from use of ML to optimize model parameters to learning the mathematical and physical models themselves or use of physics to constrain and guide an ML model. These approaches offer the opportunity to enhance computational approaches used to design and discover materials in a number of ways. Below are listed some examples of how a successful program might incorporate this thrust area:

- i. Implementation of ML-augmented physics-based models to study critical physics mechanisms. In particular, with respect to ballistics applications, it is difficult to study transient and highly-localized phenomena that occur in high-rate and ballistic events (e.g. shocks, heating, deformation, and failure), especially for complex system-level geometries and interfaces. It is further difficult to incorporate these models into a wider workflow of research. Critical physics model development, coupled with ML approaches that can accelerate the runtime of these traditionally-slow models and allow their incorporation into a high-throughput and potentially autonomous workflow, are likely important to a successful program. Many of these behaviors are still modeled with approaches that cannot collect the full generalized behavior of these transient events with the required predictive accuracy for materials design requirements. In addition to developing models that offer improved speed in a workflow, developing models that can fill in critical gaps in behavioral understanding and quantitative prediction of these transient events is also a challenge that remains outstanding.
- ii. Scale-bridging considerations remain a challenging area in computational materials. The approaches used in Physics-driven ML models may offer solutions to bridge the scales from molecular interaction models developed in smaller scales up to the performance of a system in its application scale and environment. There are a great number of implementations possible that would benefit from accelerated queries for data from databases to accelerated physics models that can be called at runtime, among others. Developing an approach to streamline and accelerate multiscale materials research is likely critical to a broader program workflow.

- iii. Training of ML models as well as the ability to condense data into functions that can be queried quickly offers substantial opportunity in collaboration and handoff of information. By providing a solution that can relatively quickly provide a predicted system structure, property, or behavior, a broader automated or collaborative research loop for materials design and discovery can be improved. Development and handoff of such links in the chain are likely to be an important part of a successful program. Further, development of data sharing and data hand-off frameworks is necessary to a successful collaboration.

- iv. Development of processing models of materials remains an area of especial challenge in the development of a materials design and discovery approach. Though no less essential to development of new best-in-class materials than composition space, the wide range in category of degrees of freedom in sequential processing steps compared to composition creates a very grand challenge in representing and executing materials design and discovery through all but the simplest of processing chains. The development of a data science-ready physics-based processing model that can be executed to explore processing space in design of ballistic materials would be a major triumph. Challenges include representing or encoding an open-ended processing history, as well as predicting the composition-dependent changes such a processing chain would engender.

- v. Development of an overarching methodology will be required as multiple data sources are used (models/characterization/ML) to predict individual features, properties, or behaviors, the consolidation of these elements into a broader model that evaluates expected macroscopic properties becomes nontrivial. The use of some combination of ML and physics-based information will likely be critical in developing a final step that brings these elements together to predict a general system performance and/or learns the landscape of system performance as a function of these elements. The eventual goal of a successful program includes design of materials with multi-objective requirements. Data will likely come from many sources, including data repositories, physics-based computational models, high-throughput or automated experiments, and ML models. The development of one over-arching umbrella that consolidates program information into a system performance model of materials will enable much greater success in discovering or designing materials with user-specified objectives.