



# Advancing Fusion with Machine Learning Research Needs Workshop Report

David Humphreys<sup>1</sup> · A. Kupresanin<sup>2</sup> · M. D. Boyer<sup>3</sup> · J. Canik<sup>4</sup> · C. S. Chang<sup>3</sup> · E. C. Cyr<sup>5</sup> · R. Granetz<sup>6</sup> · J. Hittinger<sup>2</sup> · E. Kolemen<sup>7</sup> · E. Lawrence<sup>8</sup> · V. Pascucci<sup>9</sup> · A. Patra<sup>10</sup> · D. Schissel<sup>1</sup>

Accepted: 9 September 2020 / Published online: 26 September 2020  
© The Author(s) 2020

## Abstract

Machine learning and artificial intelligence (ML/AI) methods have been used successfully in recent years to solve problems in many areas, including image recognition, unsupervised and supervised classification, game-playing, system identification and prediction, and autonomous vehicle control. Data-driven machine learning methods have also been applied to fusion energy research for over 2 decades, including significant advances in the areas of disruption prediction, surrogate model generation, and experimental planning. The advent of powerful and dedicated computers specialized for large-scale parallel computation, as well as advances in statistical inference algorithms, have greatly enhanced the capabilities of these computational approaches to extract scientific knowledge and bridge gaps between theoretical models and practical implementations. Large-scale commercial success of various ML/AI applications in recent years, including robotics, industrial processes, online image recognition, financial system prediction, and autonomous vehicles, have further demonstrated the potential for data-driven methods to produce dramatic transformations in many fields. These advances, along with the urgency of need to bridge key gaps in knowledge for design and operation of reactors such as ITER, have driven planned expansion of efforts in ML/AI within the US government and around the world. The Department of Energy (DOE) Office of Science programs in Fusion Energy Sciences (FES) and Advanced Scientific Computing Research (ASCR) have organized several activities to identify best strategies and approaches for applying ML/AI methods to fusion

energy research. This paper describes the results of a joint FES/ASCR DOE-sponsored Research Needs Workshop on Advancing Fusion with Machine Learning, held April 30–May 2, 2019, in Gaithersburg, MD (full report available at [https://science.osti.gov/-/media/fes/pdf/workshop-reports/FES\\_ASCR\\_Machine\\_Learning\\_Report.pdf](https://science.osti.gov/-/media/fes/pdf/workshop-reports/FES_ASCR_Machine_Learning_Report.pdf)). The workshop drew on broad representation from both FES and ASCR scientific communities, and identified seven Priority Research Opportunities (PRO's) with high potential for advancing fusion energy. In addition to the PRO topics themselves, the workshop identified research guidelines to maximize the effectiveness of ML/AI methods in fusion energy science, which include focusing on uncertainty

---

*Disclaimer* This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

---

✉ David Humphreys  
humphreys@fusion.gat.com

<sup>1</sup> General Atomics, San Diego, CA, USA

<sup>2</sup> Lawrence Livermore National Laboratory, Livermore, CA, USA

<sup>3</sup> Princeton Plasma Physics Laboratory, Princeton, NJ, USA

<sup>4</sup> Oak Ridge National Laboratory, Oak Ridge, TN, USA

<sup>5</sup> Sandia National Laboratories, Albuquerque, NM, USA

<sup>6</sup> Massachusetts Institute of Technology, Cambridge, MA, USA

<sup>7</sup> Princeton University, Princeton, NJ, USA

<sup>8</sup> Los Alamos National Laboratory, Los Alamos, NM, USA

<sup>9</sup> University of Utah, Salt Lake City, UT, USA

<sup>10</sup> University at Buffalo, SUNY, Buffalo, NY, USA

quantification, methods for quantifying regions of validity of models and algorithms, and applying highly integrated teams of ML/AI mathematicians, computer scientists, and fusion energy scientists with domain expertise in the relevant areas.

**Keywords** Fusion science · Artificial intelligence · Machine learning · Scientific discovery

## Executive Summary

### Background and Motivation

The pursuit of fusion energy has required extensive experimental and theoretical science activities to develop the knowledge needed that will enable design of successful fusion power plants. Even today, following decades of research in many key areas including plasma physics and material science, much remains to be learned to enable optimization of the tokamak or other paths to fusion energy. Data science methods from the fields of machine learning and artificial intelligence (ML/AI) offer opportunities for enabling or accelerating progress toward the realization of fusion energy by maximizing the amount and usefulness of information extracted from experimental and simulation output data. Jointly supported by the Department of Energy Offices of Fusion Energy Science (FES) and Advanced Scientific Computing Research (ASCR), a workshop was organized to identify Priority Research Opportunities (PRO's) for application of ML/AI methods to enable accelerated solution of fusion problems. The resulting “Advancing Fusion with Machine Learning Research Needs Workshop,” held in Gaithersburg, MD, April 30–May 2, 2019, brought together ~ 60 experts in fields spanning fusion science, data science, statistical inference and mathematics, machine learning, and artificial intelligence, along with DOE program managers and technical experts, to identify key PRO's.

### Priority Research Opportunities

The goals of the ML workshop were to assess the potential for application of ML/AI methods to achieve transformative impacts on FES research, and identify research needs, opportunities, and associated gaps in ML and AI areas that would help address fusion energy problems through targeted partnerships between fusion scientists and applied mathematicians or computer scientists.

Seven PROs were identified, including three in each of two broad categories: Accelerating Science, and Enabling Fusion (see the PRO summary table). The seventh, cross-cutting, PRO consists of research and development to provide computational and database resources that support the execution of the other six PRO's. The PRO's identified are:

PRO 1: Science Discovery with Machine Learning includes approaches to bridge gaps in theoretical understanding through identification of missing effects using large datasets; accelerating hypothesis generation and testing; and optimizing experimental planning to help speed up progress in gaining new knowledge. This approach to supporting and accelerating the scientific process itself has already proven to be among the most successful applications of ML/AI methods in many fields. PRO 1 research activities may result in theory-data hybrid models describing such important physics areas as tokamak confinement, resistive magnetohydrodynamic stability, and plasma-wall interactions. Priority planning of magnetic confinement experiments can help maximize the effective use of limited machine time.

PRO 2: Machine Learning Boosted Diagnostics involves application of ML methods to maximize the information extracted from measurements, enhancing interpretability with data-driven models, systematically fusing multiple data sources, and generating synthetic diagnostics that enable the inference of quantities that are not directly measured. For example, the additional information extracted from diagnostic measurements may be included as data input to supervised learning (e.g. classification) methods, thus improving the performance of such methods for a host of cross-cutting applications. Examples of potential research activities in this area include data fusion to infer detailed 3D MHD modal activity from many diverse diagnostics, enhancing 3D equilibrium reconstruction fidelity, extracting meaningful physics from extremely noisy signals, and automatically identifying important plasma states and regimes for use in supervised learning.

PRO 3: Model Extraction and Reduction includes construction of models of fusion systems and plasmas for purposes of both enhancing our understanding of complex processes and the acceleration of computational algorithms. Data-driven models can help make high-order behaviors intelligible, expose and quantify key sources of uncertainty, and support hierarchies of fidelity in computer codes for whole device modeling. Furthermore, effective model reduction can shorten computation times for multi-scale/multi-physics simulations. Work in the areas covered by this PRO may enable faster than real-time execution of tokamak simulations, derivation and improved understanding of empirical turbulent transport coefficients, and derivation of accurate but rapidly-executing models of

plasma heating and current drive effects for RF and other sources.

**PRO 4: Control Augmentation with Machine Learning** identifies three broad areas of plasma control research that will benefit significantly from augmentation through ML/AI methods. Control-level models, an essential requirement of model-based design for plasma control, can be improved through data-driven methods, particularly where first-principle physics descriptions are insufficient. Real-time data analysis algorithms designed and optimized for control through ML/AI can enable such critical functions as evaluation of proximity to MHD stability boundaries and identification of plasma responses for adaptive regulation. Finally, the ability to optimize plasma discharge trajectories for control scenarios using algorithms derived from large databases can significantly augment traditional design approaches. The combination of control mathematics with ML/AI approaches to managing uncertainty and ensuring operational performance may enhance the already central role control solutions play in establishing the viability of a fusion power plant.

**PRO 5: Extreme Data Algorithms** includes two principal research and development components: methods for in situ, in-memory analysis and reduction of extreme scale simulation data, and methods for efficient ingestion and analysis of extreme-scale fusion experimental data into the new Fusion Data ML Platform (see PRO 7). These capabilities, including improved file system and preprocessing capabilities, will be necessary to manage the amount and speed of data that is expected to be generated by the many fusion codes that use first-principle models when they run on exascale computers. In particular, the scale of data generation in ITER is anticipated to be several orders of magnitude greater than encountered today. ML-derived preprocessing algorithms can increase the throughput and efficiency of collaborative scientific analysis and interpretation of discharge data as it is produced, while further enhancing interpretability through optimized combination with simulation results.

**PRO 6: Data-Enhanced Prediction** will develop algorithms for prediction of key plasma phenomena and plant system states, thus enabling critical real-time and offline health monitoring and fault prediction. ML methods can significantly augment physics models with data-driven prediction algorithms to provide these essential functions. Disruptions represent a particularly essential and challenging prediction requirement for a fusion power plant, since without effective avoidance or effects mitigation, they can cause serious damage to plasma-facing components. Prevention, avoidance, and/or mitigation of disruptions will be enabled or enhanced if the conditions leading to a disruption can be reliably predicted with lead time sufficient for effective control action. In addition to

algorithms for real-time or offline plasma or system state prediction, data-derived algorithms can be used for projection of complex fault and disruption effects for purposes of design and operational analysis, where such effects are difficult to derive from first principles.

**PRO 7: Fusion Data Machine Learning Platform** constitutes a unique cross-cutting collection of research and implementation activities aimed at developing specialized computational resources that will support scalable application of ML/AI methods to fusion problems. The Fusion Data Machine Learning Platform is envisioned as a novel system for managing, formatting, curating, and enabling access to fusion experimental and simulation data for optimal usability in applying ML algorithms. Tasks of this PRO will include the automatic population of the Fusion Data ML Platform, with production and storage of key metadata and labels, as well as methods for rapid selection and retrieval of data to create local training and test sets.

## Foundational Activities and Conclusion

In addition to these seven PROs, a set of foundational activities and resources were identified as essential to the execution of effective ML/AI research that would address fusion problems. These foundational activities and resources include experimental fusion facilities and ongoing research, advances in theoretical fusion science and computational simulation, high performance and exascale computing resources, established and supported connections among university, industry, and government expert groups in the relevant fields, and establishing connections to ITER and other international fusion programs.

The set of high-impact PROs identified in the Advancing Fusion with Machine Learning Research Needs Workshop, together with the foundational activities highlighted, will significantly accelerate and enhance research towards solving outstanding fusion problems, helping to maximize the rate of knowledge gain and progress toward a fusion power plant.

Summary of Priority Research Opportunities identified in Advancing Fusion with Machine Learning Research Needs Workshop.

## Introduction

### Background and Motivation

Dealing with all aspects associated with the generation, movement, and analysis of large sets of data (“big data”) has emerged as a critical issue for fusion and plasma science research. This need is driven by new modes and opportunities of research coupled with the emergence of

more powerful computers, and has led to rapid growth in the adoption of artificial intelligence (AI) techniques and methodologies, including Machine Learning (ML), in the research areas supported by the Department of Energy (DOE) Office of Science (SC) program in Fusion Energy Sciences (FES). Examples of big data science drivers for FES include:

- Collaborations by U.S. scientists on a new generation of overseas superconducting fusion experiments whose pulse lengths are at least an order of magnitude longer than those of current experiments and with additional diagnostic capabilities, leading to a considerable increase of the volume of experimental data, culminating with the anticipated initial operation of the world's first burning plasma experiment—ITER—in 2025;
- Increases in the repetition rate of powerful laser systems coupled to x-ray drivers in the area of high energy density laboratory plasmas (HEDLP). The upgrade of the Linac Coherent Light Source (LCLS) at SLAC will increase the repetition rate from 1 kHz to 1 MHz. Future experiments at the Matter in Extreme Conditions (MEC) instrument will have to deal with big data acquired at a rate of terabytes per second compared to the current rate of megabytes per second.
- Increases in the fidelity and level of integration of fusion and plasma science simulations needed to resolve multiphysics and multiscale problems, which are enabled by advances in high performance computing hardware and associated progress in computational algorithms, and which are accompanied by orders of magnitude increases in the volume of generated data. This need is also expected to increase as the fusion energy sciences program focuses on the development of modeling capabilities and preparing to take advantage of the soon-to-be available exascale computing systems.
- The potential of ML methodologies to address critical challenges in fusion energy science, such as the prediction of potentially disruptive plasma phenomena in tokamaks.
- The potential of ML and AI to optimize the performance of fusion experiments using real-time analysis of diagnostic data, and through expanded integration of first principles and reduced plasma models into advanced control algorithms.

At the same time, the DOE SC program in Advanced Scientific Computing Research (ASCR) has been supporting foundational research in computer science and applied mathematics to develop robust ML and AI capabilities that address the needs of multiple SC programs.

Because of their transformative potential, ML and AI are also among the top Research and Development (R&D)

priorities of the Administration as described in the July 31, 2018, Office of Management and Budget (OMB) Memo on the FY 2020 Administration R&D Budget Priorities [51], where fundamental and applied AI research, including machine learning, is listed among the areas where continued leadership is critically important to our nation's national security and economic competitiveness.

Finally, the Fusion Energy Sciences Advisory Committee (FESAC), in its 2018 report on “Transformative Enabling Capabilities (TEC) for Efficient Advancement Toward Fusion Energy,” [15], included the areas of mathematical control, machine learning, and artificial intelligence as part of its Tier 1 “Advanced Algorithms” TEC recommendation.

### Purpose of Workshop and Charge

The fusion and plasma science communities, recognizing the potential of ML/AI and data science more broadly, have organized a number of information-gathering activities in these areas. These include IAEA Technical Meetings on Fusion Data Processing Validation and Analysis, Scientific Discovery through Advanced Computing (SciDAC) project meetings focused on ML, and a mini-conference on “Machine Learning, Data Science, and Artificial Intelligence in Plasma Research” held during the 2018 meeting of the APS Division of Plasma Physics (DPP). However, a need remained to assess the potential for ML/AI impacting key research problems in these communities in order to identify gaps and opportunities, and derive maximum benefit from synergies. This report describes the results of a joint FES/ASCR DOE-sponsored Research Needs Workshop on Advancing Fusion with Machine Learning, held April 30–May 2, 2019, in Gaithersburg, MD.

The objectives of the workshop were to:

- Identify areas in the fusion science supported by FES (including burning plasma science/materials science, and discovery plasma science) where application of ML and AI can have transformative impacts;
- Identify unique needs, research opportunities, and associated gaps in ML and AI that can be addressed through targeted partnerships between fusion and plasma scientists on the one hand and applied mathematicians or computer scientists on the other, to broaden the applicability of ML/AI solutions across all areas under the FES mission;
- Identify synergies and leverage opportunities within SC and DOE and also outside DOE, including private industry; and
- Identify research principles for maximizing effectiveness of applying ML methods to fusion problems.

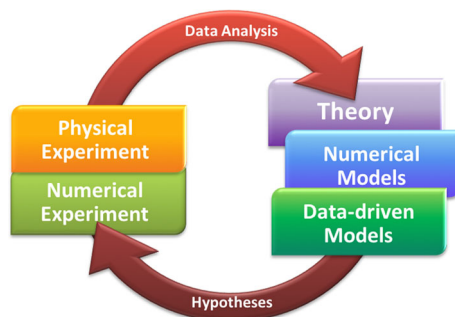
The workshop identified a set of seven Priority Research Opportunities (PROs) that can inform future research efforts in ML/AI and build a community of next-generation researchers in this area. Section 3 following describes these PRO's in detail. Section 4 describes foundational and supporting programmatic activities that are essential to enable effective application of ML/AI methods to research areas identified. Section 5 provides a summary of key results and conclusions from the study.

## Priority Research Opportunities

The Priority Research Opportunities identified in the workshop and below are areas of research in which advances made through application of ML/AI techniques can enable revolutionary, transformative progress in fusion science and related fields. These research areas balance the potential for fusion science advancement with the technology requirements of the research. Each description includes a summary of the research topic, a discussion of the fusion problem elements involved, identification of key gaps in relevant ML/AI methods that should be addressed to enable application to the PRO, and identification of guidelines that can help maximize the effectiveness of the research in each case.

### PRO 1: Science Discovery with Machine Learning

The scientific process constitutes a virtuous cycle of data interpretation to generate models and hypotheses, application of models to design experiments, and experimental execution to generate the data to test hypotheses and revise models. The introduction of ML and AI into the scientific process for hypothesis generation and the design of experiments promises to significantly accelerate this cycle. Traditionally, bottlenecks in the scientific process have



**Fig. 1** Scientific discovery with machine learning includes approaches to bridging gaps in theoretical understanding through the identification of missing effects using large datasets; accelerating hypothesis generation and testing; and optimizing experimental planning to help speed up progress in gaining new knowledge

included insufficient data, insufficient access to experimental facilities, and the speed at which data can be analyzed to generate revised models and the next hypotheses. Machine learning has already demonstrated promise in accelerating the analysis of data and the generation of data-driven models (e.g. [20, 49]), but the anticipated increase in our ability to generate data and the latency of human-in-the-loop hypothesis generation and experimental design will continue to limit our scientific throughput. The vision of this PRO is an integrated process that helps guide, optimize, automate and improve the effectiveness of laboratory and numerical experiments, and augments data analysis and hypothesis generation to accelerate scientific discovery (Fig. 1).

### Fusion Problem Elements

The ultimate goal of plasma confinement research is the attainment of the conditions needed for sustained, controlled fusion. Unfortunately, the high-dimensional space of parameters describing possible plasma confinement conditions makes the optimization of performance difficult, and our incomplete understanding of the many competing physical processes further hampers our progress. The scientific process of making a hypothesis, performing experiments to test it, and improving the understanding of the underlying physics based on the results is well established. However, in fusion, performing experiments is often costly and rare, and although experimentation is guided by simulations and physics knowledge, the characteristic time scale of each cycle of the scientific process can be weeks to years. The ability to prioritize and plan experiments in order to maximize potential knowledge gain and optimize the effectiveness of costly operations can significantly accelerate convergence toward viable fusion energy. Integration of human assessments and statistical inference from data mining has accelerated progress in the confinement of merged compact toroid plasmas [5]. Machine learning approaches applied to design, selection, and steering of tokamak experiments hold promise for similar advances in key plasma performance metrics.

The process of science discovery in fusion (and beyond) is frequently challenged by gaps between the current understanding of first principles and the observed behavior of experimental systems. Machine learning methods can help bridge such gaps by identifying aspects of missing physics and producing hybrid models that can be used in both guiding experiment and completing theory. Gaps in understanding in fields ranging from plasma transport theory to resistive MHD instabilities may benefit from such approaches.

Progress could also be greatly accelerated if analysis and interpretive simulation could be performed in closer to real



time. By building data-driven models based on prior experiments, existing simulations, and incoming experimental results that could be refined iteratively, we could more efficiently explore the space of possible plasma confinement conditions. Experimentalists would be able to make better use of the limited facility availability by learning on the fly and adjusting their experimental conditions with greater speed and intent. By taking a step further and incorporating AI into the process, experimentation could be accelerated even more through the freedom of intelligent agents to pose new hypotheses based on the entirety of available data and even to run preliminary experiments to further aid experimentalists.

### Machine Learning Methods

ML and related fields, like data mining, provide methods for classification, regression, clustering, feature identification, and dimensionality reduction. Fusion experimental campaigns could benefit greatly from the ability to automatically classify or cluster experimental results; build predictive models from data that can be interrogated to identify fruitful future directions for experiments; identify features or anomalies that correlate with improved fusion performance; and identify lower dimensional (invariant manifold) structure in high dimensional data.

For example, random forest algorithms have been applied to the development of porous materials to identify effective input variables and accelerate progress toward desirable performance [49]. Deep learning techniques have also amplified small signals in large datasets to illuminate correlations and trends otherwise not visible [33, 66].

Machine learning has already exhibited utility in optimizing fusion experiments, with demonstrated improvements to fusion performance in both magnetic [5] and inertial [20] confinement approaches. Techniques applied have included algorithms for directing the experimental conditions themselves, as well as hybrid ML models that combine simulation and experimental data and which were sufficiently generalizable to guide the optimization of new experiments. Statistical and reinforcement learning techniques [47] for reasoning and planning could be leveraged to take these existing efforts to the next level, where intelligent agents aid the experimentalist by taking learned features, correlations, etc. and proposing the most beneficial next experiments—perhaps even conducting some preliminary experiments without human intervention.

Beyond accelerating the experimental process, ML techniques could also uncover the physics of the gap between physics models and experimental data. The use of hybrid models, i.e., transformation layers that map simulation predictions to experimental reality [17, 34], can yield insights into the physics that is missing from simulations

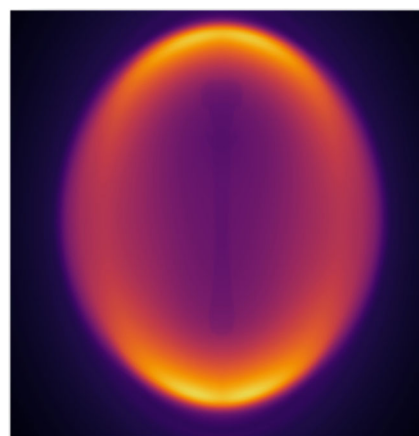
[20]. Separately, one could use ML to reveal correlations in the differences between modeled and measured data to discover features that provide insight into physics missing in existing models.

A range of ML algorithms for feature identification and selection and for data reduction will be applicable here. These include, for example, Principal Component Analysis (PCA) and its variants, random forests, and neural network-based autoencoders, with the objective of identifying correlations between identified features. For example, autoencoders could be combined with physically meaningful latent variables in order to generate falsifiable hypotheses, e.g., to test scaling laws, which in turn can be used to generate experiments.

### Gaps

Perhaps the biggest obstacle in applying new techniques from data science for hypothesis generation and experimental design is the availability of data. First, despite the sense that we are awash in data, in reality, we often have too little scientific data to properly train models. In fusion, experimental data is limited by available diagnostics, experiments that cannot be reproduced at a sufficient frequency, and a lack of infrastructure and policies to easily share data. Furthermore, even with access to the existing data, there is still the obstacle that these data have not been properly curated for easy use by others (see PRO 7).

Beyond building the necessary infrastructure, there are other beneficial circumstances that will help apply data science to hypothesis generation and experimental design. First, new sensors and experimental capabilities, like high repetition rate lasers, promise to increase the experimental



**Fig. 2** A supervised machine learning algorithm trained on a multi-petabyte dataset of inertial confinement fusion simulations has identified a class of implosions predicted to robustly achieve high yield, even in the presence of drive variations and hydrodynamic perturbations [57]

throughput rate. ML techniques like transfer learning are also showing some promise in fine-tuning application-specific models on the upper layers of neural networks developed for other applications where data is abundant [17, 34]. Finally, fusion has a long history of numerical simulation, and there have been recent successes in using simulation data constrained by experimental data to build data-driven models [57] (Fig. 2).

Aside from the availability of data, there is still much work to be done in ML and AI to improve the techniques and make their use more systematic. It is well known that many ML methods have low accuracy and suffer from a lack of robustness—that is, they can be easily tricked into misclassifying an image or fail to return consistent results if the order of data in the training process is permuted [50, 69]. Furthermore, most learning methods fail to provide an interpretable model—a model that is easily interrogated to understand why it makes the connections and correlations it does [14]. Finally, there are still open questions as to the best way to include known physical principles in machine learning models [32, 60, 75]. For instance, we know invariants such as conservation properties, and we would want any learned model to respect these laws. One approach to ensuring this might be to add constraints to the models to restrict the data to certain lower-dimensional manifolds. How this can be done efficiently in data-driven models without injecting biases not justified by physics is an important open research question that still needs to be addressed.

### Research Guidelines and Topical Examples

The principal guideline for ML and AI hypothesis generation and experimental design is caution. As noted above, there are still many gaps in our knowledge and methods, and skepticism is healthy. Like their simulation counterparts, the data-driven models must be verified to ensure proper implementation before any attempt at validation is done, and numerical simulation to generate the training data can play a role in this process since the numerical model is known. Validation against independent accepted results must be done before using the developed techniques to ensure their validity; it is important to verify that the correlations learned by (for example) the autoencoder are real and meaningful. Finally, the uncertainty of data-driven models, a function of the uncertainty of the data and the structure of the assumed model, should also be quantified.

It is important that the community develops metrics to measure the success and progress for research under this PRO related to the new science delivered. Certainly, the rate and quality of data generated through ML- and AI-enabled hypothesis generation and experimental design are two important metrics that can be quantified. It will be

more difficult to measure the transformation of that data into knowledge, since this step will still involve human insight and creativity.

Modern ML and AI techniques are not sufficiently mature to be treated as black box technologies. Both the fusion and the data science communities will benefit from close collaboration. In particular, since statistics is foundational to both data science and experimental design, there is an urgent need to engage more statisticians in bridging the communities, as well as to advance the techniques in ways that accommodate the characteristics of the fusion problems. There will also be an immediate need for data engineers who can help build the infrastructure to efficiently and openly share data in the fusion community, as it exists, e.g., in the high energy density (HED) and climate communities. Such an effort by its very nature will require close collaboration between the fusion scientists and computer scientists with expertise in data management.

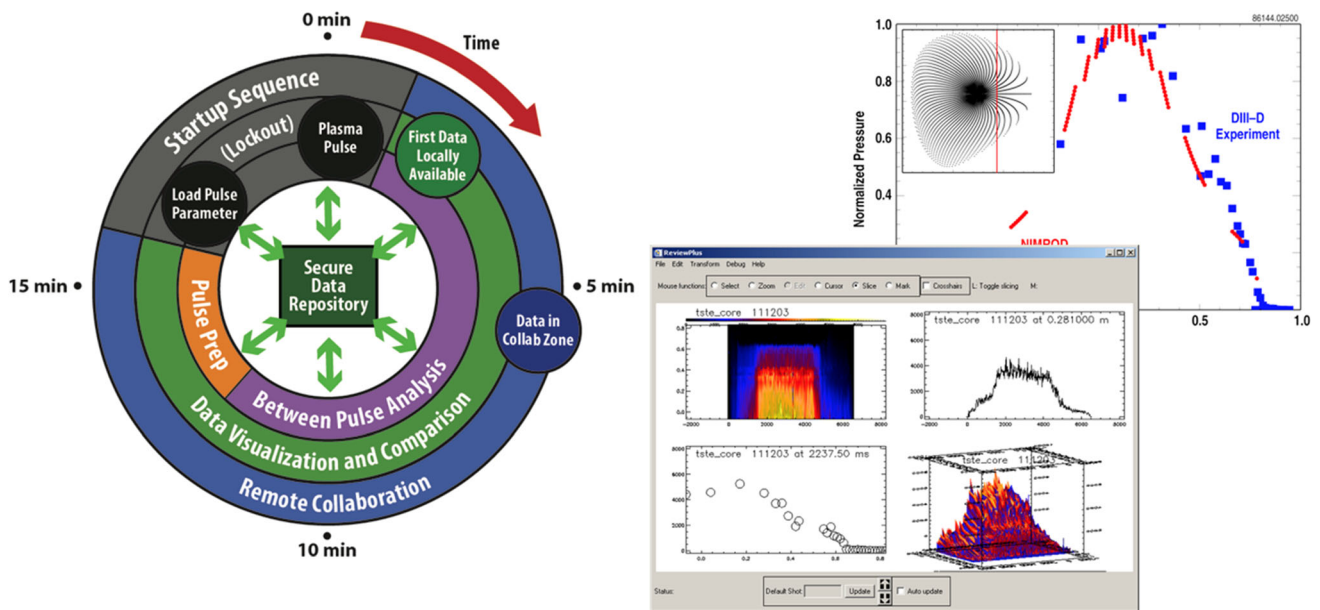
Data engineering will continue to be a challenge for the fusion community, which has not yet defined community standards or adopted completely open data policies. This PRO will require large amounts of both simulation and experimental data to have significant impact. These problems are not unique to the PRO, however (nor to the fusion community), and will need to be addressed if ML and AI are to deliver their full promise for the Fusion Energy Sciences.

### PRO 2: Machine Learning Boosted Diagnostics

Accurately and rapidly diagnosing magnetic confinement plasmas is extremely challenging, and will become more challenging for burning plasmas and power plants due primarily to increased neutron flux and reduced access. Applications of machine learning methods can “boost,” or maximize the information extracted from measurements by augmenting interpretation with data-driven models, accomplishing systematic integration of multiple data sources, and generating synthetic diagnostics (i.e. inference of quantities that are not or cannot be directly measured). Among the additional information extractable from diagnostic measurements through machine learning classifiers are metadata features and classes that can enable or improve the effectiveness of supervised learning for a host of cross-cutting applications (synergistic with a machine learning data platform: PRO 7).

### Fusion Problem Elements

The advancement of fusion science and its application as an energy source depends significantly on the ability to diagnose the plasma. Thorough measurement is needed not only to enhance our scientific understanding of the plasma



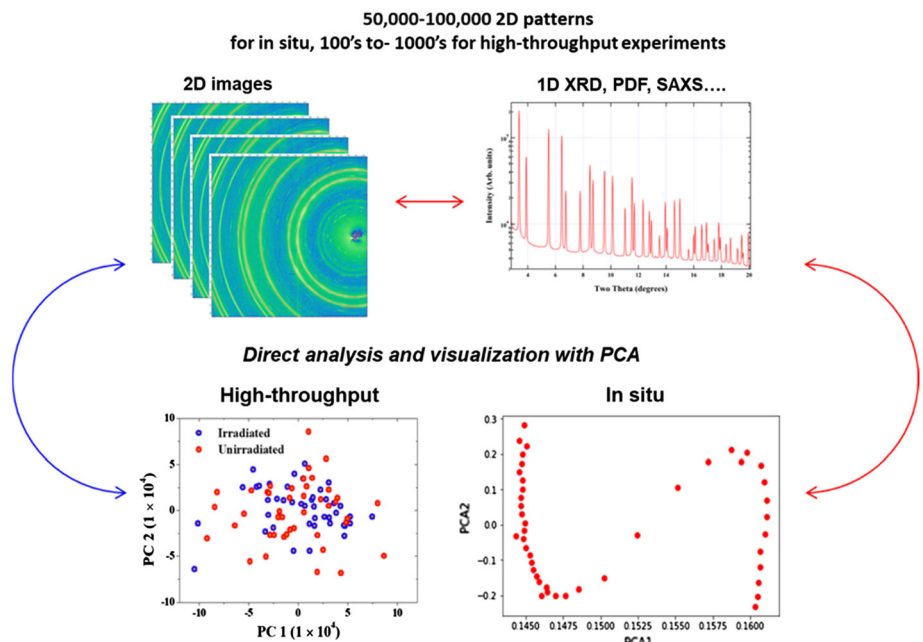
**Fig. 3** The shot cycle in tokamak experiments includes many diagnostic data handling and analysis steps that could be enhanced or enabled by ML methods. These processes include interpretation of

profile data, interpretation of fluctuation spectra, determination of particle and energy balances, and mapping of MHD stability throughout the discharge

state, but also to provide the necessary inputs used in control systems. Diagnosing fusion plasmas becomes ever more challenging as we enter the burning plasma era, since the presence of neutrons and the lack of diagnostic access to the core plasma make the set of suitable measurements available quite limited. Hence there is a need in general to maximize the plasma and system state information extracted from the available diagnostics.

Thorough measurements of the intrinsic quantities (pressure, density, fields) and structures (instability modes, shapes) of a fusion plasma are essential for validation of models, prediction of future behavior, and design of future experiments and facilities. Many diagnostics deployed on experimental facilities, however, only indirectly measure the quantities of interest (QoI), which then have to be inferred (Fig. 3). The inference of the QoI has traditionally relied on relatively simple analytic techniques, which has

**Fig. 4** Maximizing the information extracted from the large number of diagnostics in tokamak experiments requires extensive real-time and post-experiment analysis that could be enhanced or enabled by ML methods (PCA principal component analysis)





limited the quantities that can be robustly inferred. For instance, x-ray images of compressed inertial confinement fusion cores are typically used only to infer the size of the x-ray emitting region of the core. Similarly, the presence of magnetohydrodynamic (MHD) activity internal to magnetically confined plasmas is inferred based on the signatures of this activity measured by magnetic sensors located outside the plasma. However, it is clear that these measurements encode information about the 3D structure of the core, albeit projected and convolved onto the image plane and thereby obfuscated (Fig. 4). The use of advanced ML/AI techniques has the potential to reveal hitherto hidden quantities of this type, allowing fusion scientists to infer higher level QoI from data and accelerating understanding of fusion plasmas in the future.

Another example is the challenge of reconstructing three-dimensional plasma equilibrium states in tokamaks and stellarators consistent with observed diagnostics. Reconstructed three-dimensional plasma profiles have a number of important uses, such as, study of 3D effects (edge-localized modes), inputs to disruption codes, real-time feedback controls and others. In the absence of accurate forward models, computationally expensive loop refinement techniques [11, 39, 64] are presently the standard techniques for plasma reconstructions. Advanced ML/AI techniques have the potential to not only discover non-trivial and complicated data-driven physics models otherwise inaccessible via traditional first-principle methods, but also accelerate the discovery process after the models are trained.

ML can also contribute to mapping from very noisy signals to meaningful plasma behaviors. These relationships are non-trivial and analytical forward models constructed using first-principle physics are rarely adequate to capture the complicated interplay of device diagnostics on the internal plasma states. ML approaches are ideal for these purposes as they do not assume a priori knowledge of inter-dependencies; on the contrary, the correlations and non-trivial dependencies of the measured diagnostics on the internal plasma state are learned from the experimental data directly using ML approaches which are designed to discover non-linear inter-dependencies and input–output mappings that cannot be constructed via traditional forward-only analytical methods. Since the number of diagnostic signals collected during experimental campaigns is very large, systematic ML analysis can help to identify the important signals and reduce the size of the critical data stream for a given process.

Diagnostic data from fusion experiments tends to lack detailed metadata and contextual information that would enable automated comprehensive analysis and integration of data from multiple devices. For example, tokamak plasma discharges typically transition among several

confinement regimes (e.g. ohmic, L-mode, H-mode, I-mode...), but the data are not routinely tagged with this information. Manual inspection is typically required for such identification. Similarly, many MHD instabilities can be excited in the course of an evolving discharge, but these are not routinely identified to tag discharge data for later large-scale analysis. ML methods can be applied to create classifiers for such phenomena, enabling complex interpretation of diagnostic signals in real-time, between discharges, or retrospectively.

This research area can make substantial use of simulations to interpret and aid in boosting diagnostics. Simulated data can be used to generate classifiers to interpret data and generate metadata, to produce models to augment raw diagnostic signals, and to produce synthetic diagnostic outputs. Thus, a key challenge related to the fusion problems addressed with this PRO is production of the relevant simulation tools to support the research process.

### Machine Learning Methods

There are a number of existing ML approaches to create new diagnostics out of noisy plasma measurement. One approach is to treat desired higher-level diagnostics as hidden random variables and then use Bayesian methods to infer the most likely values of these variables given the noisy measurements [68]. Such higher-level diagnostics may include physics variables that are not well measured by current diagnostics. Having such physics variables is critical to understanding complex plasma behaviors.

In some cases, the forward models that could be used to interpret the measured signals are quite expensive, which has the practical impact that reconstructions of the physics parameters is a specialized activity and only a small fraction of the data is fully analyzed. Machine learning could aid this analysis through generating surrogates for these models, and through amortized inference techniques that would accelerate the inverse mapping between observables and the quantities of interest.

Even in traditional analysis of detector data, low resolution and signal quality can be a barrier to analysis. Noise reduction and super-resolution techniques can prove invaluable in this sphere. The typical workflow for noise reduction and super-resolution is a semi-supervised learning approach, where synthetic clean and high-resolution data is created, and downsampled and corrupted in preparation for training [13].

Generation of useful metadata from raw diagnostic signals requires applying varying levels of algorithmic sophistication. Some features that are not yet identified in existing datasets can be extracted and associated with relevant signals through application of well-defined algorithms or heuristics. For example, determination of

tokamak plasma temperature edge pedestal characteristics can be done with automated heuristic algorithms that identify the location of the pedestal and fit local profiles to selected nonlinear functions. Other features may be available in electronically-accessible logs or online documentation. However, identification of many features of interest now requires human inspection and/or complex processing of multiple signals. For example, Thomson scattering spectra often must be examined for validity, high signal-to-noise, and low contamination from other light sources before fitting to extract electron temperature and density. Identifying the presence of a tearing mode island benefits from human inspection of multiple measurements to correlate magnetic signatures, profile measurements, confinement impacts, and other characteristics that reflect mode growth.

Machine learning (ML) methods can enable automation of such complex inspection procedures by emulating or replacing the assessment skills of human analysts. Sophisticated mathematical algorithms that may execute a series of analyses to arrive at a classification decision can be encapsulated in more efficient, faster executing forms that will enable large-scale application to very large fusion datasets. Application of specific classification algorithms can produce critical metadata that in turn will enable extensive supervised learning studies to be accomplished. Unsupervised learning methods will also allow identification of features of interest not yet apparent to a human inspector. For example, tokamak safety factor profiles of certain classes may be signatures of tearing-robust operating states. A currently unrecognized combination of pressure profile and current density profile may correlate with an improved level of confinement.

Development and application of such algorithms to US fusion data across many devices provides the opportunity to study and combine data from a wide range of plasma regimes, size and temporal scales, etc.... This capability will accelerate and improve access to machine-independent fundamental understanding of relevant physics.

In addition to formal machine learning methods, sophisticated simulations will be required to produce the classifiers and models needed to augment and interpret raw diagnostic signals, as well as to produce synthetic diagnostic outputs. A key aspect of the research in this PRO will therefore include coordinating the use of appropriate physics model-based simulation datasets with application of machine learning tools.

### Gaps

The fundamental issue in ML/AI boosting of diagnostics is the difficulty of validating the inferred quantities and physics phenomena. This emphasizes the importance of

formulating the ML/AI problem in a way that is physically constrained (i.e. informed by some level of understanding of the relevant physical dynamics). Additionally, since the relevant ML models must often be trained on simulation data, which are themselves approximations of physical reality, it is crucial to realize that any phenomena not accounted for in the simulations cannot be modeled by these methods.

Specialized tools are needed for the many types of signal processing required for this PRO. These will require dedicated effort due to the range of solutions and level of specialization involved (e.g. tailoring for data interpretation algorithms for each device individually). Providing the ability to fuse multiple signals reflecting common phenomena and verify such mappings, as well as interpreting and fusing data from multiple experimental devices, will also be required.

In particular, enabling cross-machine comparisons and data fusion will require standardizing comparable data in various ways. Exploiting previous efforts (e.g. IMAS and the ITER Data Model [26]), or developing new approaches to standardizing data/signal models is an important enabler of data fusion and large scale multi-machine ML analysis. Normalization of quantities to produce machine-independence is one approach likely to be important.

### Research Guidelines and Topical Examples

Effectiveness of research in this PRO can be maximized by:

- Checking proposed approaches by application to artificial known data such as that produced by simulations
- Developing and confirming methods for identifying physically interesting phenomena and screening for unphysical outputs
- Developing tools for and generating a large number of simulations to enable boosting of data analysis, enable generation of synthetic diagnostics, and contribute to deriving classifiers for automated metadata definition
- Addressing specific needs of fusion power plant designs, with relevant limitations to diagnostic access and consistency with high neutron flux impacts on diagnostics. Quantifying the degree to which machine learning-generated and/or other models can maximize the effectiveness of limited diagnostics for power plant state monitoring and control.

### PRO 3: Model Extraction and Reduction

Machine learning methods have the potential to accelerate and, in some cases, extend our simulation-based modeling approaches by taking advantage of the large quantity of

data generated from experiments and simulations. In this context, extraction methods may discover models to understand complex data generating mechanisms, e.g., development of macroscopic descriptions of microscopic dynamics. Similarly, the goal of model reduction is to isolate dominant physical mechanisms for the purpose of accelerating computation. These reduced models can be used to accelerate simulations for scale bridging, complex multicomponent engineered systems (e.g. tokamaks), uncertainty quantification, or computational design.

### Fusion Problem Elements

Progress in fusion research and engineering relies heavily on complex modeling for analysis and design of experiments as well as for planning and designing new devices. Modeling for fusion is particularly challenging due to the vast range of coupled physics phenomena and length/time scales. The multi-scale/multi-physics modeling results in significant computational burdens, leading first-principle modeling efforts down the path to exascale computing and motivating the development of reduced models to make applications more practical. In order to maintain high fidelity, these reduced models are still typically quite computationally intensive, making activities like design optimization and uncertainty quantification challenging. Furthermore, gaps in theory exist that make direct application of first-principle modeling difficult, leading to the need for empirical models for certain phenomena.

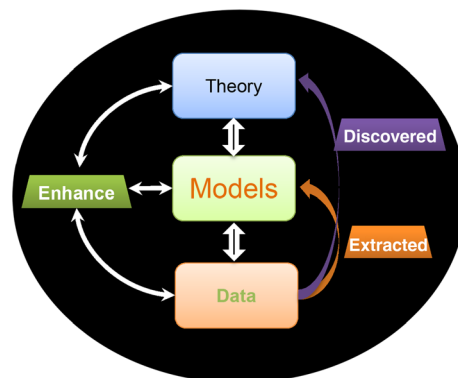
The broadest fusion modeling approach, referred to as whole device modeling, aims to perform time-dependent predictive device modeling to assess performance for physics and engineering design, as well as to provide interpretive analysis of experimental results, combining models and diagnostics to estimate the state of the system during a discharge. These applications require uncertainty quantification and often numerical optimization. Due to the range of applications and requirements (e.g., scoping future machines, planning a specific plasma discharge, real-time forecasting of plasma behavior) there are a variety of accuracy and calculation time requirements, motivating the development of model hierarchies: targeting high fidelity accuracy to faster-than-real-time execution.

Model-reduction aims to lower the computational cost of models while still capturing the dominant behavior of the system. This can be used to facilitate scale bridging and time scale separation, e.g., by generating fast surrogate models of phenomena at small spatial/temporal scales that can be used within models for larger spatial/temporal scales. As an example, in the area of studying plasma-driven degradation of divertor and first wall materials in tokamaks, molecular dynamics simulations provide insight into microscopic mechanisms and can be linked to ab initio

simulations through interatomic potentials, which are surrogate models for the energy of the quantum mechanical system (see [73]). More generally, surrogate models can be used to generate model closures for microscopic descriptions. These are conventionally constructed by phenomenological constitutive relations. However, coupling to higher fidelity codes through the use of surrogate models could yield better, faster solutions to the closure problem. While typical methods of model reduction often require careful consideration of the trade-off between computational cost and the accuracy lost by neglecting terms in a model, machine learning tools provide efficient methods for fitting and optimizing reduced models based on data developed by high fidelity codes, which can in many cases enable reduced models with significantly less computational cost while maintaining high fidelity.

Despite significant advances in theory-based modeling, gaps in understanding exist that could, in some cases, be filled with dynamic models or model parameters derived from experimental data. For example, empirical models for turbulent transport coefficients, fast ion interactions, and plasma boundary interactions could enable locally accurate modeling despite the difficulty of accurately modeling these coupled phenomena from first principles. This activity, often referred to as model extraction or discovery, can take the form of parametric models, e.g., fitting coefficients of linear models, or non-parametric models, e.g., neural networks. Extracted models are a key aspect of the scientific method and, interpretability of the resulting model can help guide experiments and physics understanding and provide a link between theory and experimental data.

Building on this idea of model extraction, ML can augment the experiment/theory scientific workflow with direct integration of models as a tool to drive innovation by bridging the data heavy experimental side and the theory side of the scientific enterprise. Figure 5 depicts this



**Fig. 5** Machine learning-driven models can augment the scientific method by bridging theory and experiment. (Image courtesy C. Michoski, UT-Austin)

interaction explicitly. Note the importance of the iteration between theory and experiment, while this is representative of the scientific method, returning to theory from a data driven model is a step often skipped but essential to generate knowledge from ML techniques.

**Machine Learning Methods**

**Model Reduction** Many of the existing and widely used machine learning tools can be directly applied to model reduction in fusion problems. The specific tools to be used will depend on the type of model and the planned applications of the reduced model. For models that are approximately static, flexible regression approaches like artificial neural networks and Gaussian process regression can readily be used. The flexibility of these approaches can enable fitting of available data to an arbitrary accuracy, such that the approach is only limited by the availability of high fidelity model data and the constraints on model complexity (computational cost to train and/or evaluate). Hyperparameter tuning methods, e.g., genetic algorithms and Markov chain Monte Carlo (MCMC) methods, can be used to optimize the trade-off of model accuracy and complexity. For high dimensional problems, it may be desirable to extract a reduced set of features from the input and/or output space, which can be accomplished with methods like Principal Component Analysis, autoencoders, or convolutional neural networks (e.g. [4, 23, 33, 40]). For developing reduced models of dynamic systems, approaches to identifying stateful models, including linear state-space system identification methods and recurrent neural networks, like long short-term memory networks can be used.

**Model Extraction** Machine learning methods also enable extraction of powerful models from experimental data. By performing advanced data analytics, new and hidden structures within the data are extracted to develop an accurate modeling framework, e.g. [27, 70]. This can lead to discovery of new physics through direct use of data to determine analytic models that generate the observed physics, e.g. [6, 42, 53, 60]. In this way parsimonious parameterized representations are discovered that minimize the mismatch between theory and data, but also potentially reveal hidden physics at play within the integrated multi-physics and engineering systems. Machine learning can also provide data-enabled enhancement [44]. In this process, ML can be used to take theoretical models and enhance them with data, or experimental data acquisition can be enhanced with theory and models. Similarly, data from empirical models can be used to enrich theoretical computational models.

This area presents a number of opportunities for development of novel methods for determining governing equations of physical phenomena. The current approach to deriving governing equations is to develop a hypothesis based on theoretical ideas. This hypothesis is checked, and challenged against experimental data. This process iterates to some informal notion of convergence. Note that typically data is not actively integrated in a way that maximizes its utility. Approaches developed in response to this PRO will allow the data to accelerate the development of an unknown model.

The specific models developed will be highly dependent on the end goal of the model itself. For instance, there may be cases where the model is used as a black box for making predictions. In this instance, the range of veracity of the model would be based on a rigorous verification/validation exercise. In other cases, the model will be used to enhance understanding of the underlying mechanisms. Here it is important that the model is interpretable so that a governing equation can be discovered. In either case, embedding known physics in the learning process as a constraint or prior will be essential to guarantee that the developed model represents a physical process.

While still in its infancy, there has been some limited development of these types of methods. Note that one intriguing potential of the methods themselves is that, in principle, they are not limited to a particular model that we want to learn (e.g. in some cases they go beyond parameter fitting), and that they should be able to extract the unknown

**Example ML techniques**

1 Sparse regression:  
Determine sparse coefficients

$$\operatorname{argmin}_{\xi} \|\partial_t U - \Theta \xi\|_2^2 + \lambda \|\xi\|_0$$

Subject to a matrix of measurements and candidate PDE terms

Library of candidate terms (R) measured from data

2 Operator Regression

Determination of Fourier coefficients

$$F[u] = \nabla \cdot \sum_i \mathcal{F}^{-1} \{ \kappa g_i(|\kappa|) \mathcal{F} \{ h_i(u) \} \}$$

Generalized Moving Least Squares (GMLS) polynomial learning

Fig. 6 ML techniques applicable to model derivation (1) sparse regression [Image courtesy of E.P. Alves and F. Fiuza, SLAC], and (2) operator regression [56]



operator directly. However, careful validation of the recovered model is necessary to guarantee physical consistency and absence of unphysical spurious effects. For ease of presentation, we reduce the approaches to two broad classes of methods; symbolic and sparse regression, and operator regression, see Fig. 6.

Sparse regression techniques use a dictionary of possible operators and nonlinear functions to determine a PDE or ODE operator that best matches the observations. It may seem that a simple regression approach (e.g. linear regression) maybe sufficient for this application. However, this type of approach may yield an unwieldy linear combination of operators whose relative combinations must be weighed against each other for successful interpretation. The key to sparse regression is to select the minimal set of terms that match the observed data. In [9], compressive sensing is used to discover the governing equations used in nonlinear dynamical systems.

An early version of the operator regression is [60] where the authors introduce a Physics Informed Neural Networks (PINNs) technology. In this approach, a neural network is trained to match observed function values with a penalization of the model residual to ensure the function satisfies a known PDE (e.g. the physics constraint), and/or a physical property (e.g. mass conservation). An alternative idea is to learn the discrete form of the operators. In this way the terms (coefficient functions, spatial operators, etc....) of the PDE are determined by an ML regression over the data. Recent work in [56] explore learning the coefficients of a Fourier expansion where the coefficients are represented by a neural network. An alternative approach to operator learning in the presence of spatially sparse experimental data is based on using Generalized Moving Least Squares (GMLS) techniques. These provide, in general, approximations of linear functionals (e.g., differential or integral operators) given measurements or known values of the action of the functional on a set of points. In the simple case of approximation of a function, where the functionals are Dirac's deltas, given a set of sparse measurements, GMLS does not provide a surrogate (e.g. a polynomial) in a closed form, but a tool to compute a value of such function at any given point. Specifically, such surrogate is a combination of polynomial basis functions with space-dependent coefficients [46].

**Machine Learned Interatomic Potentials** Interatomic potentials (IAP) represent an important advance for improving the quality of microscopic models used in fusion device simulation. The growth of computational power and algorithmic improvements have greatly increased our ability to accurately calculate energies and forces of small configurations of atoms using quantum electronic structure methods (QM) such as Density Functional Theory (DFT).

Nonetheless, the  $O(N^3)$  scaling of these methods with the number of electrons makes it impractical to apply QM methods to systems bigger than a few hundred atoms. Molecular dynamics simulations retain linear scaling in the number of atoms by writing the energy and forces as some kind of tractable function of the local atomic environment. Conventional potentials use model forms based on particular chemical and physical concepts (Embedded Atom Method, Bond Order). These produce compact computationally-efficient force models that provide good qualitative models, but they cannot match QM results over a reasonably broad range of configurations. In recent years, machine-learning data-driven approaches have emerged as an alternative. A very general set of local descriptors, rather than a strict mathematical function, is used as the model form. This means that the limiting source of error is the availability of high-accuracy training data rather than the functional form that describes the IAP. The ML-IAP approach is especially useful for systems involving strong electronic bonding interactions between atoms of different chemical elements, because such interactions are difficult to capture using simple IAP models designed for pure elements. A rapid exploration of different regression methods (ANNs, GP, parametric regression), different local descriptors (Two-body and Three-body Symmetry Functions, Moment Tensors, Fourier Invariants), and different kinds of training data has occurred, and many promising approaches have been identified (GPA, ANN, SNAP, Moment Tensors).

One potential approach is to use machine learning techniques to construct these surrogates. Two classes of technique have been identified, first a top-down training technique that solves an inversion problem for IAP parameters from physical properties. In the bottom-up training, the ML model is trained to reproduce higher-accuracy results from a more expensive calculation (e.g. Density Functional Theory).

### Gaps

The approaches discussed above demonstrate that ML offers powerful new tools to tackle critical problems in fusion plasma research. Thus, from both the computational modeling, and fusion energy science perspectives, this is an exciting new research topic that remains largely unexplored for basic plasma and fusion engineering. Ultimately this could prove to revolutionize the process of scientific discovery and improve the fusion communities' toolset by better integrating data from disparate components of an engineered device. Given these opportunities, there are a number of challenges that must be addressed.

An often-neglected task is the analysis of robustness and numerical convergence of the surrogate. As in any other



discretization technique, we need to make sure that the surrogate (1) matches expected results (e.g. linear solution of a diffusion problem for constant source terms), (2) converges to “exact” or manufactured solutions as we add training points to the learning set, or increase the complexity of the surrogate (e.g. increase the number of layers in a neural network). However, this is an idealized scenario; in fact, we cannot guarantee convergence of a neural network with respect to hyper parameters. In this regard, a sensitivity analysis to network parameters (e.g., layers, nodes, bias) would improve the understanding of how such parameters affect the learning process and interact with each other. Further, understanding how the machine learning models are robust with respect to noisy and uncertain data will be essential. Quantifying the effect of this model and being able to answer the question of “how much data is required” will guide the application of these methods to regimes where they can be most effective.

A way to achieve these goals is to test the ML algorithm on manufactured solutions and compare consistency tests and numerical convergence analysis with standard discretization methods, for which the behavior is well understood. An even better approach would be a mathematical analysis of the algorithm that would provide rigorous estimates of the approximation errors.

According to the ML method being used, the enforcement of physics constraints is pursued in different ways. For PINNs, they are added as penalization terms to the “loss” function so that the mathematical model and other physical properties are (weakly) satisfied through optimization. For GMLS, the physics can be included in the functional basis used for the approximation, e.g. basis functions could be divergence free in case of incompressible flow simulations. In the Fourier learning approach conservation properties can be enforced using structural properties when the model form is chosen. Simply learning the divergence of the flux, as opposed to the ODE source effectively enforces conservation. Additionally, the loss function can also be modified to enforce that physical properties are weakly satisfied.

When the set of model parameters to be estimated becomes large, the learning problem could become highly ill-conditioned or ill-posed (as for PDE constrained optimization); this challenge can be overcome by adding regularization terms to the loss function (in case of optimization-based approach), by increasing the number of training points or improving their quality (e.g. better location). Moreover, with larger parameter sets, and physics defined on large 3D and 2D domains, computational cost for the training can be a significant factor. These challenges are familiar to PDE-constrained optimization that can suffer from very long runtimes because of the need to solve the underlying PDE multiple times. This is

especially relevant for simulations of transient dynamics where each forward simulation itself can be a large computational burden.

Finally, given the intrinsic need for data in generating these models, the quality and quantity of fusion science data is critical for designing and applying methods. A critical piece will be to instrument diagnostics with good meta data and precisely record the type of data being recorded, the particulars of the experiment (see PRO 2, ML Boosted Diagnostics, and PRO 5, Extreme Data Algorithms), the frequency of collection, and other relevant descriptions of the data.

### Research Guidelines and Topical Examples

Guidelines to help maximize effectiveness of research in this area include:

- Tools must be made for a broad community of users, with mechanisms for high availability, open evaluation, cooperation, and communication
- Explicit focus on assessments and well-defined metrics for: model accuracy, stability and robustness, regions of validity, uncertainty quantification and error bounds
- Methods for embedding and controlling relative weighting on physics constraints should be addressed
- Methods for managing uncertainty when combining data
- Tools for validation and verification for models
- Open-data provision for trained models used in publication
- Incorporation of end-user demands for interpretability, ranges of accuracy
- Development of benchmark problems for initiating cross-community collaboration

Candidate topics for model extraction and reduction research include:

- Extraction of model representations from experimental data, including turbulent transport coefficient scalings
- Generation of physics-constrained data-driven models and data-augmented first-principle models, including model representations of resistive instabilities, heating and current drive effectiveness, and divertor plasma dynamics
- Reduction of time-consuming computational processes in complex physics codes, including whole device models
- Determination of interatomic potential models and materials characteristics

### PRO 4: Control Augmentation with Machine Learning

The achievement of controlled thermonuclear fusion critically depends on quantifiably effective and robust control of plasma operating characteristics to maintain an optimal balance of stability and fusion performance. This optimal operating point must be sustained for months in a successful power plant, with vanishingly small probability of termination from system faults, plasma fluctuations, and plasma instabilities. Such a demanding level of performance and reliability in a mission-critical fusion power plant can only be provided by the methods of mathematical control design, which can be significantly augmented by machine learning. This research opportunity involves the development of ML methods and their application to the derivation of models required for high reliability control; development of real-time data analysis/interpretation systems to optimize measurements for control; and the design of optimized trajectories and control algorithms.

#### Fusion Problem Elements

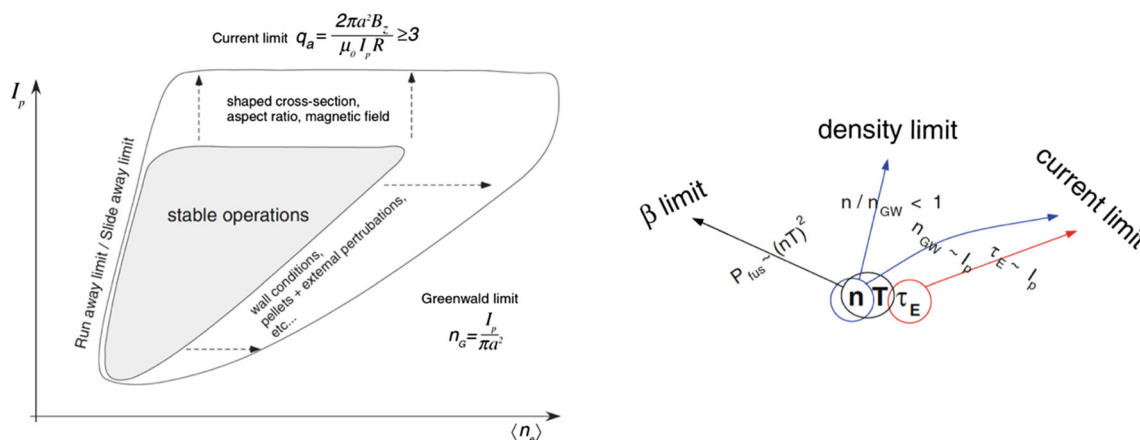
For an economical fusion power plant, fusion gain  $Q$ —defined as output fusion power divided by the input power—has to be large. Unfortunately, any attempts to increase  $n$  (density),  $T$  (temperature), or  $\tau_E$  (confinement time), render plasma less stable and we hit stability limits as illustrated in Fig. 7. The achievement of a commercially feasible fusion power plant requires the optimization of the properties of the plasma by controlling it to high performance and away from instabilities which can be damaging to the plant. Achieving these two goals (high performance and stable operations) simultaneously requires a control system that can apply effective mathematical algorithms to

all the diagnostic inputs in order to precisely adjust the actuators available to a fusion power plant.

In addition to regulation of nominal scenarios and explicit stabilization of controllable plasma modes, proximity to potentially disruptive plasma instabilities must be monitored and regulated in real-time. Robust control algorithms are required to prevent reaching unstable regimes, and to steer the plasma back into the stable regimes should the former control be unsuccessful. On the rare occasion when the plasma passes these limits and approaches disruption, the machine investment must be protected with a robustly controlled shut-down sequence. These tasks all require robust real-time control systems, whose design is complex. The lack of a complete forward model for which a provably stable control strategy can be designed makes the task harder and motivates the use of reduced models that capture the essential dynamics of complex physics (e.g., the effect of plasma microturbulence on profile evolution). ML-based approaches have great potential in the development of real-time systems that incorporate fast real-time diagnostics in decision making and control of long-pulse power plant-relevant conditions.

Parallels to many of these challenges may be seen in the development of controllers in autonomous helicopters, snake robots, and more recently self-driving cars. In each case, a complex dynamical system with limited models and complex external influences has benefited from a variety of ML-based controller development [2, 12, 21, 35, 45, 65, 71].

A significant challenge to achievement of effective fusion control solutions is optimal exploitation of diagnostics data. In order to extract maximum information from the diagnostic signals available in a fusion experiment or power plant, a large volume of data must be processed on the fast time scales needed for fusion control. This implies the need for tools for fast real-time synchronous acquisition



**Fig. 7** Left: high performance for tokamak is achieved at the edge of the stable operation regime shown for a standard tokamak. Right: various plasma stability limits are reached when components of fusion gain ( $Q$ ), are increased [25]

and analysis of large data sets. Sending all the information gathered from available diagnostics to a central processing system is not feasible due to bandwidth issues. ML offers methods to analyze the large quantities of data locally, thus sending along only the relevant information such as physics parameters or state/event classification (labels).

Another fusion control problem amenable to ML techniques is the need for appropriate control-oriented models for control design. Such models reduce the complexity of the representation while capturing the essence of the dynamic effects that are relevant for the control application. Plasma models created to gain physics insight into fusion plasma dynamics are generally complex and not architected appropriately for control design purposes. They usually include a diversity of entangled physics effects that may not be relevant for the spatial and temporal scales of interest for control while omitting the ones that are. Additionally, the underlying differential equations of these complex models are not written in a form suitable for control design. Uncertainty measures, crucial for building and quantifying robustness in control design, are also generally missing from such physics models. Plasma and system response models for control synthesis in general should have the following characteristics:

- Should be dynamic models that predict the future state of a plasma or system given the current state and future control inputs
- Should have quantified uncertainty and uncertainty propagation explicitly included
- Should span the relevant temporal and spatial scales for actuation and control
- Should be fully automatic in nature (no need for physicists to adjust parameters to converge or give reasonable results)
- Should be lightweight in computational and data needs since they may have to execute often with limited real-time resources

There are several approaches to developing control-oriented models, ranging from completely physics-based (white box) models to completely empirical models extracted from input–output data (black box models), with models in between that use a combination of some physics and empirical elements (grey box models). ML methods can contribute significantly to both generation and reduction of control-oriented models, augmenting available techniques for deriving both control dynamics and uncertainty quantification. However, recent work [43] suggests that such models need to preserve inherent constraints such as mass or energy conservation requiring the development and analysis of appropriate constrained training and inference.

A third challenge to fusion control with strong potential for application of ML solutions is the design of optimized trajectories and control algorithms. High dimensionality, high uncertainty, and potentially dynamically varying operational limits (e.g. detection and response to an anomalous event), complicate these calculations. For both fusion trajectory and control design, information is combined from a wide range of sources such as MHD stability models, empirical stability boundaries (e.g. Greenwald limit), data-based ML models for plasma behavior, fluid plasma models, and kinetic plasma models. ML techniques are effective in finding optimal trajectories and control variables in these high-dimensional global optimization problems with a diverse set of inputs, and may thus be advantageous for fusion's unique control challenges.

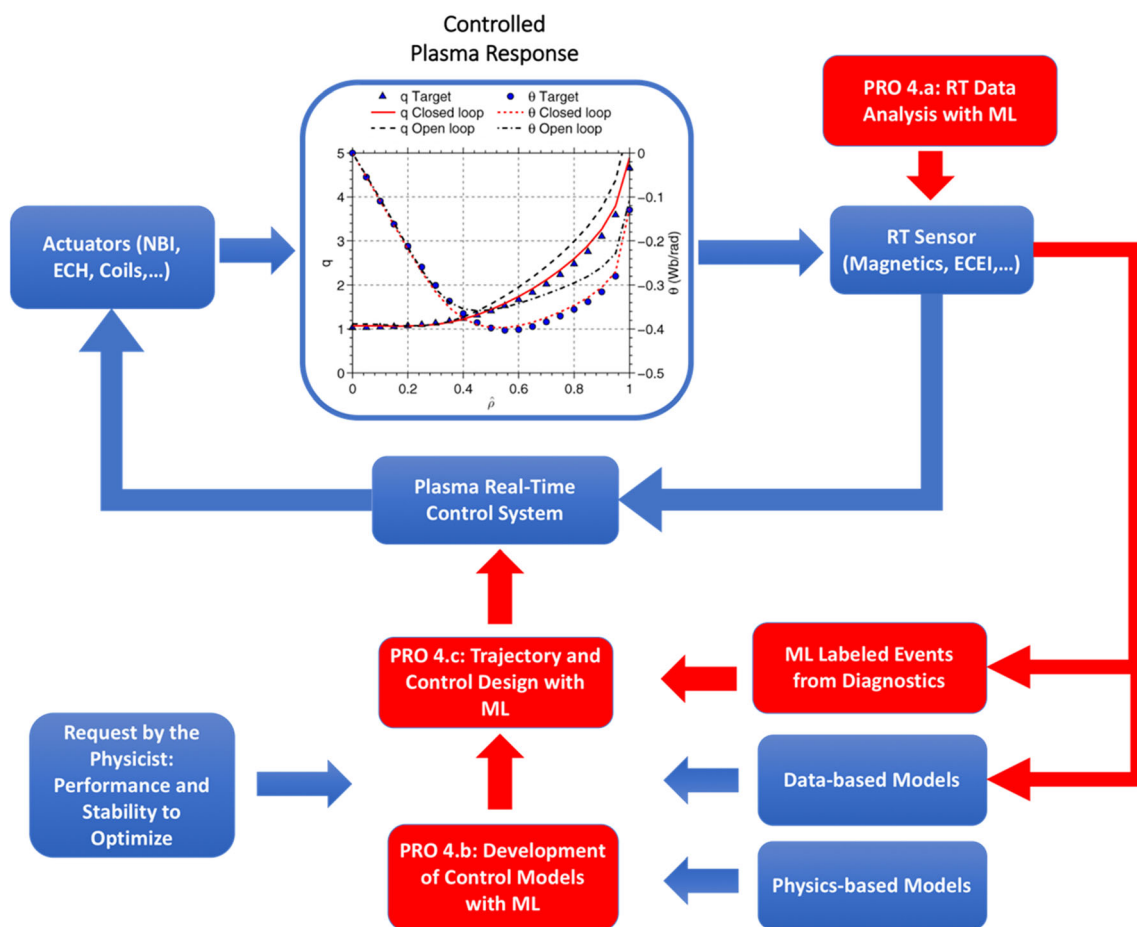
Not all fusion or plasma control problems require continuous real-time solutions. A significant challenge to sustained high performance tokamak operation is determination of exception handling algorithms and scenarios. Exceptions (as the term is used for ITER and power plant design) are off-normal events that can occur in tokamaks and require some change to nominal control in order to sustain operation, prevent disruption, and minimize any damaging effects to the device [24]. Because chains of exceptions may occur in sequence due to a given initial fault scenario, there is potential for combinatorial explosion in exception handling actions, complicating both design and verification of candidate responses. Data-driven methods may enable statistical and systematic approaches to minimize such combinatorial problems.

### Machine Learning Methods

Potentially effective ML approaches to the fusion control issues addressed in this PRO include: (1) hardware/software combinations for fast analysis of sensor data at the source using ideas of Edge-computing for ML, custom computing using FPGAs etc.; (2) analysis of sensor data using ML and physics knowledge to abstract physics information in real or near real time; and (3) ML-based reduction of the data in various ways to reduce the data transfer. Figure 8 illustrates the elements of PRO 4 and their relationship to the plasma control design and implementation process.

### Gaps

Several gaps exist in capability and infrastructure to enable effective use of ML methods for control physics problems. These include availability of appropriate data, and structuring of valuable physics codes to enable application to control-level model generation.



**Fig. 8** Schematic of the interactions among different elements of PRO 4 and the roles they play in the plasma control process

A key limitation that needs to be addressed to enable ML approaches for control is the annotation and labeling of data for application of relevant supervised regression problems. It is anticipated that a nearly automated labeling method for many events of interest is going to be developed. However, the sensitivity of the ML methods to incorrect labeling may be problematic. Frameworks to sanity check the ML mappings against expected physics trends are not necessarily available and need to be developed. Depending on the ML representation used, the navigation problem may or may not have an effective off-the-shelf algorithm, which might warrant modification of optimization/ML/AI approaches.

Because control design fundamentally relies on adequate model representations, the nature of codes used in generating such models sensitively determines the effectiveness of model-based control. In the case of data-driven approaches, codes and simulations that provide data for derivation of control-oriented models must be configured to enable large-scale data generation, and appropriate levels of description for control models that respect constraints from physics [60]. The relevant codes should also

generate uncertainty measures at the same time, a capability missing from many physics codes and associated post-processing resources at present. Methods of obtaining control with quantified stability margins and robustness for these types of combined continuous and discrete systems needs to be employed.

### Research Guidelines and Topical Examples

Effectiveness of research in this PRO can be maximized by:

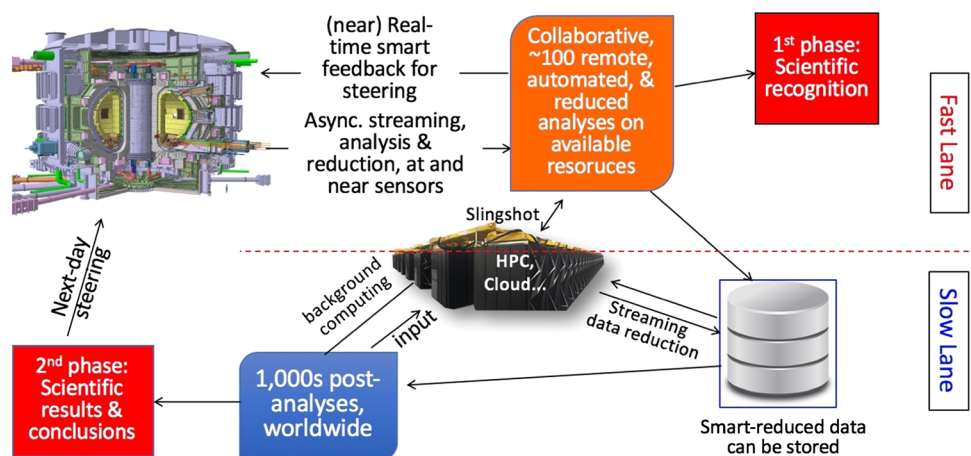
- Ensuring research focuses on interpretations of measurements that maximize the specificity of control-related phenomena. For example, models or predictive algorithms that provide outputs specific to particular instabilities or plasma phenomena to be controlled are most likely to enable effective control [7, 18].
- Providing variables that quantify relative stability, controllability, or proximity to operational boundaries
- Creating real-time calculable quantities wherever possible, that provide sufficient lead time for control action [16]

- Linking derived results to specific relevant temporal and spatial scales for actuation and control that will lead to well-defined control actions
- Following machine learning procedures that enable physics constrained extrapolation to different operating regimes, system conditions, or fusion devices
- Developing robust ML training methods with quantified stability margins and uncertainty that would perform robustly under the dynamic nature of the fusion plasma

## PRO 5: Extreme Data Algorithms

There are two components to the Extreme Data Algorithm priority research opportunity: (a) in situ, in-memory analysis and reduction of extreme scale simulation data as part of a federated, multi-institutional workflow, and (b) ingestion into the new Fusion Data ML Platform (PRO 7) and analysis of extreme scale fusion experimental data for real- or near-real time collaborative experimental research. The former research (a) is required because multiple fusion codes are expected to use first-principle models on Exascale computers with the size and the speed of the data generation being beyond the capability of the filesystem, rendering post-processing problematic. The workflow applying these multiple fusion codes will involve multiple, distributed, federated research institutions, requiring substantial coordination (see Fig. 9). The latter research (b) is needed because the amount and speed of the data generation by burning plasma devices, such as ITER in the full operation DT phase, are anticipated to be several orders of magnitude greater than what is encountered today. Intelligent ingestion into the Data ML Platform, not only the storage of data but also for streaming and subsequent analysis, can allow rapid scientific feedback from the world-wide collaborators to guide experiments, thereby accelerating scientific progress.

**Fig. 9** Federated fusion research workflow system among a fusion power plant experiment, real-time feedback and data analysis operations, and extreme-scale computers distributed nonlocally. A smart machine-learning system needs to be developed to orchestrate the federated workflow, real-time control, feature detection and discovery, automated simulation submission, and statistically combined feedback for next-day experimental design



## Fusion Problem Elements

- In-situ, on-memory analysis and reduction of extreme scale simulation data

Exascale fusion codes, studying physics at first-principle level, will produce massive amounts ( $\sim$  exabytes) of data per run. It is well-known that this volume and velocity of data cannot be streamed to the filesystem or located on permanent storage for post-processing. Therefore, visualizing and interpreting the data as much as possible concurrently from the same HPC memory (in situ), or other network connected HPC memory, is required. Critical data components can be identified, reduced, indexed and compressed to fit the storage requirement and to allow for post-processing. Otherwise, only a small fraction of the data can be written to the file system and moved to storage. If this data is not intelligently handled and critical pieces were not saved, it is possible that the costly simulation would need to be repeated. Automated AI/ML algorithms need to be relied on.

Moreover, extreme scale first-principle simulations will be needed to predict the evolution of the plasma profiles. ML training utilizing the profile-evolution data from simulations on the present tokamaks may not be ideal in predictive profile evolution in possibly different physics regimes, e.g. ITER plasmas. ML-analyzed and reduced data on in situ compute memory must be utilized for this purpose as well.

At the present time, physicists typically save terabytes of data for post-processing, subjectively chosen from previous experience. Scientific discoveries are often made from data that are not experienced before. Moreover, data for deeper first-principle physics—such as turbulence-particle interaction—are often too big to write to file systems and permanent storage.

Fundamental physics data from extreme-scale high-fidelity simulations can also be critically useful in improving



reduced model fluid equations. The greatest needs for the improvement of the fluid equations are in the closure terms—such as the pressure tensor and Reynolds stress. Fundamental physics data from extreme-scale simulations saved for various dimensionless plasma and magnetic field conditions can be used to train ML that could lead to improved closure terms.

- (b) Extreme-scale fusion experimental data for real- or near-real-time collaborative research

The amount of raw data generation by ITER in the full DT operation phase is estimated to be 2.2 PB per day ( $\sim 50$  GB/s), on average, and 0.45 EB of data per year [58]. Clearly, highly challenging data storage and I/O speed issues lie ahead. Without developing a proper global framework, the ability to efficiently store data for post-processing and also for critical near-real-time feedback to the control room, will not be possible. The sheer size and velocity of the data from ITER or a DEMO burning plasma device may not allow sufficiently rapid physics analysis and feedback to the on-going experiment with present methods. Data retrieval and in-depth physics study by world-wide scientists may take months or years before influencing the experimental planning.

A well-designed methodology to populate the Fusion Data ML Platform can allow analysis of diagnostic data streams for feature detection and importance indexing, reduce the raw data accordingly, and compress them at or near the instrument memory. The data also need to be sorted, via indexing, into different tier groups for different level analysis. While the reduced and compressed data are flowing to the Data ML Platform, a quick and streaming AI/ML analysis can be performed. Some scientists could even be working on the virtual experiment, running in parallel to the actual physical experiment. Various quick analysis results can be combined into a comprehensive interpretation via a deep-learning algorithm for presentation to ITER control room scientists.

Once the reduced and compressed data are in the ML Data Platform, more in-depth studies will begin for further scientific discoveries. Many of these studies may utilize numerous quick simulations for statistical answers as well as extreme scale computers for a high-fidelity study. New scientific understanding and discovery results can be presented to experimental scientists for future discharge planning.

Various AI/ML, feature detection, and compression techniques developed in the community can be utilized in the Data ML Platform. The federation system requires close collaboration among applied mathematicians, computer scientists, and plasma physicists. Prompt feedback for ITER experiments could accelerate the scientific progress of ITER and shorten the achievement time of its goal of

ten-times more energy production than the input energy to the plasma.

### Machine Learning Methods

Both topics require fast detection of physically important features. Therefore, data generation must be followed by importance indexing, reduction of raw data, and lossy compression while keeping the loss of physics features below the desired level. Both supervised and unsupervised ML methods can be utilized. Correlation analysis in the  $x$ - $v$  phase space and spectral space can be a useful tool, not often discussed in the ML community but amenable to ML approaches. A key goal is to develop ML tools to meet the operational (during and between-shot) real-time requirements. Highly distributed and GPU-local ML is required. For time-series ML of multiscale physics, utilization of large size high bandwidth memory (HBM), address-shared CPU memory, or non-volatile memory-express (NVMe) can be beneficial. Close collaboration with applied mathematicians and computer scientists is another essential requirement. The wide range of AI/ML solutions developed for other purposes in the community should also be utilized or linked to extreme data algorithm solutions as much as possible.

### Gaps

The main gaps are: (1) the identification and development of various ML tools that can be utilized for fusion physics and that are fast enough for real-time application to streaming data at or near the data source, (2) the development of the Data ML Platform that can utilize ML tools for real-time data analysis, and suggest intelligent ML decision from various real-time analysis results, (3) in the case of the extreme-scale simulations, the predictive AI/ML capability for accomplishing accurate plasma profile evolution on experimental transport time scales based on instantaneous flux information (current approaches are mostly based on trained data from numerous simulations on present experiments, which may not be applicable for extrapolation into unknown physics regimes, such as may be expected in ITER), (4) derivation of training data from a small number of large-scale simulation runs, and (5) increasing the collaboration among fusion physicists, applied mathematicians and computer scientists, which will be critical for the success of this PRO.

### Research Guidelines and Topical Examples

As discussed in this section, real-time analysis of streaming data will require fast and distributed ML and compression algorithms, which necessitates a team effort including

fusion physicists, applied mathematicians and computer scientists. The International Collaboration Framework for Extreme Scale Experiments (ICEE) is an example of such a collaborative research activity [10]. These types of collaborative multidisciplinary activities need to be expanded and extended to accelerate progress. A close collaboration with ITER data scientists is also recommended.

Some specific examples include:

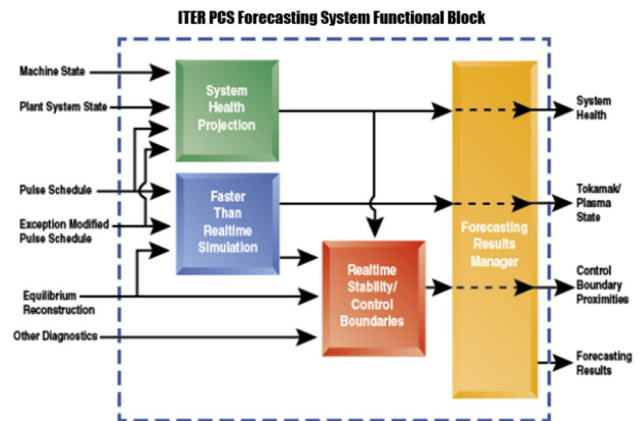
- Extracting parameter dependent families of features to enable flexible post-processing
- The ability to utilize small representative data sets for accurate and robust ML/AI training
- Robust and verifiable ML to ensure generalizability and stability of detection
- Parallelizing ML/AI models in a distributed environment for rapid training
- ML algorithms that can rapidly/accurately detect desired features in situ for data triage
- In-situ systems that can generate privileged metadata (e.g. uncertainty, provenance)
- Addition of first-principle simulation data to the present experimental data for predictive ML on ITER and DEMO performance

## PRO 6: Data-Enhanced Prediction

All complex, mission-critical systems such as nuclear power plants, or commercial and military aircraft, require real-time and between-operation health monitoring and fault prediction in order to ensure high reliability sustained fault-free operation. Prediction of key plasma phenomena and power plant system states is similarly a critical requirement for achievement of fusion energy solutions. Machine learning methods can significantly augment physics models with data-driven prediction algorithms to enable the necessary real-time and offline prediction functions. Disruptions represent a particularly essential and demanding prediction challenge for fusion burning plasma devices, because they can cause serious damage to plasma-facing components, and can result from a wide variety of complex plasma and system states. Prevention, avoidance, and/or mitigation of disruptions will be enabled or enhanced if the conditions leading to a disruption can be reliably predicted with sufficient lead time for effective control action.

### Fusion Problem Elements

Fusion power plants must have sophisticated forecasting systems to enable sufficiently early prediction of fault conditions to enable prevention or correction of such faults and sustainment of high-performance operation.



**Fig. 10** The ITER plasma control system (PCS) forecasting system will include functions to predict plasma evolution under planned control, plant system health and certain classes of impending faults, as well as real-time and projected plasma stability/controllability including likelihood of pre-disruptive and disruptive conditions. Many or all of these functions will be aided or enabled by application of machine learning methods

Asynchronous plasma or system events that can occur under fault conditions and require some active control change are called “exceptions” in ITER and elsewhere. Exceptions that threaten steady operation include impending violation of controllability limits, problematic variance in plasma or system performance, and failure of key subsystems. While many exceptions can be detected as they occur and trigger an “exception handling” response, many must be identified with significant look-ahead capability to enable an effective response. The solution envisioned in the ITER control forecasting system includes Faster-than-Real-Time-Simulation (FRTS), as well as direct projection with reduced models and extrapolation algorithms to predict problematic plasma and system states (see Fig. 10) [24]. It is expected that a viable and economical power plant will require similar functions. Many of these predictor algorithms can be enabled or enhanced by application of machine learning methods.

Nominal, continuous plasma control action (see PRO 4) must reduce the probability of exceptions requiring early termination of a discharge to a very low level (below  $\sim 5\%$  of discharges) in ITER, and to a level comparable to commercial aircraft or other commercial power sources in a fusion power plant ( $\ll 10^{-9}/s$ ). In addition to these levels of performance, effective exception detection, prediction, and handling are required to enable ITER to satisfy its science mission, as well as to enable the viability of a fusion power plant. The complexities of the diagnostic, heating, current drive, magnets, and other subsystems in all burning plasma devices make these kinds of projections highly amenable to predictors generated from large datasets. At the same time, the importance of the envisioned

applications to machine operation and protection place high demands on uncertainty and performance quantification for such predictors.

Prediction of plasma or system state evolution is essential to enable application of controllability and stability assessments to the projected state. While FRTS computational solutions may enable such projection sufficiently fast integrated/whole device models, this function may be aided or enabled by machine learning methods.

Prediction of plasma states with high probability of leading to a disruption, along with uncertainty quantification (UQ) to characterize both the probabilities and the level of uncertainty in the predictive model itself, is a particularly critical requirement for effective exception handling in ITER and beyond [28]. ITER will only be able to survive a small number of unmitigated disruptions at full performance. Potential consequences of unmitigated disruptions include excessive thermal loads at the divertor strikepoints, damaging electromagnetic loads on conducting structures due to induced currents, and generation of high current beams of relativistic electrons which can penetrate and damage the blanket modules. Methods to mitigate disruptions have been and are being tested, but a minimum finite response time of  $\sim 40$  ms is inherent in the favored mitigation methods. This means that a highly credible warning of an impending disruption is required, with at least a 40 ms warning time [41]. Although invoking the mitigation action will minimize damage to a tokamak, it will also interrupt normal machine operation for an appreciable time, and therefore it is highly desirable to avoid disruptions, if possible, and only invoke mitigation actions if a disruption is unavoidable. This requires knowledge of the time-to-disrupt, along with effective exception handling actions that should be taken.

The principal problems are to determine in real time whether or not a plasma discharge is on a trajectory to disrupt in the near future, what is the uncertainty in the prediction, what is the likely severity of the consequences, and which input features (i.e. diagnostic signals) are primarily responsible for the predictor output. In the event that a disruption is predicted, a secondary problem is to determine the time until the disruption event.

The physics of tokamak disruptions is quite complicated, involving a wide range of timescales and spatial scales, and a multitude of possible scenarios. It is not possible now, or in the foreseeable future, to have a first-principle physics-based model of disruptions. However, it is believed that most disruptions involve sequential changes in a number of plasma parameters, occurring over timescales much longer than the actual disruption timescale. These changes may be convolved among multiple plasma parameters, and not necessarily easy to discern. Very large sets of raw disruption data from a number of

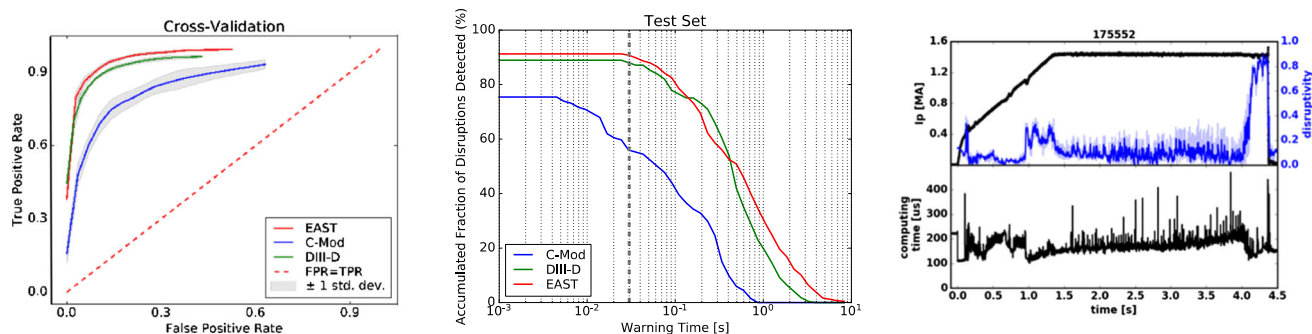
tokamaks already exist, from which refined databases can be constructed and used to train appropriate AI algorithms to predict impending disruptions.

## Machine Learning Methods

Prediction of plasma and system state evolution, and in particular probability of key exceptions and potentially disruptive conditions, will benefit from a broad swath of machine learning techniques. The basic problem is quite general: given the data on the current state and evolutionary history of the system, predict the state evolution over a specified time horizon with the times and probabilities of key events that may occur up to that horizon. Many approaches to classification and regression have potential for successful application. Some state-of-the-art examples include random forests, neural networks, and Gaussian processes, although many approaches are applicable.

Predictions from machine learning models trained on large data sets have been employed in fusion energy research since the early-1990s. For example, Wroblewski et al. [74] employed a neural network to predict high beta disruptions in real-time from many axisymmetric-only input signals, Windsor et al. [72] produced a multi-machine applicable disruption predictor for JET and ASDEX-UG, Rea et al. [61] and Montes et al. [48] demonstrated use of time series data and explicit look-ahead time windows for disruption predictability in Alcator C-Mod, DIII-D, and EAST (see Fig. 11), and Kates-Harbeck [33] demonstrated use of extensive profile measurements in multi-machine disruption prediction for JET and DIII-D with convolutional and recurrent neural networks. Even with the growing use of ML predictions for fusion energy science applications, very little attention has been given to uncertainty quantification. Due to the inherent statistical nature of machine learning algorithms, the comparison of model predictions to data is nontrivial since uncertainty must be considered [67]. The predictive capabilities of a machine learning model are assessed using both the model response as well as the uncertainty, and both aspects are critical to effectiveness of both real-time and offline applications.

Predicting plasma state evolution and the resulting consequences are challenging tasks that typically require the use of computationally expensive physics simulations. The task will benefit from machine learning approaches developed for making inference with such simulations. Emulation is a broad term for machine learning approaches that build approximations or surrogate models that can predict the output of these simulations and do so over many orders of magnitude. These emulators can then be used to make fast predictions (e.g. what are the consequences of a fault or disruption with some set of initial conditions) or



**Fig. 11** The left two plots compare the performances of machine-specific disruption predictors on 3 different tokamaks (EAST, DIII-D, C-Mod). The rightmost plot shows the output of a real time predictor

solve inverse problems (e.g. what are the initial conditions that likely caused some particular disruption). Deep neural networks, Gaussian process, and spline models are state-of-the-art approaches to emulation. General optimization and Bayesian techniques are used for solving inverse problems [22, 31, 36].

### Gaps

Data availability is a noted gap in development of predictor solutions. There are a number of data repositories, but no standardized approach to data collection or formatting. Further, much of the data is incompletely labelled (see PROs 2 and 7).

There are a number of major gaps in the mathematical/ML understanding. First, there is no accepted approach for incorporating physics knowledge into machine learning algorithms. Current solutions are mostly ad hoc and problem specific. Second, there is no unified framework for incorporating uncertainty into certain types of models. Models based on probability do this naturally, but other approaches, deep neural networks, lack this basis. Ad hoc solutions, such as dropout in neural networks, only partially address this issue and may not provide desired results. In addition, there is no solid framework for ensuring success in extrapolation. Indeed, this may prove impossible [52].

A fundamental challenge to application of ML techniques to predictors for ITER and beyond is the ability to extrapolate from present devices (or early operation of a commissioned burning plasma device) to full operational regimes. The ability to extrapolate classifiers and predictors beyond their training dataset and quantify the limitations of such extrapolation are active areas of research, and will certainly require coupling to physics understanding to maximize the ability to extrapolate. In addition to advances in ML mathematics, careful curation of data and design of algorithms based on physics understanding (e.g. use of

installed in the DIII-D plasma control system, demonstrating an effective warning time of several hundred ms before disruption [48])

appropriate dimensionless input variables, hybrid model generation) are expected to help address this challenge.

### Research Guidelines and Topical Examples

Approaches that incorporate uncertainty quantification, such as those built on probabilistic models, have particularly strong potential. These predictions will typically be used to make decisions about control and mitigation. Well-calibrated prediction probabilities and credible intervals are needed for decision makers and algorithm to balance the likelihood of events with the potential severity of consequences.

Approaches that provide some hope of extrapolation are also important. Machine learning models for these problems will necessarily be developed on small-scale plasma and tokamak experiments and then applied to large scale machines where training data is nonexistent and failure is nearly unacceptable. Although extrapolation is always a difficult task, approaches that are physics-informed give greater confidence of success. Incorporating physical principles and insight from physics simulations is a key ingredient.

Finally, real-time performance for prediction is an important requirement. The highest impact predictor algorithms for real-time application will directly inform control action, e.g. enabling regulation of relevant states for real-time continuous disruption prevention or enabling effective asynchronous disruption avoidance. Predictors used for triggering of machine protection disruption effects-mitigation techniques must ultimately be sufficiently fast and reliable to facilitate this function in real-time.

Any research that addresses the needs and guidelines presented above is worthy of consideration. An inclusive approach should be used when considering proposed research topics. Given that, there are a number of specific examples that could be envisioned.



Physics-informed machine learning is a broad area of current research that seeks to incorporate physical principles into machine learning approaches. As an example, these principles could be used to design the structure of a neural network or the covariance function of a Gaussian process. Incorporating physical principles would give greater confidence in the robustness of machine learning approaches and their ability to extrapolate.

Interpretable machine learning is also another broad area of current research. This area seeks to develop machine learning methods in which the “reasoning” behind the predictions is understandable to the human user. Here again, these approaches should give greater confidence in robustness and extrapolations. Interpretable methods may also make disruption mitigation more feasible by providing clues to what signals the ML algorithm is using to make prediction.

Uncertainty quantification is one of the guiding principles and could be a major research topic in this area. Many of the most flexible approaches to machine learning, particularly deep neural networks, have poorly developed notions of predictive uncertainty. Research into this area has great potential because quantified prediction uncertainty is an important component of the decision-making process that the predictions feed.

Many machine learning methods are not robust to perturbations in inputs and thus give unstable predictions. Research, such as the ongoing work in adversarial training, is needed to address this issue [19].

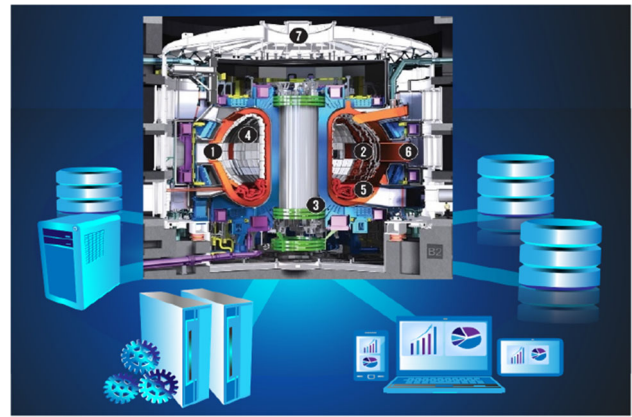
Research into automated feature and representation building is important as it connects with all aspects of this problem. First, it makes prediction considerably simpler when the features themselves are most informative. It also has the potential to improve robustness and interpretability.

Bayesian approaches do a good job at quantifying uncertainty, but work is needed to accelerate these methods with modern estimation schemes like variational inference [8]. This is particularly true for Bayesian inversion tasks to solve for things like initial and boundary conditions that are crucial in disruption mitigation.

Multi-modal learning is concerned with machine learning approaches that combine disparate sources of information [29] This kind of work is crucial in the disruption problem where data from many sensors is combined to make predictions.

## PRO 7: Fusion Data Machine Learning Platform

The majority of present experimental data repositories for fusion are designed for simultaneously visualizing relatively small amounts of data to support both effective consumption in the control room as well as post-experiment studies. Existing data repositories are small on the



**Fig. 12** Vision for a future fusion data machine learning platform that connects tokamak experiments with an advanced storage and data streaming infrastructure that is immediately queryable and enables efficient processing by ML/AI algorithms

scale of ITER’s anticipated needs and for Exascale simulations (there are no such general repositories for simulation data). ML/AI workflows need to read entire data repositories rapidly, something that present systems are not designed to efficiently accomplish. Therefore, this PRO addresses the need for a Data Platform dedicated to the needs of the Fusion community for ML/AI workflows. The vision is that this system (see Fig. 12) will provide unified management of both experimental and simulation data, deal intelligently with compression, allow rapid parallel and distributed access for remote/collaborative use, enable selective access for ML and analytics, and contain all the required metadata to maximize the ability to perform large scale supervised learning.

## Fusion Problem Elements

There is presently an enormous amount of data already available from the many fusion experiments that have been operating worldwide for decades. The devices presently in operation continue to produce amounts of data that increase steadily and yield richer resources for data-driven methods year by year. The advent of very long pulse devices, both already operating and soon to be operating, with even larger diagnostic arrays than historically available, will rapidly multiply the amount of fusion data produced per year worldwide. The large amount of experimental data presently available and growing rapidly has been demonstrated by the long history of ML/AI applications in fusion to be sufficient to enable significant knowledge amplification for accelerated advancement of fusion. Perhaps even more exciting than the potential of knowledge-extraction from experimental data alone, the emerging ability to produce simulation data on extreme scales will increasingly offer opportunities to augment this large



experimental data supply in targeted ways. In particular, the production of simulation data targeted at known gaps in physics understanding will enable integration with experimental data to bridge these gaps and produce specific new understanding in an extremely efficient way.

However, currently the data available to fusion scientists is not ‘shovel ready’ for doing ML/AI. The causes of this are varied, but include differing formats, distributed databases that are not easily linked, different access mechanisms, and lack of adequately labeled data. In the present environment, scientists performing ML/AI research in the Fusion community spend upwards of 70% of their time on data curation. This curation entails finding data, cleaning and normalization, creating labels, writing data in formats that fit their needs, and moving data to their ML platforms. In addition, ML data access patterns perform poorly on existing fusion experimental data ecosystems (including both hardware and software) that are designed to support experiments and therefore present a major bottleneck to progress.

In contrast to centralized experimental data repositories, fusion simulation data is typically organized by individuals or small teams with no unified method of access or discovery. Therefore, performing any ML analysis on simulation data often requires seeking out the simulation scientist or having the ML analyst run their own simulations. Making integrated use of data across different simulation codes is impractical because of the critical barriers of data mapping and non-interoperability of codes and results databases.

An attempt at standardization of data naming conventions in the fusion community is being addressed by the ITER Integrated Modeling and Analysis Suite (IMAS), which provides a hierarchical organization of experimental and modeling data. This convention is rapidly becoming a standard in the community, adopted by international experiments as well as international expert groups such as ITPA. The ITER system provides a partial technical solution for on-the-fly conversion of existing databases to the IMAS format, although with significant limitations and inefficiencies. The effort to map US facilities experimental data to IMAS is only in the early stages, but it provides a candidate abstraction layer that could be used to access data from US fusion facilities in a uniform way.

Data ecosystems and computing environments are critical enabling technologies for data-driven ML/AI efforts. Adherence to high performance computing (HPC) best practices and provisioning of up-to-date, modern software stacks will facilitate effective data ecosystems and computing environments that can leverage state-of-the-art ML/AI tools [59]. At present, no US fusion facility provides a computing hardware infrastructure sufficient for the ML

use case and development of new data ecosystems and computing environments is imperative.

### Machine Learning Methods

Development of an automatically populated Fusion Data ML Platform for both experimental facilities and simulation codes requires proper consideration of the use cases and requirements of AI algorithms for fusion applications. These include production and storage of required metadata and labels. As new feature definitions or extraction algorithms are defined, re-processing existing data may be necessary. Presently, the generation of labels and provenance contexts for data sets is an extremely labor-intensive effort. As a result, it tends to be forgone for most fusion data, with negative consequences for practical application of ML techniques. To address this problem, development of the data platform should consider applying ML approaches such as supervised and unsupervised classifiers or surrogate model generation as a way to provide automated partial metadata information for the archived data. Practically, this should include freely available libraries such as TensorFlow [1], PyTorch [55], Keras [3], etc. Such approaches include:

- Unsupervised anomaly detection
- Long-Short-Term Memory (LSTM) time series regression and clustering
- Convolutional Neural Networks (CNN) for image data
- Natural Language Processing (NLP) methods to derive labels from user annotations and electronic log books
- Synthesis of heterogeneous data types over a range of time scales
- ML-based compression of data

Because there may be no one-size-fits-all solution for storing and accessing fusion data, it is expected that efficient retrieval of data will require advanced algorithms to determine the best manner to plan execution, much like the query planner in a traditional relational database. This could take the form of a heuristic determination, but could also be implemented as an ML algorithm that learns how to determine the most efficient access patterns for the data. Still, the development of the Fusion Data ML Platform will be founded on a number of general data access patterns upon which more advanced data retrieval can be developed. These include, for example, selective access to data sub-regions, retrieval of coarse representations, or direct access to dynamic models with adaptive time sampling and ranges as well as data with limited (bounded) accuracy. A key characteristic is that in all cases data movements should be limited to the information needed, and data transfers should adopt efficient memory layout that avoids wasteful data access, which has the potential to penalize

performance and energy consumption dramatically in modern hardware architectures.

## Gaps

Today research in Fusion for ML/AI is hampered by incomplete datasets and lack of easy access, both by individuals and communities. Current data sets for experimental and simulation data do not have sufficiently mature metadata and labeling, which is required for machine learning applications. Attaching a label to a dataset is a critical step for using the data, and the label will be task-specific (e.g. plasma state is or is not disrupting). Furthermore, there is no centralized and federated data repository from which well-curated data can be gathered. This is true for both domestic facilities and simulation centers, as well as international facilities. This gap prohibits the use of the large volumes of fusion data to be analyzed with ML-based methods. A targeted effort is needed to make it easy to create new labeling for existing and future data collection and to be able to index the labeling information efficiently.

Functional requirements and characteristics of gathering simulation data tend to be very different from those governing access of experimental data. While most experimental data are available in a structured format (e.g. MDSplus), simulation data is often less standardized and frequently stored in user directories or project areas on HPC systems. Mechanisms for identifying relevant states of the simulation code (e.g. versions, corresponding run setups) are likely required to group simulation output from each “era” of the physics contained in that code and input/output formats read and written by the code. Ultimately, this is not fundamentally different from experiments, which often undergo periodic “upgrades,” configuration changes, or diagnostic recalibrations. In both cases, proper interpretation of the data stored in a given file or database requires the appropriate provenance information both for proper encoding/decoding and for informed interpretation (e.g., after recalibration, the same data may be stored in the same way but should be understood differently).

One potential hurdle for producing a large fusion data repository is the incentive to deposit the most recent, high quality experimental and simulation data into a data center shortly after data creation. There are data protection issues for publication and discovery, maturity and vetting of data streams for errors that would need to be re-processed, or potential proprietary status considerations. Incentives must be provided for the community to share data in ways that serve both the individual and community as a whole. Demonstrating the usefulness for a community-wide data sharing platform, and drawing from experience in other

communities (e.g. climate science [63] or cosmology), are potential means to address this gap.

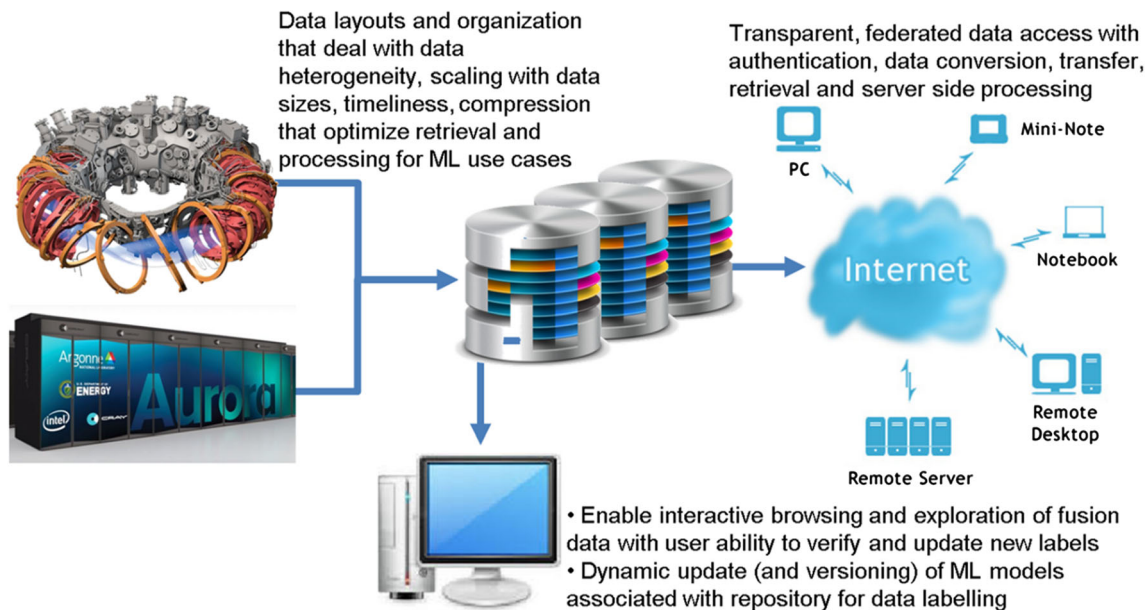
Existing hardware deployments are not designed for the intense I/O and computation generally required by ML/AI [38]. Currently most large-scale computing hardware is optimized for simulation workloads and heavily favors large parallel bulk writes that can be scheduled predictably. This leads to unexpected bottlenecks when creating and accessing the data. This gap needs to be addressed for a successful usage of ML/AI science to support fusion research. As a commonly-encountered co-design problem in which the interaction between the software/algorithmic stack and the physical hardware are tightly coupled, this gap should be solvable with approaches that properly account for their complementary roles in such use cases.

## Research Guidelines and Topical Examples

The workshop highlighted several primary research directions that will help the fusion community fully exploit the potential of ML/AI technology. This section summarizes several key research guidelines for the development of a community Fusion Data ML Platform (see Fig. 13).

Data layouts, organization, and quality are fundamental to storing, accessing, and sharing of fusion data. While smaller metadata may be stored in more traditional databases, large simulation and imaging data require specific work for separate storage to make sure that they do not overwhelm the entire solution. Key aspects to consider involve addressing a variety of data access patterns while avoiding any unnecessary and costly data movement as demonstrated in the PIDX library for imaging and simulation data [37]. For example, access to local and/or reduced models in space, time, or precisions should be achieved without transferring entire data that is simplified only after full access. The heterogeneity of the data may require building different layouts for different data models. Scaling to large data models and collection has to be embedded as a fundamental design principle. Data quality will also be an essential research direction since each data model and use case can be optimized with proper selection of a level of data quality and relative error bounds. Such a concept should also be embedded as a fundamental design principle of the data models and file formats.

Interoperability with standardized data as it is being pursued under the ITER project is a major factor. Unfortunately, the ITER process does not yet fully address the relevant use cases with the development of the data management infrastructure. Specific activities will be needed to develop an API that is compatible with IMAS and allows effective interoperation while maintaining internal storage that is amenable to high performance ML algorithms.



**Fig. 13** Unified solution for experimental and simulation fusion data. Scalable data streaming techniques enable immediate data reorganization, creation of metadata, and training of machine learning models

Interactive browsing and visualization of the data will be a core capability that can enable, for example, a user to verify and update labels generated automatically in addition to creating manual labels. In addition, ML applications greatly benefit from interactive data exploration capabilities. A key advantage is the adjustment and “debugging” of ML models based on the exploration of the data involved in different use cases. In fact, this will be a critical enabling technology for the development of interpretable models via interactive exploration of uncertainties and instabilities in the outcomes based on variations of the input labels [30].

Streaming computation and data distribution are capabilities that are at the core of the community effort. It can allow local development of modern repositories that can be federated as needed and as permitted by data policies without the hurdle of building any mandatory centralized storage. Software components will need to be developed to expose the storage and facilitate the development of common interfaces (e.g. RESTful API). This would not preclude specific efforts that may require more specialized interfaces. Easily deployable and extensible technologies such as Python and Jupyter can be used as scripting layers that expose advanced, efficient components developed in more traditional languages such as C++ [54].

Transparent data access with authentication are technologies that are essential for practical use in a federated environment. Automated data conversion, transfer, retrieval, and potential server-side processing are capabilities that can become major performance bottlenecks and therefore need to be addressed with a specific focus. Secure

data storage, based on variable requirements, must be provided, but authentication systems cannot block scientists and engineers from performing their work. Similarly, the latencies due to data transfers need to be limited. Server-side or client-side execution of queries need to be available and selected intelligently to maximize the overall use of the available community infrastructure.

Reproducibility will be a core design aspect of the data platform. Careful versioning of all the data products, for example, will allow the use of the same exact data for verification of any computation. Complete provenance tracking of the data processing pipelines will also make sure that the same version of a computational component is used. While absolute reproducibility may lead to extreme, unrealistic solutions, the fusion community will need to advocate for proper tradeoffs that allow for sufficient ability to compare results over time (e.g., with published and versioned code and datasets) while maintaining an agile infrastructure that allows fast-paced progress.

## Foundational Resources and Activities

Effective large-scale application of machine learning methods to fusion challenges relies on a substantial amount of infrastructure, foundational resources, and supporting activities. These include experimental fusion facilities and ongoing research in this area; advancement in theoretical fusion science and computational simulation; supported connections among university, industry, and government

expert groups in the relevant fields; and establishing connections to ITER and other international fusion programs.

High performance computing and exascale computing resources underpin virtually all activities identified in this report. Continued development of ML/AI based methods will further increase the requirements associated with these resources, for example, at NERSC (capacity) and OLCF/ALCF (capability). It is difficult to estimate the real future needs, which will often be specific to the problem/question at hand. Capacity increases will be critical for low-resolution simulations, while the higher resolution ones will require improvements in capabilities. With approaches that are data driven, a lot will depend on how much training is done and at what scale. The other drive for the increase in computing demands may be simply the ease of application of UQ and design optimization techniques that were previously out of reach.

Because the science acceleration made possible by ML/AI mathematics depends on data-driven or derived algorithms, strong experimental programs that produce large quantities of specifically useful data are critical to the effort. Close engagement between ML/AI efforts and relevant components of experimental programs can maximize the efficiency and effectiveness of ML/AI applications in extracting additional knowledge. Efforts to format and curate experimental data at point of generation are important for the envisioned Fusion Data ML Platform to best enable use of datasets for large scale analysis.

Specific research programmatic connections between fusion experimental and theoretical programs organized around and dedicated to exploiting ML methods are essential to enable effective use of these transformational approaches. Close community coordination will be instrumental in ensuring data and physics representations are common to experimental efforts and theory/modeling community simulations. A focus of combined teams on ML/AI applications with common science goals will enhance and exploit synergies available.

Advancing fusion through ML/data science is a complex enough endeavor that it requires cross-disciplinary teams of applied mathematicians, machine learning researchers, and fusion scientists. One reason is that the disparate tasks of modeling, experimentation, and analysis of data coming from experiments and model outputs require an iterative process that may eventually lead to new discoveries. With the increasing emphasis on algorithmic data analysis by researchers from various fields, machine learning researchers regularly call for clarifying standards in research and reporting. A recent example is the Nature Comment by Google's Patrick Riley [62], which focuses on three of the common pitfalls: inappropriate splitting of data (training vs test), inadvertent correlation with hidden variables, and incorrectly designed objective functions.

Negative effects of such pitfalls can be minimized in various ways, all requiring awareness of potential vulnerabilities, and diligence in seeking mitigation methods. For example, data splitting into test and training sets typically depends on the randomness of the process, but is frequently done in ways that preserve trends and undesirable correlations. Splitting in multiple, independently different ways can help limit such problems. Consideration of the effects being studied can also help guide a process of data selection that is not necessarily random, but specifically seeks to produce appropriate diversity in the data sets. Meta-analysis of initial training results can serve to illuminate hidden variables that mask the desired effects. Iterative approaches that challenge the choice of objective functions and pose alternatives can help reduce fixation on objectives that address an entirely different problem from that intended. Higher-order approaches to help avoid such pitfalls include applying ML methods in teams made up of experts in the relevant fusion science areas and experts in the relevant areas of mathematics, and challenging results with completely new data generated separately from both training and test suites.

Fusion energy R&D provides unique opportunities for data science research by virtue of the amount of data generated through experiments and computation. Most of the datasets studied and used by ML/AI come from non-scientific applications, and so the data produced by FES are unique. The computing resources available at the various DOE laboratories provide additional appeal to university researchers.

Direct programmatic connections from ML/AI efforts to the ITER program and other international fusion efforts are essential to make best use of emerging data in the coming burning plasma era. In addition to ITER, strong potential exists for synergies through program connections with international long pulse superconducting devices including JT-60SA, EAST, KSTAR, and WEST. Other developing fusion burning plasma designs and devices offering potentially important connections include SPARC, CFETR, and a possible US fusion pilot plant.

## Summary and Conclusions

Machine learning and artificial intelligence are rapidly advancing fields with demonstrated effectiveness in extracting understanding, generating useful models, and producing a variety of important tools from large data systems. These rich fields hold significant promise for accelerating the solution of outstanding fusion problems, ranging from improving understanding of complex plasma phenomena to deriving data-driven models for control design. The joint FES/ASCR research needs workshop on



“Advancing Fusion Science with Machine Learning” identified several Priority Research Opportunities (PROs) with high potential impact of machine learning methods on addressing fusion science problems. These include Science with Machine Learning, Machine Learning Boosted Diagnostics, Model Extraction and Reduction, Control Augmentation with Machine Learning, Extreme Data Algorithms, Data-Enhanced Prediction, and Fusion Data Machine Learning Platform. Together, these PROs will serve to accelerate scientific efforts, and directly contribute to enabling a viable fusion energy source.

Successful execution of research efforts in these areas relies on a set of foundational activities and resources outside the formal scope of the PROs. These include continuing support for experimental fusion facilities, theoretical fusion science and computational simulation efforts, high performance and exascale computing resources, programs and incentives to support connections among university, industry, and government experts in machine learning and statistical inference, and explicit connections to ITER and other international fusion programs.

Investigators leading research projects that strongly integrate ML/AI mathematics and computer applications must remain vigilant to potential pitfalls that have become increasingly apparent in the commercial ML/AI space. In order to avoid errors and inefficiencies that often accompany steep learning curves, these projects should make use of highly-integrated teams including mathematicians and computer scientists having high levels of ML domain expertise with experienced fusion scientists in both experimental and theoretical domains. In addition to personnel and team design, projects themselves should be developed with explicit awareness and mitigation of known potential pitfalls. Development of training and test sets in general should incorporate methods for confirming randomization in relevant latent spaces, supporting uncertainty quantification needs, and enabling strong interpretability where appropriate. Specific goals and targets of each PRO should be well-motivated by need to advance understanding, development of operational solutions for fusion devices, or other similar steps identified on the path to fusion energy.

Burning plasma experiments such as ITER, and eventually engineering test or pilot plant reactors, introduce a combination of particular challenges to the application of ML/AI methods. For example, how can one apply data-driven methods before specific data is available on a device such as ITER? Can actual operational solutions be applied to a reactor, if data-driven algorithms are very sensitive to the details of a particular installation? There are several potential approaches for enabling use of such methods before and as data becomes available. Gradual evolution of models and operational solutions during the ITER research

plan (which develops over more than a decade before DT  $Q = 10$  scenarios are expected to be explored) may be possible, along with similar potential for transfer learning development of relevant models in the commissioning process of a reactor. In addition, development of dimensionless, machine-independent results that can be transferred to new devices may be possible with minimal need for transfer learning (augmentation of an initial model). Establishing the actual reliability of such approaches will require significant research before application to mission-critical environments (and it is possible such approaches will not prove sufficiently reliable, and more traditional operations and control solutions will be required for operating reactors).

The high-impact PROs identified in the Advancing Fusion with Machine Learning Research Needs Workshop, relying on the highlighted foundational activities, have strong potential to significantly accelerate and enhance research on outstanding fusion problems, maximizing the rate of US knowledge gain and progress toward a fusion power plant.

**Acknowledgements** Work supported by U.S. Department of Energy under DE-FC02-04ER54698, DE-AC52-07NA27344, and DE-NA0003525. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

## Compliance with Ethical Standards

**Conflict of interest** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Glossary of Machine Learning Terminology

Artificial intelligence (AI)	The study of computational and algorithmic methods for emulating functions generally viewed as requiring (human) intelligence
------------------------------	---



Big data	A general term referring to the phenomenon of large quantities of data available in various fields along with rapidly growing processing and memory power, enabling use of algorithms and data management approaches that exploit such large data quantities (and quality) for statistical inference	Data driven algorithms	Mathematical algorithms whose behavior is defined on the basis of data analysis (e.g. determination of neural network weights and thresholds from training data) and are intended to operate on additional/separate data (first test data, then application data) to produce some desired analytical outcome
Classification	In the context of machine learning the mathematical process of separating a (large) parameter space into regions (classes) that share common characteristics. Also, the generation and operation of an algorithm that performs such classification for new input data	Deep learning	A general term referring to use of several performance-enhancing approaches to machine learning typically including many-layered (“deep”) neural networks, recurrent neural networks, and convolutional neural networks
Convolutional neural network (CNN)	A type of neural network that performs a mathematical convolution filtering function on the data before passing the results to succeeding layers of the network. Typically refers to a multi-layer perceptron architecture which embeds such convolution filtering processes	Extreme data	A general term referring to amounts and behavior of data streams that occur at challenging levels for the present capability of networks computer platforms, acquisition systems, and human systems, often including limitations imposed by human organizations and collaboration. See also Extreme Scale
Cost function	A mathematical function that represents a penalty or “cost” whose minimization is used as a metric for optimization of some algorithm. Minimizing cost functions that represent mapping errors or worsening performance in some process can produce algorithmic solutions with high accuracy or otherwise desired levels of performance	Extreme scale	Computer algorithms data sizes, and data stream scales that challenge the present state of the art for data handling and computing, often including limitations imposed by human organizations and collaboration. See also Extreme Data
		Hyperparameters	A parameter used to control the learning process in developing a machine learning algorithm. For example the width and depth chosen for a neural network (i.e. number of layers, and number of neurons in each layer) constitute important hyperparameters for typical neural network problems

Loss (or loss function)	A function that maps the results of an operation to a new quantity representing some performance metric or “cost” associated with the operation. Often the loss is a measure of the difference between the true value of (output) data from a data set and an algorithm’s predicted value of that output. The loss (function) characterizes the success of the algorithm for a given set of data and thus enables optimization of the algorithm on that basis	Reinforcement learning	A type of machine learning in which a software agent takes action based on a system environmental state in order to maximize a cumulative reward. Reinforcement learning-trained algorithms have proven extremely successful in navigating completely-specified environmental domains including the games of chess, shogi, and Go, as well as many video games
Machine learning (ML)	The field of computational mathematics that deals with algorithms whose behavior is determined by data rather than explicit programming. It thus encompasses methods of data analysis that automate the generation of algorithms on the basis of data and defined performance metrics	Receiver operating characteristic (ROC) curve	A curve characterizing the performance of an algorithm by plotting the True Positive rate versus the False Positive rate of the process. Originally applied to radar signal interpretation and decision-making the ROC is a fundamental way to characterize performance of classification algorithms
Multi-layer perceptron (MLP)	A type of neural network consisting of multiple layers of neurons that act as perceptrons (i.e. threshold-activated functions that perform a binary classification operation) in processing signals and passing the result to succeeding layers	Supervised learning	A form of machine learning in which a mapping algorithm is determined from a set of labeled example input–output pairs. A supervised learning algorithm typically analyzes a set of training data to produce a desired mapping and applies the resulting mapping to a specified test set to quantify performance outside the training data. See also Unsupervised Learning
Neural network (NN)	A network of interconnected neurons either biological or artificial. Artificial neural networks typically consist of layers of neuron functions, each of which performs an “activation function” operation based on its input signals that results in a threshold behavior of the output, which is passed to succeeding neurons or layers of neurons	Test set	Data set used to test the performance of an algorithm trained on a training set of data. Performance metrics often used include accuracy and precision of prediction as well as extrapolability beyond the input data domain used for a training set

Training set	Data set used to train an algorithm to accomplish a given goal (e.g. minimization of a cost or loss function classification of a data space, matching of input–output characteristics)
Transfer learning	An application of machine learning methods that seeks to apply knowledge gained in solving one problem to solving a second problem that is (typically) related in some way. Examples include adapting and extending a neural network trained on images of one type (e.g. cats) in order to enable identification of images of a second type (e.g. dogs)
Universal approximation theorem	Theorem that demonstrates that a neural network of sufficient width in number of neurons and (at least) one hidden layer, can fit an arbitrary continuous function provided the inputs are limited to a finite range
Unsupervised learning	A form of machine learning that seeks to identify previously-unrecognized patterns in an unlabeled dataset with minimal human supervision in the process. See also Supervised Learning

## References

1. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, R. Jozefowicz, Y. Jia, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, M. Schuster, R. Monga, S. Moore, D. Murray, C. Olah, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org
2. P. Abbeel, A. Coates, M. Quigley, A.Y. Ng, An application of reinforcement learning to aerobatic helicopter Flight, in *NIPS* (2007)
3. R. Atienza, *Advanced Deep Learning with Keras: Apply Deep Learning Techniques, Autoencoders, Gans, Variational Autoencoders, Deep Reinforcement Learning, Policy Gradients, and More* (Packt Publishing, Birmingham, 2018)
4. J.A. Atwell, B.B. King, Proper orthogonal decomposition for reduced basis feedback controllers for parabolic equations. *Math. Comput. Model.* **33**(1–3), 1–19 (2001)
5. E.A. Baltz, E. Trask, M. Binderbauer, M. Dikovsky, H. Gota, R. Mendoza, J.C. Platt, P.F. Riley, Achievement of sustained net plasma heating in a fusion experiment with the optometrist algorithm. *Nat. Sci. Rep.* **7**, 1 (2017). <https://doi.org/10.1038/s41598-017-06645-7>
6. J. Berg, K. Nystrom, Data-driven discovery of PDEs in complex datasets. *J. Comput. Phys.* **384**, 239–252 (2019)
7. J. Berkery et al., A reduced resistive wall mode kinetic stability model for disruption forecasting. *Phys. Plasmas* **24**, 056103 (2017)
8. D.M. Blei, A. Kucukelbir, J.D. McAuliffe, Variational inference: a review for statisticians. *J. Am. Stat. Assoc.* **112**(518), 859–877 (2017)
9. S.L. Brunton, J.L. Proctor, J.N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci.* **113**(15), 3932–3937 (2016)
10. J. Choi et al, “ICEE: Wide-area In Transit Data Processing Framework For Near Real-Time Scientific Applications,” Report from 2013 Supercomputing Workshop, <http://sc13.supercomputing.org/sites/default/files/WorkshopsArchive/pdfs/wp148s1.pdf>
11. M. Cianciosa, A. Wingen, S.P. Hirshman, S.K. Seal, E.A. Unterberg, R.S. Wilcox, P. Piovesan, L. Lao, F. Turco, Helical core reconstruction of a DIII-D hybrid scenario tokamak discharge. *Nucl. Fusion* (2017). <https://doi.org/10.1088/1741-4326/aa6f82>
12. A. Coates, P. Abbeel, A.Y. Ng, Learning for control from multiple demonstrations, in *ICML* (2008)
13. C. Dong et al., Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(2), 295–307 (2016)
14. F. Doshi-Velez, *Towards a Rigorous Science of Interpretable Machine Learning* (2017), [arXiv:1702.08608](https://arxiv.org/abs/1702.08608)
15. FESAC, *Transformative Enabling Capabilities (TEC) for Efficient Advancement Toward Fusion Energy*, Fusion Energy Sciences Advisory Committee TEC Panel (2017), [https://science.energy.gov/~media/fes/fesac/pdf/2018/TEC\\_Report\\_IFeb20181.pdf](https://science.energy.gov/~media/fes/fesac/pdf/2018/TEC_Report_IFeb20181.pdf)
16. Y. Fu, D. Eldon, K. Erickson, K. Kleijwegt, L. Lupin-Jimenez, M.D. Boyer, N. Eidiatis, N. Barbour, O. Izacard, E. Kolemen, Real-time machine learning control at DIII-D for disruption and tearing mode avoidance, *Phys. Plasmas* (2019) (in review)
17. J. Gaffney et al., Making inertial confinement fusion models more predictive. *Phys. Plasmas* **26**(8), 082704 (2019)
18. A.S. Glasser, E. Kolemen, A.H. Glasser, A Riccati solution for the ideal MHD plasma response with applications to real-time stability control. *Phys. Plasmas* **25**, 032507 (2018)
19. I. Goodfellow et al., Generative adversarial nets, in *Advances in Neural Information Processing Systems* (2014)
20. V. Gopalaswamy et al., Tripled yield in direct-drive laser fusion through statistical modelling. *Nature* **565**, 581 (2019)
21. T. Haarnoja, A. Zhou, P. Abbeel, S. Levine, Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor, in *International Conference on Machine Learning (ICML)* (2018)
22. D. Higdon et al., Computer model calibration using high-dimensional output. *J. Am. Stat. Assoc.* **103**(482), 570–583 (2008)
23. G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks. *Science* **313**(5786), 504–507 (2006)
24. D. Humphreys, G. Ambrosino, P. de Vries, F. Felici, S.H. Kim, G. Jackson, A. Kallenbach, E. Kolemen, J. Lister, D. Moreau, A. Pironi, G. Raupp, O. Sauter, E. Schuster, J. Snipes, W.

- Treutterer, M. Walker, A. Welander, A. Winter, L. Zabeo, Novel aspects of plasma control in ITER. *Phys. Plasmas* **22**, 021806 (2015)
25. V. Igochine, *Active Control of Magnetohydrodynamic Instabilities in Hot Plasmas. Springer Series on Atomic, Optical, and Plasma Physics* (Springer, Berlin, 2015). ISBN 978-3-662-44221-0
  26. F. Imbeaux, S.D. Pinches, J.B. Lister, Y. Buravand, T. Casper, B. Duval, B. Guillerminet, M. Hosokawa, W. Houlberg, P. Huynh, S.H. Kim, G. Manduchi, M. Owsiak, B. Palak, M. Plociennik, G. Rouault, O. Sauter, P. Strand, Design and first applications of the ITER integrated modeling and analysis suite. *Nucl. Fusion* **55**, 123006 (2015)
  27. R. Iten, T. Metger, H. Wilming, L. del Rio, R. Renner, *Discovering Physical Concepts with Neural Networks* (2018), [arXiv:1807.10300](https://arxiv.org/abs/1807.10300)
  28. ITER Physics Basis, Y. Gribov et al., Chapter 8: plasma operation and control. *Nucl. Fusion* **47**, S385 (2007)
  29. D. Jeske, M. Xie (eds.), *Applied Stochastic Models in Business and Industry Special Issue on Data Fusion*, vol. 34, no. 1 (Wiley, New York, 2018)
  30. C. Johnson, C. Hansen, *The Visualization Handbook* (Academic Press Inc, New York, 2004)
  31. D.R. Jones, M. Schonlau, W.J. Welch, Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **13**(4), 455–492 (1998)
  32. A. Karpatne et al., *Physics-Guided Neural Networks (PGNN): An Application in Lake Temperature Modeling* (2017), [arXiv preprint arXiv:1710.11431](https://arxiv.org/abs/1710.11431)
  33. J. Kates-Harbeck, A. Svyatkovskiy, W. Tang, Predicting disruptive instabilities in controlled fusion plasmas through deep learning. *Nature* **568**, 526 (2019)
  34. L. Kegelmeyer, Image analysis and machine learning for NIF optics inspection, Technical report, Lawrence Livermore National Lab. (LLNL), Livermore, CA (United States) (2018)
  35. A. Kendall, J. Hawke, D. Janz, P. Mazur, D. Reda, J.-M. Allen, V.-D. Lam, A. Bewley, A. Shah, *Learning to Drive in a Day* (2018), [arXiv:1807.00412](https://arxiv.org/abs/1807.00412)
  36. M.C. Kennedy, A. O'Hagan, Bayesian calibration of computer models. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **63**(3), 425–464 (2001)
  37. S. Kumar, J. Edwards, P.-T. Bremer, A. Knoll, C. Christensen, V. Vishwanath, P. Carns, J.A. Schmidt, V. Pascucci, Efficient I/O and storage of adaptive-resolution data, in *SC14: International Conference for High Performance Computing, Networking, Storage and Analysis* (2014)
  38. T. Kurth, S. Treichler, J. Romero, M. Mudigonda, N. Luehr, E. Phillips, A. Mahesh, M. Matheson, J. Deslippe, M. Fatica, Prabhat, M. Houston, Exascale deep learning for climate analytics, in *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC '18)* (IEEE Press, 2018)
  39. S.A. Lazerson et al., Three-dimensional equilibrium reconstruction on the DIII-D device. *Nucl. Fusion* **55**, 023009 (2015)
  40. K. Lee, K. Carlberg, *Model Reduction of Dynamical Systems on Nonlinear Manifolds Using Deep Convolutional Autoencoders* (2018), [arXiv preprint arXiv:1812.08373](https://arxiv.org/abs/1812.08373)
  41. M. Lehnen et al., Disruptions in ITER and strategies for their control and mitigation. *J. Nucl. Mater.* **463**, 39 (2015)
  42. Z. Long, Y. Lu, X. Ma, B. Dong, *PDE-Net: Learning PDEs from Data* (2017), [arXiv preprint arXiv:1710.09668](https://arxiv.org/abs/1710.09668)
  43. B. Maboudi Afkham, J.S. Hesthaven, Structure preserving model reduction of parametric Hamiltonian systems. *SIAM J. Sci. Comput* **39**(6), A2616–A2644 (2017)
  44. C. Michoski, M. Milosavljevic, T. Oliver, D. Hatch, *Solving Irregular and Data-Enriched Differential Equations Using Deep Neural Networks* (2019), [arXiv:1905.04351](https://arxiv.org/abs/1905.04351)
  45. V. Minh, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, M. Riedmiller, Playing Atari with deep reinforcement learning, in *NIPS* (2013)
  46. D. Mirzaei, R. Schaback, M. Dehghan, On generalized moving least squares and diffuse derivatives. *IMA J. Numer. Anal.* **32**(3), 983–1000 (2012)
  47. V. Minh et al., Human-level control through deep reinforcement learning. *Nature* **518**, 529 (2015)
  48. K.J. Montes, C. Rea, R.S. Granetz, R.A. Tinguely, N.W. Eidietis, O. Meneghini, D. Chen, B. Shen, B.J. Xiao, K. Erickson, Machine learning for disruption warnings on Alcator C-Mod, DIII-D, and EAST. *Nucl. Fusion* **59**, 096015 (2019)
  49. S.M. Moosavi et al., Capturing chemical intuition in synthesis of metalorganic frameworks. *Nat. Commun.* **10**, 539 (2019)
  50. A. Nguyen et al., Deep neural networks are easily fooled: high confidence predictions for unrecognizable images, in *CVPR* (2015)
  51. <https://www.whitehouse.gov/wp-content/uploads/2018/07/M-18-22.pdf>
  52. D. Osthus et al., Prediction uncertainties beyond the range of experience: a case study in inertial confinement fusion implosion experiments. *SIAM/ASA J. Uncertainty Quantif.* **7**(2), 604–633 (2019)
  53. G. Pang, L. Lu, G.E. Karniadakis, *fPINNs: Fractional Physics-Informed Neural Networks*, [arXiv preprint arXiv:1811.08967](https://arxiv.org/abs/1811.08967) (2018)
  54. V. Pascucci, P.-T. Bremer, A. Gyulassy, G. Scorzelli, C. Christensen, B. Summa, S. Kumar, Scalable visualization and interactive analysis using massive data streams, in *Advances in Parallel Computing: Cloud Computing and Big Data*, vol. 23, ed. by C. Catlett, W. Gentzsch, L. Grandinetti (IOS Press, Amsterdam, 2013), pp. 212–230
  55. A. Paszke, S. Gross, C. Soumith, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, A. Lerer, Automatic differentiation in PyTorch, in *Proceedings of NIPS-W* (2017)
  56. R.G. Patel, O. Desjardins, *Nonlinear Integro-Differential Operator Regression with Neural Networks*, [arXiv preprint arXiv:1810.08552](https://arxiv.org/abs/1810.08552) (2018)
  57. J.L. Peterson et al., Zonal flow generation in inertial confinement fusion implosions. *Phys. Plasmas* **24**(3), 032702 (2017)
  58. S.D. Pinches, “ITER’s Strategy for Diagnostic Data Analysis,” presented at the 3rd IAEA Technical Meeting on Fusion Data Processing, Validation, and Analysis, Vienna (2019)
  59. N. Polyzotis et al., Data management challenges in production machine learning, in *SIGMOD’17*, May 14–19, Chicago, IL, USA (2017)
  60. M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019). <https://doi.org/10.1016/j.jcp.2018.10.045>
  61. C. Rea, R.S. Granetz, K. Montes, R.A. Tinguely, N.E. Eidietis, J.M. Hanson, B. Sammuli, Disruption prediction investigations using machine learning tools on DIII-D and Alcator C-Mod. *Plasma Phys. Control. Fusion* **60**, 084004 (2018)
  62. P. Riley, Three pitfalls to avoid in machine learning. *Nature* **572**, 27 (2019)
  63. D. Rolnick et al., *Tackling Climate Change with Machine Learning*, <https://arxiv.org/pdf/1906.05433.pdf> (2019)
  64. S.K. Seal, M.R. Cianciosa, S.P. Hirshman, A. Wingen, R.S. Wilcox, E.A. Unterberg, Parallel reconstruction of three dimensional magnetohydrodynamic equilibria in plasma confinement devices, in 2017 46th International Conference on Parallel

- Processing (ICPP) (IEEE Computer Society, Los Alamitos, 2017), pp. 282–291. <https://doi.org/10.1109/ICPP.2017.37>
65. D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, D. Hassabis, Mastering the game of Go without human knowledge. *Nature* **550**, 354–359 (2017)
  66. V. Slon et al., The genome of the offspring of a Neanderthal mother and a Denisovan father. *Nature* **561**, 113 (2018)
  67. R.C. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications* (SIAM, Philadelphia, 2014)
  68. J. Svensson, A. Werner, Large scale Bayesian data analysis for nuclear fusion experiments, in *IEEE International Symposium on Intelligent Signal Processing* (2007), pp. 1–6
  69. C. Szegedy et al., Intriguing properties of neural networks, in *International Conference on Learning Representations* (2014)
  70. A. Tartakovsky, C. Marrero, P. Perdikaris, G. Tartakovsky, D. Barajas-Solano, [arXiv:1808.03398](https://arxiv.org/abs/1808.03398) (2018)
  71. M. Tesch, J. Schneider, H. Choset, Expensive function optimization with stochastic binary outcomes, in *International Conference on Machine Learning* (2013)
  72. C.G. Windsor, G. Pautasso, C. Tichmann, R.J. Buttery, T.C. Hender, JET EFDA Contributors and the ASDEX Upgrade Team, A cross-tokamak neural network disruption predictor for the JET and ASDEX Upgrade tokamaks. *Nucl. Fusion* **45**, 337 (2005)
  73. M.A. Wood, M.A. Cusentino, B.D. Wirth, A.P. Thompson, Data-driven material models for atomistic simulation. *Phys. Rev. B* **99**, 184305 (2019)
  74. D. Wroblewski, G.L. Jahns, J.A. Leuer, Tokamak disruption alarm based on a neural network model of the high-beta limit. *Nucl. Fusion* **37**(6), 725 (1997)
  75. Y. Zhu et al., Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *J. Comput. Phys.* **394**, 56–81 (2019)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.