# PMDT: AI-Enabled Predictive Maintenance Digital Twins for Advanced Nuclear Reactors

Final Report

## ARPA-e GEMINA PMDT program

Ms. Helena Goldfarb[a], Dr. Hao Huang[a], Dr. Naresh Iyer[a], Dr. Subhrajit Roychowdhury[a], Dr. Abhinav Saxena[a*], Dr. Arun Subramaniam[a], Dr. Liang Tang[a], Dr. Nurali Virani[a],
Dr. Alex Huning[b], Dr. Michael Muhlheim[b], Dr. Pradeep Ramuhalli[b], Dr. Askin Yigitouglu[b],
Dr. Jamie Coble[c], Mr. Jorge L. Narvaez[c], Dr. Richard Wood[c],
Mr. John Korsedal[d], Dr. Larry Nutt[d], Ms. Emily Stewart[d],
Mr. Jeffrey Clark[e], Ms. Amy Ferko[e],
*Note: listed in alphabetical order by author last name*

[a]GE Research, Niskayuna, NY
[b]Oak Ridge National Laboratory, Oakridge, TN
[c]University of Tennessee, TN
[d]GE Hitachi Nuclear (consulting OEM partner)
[e]Constellation Nuclear (consulting utility partner)

*Corresponding author: asaxena@ge.com; iyerna@ge.com
Contract: DE-AR0001290
March 3, 2025

# Table of Contents



# List of Figures

## List of Tables

## Acronyms

| | |
|---|---|
| AI | Artificial Intelligence |
| APM | Asset Performance Management (a monitoring software solution offered by GE) |
| AR | Advance Reactor |
| ARPA-E | Advance Research Project Agency-Energy |
| BOP | Balance of Plant |
| BWR | Boiling Water Reactor |
| CM | Corrective Maintenance |
| CP | Conformal Predictions |
| CPI | Conformal Prediction Interval |
| DT | Digital Twin |
| FMCRD | Fuel Motion Control Rod Drive (Mechanisms if FM-CRDM) |
| FPR | False Positive Rate |
| FWP | Feedwater Pump |
| GE | General Electric |
| GER | GE Research |
| GEH | GE Hitachi |
| GEMINA | Generating Electricity Managed by Intelligent Nuclear Assets |
| HDT | Health Digital Twin |
| HPIS | High Pressure Injection System |
| INPO | Institute of Nuclear Power Operations |
| IR | Incident Reports |
| LSTM | Long Short Term Memory deep learning models |
| ORNL | Oak Ridge National Laboratory |
| PRA | Probabilistic Risk Analysis |
| MIT | Massachusetts Institute of Technology |
| ML | Machine Learning |
| NLP | Natural Language Processing |
| NPP | Nuclear Power Plant |
| NRC | Nuclear Regulatory Commission |
| O&M | Operations and Maintenance |
| OEM | Original Equipment Manufacturer |
| OPRA | Operational Performance Risk Assessment |
| OSP | Optimal Sensor Placement |
| PHM | Prognostics and Health Management |

| | |
|---|---|
| PM | Preventive Maintenance |
| PMx | Predictive Maintenance |
| PMDT | Predictive Maintenance Digital Twin |
| R&D | Research and Development |
| RNN | Recurrent Neural Network |
| SAE | Simulated Assisted Engineering (simulation) |
| SME | Subject Matter Expert |
| TPR | True Positive Rate |
| TS | Technical Specification |
| UQ | Uncertainty Quantification |
| WO | Work Order |

# Abstract / Executive Summary

Our team made substantial technical progress on various fronts during the course of the program. Multiple milestones were geared towards demonstrating the feasibility of machine learning based predictive maintenance digital twins towards reducing O&M costs, whereas some other milestones actually focused on identifying technical gaps and developing technologies such as humble AI to provide necessary robustness to the ML-based models.

We were able to demonstrate in many cases that Machine learning-based methods can be successfully adapted for Nuclear plant environments especially for remote monitoring applications. Detailed analyses were carried out with plant and full scope simulation data along with capabilities of enhanced analytics to assess and set realistic expectations on cost reductions in O&M. These assessments are paving the way for investments towards reactor design improvements as well project planning for SMR projects as they develop and mature in the next few years. Technology developed under this program got direct visibility to GE Hitachi and their utility customers and resulted in positive intents to deploy some of the elements from design phase. The project additionally resulted in several reports, publications, software and data generation that will be useful in deployment and O&M services for BWRX300 fleets.

# 1 Summary of Project Outcomes

## 1.1 Background: Proposed Scope and Objectives

This research and development (R&D) project targeted development and demonstration of AI-based technologies towards achieving a substantial reduction in Operations and Maintenance (O&M) costs through Predictive Maintenance Digital Twins (PMDT) by moving from time-based to condition-based maintenance. The goal of this R&D was the establishment and demonstration of technologies that can contribute to enhanced O&M efficiencies through condition-aware maintenance and risk-informed operational decisions. GEH's BWRX-300 Advanced Reactor (AR) was used as the reference design for development and demonstration activities. Three levels of PMDTs—Operational, Health, and Decision digital twins—were considered for development for a subset of safety-critical components in and near the reactor vessel that are identified as high-risk, high-reward targets in the BWRX-300. PMDT technologies developed in the project were demonstrated through simulation or hardware flow-loops in realistic scenarios and by using utility field data within an established asset management environment. Scope of the development focused on a subset of critical systems and their failure modes, digital twins and AI technologies developed under this program were developed to be generalizable to other components, subsystems, and potentially other reactor designs.

## 1.2 Summary of Status at Completion

During the course of this program our teams collaboratively worked on twenty technical milestones (M2.1 – M6.1) distributed over 5 technical tasks (T2 – T6) in addition to the non-technical project management task T1. In this report we summarize outcomes and accomplishment from all these milestones and refer back to detailed discussions on each of these milestones in previous quarterly reports. Table 1 below summarizes references to quarterly reports for each milestone reporting as minor updates as they were worked followed by major update at completion and follow ups in some cases.

**Table 1. References to Program's Quarterly Reports for each Milestone's Technical Details.**

| | | | FY20 | FY2021 | | | | FY2022 | | | | FY2023 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Task | M# | Milestone Definition | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 |
| 1 | M1.1 | GO/No-Go: Workplan Finalized | | | | | | | | | | | | |
| | M1.2 | Software Development Plan | | | | | | | | | | | | |
| | M1.3 | Draft Impact Sheet | | | | | | | | | | | | |
| | M1.4 | Updated Impact Sheet | | | | | | | | | | | | |
| 2 | M2.1 | Target Component/subsystem Selection Report | | | | | | | | | | | | |
| | M2.2 | Reactor Components/Subsystems Operational twins developed | | | | | | | | | | | | |
| | M2.3 | Quantitative Performance Targets for all workstreams defined | | | | | | | | | | | | |
| | M2.4 | Targets for maximally informative sensor placement for fault detection and diagnosis identified | | | | | | | | | | | | |
| 3 | M3.1 | GO/NOGO: Initial set of health DTs implemented and proven to meet Level 1 targets | | | | | | | | | | | | |
| | M3.2 | Full set of DTs implemented and proven to meet Level 2 targets | | | | | | | | | | | | |
| | M3.3 | DBN integrated with health and decision twins | | | | | | | | | | | | |
| | M3.4 | Humble AI integrated into surrogate twins | | | | | | | | | | | | |
| | M3.5 | Humble AI integrated into diagnostic/predictive health twin | | | | | | | | | | | | |
| | M3.6 | Humble AI generalized and integrated into all types of health twins | | | | | | | | | | | | |
| | M3.7 | Initial set of calibration drift detection algorithms developed | | | | | | | | | | | | |
| | M3.8 | Improved performance for calibration drift detection and autocalibration algorithms achieved | | | | | | | | | | | | |
| | M3.9 | Autocalibration integrated with flow loop Instrumentation and Control (I&C) | | | | | | | | | | | | |
| 4 | M4.1 | Baseline Maintenance schedules identified and demonstration scenarios for opportunistic maintenance | | | | | | | | | | | | |
| | M4.2 | Maintenance planning tool demo to achieve reduction in O&M costs over baseline schedule | | | | | | | | | | | | |
| | M4.3 | Mechanized decision framework for risk-informed management of operations and maintenance at ARs | | | | | | | | | | | | |
| | M4.4 | Large scale WO automation implemented | | | | | | | | | | | | |
| 5 | M5.1 | Assessment of current regulatory guidance applicable to PMx completed | | | | | | | | | | | | |
| | M5.2 | Comprehensive assessment of regulatory impact to AI-based AR O&M completed | | | | | | | | | | | | |
| 6 | M6.1 | Selected DTs for existing BWR fleets deployed in APM cloud | | | | | | | | | | | | |

Major Update
Minor Update
Follow up

**Task 2** was aimed at analyzing historical operational and maintenance burden data to identify most critical target components for development of health twins. Specifically, in *Milestone M2.1* using a 4-Leg analysis we identified top ten systems that incur the highest maintenance burden (Figure 1). The three legs included (1) NRC's Licensee Event Reports (LERs), (2) Analysis of cases and impacts of lost production available from INPO reports, and (3) maintenance burden data from current fleet of multiple BWR plants over a period of several years. These data sources helped us identify top 10 components and systems that led to most cost burdens overall. As Leg 4, we also considered BWRX300 design to identify systems new and unique to BWRX300 to include in our DT development. We further analyzed most reported failures in these high ranked systems to focus our attention. This allowed us to focus on key analytics (code and data) that will be useful during commissioning and initial phases of operation.

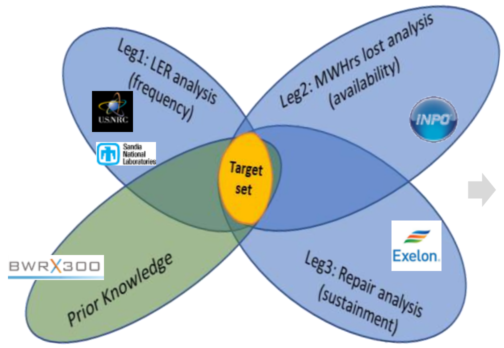


**Figure 1. Four-Leg Analysis and Critical Component Selection for Health Digital Twin Development.**

*Milestone M2.2* focused on building two operational twins. Since actual reactor does not exist yet, we needed simulators (digital twins) with enough fidelity to demonstrate a number of concepts and build representative analytics. With ORNL expertise in modeling we built a TRANSFORM modellica based simulation to mimic BOP of the BWRX300. This system was later used to demonstrate Decision twin namely – short term reconfiguration using OPRA towards avoiding a trip setpoint and keep the plant operational in M4.3. Additionally, the plant BOP simulation was used to demonstrate the concept of soft reconfiguration in the event of pump degradation.

Another operational twin developed under this milestone included that of servo motor driven FMCRDs (Figure 2). In this first of a kind use of servo motors, fault detection, diagnosis and severity estimation posed a unique challenge due to short intermittent operation of the motors, giving a limited opportunity for detecting any deterioration, that too under transient operation which is not much investigated in the open literature. Data generated from this operation twin modeled a variety of relevant external and internal faults and was used to develop a variety of health twins for several milestones under task 3. Later data were also open sourced through publishing on PHM Society degradation data repository, with an intent to open source the simulation as well for the wider research community.

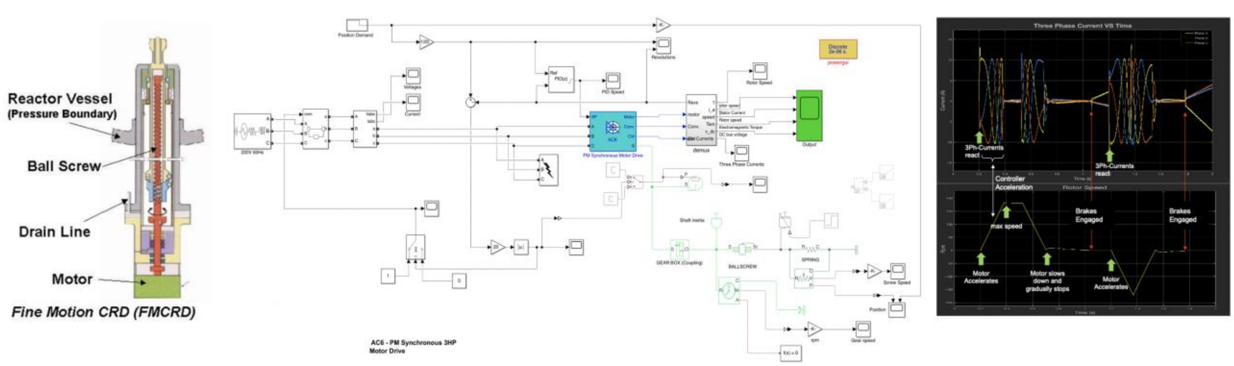


**Figure 2. FMCRD Operational Twin.**

*Milestone M2.3* performed analysis using a cost impact calculator developed as part of the project to quantify the performance targets to achieve maximum possible O&M cost reductions based on benefits from health twins. Specifically, sensitivity analyses were carried out under a range of health twin performance ranges (false positives, false negatives, etc.) and their impact on frequency (and consequently cost) of corrective maintenance (CM). All assumptions were analyzed in details and it was determined that use of Humble AI based model competence can yield up to 38% savings on current fixed O&M baselines as shown in Figure 3. These savings are estimated to result in $6.2/Mwh from current baseline of $18/MWhr. This milestone also set health twins performance targets (see Table 2) that we strived to achieve for all health twins developed under Task 3.

**Table 2. Performance targets for health twins set in M2.3.**

| | Level1 Targets | | Level2 Targets | |
|---|---|---|---|---|
| | TPR | FPR | TPR | FPR |
| Base PMDT | >= 82% | <= 5% | >= 88% | <= 2% |
| Humble AI PMDT | >= 92% | <=4% | >= 99% | <=0.5% |

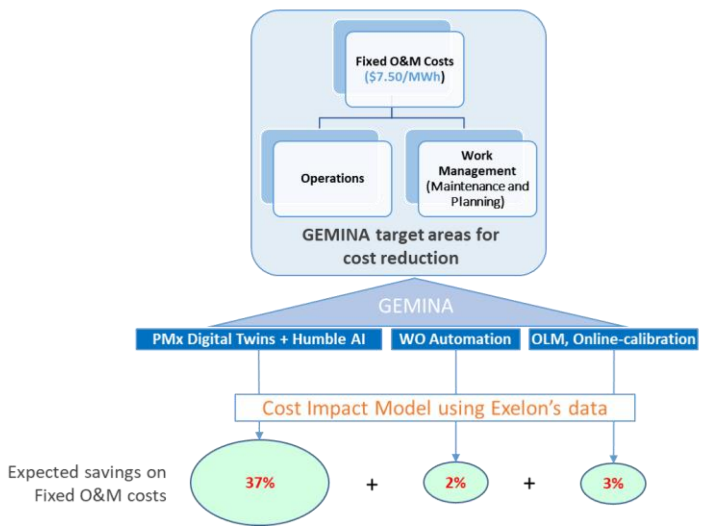


**Figure 3: Estimated Reduction in Fixed O&M costs from GEMINA thrusts.**

*Milestone M2.4* was intended to carry out a design analysis in identifying optimal sensor suite to maximize health observability into BWRX300. Although we had to reschedule this milestone to the last year in the program, we developed a generic tool that can leverage simulation data to determine best sensor suite combination along with optimizing factors such as sensor costs, sensor lifecycle and sustainment costs, fault coverage, ease of installation, reliability, etc. This allows carrying out trade studies during the design phase and make necessary design improvements at early stages. We also filed invention disclosure on this concept followed by a peer reviewed conference paper and technical demonstration at the PHM conference to solicit independent external feedback.

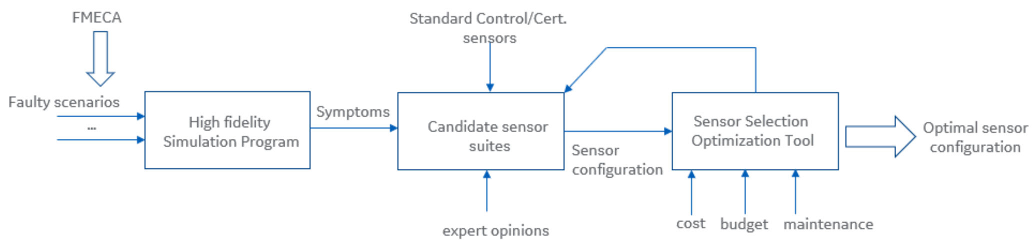


**Figure 4. The OSP process.**

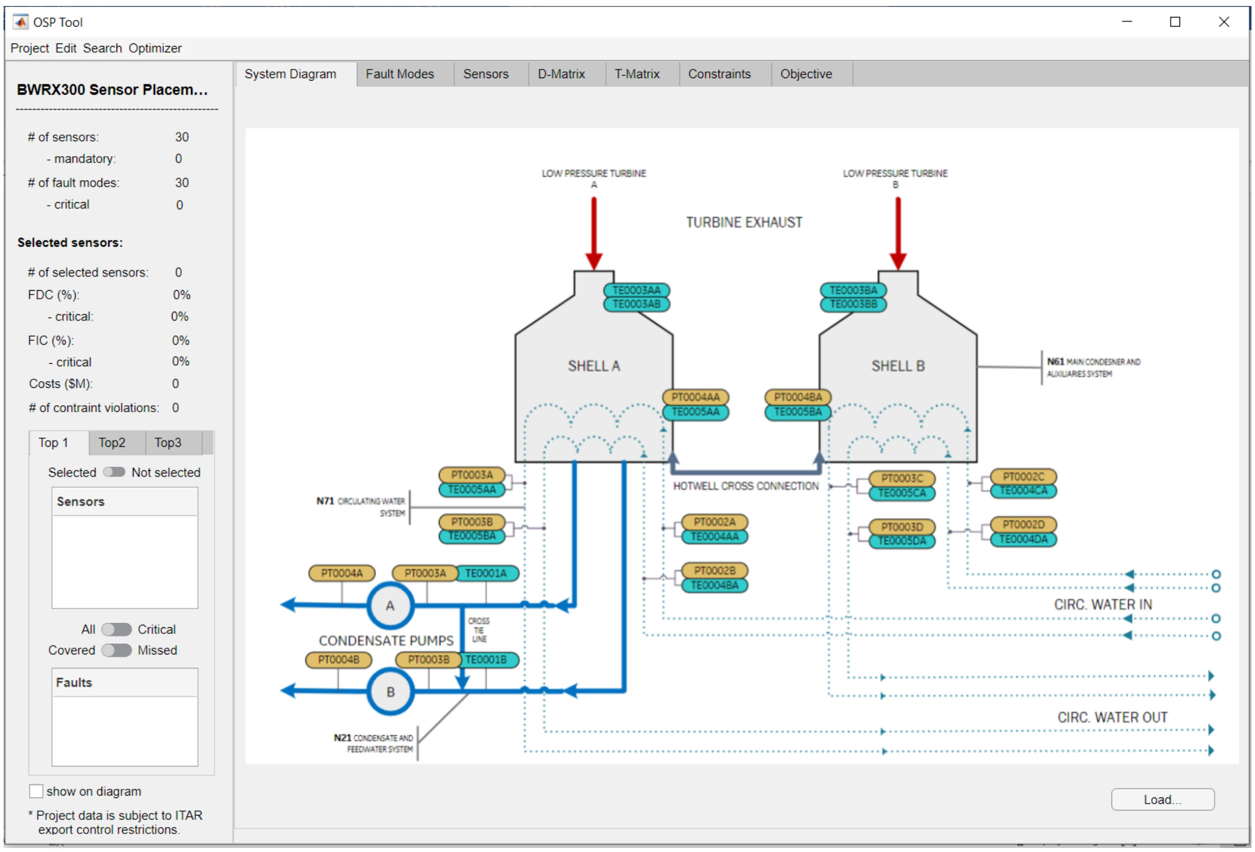


**Figure 5. Preliminary OSP tool.**

**Task 3** was focused on developing Machine learning-based health twins and algorithms for uncertainty quantification and model prediction competence assessment. This formed the core of the program and resulted in significant code base, demonstrations, and peer reviewed publications, which we plan to leverage as we continue to work towards implementing successful techniques for BWRX300.

Under *milestone M3.1* the research team pursued development of a number of health twins (see Table 3). Specifically, a fault detection and diagnosis model for FMCRD mechanism faults (motor short circuit, step load increase), a detection and severity estimation model for control rod binding faults, a model for feedwater pump anomaly detection, and a fouling detection for heat exchanger/condenser system was developed. These health twins were built using data from existing BWR fleets, and operational twins (simulations). Validation was carried out with in-depth review by domain experts, available maintenance history, or plant personnel with knowledge of asset histories. While a completion of validation will require implementing and testing of these models on actual data from BWRX-300, based on results so far we conclude that machine learning-based models along with domain physics information allow us to develop health digital twins to meet level 1 performance targets. Incorporation of humble-AI will further allow improving these twins for runtime-robustness and reducing maintenance costs.

**Table 3. Summary of health twins developed for milestone 3.1.**

| System/component | Fault/failure | Health Twin | Data Source | Expected Outcome |
|---|---|---|---|---|
| FMCRD Servo Motor | Motor short circuit, open circuit, mechanical jam | Fault detection | Physics-based simulator | Investigation of full range of health twins for FMCRDs, sensitivity to noise, and development of humble AI for all health twins |
| FMCRD Servo Motor | | Fault Diagnosis | Physics-based simulator | |
| FMCRD Servo Motor | Mechanical resistance to movement | Degradation Prognostics | Physics-based simulator | |
| Condenser Fouling | Fouling | Fault detection | BWRX-300 SAE simulator | Fouling detection under external and operational variations |
| Feedwater pump | High bearing vibration | Fault detection and early warning | BWR Plant data | Demonstration of early warning and reduction in false alarms |
| I&C (covered under M3.7) | Calibration drifts | Drift detection | Simulated data from testbed + Plant data | Demonstration of online drift detection |

As part of *milestone M3.2* we showed that health twins developed in this project were easily able to meet level 2 performance target of true positive rates, however, in realistic situations meeting a low false positive rate is more challenging. This observation is consistent with past industry experience across several industrial domains. This justifies the need for runtime-time robustness methods to assess individual prediction reliability. As part of the discussion in this report we provided improvement in results from using humble AI.

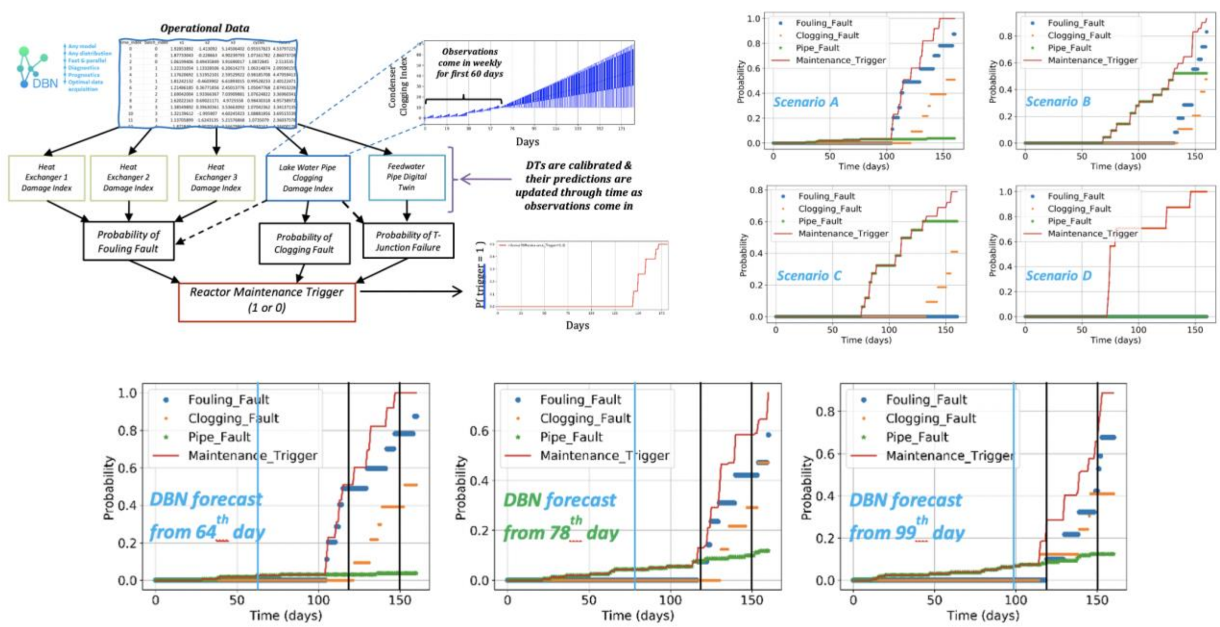


**Figure 6. A schematic overview of the DBN developed for long term decision. A demonstration shown on four different scenarios based on the selected hypothetical operational policy.**

Under *milestone M3.3* we accomplished a modular structure of DBN framework (Figure 6) where individual health twins, potentially operating at different time resolutions can be combined into a single risk roll up at the system level. As shown through various scenarios, a system level maintenance trigger can be used to inform opportunistic maintenance tool (in M4.2) to optimize outage planning. We worked synergistically with the MIT GEMINA performer team and integrated the high-fidelity T-junction corrosion monitoring twin into our framework of health twins.

Health twins reported under M3.1 were updated with humble-AI under *Milestones M3.4 - M3.6*. Multiple approaches were developed and investigated to compute justification for model competence and consequently expressing an evidence-based confidence at the runtime that allows different actions based on predictions at runtime. Key premise here was that predictions with evidence of strong competence can be dealt with automated processes, whereas weak competence scenarios can be referred to domain experts. This reduces the risk of false positives and false negatives, and at the same time offers to minimize monitoring load on plant operational staff. These concepts were demonstrated with Condenser and FMCRD health twins. Most notably three quantitative approaches were developed and compared for robust model output uncertainty quantification of regression-based health twins (see Figure 7).

| | Conformal Prediction Interval | Jackknife+ Prediction Interval | Ensemble Quantile Prediction |
|---|---|---|---|
| Validity ($\alpha = 0.1$) | $90.027 \pm 1.11\%$ | $90.8 \pm 1.4\%$ | $83.33 \pm 2.9\%$ |
| Theoretical guarantees | Coverage guarantee $\mathcal{P}(y_{n+1} \in \tau(x_{n+1})) \geq 1-\alpha)$ | Coverage guarantee $\mathcal{P}(y_{n+1} \in \mathbb{C}_{\alpha}(x_{n+1})) \geq 1-\alpha)$ | Good empirical performance, no theoretical guarantees |
| Model applicability | Any ML regression model | Any ML regression model | Ensemble of any ML model |
| Calibration | Calibrated to be valid | Calibrated to be valid | Uncalibrated |
| Challenges | • Suitable distance metric needs to be defined in input/latent space<br>• Works for single output models | | Computationally expensive training and inference for large models |

**Figure 7. Qualitative and Quantitative Comparison of UQ Approaches for model output uncertainty in regression-based FMCRD health twins.**

As part of *milestone M3.6* we summarized various HAI capabilities developed over the course of our project. We report that a number of HAI methods were developed to infer model competence in nominal (in relation to training set) and off-nominal scenarios (gradual or abrupt drift in exogenous/controlled inputs or in input/output relationship) regardless of system fault or no fault conditions for a variety of learning-based tasks. Methods to assess competence using input-space out of experience detection, latent space anomaly-detection, and model output-space conformal-prediction intervals were developed and applied to a number of component health twin applications which involved classification, regression, as well as anomaly detection settings that covers a majority of health twin applications (Figure 9 and Table 4).

**Table 4. Summary of various HAI methodologies as applied to relevant machine learning-based health twins.**

| # | Types of Support | Guard against | Machine Learning Problems | | |
|---|---|---|---|---|---|
| | | | Classification | Regression | Anomaly Detection |
| 1 | Input anomaly | • Input drift<br>• Extrapolation | Density estimation:<br>• parametric (GMM)<br>• nonparametric (kernel density) | Density estimation:<br>• parametric (GMM) [TRACG Surrogate]<br>• nonparametric (kernel density) [TRACG Surrogate] | Density estimation:<br>• parametric (GMM)<br>• nonparametric (kernel density) [Condenser Fouling] |
| 2 | Reconstruction error | • Input drift<br>• Extrapolation | Self supervised learning with autoencoders | Self supervised learning with autoencoders [Condenser Fouling] | Self supervised learning with autoencoders [TEP, Condenser Fouling] |
| 3 | Geometric neighborhood in embeddings | • Extrapolation<br>• Ambiguity<br>• Adversarial manipulations | Epistemic classification [TEP] | • KNN regression from latent space with conformal PI<br>• Prediction error estimate from latent space [FMCRD] | Farthest k-neighbor distance [TEP] |
| 4 | Output uncertainty anomaly | • Input drift<br>• High process uncertainty | - | Conformal outlier detection over size of PI<br>• Conformal PI, jackknife+ [FMCRD]<br>• Bayesian LSTM [Sensor calibration]<br>• GP | • Conformal outlier detection over size of PI<br>• PI from: conformal PI, jackknife+, GP, or Bayesian LSTM |
| 5 | Residual error drift | • Concept drift | N/A | • Physics-informed regression [Condenser Fouling] | N/A |

KNN – k-nearest neighbors, PI – prediction interval, GP – Gaussian process, GMM – Gaussian Mixture Models

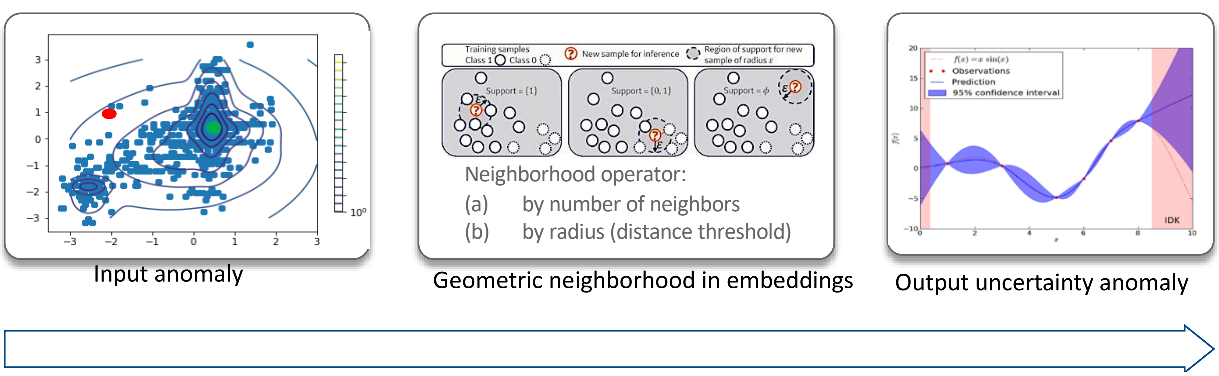


**Figure 8. Summary of Humble AI methodologies developed.**

Under *milestone M3.7*, an initial set of calibration drift detection algorithms were developed for the purpose of detecting drift and correcting for it through an online automated recalibration method. The methods described here were developed to target model prediction and drift correction accuracy exceeding 80%. The reported results indicate that the initial approaches to online recalibration were successfully able to achieve model prediction and drift correction accuracies exceeding 80%.

Under *Milestone M3.8* the calibration-drift detection and autocalibration algorithms reported above were updated to quantify prediction and autocalibration uncertainty.

Multiple approaches were examined for uncertainty quantification. Results from the inclusion of a Bayesian uncertainty quantification approach indicate potential for determining the confidence in the autocalibration result (Figure 9). Other approaches to evaluating the confidence in the result that leverage internal state information of the LSTM model were identified and are presently under evaluation. Additional studies conducted hyperparameter optimization for the models, identifying a semi-optimal configuration that appears sufficiently complex to represent the underlying complexity in the data while maintaining the high prediction accuracy. Collectively, these approaches to improving the performance in the calibration-drift detection and autocalibration approaches increase confidence in online calibration monitoring, and address elements that are expected to be necessary for regulatory approval of the technology.

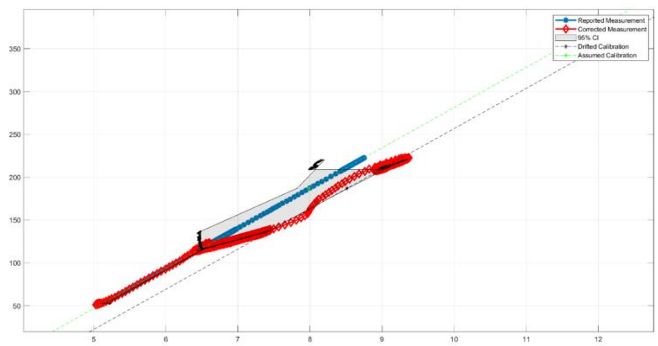


**Figure 9. Example of Drift Detection with Uncertainties.**

*Milestone M3.9* built on the previous work and presented initial results of integration of the autocalibration algorithms with advanced reactor-relevant flow loops for demonstration purposes and describes the integration of uncertainty quantification (UQ) approaches for the autocalibration.

**Task 4** focused on decision making in the plant O&M based on component and system health as determined by health twins developed in task 3. As shown in Figure 10, Milestones were designed to investigate and develop methods for both long term planning for maintenance outage and short-term risk-informed planning to avoid trip set point through hard- and soft-reconfiguration.

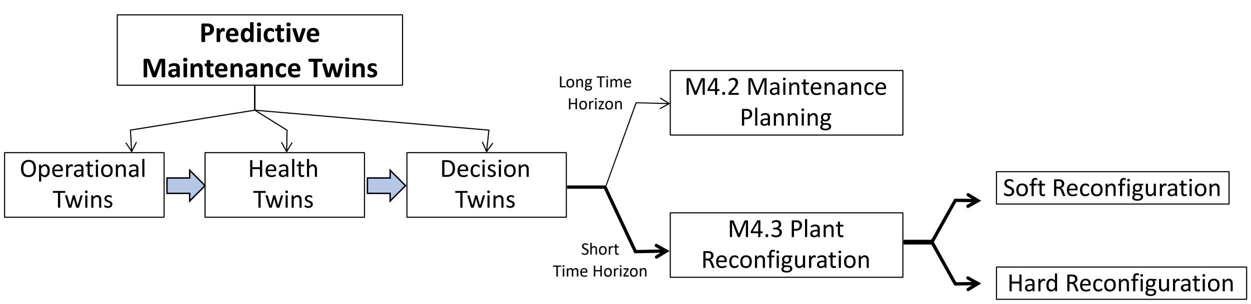


**Figure 10. Digital twin-based hierarchical decision-making concept.**

Under *milestone M4.1*, we collaborated with another group at Constellation Energy and an operations research team at University of Tennessee Knoxville to identify an approach to optimizing maintenance and outage planning. A baseline schedule was obtained, and benchmark cases were identified, that were used to evaluate effectiveness of opportunistic maintenance approach developed and demonstrated under M4.2. Since this task required a detailed knowledge of various tasks, activities, and associated resource requirements that were not available for BWRX300 yet, our team utilized an existing model for a high pressure injection system (with over 130 activities) as a generic yet representative nuclear plant subsystem and focused on developing maintenance schedule/plan optimization method which was the core focus for the milestone. All developments were done as generic modules such that they can be adapted with relative ease when information from actual system becomes available.

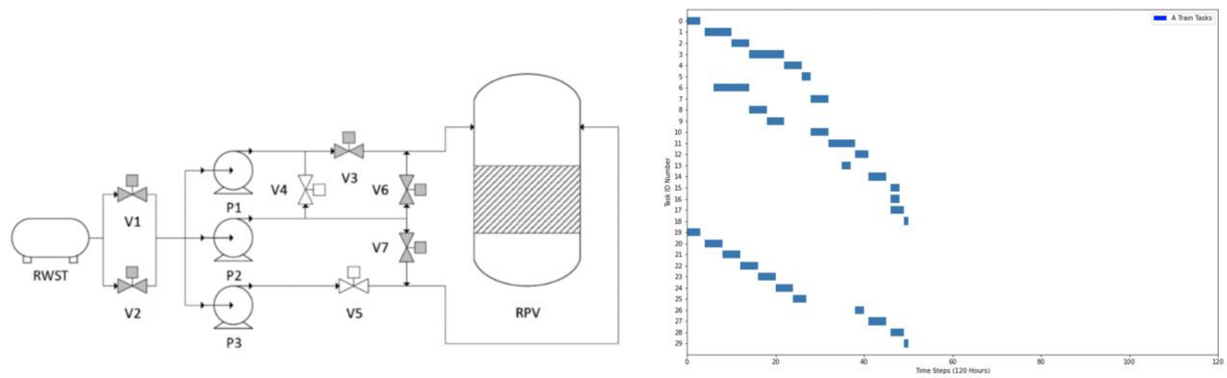


**Figure 11. Baseline Maintenance Schedule for Train A Maintenance Activities for a generic High Pressure Injection System.**

*Milestone M4.2* demonstrated potential for significant cost savings in terms of labor even when applied to a limited set of component scenarios. This demonstration did not consider other costs that may be avoided by safely deferring maintenance. We also did not consider any potential impacts on outage duration or personnel requirements. With minimal adaptations, the proposed PIMO scheduling system could be used to determine the minimum maintenance staff needs to complete tasks within the allocated outage duration.

*Milestone 4.3*, as shown in Figure 10, focused on shorter term plant reconfiguration to avoid plant trip setpoint. As summarized in Figure 12 below there is a continuum of reconfigurability that together achieves plant uptime and reliability. For our demonstration we focused on strategic and physical reconfigurations. BOP operational twin developed under M2.2 was utilized to incorporate effects of transients initiated due to fault modes and reconfigurations were shown as a means to avoid trip setpoint.

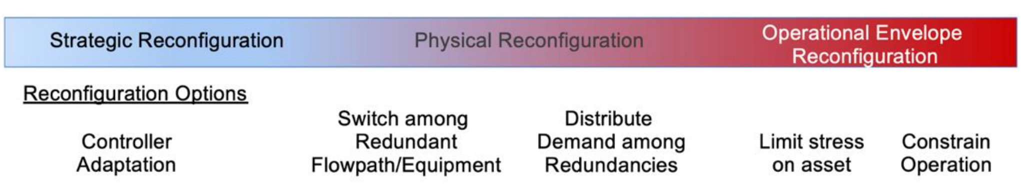


**Figure 12. Reconfigurability continuum.**

At the higher supervisory level, OPRA-based risk assessment was utilized to identify alternative success paths similar to what experienced plant operators apply. This can be seen as documenting experiential operational knowledge along with dynamic assessment of risk from component health in real-time in determining the best course of action. We envision a methodology like this to work as operator assistance tool in helping them in a more risk informed decision making supported by data as nuclear plants move towards more plant automation.

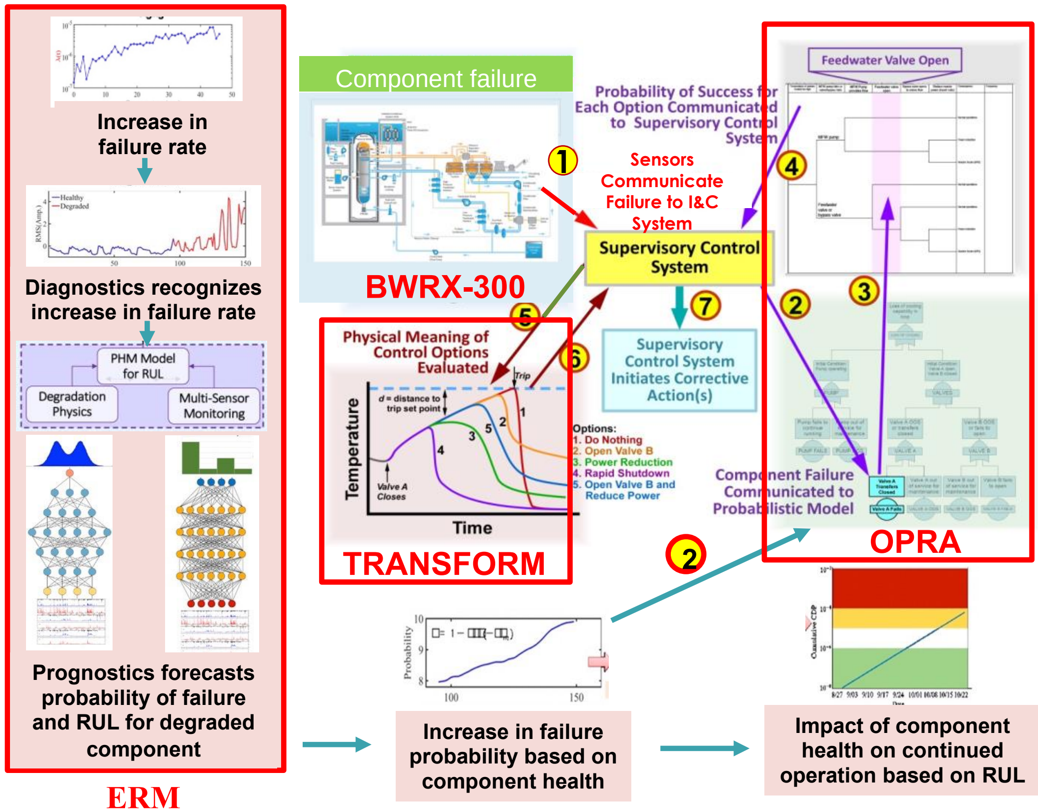


**Figure 13. OPRA and its Interfaces.**

*Milestone M4.4* was by far the most challenging in this program due to the inability to get access to the originally planned dataset from Constellation. Our team, together with the innovation team from Constellation (then Exelon) had initially identified and prioritized a number of tasks that would potentially benefit from automation and help improve cost efficiency. One such task was to develop a system for action recommendations. Since we were not successful in getting the required data from Constellation, we pivoted to work with unstructured text data from gas and steam power plants, which was made available to us by the GE Digital team. With these data we developed a recommendation system pipeline. However, given the inconsistent quality of the unstructured data and the lack of curated ground truth, our model generation, evaluation and iterative improvement has been challenging and time consuming. To alleviate this issue, we followed a systematic approach to normalize the effects of experts' bias while collating SME opinions to prepare benchmark data.

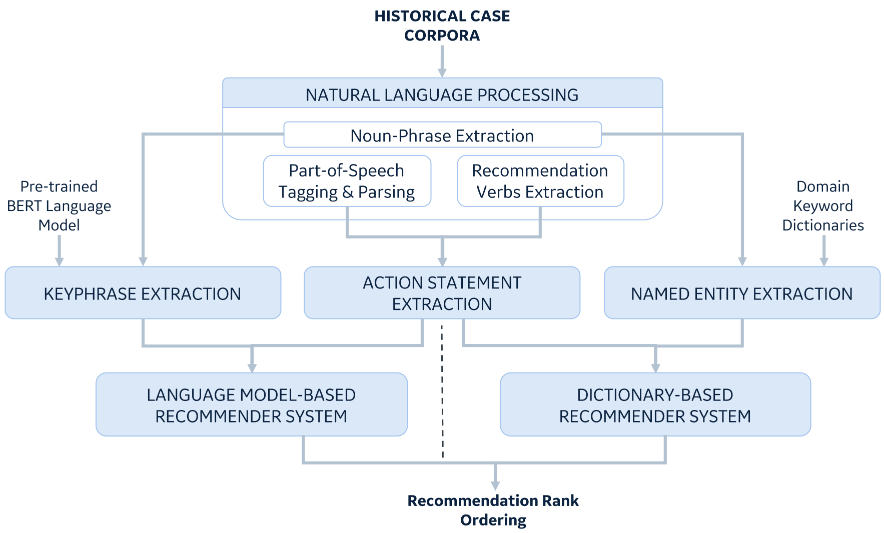


**Figure 14. Case Knowledge Curation Engine Pipeline for Keyterms-based Recommender System.**

Since our development using natural language processing methods including language models such as BERT, success of large language models (LLMs) has taken a center stage that will potentially outperform most NLP tasks. However, we believe that a generic LLM may not fully address industrial domains such as a maintenance recommended action set as is and will need finetuning with curated historical data. Therefore, we strongly believe that recommendation ordering method developed in this program will still be useful in automatically processing historical data and curate training data to be used for updating LLMs. This has become a topic for current and ongoing research, and we hope to leverage the pipeline developed here in those endeavors.

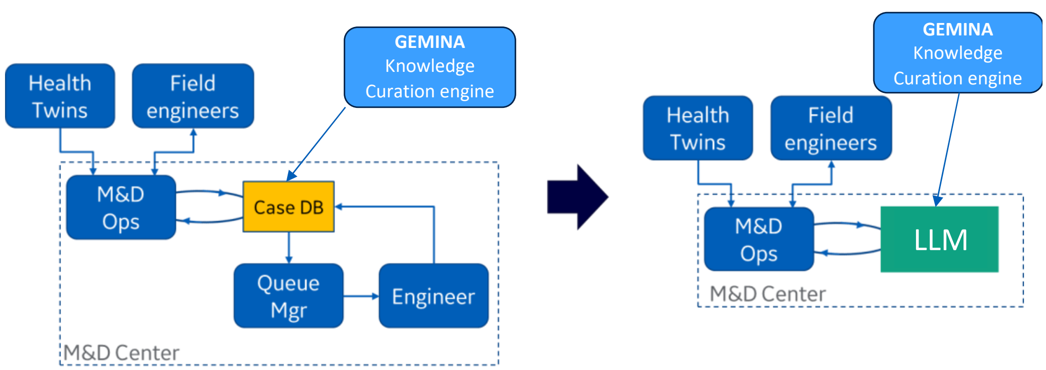


**Figure 15. Proposed Concept for Leveraging Knowledge Curation Engine to be used with LLMs.**

Under **Task 5** the objective was to identify the potential regulatory pathways and associated requirements for incorporation of preventative maintenance of digital twin (PMDT) technology for operations and maintenance (O&M) of an advanced reactor (AR). This analysis task was completed in two parts (M5.1 and M5.2).

*Milestone M5.1* provided a review of the current regulatory guidance relevant to application of PMx, artificial intelligence (AI), and automation. The outcomes of this review included determination of constraints on the application of such technology, identification of any regulatory gaps or uncertainties, and clarification of anticipated technical basis information likely to be important for regulatory acceptance of PMx, AI, and automation. Detailed insights into a number of DT related aspects were discussed where current regulatory framework will likely have an impact on development and deployment of DTs in AR applications.

*Milestone 5.2* supplemented he report from M5.1 and described the research results to date to identify regulatory implications of AI technologies and their uses. Specifically, this report reviewed current regulatory guidance relevant to the application of predictive maintenance twins, artificial intelligence (AI), and automation. The focus of this review included determination of constraints on the application of AI technology, identification of any regulatory gaps or uncertainties, and clarification of anticipated technical basis information likely to be important for regulatory acceptance of these technologies. The research included review of topical reports and other submissions to the US Nuclear Regulatory Commission (NRC) on technologies relevant to using AIs for predictive maintenance, NRC safety evaluation (SE) reports, and other relevant literature to identify specific regulatory concerns. Outcomes are briefly summarized below for quick reference.

- Use of DTs span a large number of potential applications throughout ARs lifecycle, and depending on DTs role (*in design, operations, and maintenance*) one or more regulatory documents may be applicable
- Further determination will be based on the level of automation DTs will achieve/enable, e.g. non-control function, control function, and/or communications function.
- DT complexity will also determine the level of regulatory considerations. Simple monitoring, the lowest complexity level, minimal consequences of failure, etc. would not require guidance on software tools or type of digital device. At the other end of the graded approach, existing guidance will apply to control those DTs that perform a control function while more scrutiny would result if the component performs a safety function.

The capability to classify the DT by safety significance, functionality, complexity and other characteristics can provide the basis for a graded approach to implementing or licensing such technologies. Further investigations are underway to establish the basis to enable such determinations. In the meantime, the following are clear conclusions about when more rigorous review will likely arise.

- If the DT is used for a safety application, it must meet the requirements of a safety-related I&C system. More specifically the DT would need to meet 10 CFR 50, Appendix B; IEEE 1012 (endorsed by Regulatory Guide 1.168); IEEE 603 etc.
- If the DT is used to support analysis that is used in a safety analysis/calculation then you are governed by 10 CFR 50.46; 10 CFR 50, Appendix K; and some form of software V&V most likely ASME NQA-1. If the DT is to be used for determining TSs, surveillance intervals, etc., 10 CFR 50.46; 10 CFR 50, Appendix K; and ASME NQA-1, and maybe some additional V&V are the governing requirements.

**Task 6** and corresponding *Milestone M6.1* intended to show proof of concept DTs for existing BWR fleets demonstrated on a fielded cloud platform, such as GE's Asset performance Management (APM). Various algorithms and analysis tools developed in prior milestones generated a substantial code base, however it was important to structure this codebase in compliance with digital delivery platforms offered by GE. Intent behind this milestone was, therefore, to get the code ready for integration with tools planned to be used for monitoring of BWRX300.

As was reported for several quarters, there was a persistent risk that the GE project team couldn't not access APM environment for Constellation BWR fleet due to inaccessibility to their on-prem APM environment. Therefore, while we completed various subtasks needed to accomplish the milestone an integrated demonstration on a deployed system couldn't be feasible as originally planned. Therefore, we estimate this milestone was completed at around 85%.

# 2 T2M Activities and Outcomes

The U.S. Commercialization plan lays the groundwork of business strategy across GE Research, GE-Hitachi, and Constellation—three entities with key interests in PMDT work.

## 2.1 Collaborations with Constellation, GEH and GEH customers, MIT

Our research team was very closely involved with AR OEM and utility partners. This quarter, we worked closely with GE Hitachi team on BWRX300 simulator data generation as well soliciting feedback on optimal sensor placement tool. We also participated in a number of customer meetings with a nuclear operator in formulating a plan for pilot demonstration project that utilizes GEMINA technologies. GE Hitachi has filed a provisional patent disclosure on the OSP tool. Moreover, there is interest in utilizing this tool with other AR developments.

## 2.2 Project Technical Outputs

Various workstreams under our project resulted in impressive 50+ public documents, reports, and presentations and our team continues to engage in similar workshops and conferences.

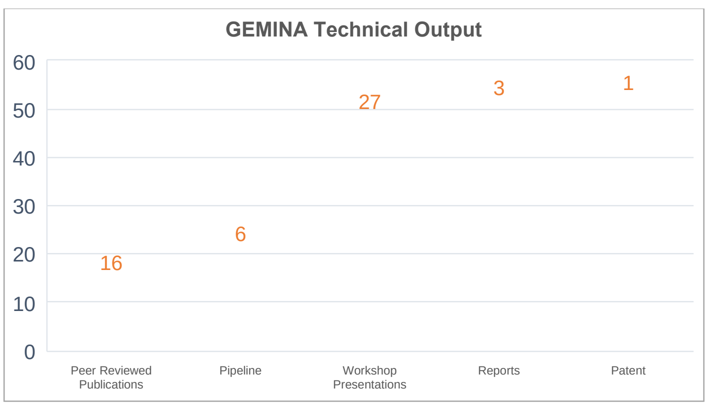


**Figure 16. Summary of Technical Outputs from this Program.**

These outcomes are enlisted below for quick reference.

### Patents/Disclosures

SYSTEM AND METHOD FOR OPTIMAL SENSOR PLACEMENT - Docket No. 8564-000451-US-01 - Tang, Evans, Saxena, Goldfarb

### Technical Reports (OSTI)

1. M. D. Muhlheim, P. Ramuhalli, A. Huning, A. Guler, R. Wood, and A. Saxena, "Status Report on Regulatory Criteria Applicable to the Use of Digital Twins," ORNL/SPR-2022/2493, June 2022
2. M. D. Muhlheim, A. Huning, P. Ramuhalli, A. Guler, R. Wood, Narvaez, J., and A. Saxena "Risk-Informed Decision-Making and Reconfiguration using OPRA", ORNL/SPR-2023/3078, September 2023 [submitted]
3. M. D. Muhlheim, P. Ramuhalli, A. Huning, A. Guler, and A. Saxena, "Status Report on Regulatory Criteria Applicable to the Use of Artificial Intelligence (AI) and Machine Learning (ML)", ORNL/SPR-2023/3072, September 2023 [submitted]

### Peer-reviewed Publications

1. P. Ramuhalli, M. Muhlheim, A. Guler Yigitoglu, N. Virani, N. Iyer, A. Saxena, "Assessment of Confidence Measures for Online Monitoring and Sensor Recalibration," Abstract submitted to ANS NPIC-HMIT 2023
2. A. Ifeanyi, A. Saxena, and J. Coble. "Within-Bank Analysis for Anomaly Detection and Condition Monitoring of Fine Motion Control Rod Drives." Submitted to Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technology Meeting (NPIC&HMIT), Knoxville, TN. July 15-21,2023
3. Michael D. Muhlheim, Pradeep Ramuhalli, Askin Guler Yigitoglu, Alex Huning, and Abhinav Saxena, "A Risk-Informed Assessment of Operational Options for Successfully Avoiding a Trip Setpoint," Abstract submitted to PSA 2023
4. Michael D. Muhlheim, Pradeep Ramuhalli, Askin Guler Yigitoglu, Alex Huning, and Abhinav Saxena, "Forecasting the probability of failure and remaining useful life of the degraded components status," Abstract submitted to PSA 2023
5. Fullilove, N., Dos Santos, D., Saxena, A., & Coble, J. (2022). Leveraging Within-Bank Comparison for Anomaly Detection, Diagnostics, and Prognostics in Advanced Nuclear Power Plants. Annual Conference of the PHM Society, 14(1). https://doi.org/10.36001/phmconf.2022.v14i1.3215
6. Peshave, A., Aggour, K., Ali, A., Mulwad, V., Dixit, S., & Saxena, A. (2022). Evaluating Vector Representations of Short Text Data for Automating Recommendations of Maintenance Cases. *Annual Conference of the PHM Society*, *14*(1). https://doi.org/10.36001/phmconf.2022.v14i1.3196 .
7. Iyer, N., Virani, N., Yang, Z., & Saxena, A. (2022). Mixed Initiative Approach for Reliable Tagging of Maintenance Records with Machine Learning. *Annual Conference of the PHM Society*, *14*(1). https://doi.org/10.36001/phmconf.2022.v14i1.3159.
8. Pradeep, R., Muhlheim, M., Guler Yigitoglu, A., Huning, A., & Saxena, A. (2022). A Methodology for Online Sensor Recalibration. *Annual Conference of the PHM Society*, *14*(1). https://doi.org/10.36001/phmconf.2022.v14i1.3274.
9. Huning, A.J., J. Rader, and P. Ramuhalli, "Development of a BWR System Fault Simulator Using TRANSFORM/Modelica," Transactions of the American Nuclear Society, 2022 Winter ANS Meeting pp: 850-853, Nov. 13 – 17, Phoenix, AZ, https://*doi.org/10.13182/T127-39883*
10. M. D. Muhlheim, P. Ramuhalli, A. Huning, A. Guler Yigitoglu, A. Saxena, and R. T. Wood, "How Digital Twins May Be Used in Design and Operations," Transactions of the American Nuclear Society, 2022 Winter ANS Meeting, pp: 250-253, Nov. 13 – 17, Phoenix, AZ, https://*doi.org/10.13182/T127-39790*
11. M. D. Muhlheim, P. Ramuhalli, A. Huning, A. Guler Yigitoglu, A. Saxena, and R. T. Wood, "Regulatory Requirements, Guidance, and Review of Digital Twins," Transactions of the American Nuclear Society, 2022 Winter ANS Meeting, pp: 254-257, Nov. 13 – 17, Phoenix, AZ, https://*doi.org/10.13182/T127-39785*
12. Iyer, N., Virani, N., Yang, Z., & Saxena, A. (2022). Mixed Initiative Approach for Reliable Tagging of Maintenance Records with Machine Learning. Annual Conference of the PHM Society, 14(1). https://doi.org/10.36001/phmconf.2022.v14i1.3159
13. Sharad Dixit, Varish Mulwad, Abhinav Saxena, "Extracting Semantics from Maintenance Records" presented at The IJCAI-21 Workshop on Applied Semantics Extraction and Analytics (ASEA) https://arxiv.org/abs/2108.05454
14. Huang, H., Subramanian, A., Saxena, A, Virani, N. ,Iyer, N. (2023). Deep Regression Network with Prediction Confidence in Time Series Application for Asset Health Estimation. Annual Conference of the PHM Society, 2023- https://doi.org/10.36001/phmconf.2023.v15i1.3556 (*Best paper candidate*).
15. Subramanian, A. (2023) Servomotor Dataset: Modeling Health in Mechanisms with Typically Intermittent Operation. Annual Conference of the PHM Society, 2023 https://doi.org/10.36001/phmconf.2023.v15i1.3580 (*Best paper candidate*).
16. Saxena, A., Goldfarb, H., Clark, J. (2023) Unsupervised Causal Deep Learning-Based Anomaly Detection in Nuclear Power Plant Applications. Annual Conference of the PHM Society, 2023. https://doi.org/10.36001/phmconf.2023.v15i1.3570 (***Best Paper Award***).

17. Tang, L., Saxena, A., Evans, S., Iyer, N., Goldfarb, H. (2023) OSPtk: Cost-aware Optimal Sensor Placement Toolkit Enabling Design-for-PHM in Critical Industrial Systems. Annual Conference of the PHM Society, 2023 - https://doi.org/10.36001/phmconf.2023.v15i1.3557.

## Submitted/Planned Publications

1. Virani, N., Iyer, N., and Saxena, A., 2023, "Towards understanding of trust in uncertainty estimates", AAAI conference on AI.
2. N. Virani, A. Saxena, N. Iyer, and A. Subramanian, "Trust in uncertainty estimates from learning-based digital twins", Frontiers in Energy Research, Special Issue on Shaping the Future of Nuclear Assets with Digital Twins, Nov 2023 (in prep.)
3. Virani, N., Iyer, N., and Saxena, A., 2023, Assessing Prediction Reliability of Anomaly Detection Methods, Engineering Applications of AI, Elsevier.
4. Roychowdhury, S. et. al., 2023, Physics-Informed Regression for Anomaly Detection.
5. N. Fullilove, D. Dos Santos, A. Saxena, and J. Coble, "Health Monitoring and Fault Detection Techniques for Electric Servomotor Using Simulated Data".
6. A. Ifeanyi, N. Fullilove, D. Dos Santos, A. Saxena, and J. Coble, "Anomaly Detection across Coincident Fine Motion Control Rod Drive Mechanisms." Submitted to International Journal of Prognostics and Health Management.

## Invited Talks and Panels

1. Iyer, N. (2023) Digital Twins for Prognostics and Health Management in Nuclear, NRC Virtual Workshop on Structural Health Management for Nuclear Power Plants Session 2: Diagnostic and Prognostic Tools for Condition Monitoring and Structural Health Management, November 2023.
2. Virani, N. (2023). Reliable Anomaly Detection for Individual Prediction Reliability, Artificial Intelligence for Robust Engineering & Science (AIRES 4), Oak Ridge National Lab.
3. Iyer, N. (2023). Cost-benefit analysis for justifying investment in Predictive Maintenance: mapping it to observability, performance targets and robustness of AI models, IEEE Conference on Artificial Intelligence (CAI), June 5-6th, Santa Clara, CA.
4. Virani, N (2023). "Machine learning for reliable health monitoring and anomaly detection", Accel AI Meeting, FermiLab, Sep 14 2023.
5. Virani, N. (2023). Reliable Anomaly Detection for Individual Prediction Reliability, Artificial Intelligence for Robust Engineering & Science (AIRES 4), Oak Ridge National Lab.
6. Iyer, N. (2023). Cost-benefit analysis for justifying investment in Predictive Maintenance: mapping it to observability, performance targets and robustness of AI models, IEEE Conference on Artificial Intelligence (CAI), June 5-6th, Santa Clara, CA.
7. Saxena, A. (2023). Role of AI in Enabling Carbon Free Energy Transition through Predictive Maintenance, IEEE Conference on Artificial Intelligence (CAI), June 5-6th, Santa Clara, CA.
8. Saxena, A., (2023). "Digital Twins for Sustanbility"Opportunities and Challenges for Digital Twins in Atmospheric and Climate Sciences, A National Academies Workshop on Digital Twins in Atmospheric, Climate, and Sustainability Sciences, published in the workshop proceedings - https://nap.nationalacademies.org/26921
9. Iyer, N., Saxena, A. (2022), "Cost Impact Model to assess economic value of digital twins for nuclear reactors", at AAAI workshop on Predictive Maintenance in Arlington VA on Nov 17.
10. Saxena, A. (2022), "Potential use and relevance of AI/ML techniques in the Nuclear Industry" Plenary at the DIET 2022 conference of the Canadian Nuclear Society on Nov 2, 2022, and also participated in a following panel discussion.

11. Saxena, A. (2022), "AIML-based Predictive Maintenance towards Industrial reliability and Operational Autonomy", Plenary keynote at ASME InterPACK conference in Anaheim CA on Oct 27, 2022
12. Saxena, A. (2022),, "AIML-based Digital twin approaches to Advanced Reactors", presented to GE Hitachi's Advanced Reactor Innovation group.
13. N. Iyer, Saxena, A. (2022), "Cost Impact Model to assess economic value of digital twins for nuclear reactors", presented to MIT nuclear engineering department.
14. P. Ramuhalli (2022), "Prognostic Health Management (PHM) Solutions for Sustainable Generation," Presented at the PG&E Data Science Seminar, July 2022.
15. A. Saxena, "AI Opportunities in Improving Industrial Reliability through Predictive Maintenance", 2022 EDGE Symposium, September 2022
16. Guler-Yigitoglu, A. (2022), "Digital Twins for Risk and Safety Assessments," Panel, 2022 American Nuclear Society (ANS) Annual Meeting, Anaheim, CA, June 12–16, 2022, https://www.ans.org/meetings/am2022/session/view-1203/, accessed 7/28/2022.
17. M. Muhlheim, P. Ramuhalli, A. Huning, A. Guler, R. Wood, and A. Saxena (2022), "Regulatory Criteria Applicable to the Use of Digital Twins," Digital Twins Working Meetings to Identify Needs, Washington DC, June 22, 2022
18. Guler-Yigitoglu (2021), A., "Artificial Intelligence for Risk Assessment of Complex Engineering Systems: Nuclear Industry Applications", panel on Risk Management and Industrial Artificial Intelligence at the INFORMS Annual Meeting 2021, Anaheim CA October 24-27
19. Coble, J. (2021), "Enhancing Risk Assessment With Greater Situational Awareness", panel on Risk Management and Industrial Artificial Intelligence at the INFORMS Annual Meeting 2021, Anaheim CA October 24-27.
20. Saxena, A. (2021), PMDT as an example of ML and AI based predictive maintenance development towards O&M cost reduction in nuclear power plants, NHA Clean Currents 2021 Conference in Atlanta GA, Oct 19-22, 2021 https://cleancurrents.org/
21. Ramuhalli, P. (2021), "Artificial Intelligence and Machine Learning Technologies: Enabling Advanced Reactor Deployment Economics ", GAIN AI/ML Technologies for Advanced Reactors Virtual Workshop (Oct 5-6, 2021)
22. Ramuhalli, P. (2021), panel discussion on Technologies for Asset Management at the 2021 Asset Management (virtual) conference.
23. Virani, N. (2021), "Humble AI for reliable machine learning-based health twins" in Digital Twin Enabling Technologies in Advanced Reactor Applications session at the NRC Workshop on Enabling Technologies for Digital Twin Applications for Advanced Reactors and Plant Modernization – September 14-16, 2021
24. Askwith, N., Saxena, A. (2021), "Digital Solutions for Nuclear Power Plants" EnergieFjorsk, Nov 17, 2021
25. Virani, N. (2021), "Humble AI concept for Digital Twins" invited presentation at AIRES2.0 workshop (Jan2021)
26. Huang, H. (2021), "Causal Modeling for Anomaly Detection (Health Twins)" Contributed Presentation at AIRES2.0 workshop (Jan2021)
27. Saxena, A. (2020), "AI Enabled Predictive Maintenance Digital Twins for Advanced Nuclear Reactors", Invited Presentation on GEMINA PMDT Project at INL DICE workshop (Dec2020)
28. Saxena, A. (2020), "AI for sustainment of critical infrastructure", Panel on sustainability at the annual Nvidia GTC conference, October 5-9, 2020.

## Conclusions

The PMDT team made substantial progress on a number of technical fronts. We utilized the BWRX-300 simulator extensively to build twins for other key components as well as for simulating more complex scenarios (e.g. simultaneous degradations in multiple plant components). We were able to demonstrate in many cases that Machine learning-based methods can be successfully adapted for Nuclear plant environments especially for remote monitoring applications. Detailed analyses were carried out with plant and full scope simulation data along with capabilities of enhanced analytics to assess and set realistic expectations on cost reductions in O&M. These assessments are paving the way for investments towards reactor design improvements as well project planning for SMR projects as they develop and mature in the next few years. Technology developed under this program got direct visibility to GE Hitachi and their utility customers and resulted in positive intents to deploy some of the elements from design phase. The project additionally resulted in several reports, publications, software and data generation that will be useful in deployment and O&M services for BWRX300 fleets.

The GEMINA PMDT project team thanks the ARPA-e team for their support and interest in this work. We strongly believe this paves the path for a more efficient O&M and extends the state-of-the-art in deploying ML-based solutions in nuclear contexts towards realizing the goal of a more economic nuclear power generation.