INL/JOU-21-64929-Revision-0

Deceptive Infusion of Data: A Novel Data Masking Paradigm for High-Valued Systems

April 2022

Arvind Sundaram, Hany Abdel-Khalik, Ahmad Y Al Rashdan

Changing the World's Energy Future

INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

INL/JOU-21-64929-Revision-0

Deceptive Infusion of Data: A Novel Data Masking Paradigm for High-Valued Systems

Arvind Sundaram, Hany Abdel-Khalik, Ahmad Y Al Rashdan

April 2022

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Nuclear Science and Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/unse20

Deceptive Infusion of Data: A Novel Data Masking Paradigm for High-Valued Systems

Arvind Sundaram, Hany Abdel-Khalik & Ahmad Al Rashdan

To cite this article: Arvind Sundaram, Hany Abdel-Khalik & Ahmad Al Rashdan (2022) Deceptive Infusion of Data: A Novel Data Masking Paradigm for High-Valued Systems, Nuclear Science and Engineering, 196:8, 911-926, DOI: 10.1080/00295639.2022.2043542

To link to this article: https://doi.org/10.1080/00295639.2022.2043542

Published online: 07 Apr 2022.

NUCLEAR SCIENCE AND ENGINEERING · VOLUME 196 · 911–926 · AUGUST 2022

DOI: https://doi.org/10.1080/00295639.2022.2043542

Deceptive Infusion of Data: A Novel Data Masking Paradigm for High-Valued Systems

Arvind Sundaram,[a]* Hany Abdel-Khalik,[a] and Ahmad Al Rashdan[b]

[a]*Purdue University, 205 Gates Road, West Lafayette, Indiana 47906*
[b]*Idaho National Laboratory, 1955 North Fremont Road, Idaho Falls, Idaho 83415*

Received November 5, 2021
Accepted for Publication February 8, 2022

Abstract — *This work addresses how analysts of a high-valued system (e.g., nuclear reactor, aircraft turbine designs) can extract findable, accessible, interoperable, and reusable scientific data for public dissemination to artificial intelligence and machine-learning (AI/ML) researchers in a manner that cannot be reverse-engineered, potentially compromising sensitive or proprietary information. State-of-the-art methods address this problem through data masking techniques, which allow access to a subset of the information while obfuscating private and potentially identifying information (e.g., personally identifying medical data). These methods are unsuitable for industrial engineering processes, where AI/ML tools need explicit access to all the data available to draw the best inference about the system to help optimize its performance and identify its vulnerabilities, etc. Our novel deceptive infusion of data paradigm provides a solution to this conundrum by developing a mathematical approach capable of concealing the identity of the system while providing full access to all the features employed by AI/ML tools to ensure their optimal performance.*

Keywords — *Data anonymity, data manipulation, mutual information, data masking.*

Note — *Some figures may be in color only in the electronic version.*

I. INTRODUCTION

Industry 4.0 has revolutionized critical infrastructure systems via digitization and data informatics [i.e., artificial intelligence and machine learning (AI/ML)] as effective tools to integrate computations, network communications, and physical processes. Examples of these systems, which are referred to as cyber-physical systems,[1] include energy-producing chemical, nuclear, oil, and gas units; electric energy distribution and transportation; air traffic control; water treatment facilities; healthcare systems; banking; etc. The success of this revolution is contingent upon the ability to unlock the full potential of AI/ML tools for the various activities associated with the design, deployment, and operation of these systems. For example, there is a strong need to employ AI/ML tools to optimally integrate computable knowledge from physics models with experimental data for a wide range of applications, such as condition monitoring, vulnerability analysis, system performance optimization, intrusion detection, and autonomous control.[2] To realize this vision, AI/ML researchers must be granted access to the system data in a manner that ensures the optimal performance of AI/ML tools. Unfortunately, owners of critical systems are often reluctant to share knowledge about their systems for various reasons.[3] For example, an AI/ML application aiming to identify signatures for equipment degradation performance (e.g., a valve or a pump) using flow rate and temperature measurements in the form of time-series data may inadvertently lead to the discovery of design faults or vulnerabilities or may be misused to reverse-engineer some

*E-mail: sundara4@purdue.edu

of the equipment proprietary details. Some of these outcomes may result in undesirable repercussions for the system owners, reducing support from both the public and investment communities. Thus, the question is how does one protect the identity of the proprietary system by masking proprietary data while also transmitting relevant information.

Historically, the need for data masking has followed the increased adoption of data warehouses in industry that hold critical business information about enterprises and make appealing targets for malicious agents, insiders, and competitors.[4,5] In an effort to protect the identity of data, data masking efforts, such as substitution, shuffling, and differential privacy, are often enacted prior to transmission.[6–8] The core idea behind such anonymization/masking practices is to hide critical information that is irrelevant to the target application in a manner that cannot be reverse-engineered (i.e., a malicious party cannot glean information about the masked data from the unmasked data). Additionally, the unmasked data must also serve their target purpose and remain usable for the intended application.[5,9]

Existing data masking techniques often rely on the assumption that the target application is insensitive to the masked data. While this may be true in the case of identifying information, such as social security numbers (SSNs), other sensitive information such as age may be a factor for the target application, especially in healthcare.[10] In such cases, the data are often encrypted prior to transmission to prevent unauthorized access. However, encryption routines often lead to massive overheads in response time and storage space, and they degrade performance.[4] Additionally, they do not prevent potential misuse by the trusted third-party AI/ML researchers after decryption, and consequently, the release of such information is often accompanied by legal agreements to prevent disclosure and inference of private data.[11]

In the context of industrial data, such as sensor data from experiments, privacy needs are significantly different from data warehouses in that it is often infeasible to completely prevent access to the private information. In this case, existing association rules and correlations within the dataset must be accessible to maximize the benefits of AI/ML techniques and to remain invariant to masking techniques. For example, if the original data describe a turbine operating at higher and lower temperatures and show distinct patterns in the two cases, the masked data must also reflect the distinction to be relevant to an analyst. The alteration or masking of the data that voids existing correlations for the sake of privacy may lead to nonoptimal or incorrect conclusions and a degradation in the quality of the AI/ML model. The present work addresses the challenge of data anonymity for industrial data by proposing a noninvertible masking deceptive infusion of data (DIOD) methodology that preserves the necessary association rules and correlations for AI/ML tools to operate, while masking proprietary information.

Figure 1 depicts the flow of data in the nuclear industry, tracing the path from their origin at nuclear power plants, followed by storage in data warehouses, provision to research organizations for the development of AI/ML-based data analytics, and culminating in the data-based insight leveraged by power plants. The DIOD methodology, depicted as a shield on public data, seeks to protect the outgoing data from the data warehouses through obfuscation while preserving their usability. In other words, DIOD anonymizes the outgoing data by removing all associations with the source, i.e., the nuclear power plant, while maintaining the inferential properties (association rules, correlations, etc.) for the AI/ML researchers. From the perspective of the nuclear industry, this enables proprietors to collaborate without fear of data misuse while providing them the necessary insight since the sensitivity of the data is no longer an issue. Other applications include but are not limited to safeguards inspection, handling sensitive images, and obfuscating security mechanisms to make them appear benign to unsuspecting adversaries.

To achieve this goal, this work explores, given a physical system (i.e., a system that behaves according to some governing laws), the feasibility of developing a data masking methodology capable of concealing the identity of the associated system while retaining all the statistical dependencies required to allow researchers to explore the use of their methods to optimally realize the highly promoted benefits for AL/ML techniques. Can the masked data be used as a benchmark preparation methodology to gauge various AI/ML techniques since a good technique is expected to be invariant to the DIOD transformations?

Departing from existing privacy preserving, data masking, and anonymization techniques,[12–14] we developed a novel data masking paradigm based on a noninvertible DIOD methodology. The methodology is designed to preserve the correlations among the benchmark datasets, representing the target information harvested by AI/ML techniques, while obfuscating the fundamental structure of the system's confidential underlying governing laws, achieved in a noninvertible manner that cannot be self-learned from the benchmark datasets. This DIOD methodology represents a paradigm shift to

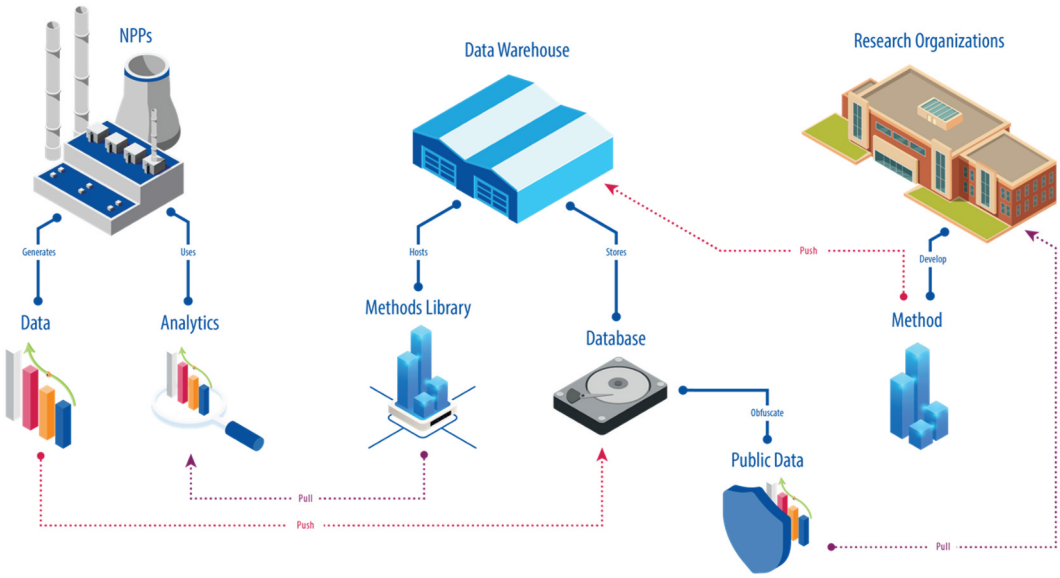

Fig. 1. Pictorial representation of the needs of industrial plants; the data generated by nuclear power plants must be obfuscated before dissemination as public data to research organizations that develop AI/ML analytical tools for additional insight.

data anonymity methods, such as k-anonymity,[15] l-diversity,[16] t-closeness,[17] and m-invariance.[18] These methods focus on limiting access to structured datasets, such as personally identifiable information and health records. To protect the data, the methods either use non-discriminative loss of information or introduce ambiguity though means such as perturbation, encryption, or suppression that changes the inference characteristics of data and impacts the application of AI/ML. These methods provide limited protection because of their core fundamental flaw, which is that they do not impact the underlying governing structure.[19] The obfuscation of data provenance can be compromised if enough data are provided through single or multiple data dissemination events.

The DIOD's new paradigm focuses on overcoming this limitation by obfuscating the underlying governing laws, without impacting the application of AI/ML methods. It expands the scope of data protection from data warehouses to scientific types of data, such as time-series data from critical experiments. We developed this DIOD methodology using a nonunique AI/ML-invariant mathematical transformation for the disseminated data. This is achieved over two steps: a decomposition step and a fusion step. In the decomposition step, a randomized range-finding algorithm[20,21] is used to decompose the benchmark multivariate data into two sets of independent metadata.

The first set is referred to as fundamental metadata, which describes the underlying governing laws (often represented by a combination of physics principles and constraints). As such, it can be tied to the identity of the system that generated the benchmark datasets and must be concealed to prevent malicious third parties from undesirably reproducing the data.

The second set is referred to as inference metadata, which is used to train the AI/ML. In the fusion step, a mathematical kernel, which employs a library of precalculated concealment operators, is used to fuse the system-specific inference metadata with another set of fundamental metadata that are representative of a similar or markedly different system. The fusion step allows one to preserve the inference metadata as required by AI/ML training while concealing the identity of the associated system to different levels of privacy as required by the user.

The invariance property ensures that the AI/ML performance is not impacted by the transformation whereas the "nonuniqueness" means that any inference/inverse analysis attempting to estimate the transformation would be infeasible (i.e., many possible system identities and possible concealment operators exist). By way of example, the information about the core inlet temperature of a nuclear reactor operating at different conditions is concealed in the current profile of a generic direct-current permanent magnet (DCPM) motor in this paper. Since the second system is generic, its fundamental metadata are well-known, thus permitting the extraction of all the relevant inference metadata for AI/ML applications. However, the inference metadata cannot be

sourced to the fundamental metadata of the nuclear reactor, granting anonymity since the inverse problem consists of finding both the fundamental metadata and the deception kernel of the system. Further obfuscation of the inference metadata is also possible and will be explored in future work.

The paper is organized as follows. Section II provides a summary of various data masking techniques and the motivation behind DIOD, followed by a mathematical description of the problem statement in Sec. III. In Sec. IV, we demonstrate the anonymization procedure by embedding the temperature profile obtained from the simulation of a nuclear power plant into the current profile of a DCPM motor. Then, we seek to validate the above claims by using statistical and AI/ML methods to distinguish the data before and after the masking process and extract the metadata. Last, Sec. V summarizes and discusses the results of the DIOD implementation.

II. BACKGROUND AND LITERATURE REVIEW

This section describes a few types of existing data masking methods. While the methods outlined in this section are applicable to data warehouses, the privacy needs of industrial data are significantly different as they are bound by physical laws and other domain-related constraints. Data masking, broadly classified into static and dynamic methods, protects sensitive information from unintended exposure by masking the environment and only providing the necessary information for the target application.[5] In static methods, sensitive information is permanently altered via substitution or removing connections between data fields (shuffling). For example, if an application requires records of the gender distribution in a company, the real names of the employees may be substituted by common names to protect their identities. Additionally, the link between their birthday and SSN may be removed by shuffling the fields prior to transmission. Since these changes are permanent, oftentimes a copy of the original database needs to be created. Dynamic methods, on the other hand, mask information as they reach the recipient. Masking out is one such example of dynamic masking where all but the last four digits of bank account numbers are obscured when requested. Dynamic methods do not require the creation of a copy database and can work in real time unlike static procedures. However, they are not suited for environments where data could be written back and corrupted or where the masking procedure is bypassed and the original data are available.[22]

In recent decades, a type of data masking known as differential privacy has emerged that allows for small statistical perturbations in data to protect sensitive information.[23] Its goal is to prevent an end user from obtaining more information by piecing together bits of information from different queries, popularly known as the reconstruction attack.[24] The notion of differential privacy is a consequence of the informal Fundamental Law of Information Recovery that states "overly accurate answers to too many questions will destroy privacy in a spectacular way."[8] It is thus inferred that some statistical distortion, such as the addition of noise, is necessary to prevent complete reconstruction (i.e., uncertainty must be induced at the individual level while preserving group parameters).

Another well-known example of local differential privacy is that of a randomized response when dealing with surveys of illegal behavior. For example, consider the case of estimating the probability of cheating in university exams. A randomized response procedure would ask survey takers to toss a coin and answer the question honestly if heads and simply answer an arbitrary response otherwise. This protects the survey taker from releasing potentially incriminating information while providing sufficient information to predict the probability of cheating during an exam.

As mentioned previously, the privacy needs of industrial data are significantly different in that existing correlations between data must be preserved, and it is often infeasible to completely mask out such information, thus rendering existing data masking paradigms unsuitable. For example, from the perspective of a consumer, with the advent of the smart grid, utility providers often collect massive amounts of consumer-related information, such as location, energy consumption, and usage patterns.[25–27] Moreover, potential unauthorized access to just energy consumption data by a target household may lead to inference about the number of individuals and their behavioral patterns, real-time surveillance, and location tracking.[28–30] While such information is sensitive, it is also relevant to AI/ML algorithms that optimize electricity generation and supply from the smart grid and cannot be masked using conventional methods.

On the proprietor end, a lack of control over data usage by third-party services, legal uncertainties over data ownership, and the potential liability due to data leaks or exposure are significant barriers in promoting data sharing and open access despite sanitization procedures.[31] For instance, recent research by Accenture has demonstrated that public exposure of sensitive data leads to an almost 10% decline in revenue for up to 6

months after the breach.[32] An emerging solution to the problem is the so-called privacy preserving computation technique that seeks to provide control of the environment that the data can be operated on, obscure the data to protect their privacy and remove identifying traits, and render the data operable while encrypted without decryption.[33]

At its core, privacy preserving computation relies on encrypting sensitive data using fully homomorphic encryption, allowing researchers to manipulate and analyze encrypted data, and only requiring decryption at the end to obtain the necessary group parameters.[34] This makes it an ideal candidate, especially in the healthcare industry, where the data of individual patients are encrypted and protected while also permitting a statistical analysis of the group of patients. However, such encryption is in its infancy and remains commercially infeasible due to slow computation speeds and potential issues with accuracy (i.e., there is a lack of scalability among these applications that is required of vast amounts of industrial data).[35] The lack of scalability is a serious concern for AI/ML applications since high encryption overhead costs hamper the generation of the vast amounts of data required for deep-learning models that require large datasets to estimate millions of trainable parameters (e.g., Resnet).[36]

The proposed DIOD methodology is significantly different from existing methods in the sense that it does not require encryption and relies on existing reduced-order modeling (ROM) techniques to decompose the data into independent sets of fundamental metadata and inference metadata. The former refers to the governing physical laws relevant to the identity of the system while the latter refers to the operational information relevant to AI/ML tools. The reduction may be across space, time, multiple sensors, etc., and is scalable to large datasets.[37] While ROM techniques may have high initial costs to obtain a low-dimensional representation of a given system, they need to be performed only once per system and offer the advantage of subsequent inexpensive computations. Furthermore, it is feasible to build a precalculated library of concealment operators that can then be deployed in an efficient and scalable manner for a wide variety of industrial systems. It may be customized to the desired level of masking, and the inference metadata can be extracted from the masked data by invariant AI/ML tools. Section III provides a mathematical basis for the proposed methodology and introduces the concept of mutual information to validate our hypotheses.

III. PROBLEM STATEMENT

This section formulates the problem statement and provides a mathematical foundation for the DIOD methodology. The first step, as described above, is the decomposition of a given industrial process into its fundamental and inference metadata to separate the underlying phenomena from the operational data. In general, this is performed using ROM techniques that have matured over the past several decades.[20,38–40] A general description of the above process is given by Eq. (1), where industrial data y are decomposed into their constituent metadata:

$$y(x,\alpha) \approx \sum_{i=1}^{r} \psi_i(x)\phi_i(\alpha) \tag{1}$$

and

$$\left| y(x,\alpha) - \sum_{i=1}^{r} \psi_i(x)\phi_i(\alpha) \right| < \in , \tag{2}$$

where ψ carries proprietary information denoted as fundamental metadata, such as the geometry of the system, material composition, and underlying differential equations, that are crucial to the identity of the system. Here, x may denote the physical position in space/time, pixel location on an image, etc. On the other hand, ϕ carries information about the operational history of the system denoted as inference metadata such as the temperature, mass flow rate, and other control parameters α that are relevant to the AI/ML applications for optimization or inference purposes.

In the above notation, the r functions $\psi_i(x)$ span the active subspace describing the dominant trends in the data to a user-defined tolerance of $\in$ as described in Eq. (2). The active subspace of the system's response $y(x,\alpha)$ may be found using the rank identification algorithm introduced in the nuclear community.[20,38] The gradient of the function at k randomly chosen points may be computed and collated into a single gradient matrix $\mathbf{G}$, which is then decomposed to find the basis of the active subspace as shown in Eq. (3), where the first $r < k$ dominant columns of the $\mathbf{Q}$ matrix form an orthonormal basis of the active subspace spanned by $\psi_i(x)$:

$$\mathbf{G} = \left[\; y|_{x_1} \;\; y|_{x_2} \ldots \;\; y|_{x_k} \right] = \mathbf{QR}. \tag{3}$$

The gradient subspace can represent all possible variations of the response to the user-defined tolerance $\in$.

The second step is the identification of a generic system that is well understood to form the basis for the concealment operators on which the inference metadata of the proprietary system are fused. Using the above ROM techniques, the generic system can also be decomposed into independent sets of metadata that describe their identity and operation as shown in Eq. (1). Depending on the desired level of masking, additional constraints may be imposed on the generic system, and invariant transformations, such as scaling, shifting, and rotation, may be applied on the inference metadata for the AI/ML application. Once the target application is identified, the concealment operators φ are computed based on the fundamental metadata of the generic system. These operators are mathematical operators that prevent the inference of the fundamental metadata of the proprietary system.

The final step is the generation of the DIOD version of the industrial data that can be transmitted to a third-party AI/ML service. This may be accomplished by a deception kernel $k(x',x)$ using the concealment operators as shown in Eqs. (4) and (5):

$$k(x',x) = \sum_{i=1}^{r} \varphi_i(x')\psi_i^*(x) \tag{4}$$

and

$$y'(x',\alpha) = \int k(x',x)y(x,\alpha)dx \ , \tag{5}$$

where

$*$ = operator on $\psi_i(x)$ that denotes the Hermitian conjugate

x' = spatial/temporal/pixel information for the generic system

$y'(x',\alpha)$ = DIOD rendition of the data $y(x,\alpha)$.

The deception kernel is a transformation that effectively overwrites the fundamental metadata of the proprietary system $\psi(x)$ with that of the generic system $\varphi(x')$, thus successfully masking the proprietary information. The invariance lies in the fact that the inference metadata $\phi_i(\alpha)$ are separated out via the ROM-based decomposition and simply mounted onto the generic system represented by $\varphi(x')$. In other words, assuming an orthonormal basis is found via ROM techniques, one considers the following decomposition of the DIOD data shown in Eq. (6):

$$y'(x',\alpha) \approx \sum_{i=1}^{r} \varphi(x')\phi_i(\alpha) \ . \tag{6}$$

The above representation indicates that the decomposition of the DIOD data yields information about the inference metadata of the proprietary system and the fundamental metadata of the generic system.

The core advantage of the DIOD approach is that the proprietary system is effectively anonymized since its fundamental metadata cannot be extracted from the dataset y'. However, an AI/ML technique recovers the same inference data $\phi_i(\alpha)$ from the datasets y and y' since the fundamental metadata of the generic system is well-known and AI/ML techniques are invariant to such transformations. The noninvertibility of the DIOD methodology is observed via the overwriting action performed by the kernel. The task of finding the original fundamental metadata $\psi(x)$ given $y'(x',\alpha)$ is ill-posed since any basis of functions may provide an active subspace of rank r; in other words, the kernel $k(x',x)$ is required to uniquely identify $\psi(x)$. Note that the kernel itself cannot be identified unless the third party is given access to the original dataset $y(x,\alpha)$. Also, the kernel $k(x',x)$ itself is invertible, which is valuable in the context of mutual information as discussed below.

Mutual information, or mutual entropy, is a measure of the dependency between two random variables that captures all functional relationships, linear or otherwise. Unlike other measures of dependency, such as correlation coefficient, mutual information is invariant under invertible transformations, such as the kernel operator, and is only affected by the entropy of the data (e.g., noise). In recent decades, the idea of mutual information has been incorporated into AI/ML techniques to generate features that maximize the mutual information between the response variable and its inputs[41] and as an attempt to explain the behavior of deep-learning models.[42] The inner product of the data with such features consequently provides information about the significance of each feature. In the context of this discussion, the features from mutual information-based AI/ML techniques are representative of the fundamental metadata while the inner product represents the inference metadata. Thus, by infusing the concealment operator with the inference metadata using the deception kernel, the mutual information of the response and the inference metadata I is the same before and after the DIOD transformation as shown in Eq. (7):

$$I(y';\phi_i(\alpha)) = I(y;\phi_i(\alpha)) \ . \tag{7}$$

To further illustrate the invariance property, consider a classifier f that trains on the original input $y(x,\alpha)$ attempting to classify new data $z(x,\alpha)$. The mutual information between the classification label and the inference metadata is the theoretical upper limit of separability between the classes in the dataset.[43] Since the labels for the DIOD data y' and the original dataset y must be the same, and the inference metadata $\phi_i(\alpha)$ are also the same based on the above methodology, the upper limit of classifiability is unchanged. The deception kernel $k(x',x)$ is thus designed to ensure that the classifier produces a similar classification performance when trained on the DIOD data $y'(x',\alpha)$ and attempting to classify new data $z'(x',\alpha)$, symbolically represented via the identity in Eq. (8):

$$f_y(z) = f_{ky}(z') = f_{y'}(z') \ . \tag{8}$$

The proposed DIOD methodology is demonstrated in Sec. IV using a Westinghouse pressurized water reactor (PWR) model representing a proprietary system and a DCPM model representing a well-known generic system. The property of invariance to AI/ML techniques is illustrated via manually crafted features based on domain knowledge (supervised), k-means clustering (unsupervised), and principal component analysis (unsupervised). The complete masking of the fundamental metadata of the PWR model is demonstrated via principal component analysis and an analysis of the correlation between multiple system responses.

IV. IMPLEMENTATION

In this section, the DIOD methodology is implemented using the simulation of a nuclear reactor and a DCPM representing a proprietary and generic system, respectively, for data masking. The inlet temperature of the reactor and the current output are the measured quantities, and the above simulations are carried out in Dymola (version 2020x) (Refs. 44 and 45). The Westinghouse four-loop PWR example from the TRANSFORM package[46] and the current-controlled DCPM from the default Modelica package are simulated under various operating conditions until steady state for 100 s. In this implementation, the inlet temperature represents sensitive information that must be masked since it can be linked to a nuclear reactor based on its temporal evolution by a knowledgeable adversary. However, the data must be masked in a manner that preserves the AI/ML-relevant inference metadata as obtained from the decomposition in Eq. (1). To this end, the current-controlled DCPM serves as an ideal canvas to mount the inference metadata since it is representative of a noncritical, generic, and well-understood system. The goal of the DIOD methodology is to fuse the inference data from the PWR with the fundamental metadata of the DCPM to create the DIOD version of the data. This is accomplished using Eqs. (1), (2), and (3) via kernel deception using the concealment operators developed for the DCPM system. The fundamental and inference metadata of the PWR physically describe the temporal evolution of the inlet temperature (parameterized by x) and the operational conditions (parameterized by α), respectively. The same holds for the DCPM with regard to its current sensor.

In the following sections, the DIOD methodology is implemented and validated via statistical and machine-learning techniques. They are organized as follows. In Sec. IV.A, the inference metadata from the PWR under two different simulation conditions are concealed within the DCPM responses. The validation process is performed using simple statistical tools to ensure that in-class and between-class separability are maintained. Here, "separability" refers to the ability of an ideal classification algorithm to distinguish between the two classes/simulation conditions (between-class) or between multiple datasets within the same class (in-class). In Sec. IV.B, we assume that the target application does not have any preexisting knowledge of the dataset and implement an unsupervised k-means algorithm to simulate the inference of AI/ML tools. The goal is to ensure that the same separability exists in the inference metadata of the PWR and the DIOD data. In Sec. IV.C, a more complex analysis, such as singular value decomposition (SVD), is performed to extract information about both the fundamental and inference metadata of the system. The validation process ensures that the DIOD data and the PWR data share the same inference metadata but different fundamental metadata to conceal proprietary information and verifies this using the mutual information of the response and the inference metadata. Last, in Sec. IV.D, two responses from each system are considered, and the correlation between these responses is computed to produce correlation curves that characterize the fundamental metadata of the system. Here, the DIOD methodology is validated by ensuring that the DIOD correlation curve exhibits significantly different behavior than that of the PWR.

IV.A. DIOD Validation Using Domain Knowledge

In this section, the the DIOD methodology is demonstrated by mounting the inference metadata of the PWR

simulations on the current-controlled DCPM as seen in Fig. 2. The inlet temperature of the PWR is simulated under two different conditions, namely, normal operation and partial pump failure as shown in Fig. 2a. For each condition, five datasets were generated by varying the mass flow rates and the final pump revolutions per minute (RPM), respectively, although only one set from each condition is plotted below for visual acuity. The goal of the paper is to anonymize the above data by mounting the inference metadata of the two operating conditions onto the current-controlled DCPM simulation as seen in Fig. 2b. Figure 3 depicts the DIOD version of the inlet temperature data from Fig. 2 after applying the methodology as described by Eqs. (1), (2), and (3).

The DIOD version of the data is similar in appearance to the operation characteristic of the current-controlled DCPM, and the visual separability is maintained within the two operating conditions: partial pump failure and normal operation. Using domain knowledge that the DCPM current follows an exponential profile, statistical analysis reveals the separation of the two classes in the DIOD data via a significantly different time constant. Additionally, the in-class separability in each class is in the final saturation value of the DIOD curves as tabulated in Table I. The fraction of

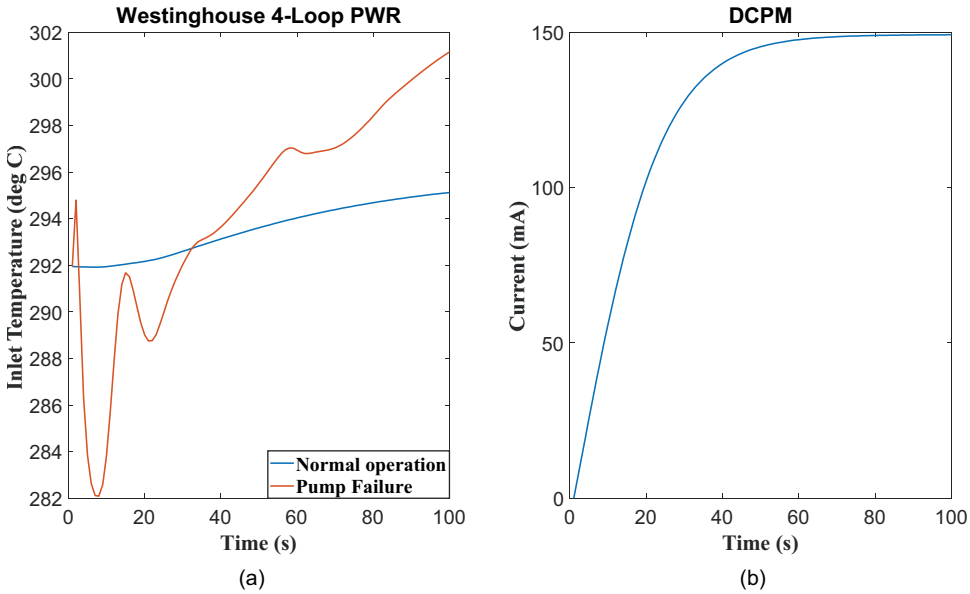

Fig. 2. Simulation data from PWR and DCPM.

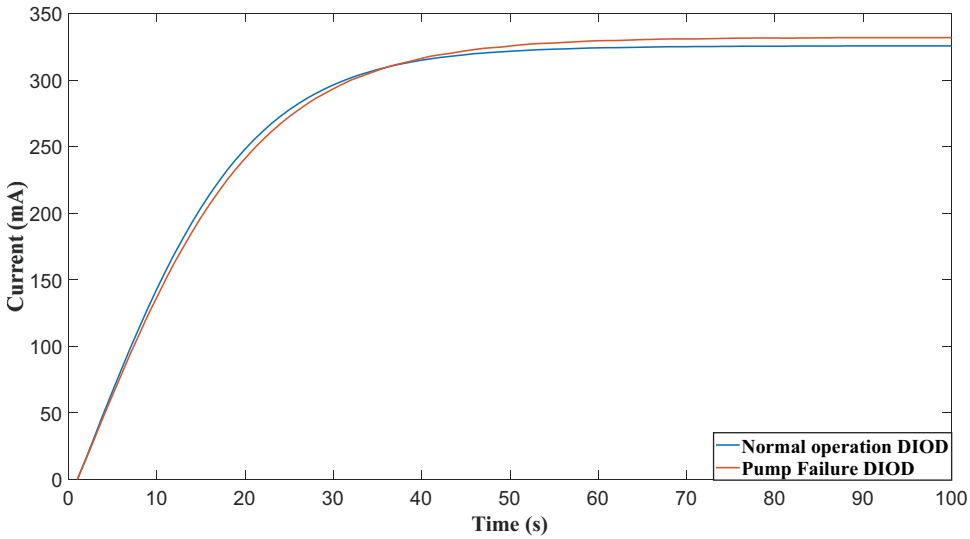

Fig. 3. The DIOD version of PWR data.

TABLE I
Statistical Analysis of DIOD Data

Operation Mode	Normal Operation Mass Flow Rate (kg/s)					Partial Pump Failure (μ)				
	450	460	470	480	490	0.2	0.25	0.3	0.35	0.4
DIOD time constant (s)	16	16	16	16	16	17	17	17	17	17
DIOD saturation current (mA)	327.8	327.2	326.7	326.2	325.7	331.9	331.4	330.8	329.6	329.2

pump power at steady state after partial pump failure is denoted by μ.

IV.B. DIOD Validation Using Unsupervised AI/ML

The next task is to use an AI/ML tool to classify the ten datasets based on their operating conditions (i.e., normal operation and partial pump failure without domain knowledge). In this section, we assume that there is no a priori knowledge of the operating conditions/classes and utilize an unsupervised learning algorithm to cluster the datasets. The PWR inlet temperature is simulated under normal operating conditions and partial pump failures by varying the mass flow rate and fraction of pump power at steady state, respectively, as shown in Sec. IV.A. The resultant time series are separated into two clusters using a k-means clustering algorithm without explicitly labeling the membership class of each dataset or the number of members in each class (unsupervised). Table II shows that the two clusters actually correspond to the normal operation and partial pump failure conditions. This can be attributed to the vastly different structures of the two curves as seen in Fig. 2a.

Then, the DIOD methodology as described by Eqs. (1), (2), and (3) is applied onto the inlet temperature data using the fundamental metadata from the DCPM. The above process is repeated, and the k-means clustering algorithm is applied on the DIOD datasets. The class separability is still maintained in the DIOD datasets and corresponds to the normal operation and partial pump failure conditions. Table II shows the results below and demonstrates that the class separability is maintained even in an unsupervised learning environment where we assumed that the tool has no domain knowledge of the system.

IV.C. DIOD Validation Using SVD

In Secs. IV.A and IV.B, the extraction of inference metadata and their separability was demonstrated using statistical and AI/ML tools. In this section, we use SVD to extract information about both the fundamental and inference metadata from the PWR, DCPM, and DIOD data. SVD is a widely used analysis tool and is defined via an orthogonal transformation of the data onto a new set of coordinate axes ordered from greatest to least variance. The given dataset $\mathbf{z}$ is decomposed as shown in Eq. (9), where $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ describe the n dominant features of the dataset and $\beta_1, \beta_2, \ldots, \beta_n$ describe their respective coefficients:

TABLE II
k-Means Clustering of DIOD and PWR Data

Operation Mode	Normal Operation Mass Flow Rate (kg/s)					Partial Pump Failure (μ)				
	450	460	470	480	490	0.2	0.25	0.3	0.35	0.4
Cluster index, PWR data	1	1	1	1	1	2	2	2	2	2
Cluster index, DIOD data	1	1	1	1	1	2	2	2	2	2

$$\mathbf{z} \approx \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \ldots + \beta_n \mathbf{u}_n \ . \tag{9}$$

The extracted $\mathbf{u}_i$ vectors describe the behavior of the physical system itself and provide some insight into the fundamental metadata of the system. Their coefficients β_i describe the operation of the system and provide information about the inference metadata of the system. The DIOD methodology is validated by preserving existing correlations among the coefficients β_i of the PWR and masking the fundamental metadata of the PWR.

Consider ten time series generated by varying the mass flow rate of the PWR under normal operation. Similar to previous sections, Eqs. (1), (2), and (3) are used to mount the inference data of the PWR onto the DCPM using the library of concealment operators developed earlier, thus generating the DIOD version of the PWR data. Using Eq. (9), the PWR data and the DIOD data are decomposed, and the first three coefficients β_1, β_2, and β_3 representing the dominant coefficients are analyzed. Figure 4 presents the correlation between the coefficients β_i obtained from the SVD of the PWR and DIOD data and shows that the correlations are preserved across the PWR and the DIOD datasets.

Next, we consider the first three dominant $\mathbf{u}$ vectors obtained from the decomposition of the PWR and its DIOD version using Eq. (4) as shown in Fig. 5. A successful implementation of the DIOD methodology completely masks the fundamental metadata of the proprietary system (PWR). The $\mathbf{u}$ vectors of the PWR containing information about the fundamental metadata do not resemble those extracted from the DIOD version, thus protecting the fundamental metadata of the PWR. Additionally, the extracted vectors of the DIOD data resemble those of the DCPM, implying that the underlying physical processes of the two are similar. Therefore, any analysis of the $\mathbf{u}$ vectors of the DIOD dataset provides insight into the generic DCPM system and not the proprietary PWR.

Last, the DIOD methodology is validated using the mutual information between the response variable and the inference metadata before and after implementing the DIOD methodology. The goal of this experiment is to ensure that the mutual information is invariant to the DIOD transformation to ensure a similar AI/ML performance as described in Eq. (7). Let the original response be $\mathbf{z}$ with dominant features $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ and the DIOD response be $\mathbf{z}'$ with dominant features $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Using the notation expressed in Eq. (9), we use the following properties of mutual information:

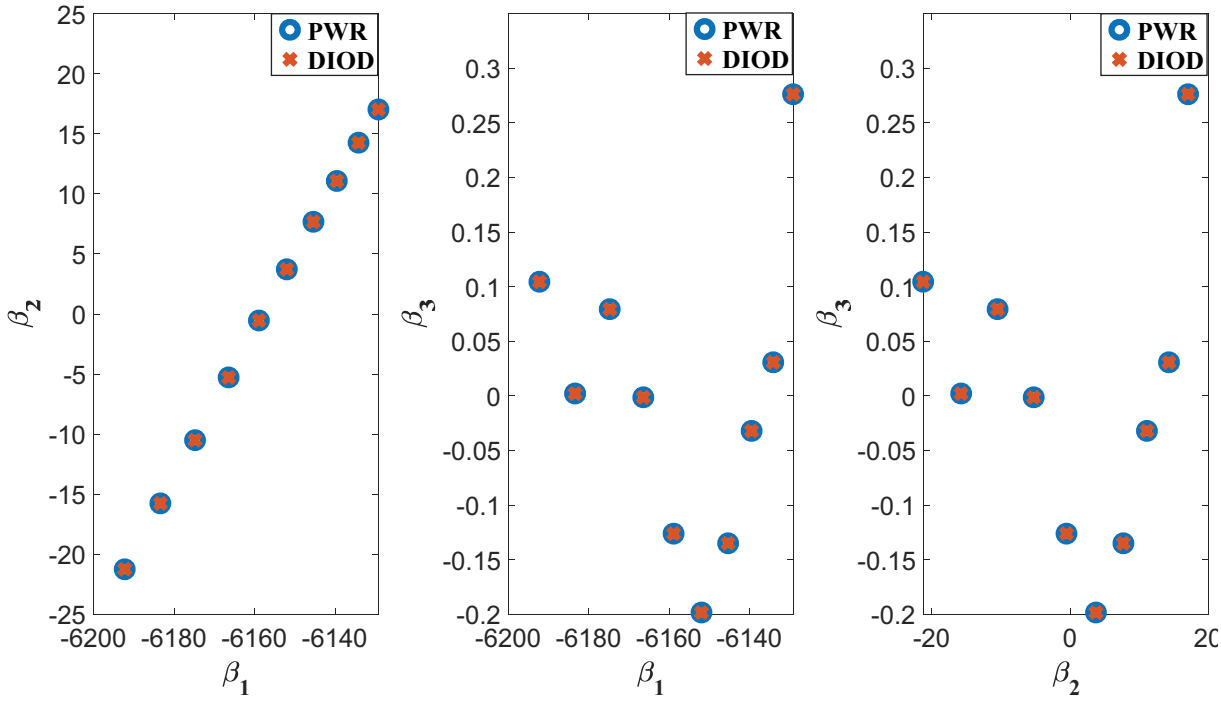

Fig. 4. Correlation between coefficients.

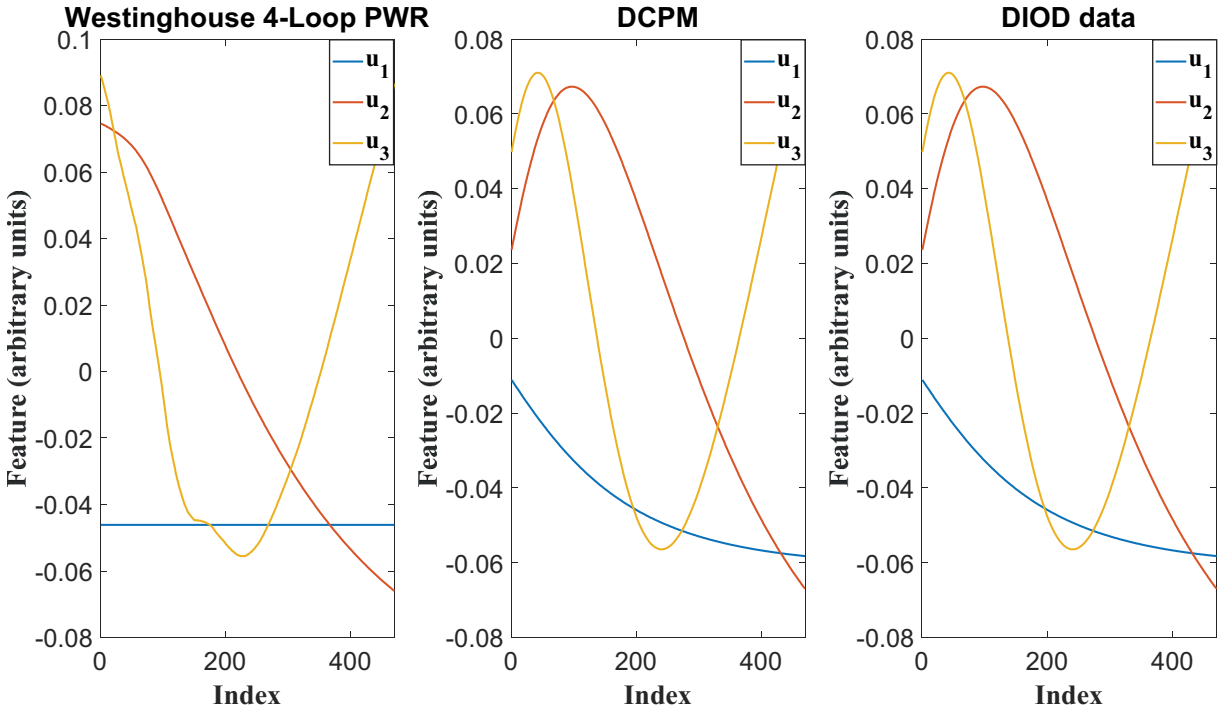

Fig. 5. First three **u** vectors from SVD.

$$I(\mathbf{z}';\beta_i\mathbf{v}_i) = I(\mathbf{z}';\beta_i) = I(\mathbf{z};\beta_i) = I(\mathbf{z},\beta_i\mathbf{u}_i) \ , \quad i = 1,2,\ldots,n \ .$$

The above holds true if there exists a smooth invertible transformation between $\mathbf{z}'$ and $\mathbf{z}$ and if $\mathbf{u}_i$ and $\mathbf{v}_i$ are constants, or more generally, only functions of β_i. From

the observed data, we see that the dataset $\mathbf{z}$ exists in the n-dimensional subspace spanned by the orthonormal set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ while $\mathbf{z}'$ exists in the n-dimensional subspace spanned by the orthonormal set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Therefore, the smooth invertible transformation between the two subspaces is given by $\mathbf{z}' = \mathbf{T}\mathbf{z}$, where $\mathbf{T} = \sum_n \mathbf{v}_i \mathbf{u}_i^T$ is a linear transformation denoting the sum of the outer product of the vectors $\mathbf{u}_i$ and $\mathbf{v}_i$. The rationale behind the invariance to $\mathbf{u}_i$ and $\mathbf{v}_i$ is the core of the kernel trick in many machine-learning algorithms as it allows for a variable to be projected to a higher-dimensional space (in this case, from a scalar β_i to a vector $\beta_i \mathbf{u}_i$) without changing the inherent information and dimensionality of the problem.[47] Also, note that $I(\mathbf{z}; \beta_i)$ is invariant to scaling or other smooth invertible transformations of $\mathbf{z}$ and β_i in general, assuming they are jointly generated from a fixed underlying model, and thus, nonlinear transformations may be used on the response variable to further obfuscate the proprietary system and mimic the properties of the generic system.

IV.D. DIOD Validation Using Response Correlation

In Sec. IV.C , the DIOD methodology was validated using SVD by preserving the correlations among the SVD coefficients of the responses representing inference metadata and masking the fundamental metadata of the PWR. In this section, the masking of fundamental metadata of the PWR is further demonstrated using the correlations among the responses themselves. Every physical system is expected to have its own set of correlations among its responses based on the underlying physical model. For example, the data from an experiment on a resistor may exhibit a linear relationship between the current and the voltage. A successful DIOD implementation masks these relationships so that any attempts at inference do not lead back to the original proprietary system. In this experiment, two responses from each system are considered, namely, the inlet and outlet temperatures from the PWR and the current and rotation speed of the DCPM. The correlation between the inlet and outlet temperatures is computed to produce correlation curves that characterize the fundamental metadata of the PWR.

Consider a DIOD implementation involving the above responses from the Westinghouse four-loop PWR and current-controlled DCPM system described at the beginning of Sec. IV. The PWR and the DCPM are simulated until steady state is achieved. Using Eqs. (1), (2), and (3), the inference metadata of the outlet temperature are mounted on the speed data while those of the inlet temperature are mounted on the current data.

The correlation curves of the PWR and DIOD datasets are shown for a representative case in Fig. 6 using the responses obtained from each system. The DIOD methodology is validated by observing that the correlation curve of the Westinghouse PWR is masked in the DIOD data, thus protecting it from discovery. Although not explored in this paper, the idea may be further extended by using invertible mathematical transformations to better fit the features of the generic system depending on the target application. For example, the DIOD inference metadata could be scaled/transformed to exhibit a relationship similar to a DCPM correlation curve. The DIOD data can thus be reasonably expected to have come from the DCPM system, achieving another level of masking if desired.

V. DISCUSSION AND CONCLUSION

The past decade has overseen the integration of AI/ML techniques with industry for optimization, intrusion detection, predictive maintenance, etc. Nevertheless, industrial companies are often reluctant to share sensitive data due to the potential implications of data leaks, reverse-engineering, ownership issues, privacy violations, etc., often resulting in loss of revenue or competitive edge. However, such sensitive information cannot be simply hidden from third-party AI/ML services if one was to maximize its benefits. Thus, there is a need to address the question of data anonymity while extracting relevant information about a proprietary system. The present paper addresses this via the DIOD methodology. The proposed methodology is scalable and overcomes many of the limitations of existing data masking techniques, such as encryption, substitution, and shuffling. In the DIOD method, the given industrial data are decomposed into their fundamental and inference metadata where the former describes proprietary information, such as the physical parameters and underlying laws, while the latter describes operational characteristics that are relevant to AI/ML algorithms. The inference metadata are then fused with the fundamental metadata from a well-known and noncritical generic system, thus permitting their extraction by invariant AI/ML algorithms. The fusion process is done using a set of carefully designed mathematical operators known as concealment operators to protect the identity of the proprietary system.

The proposed DIOD methodology is demonstrated via Dymola simulations of a Westinghouse four-loop PWR (proprietary) and a current-controlled DCPM (generic). The effectiveness of AI/ML tools is also demonstrated using statistical inferences and unsupervised k-means classification

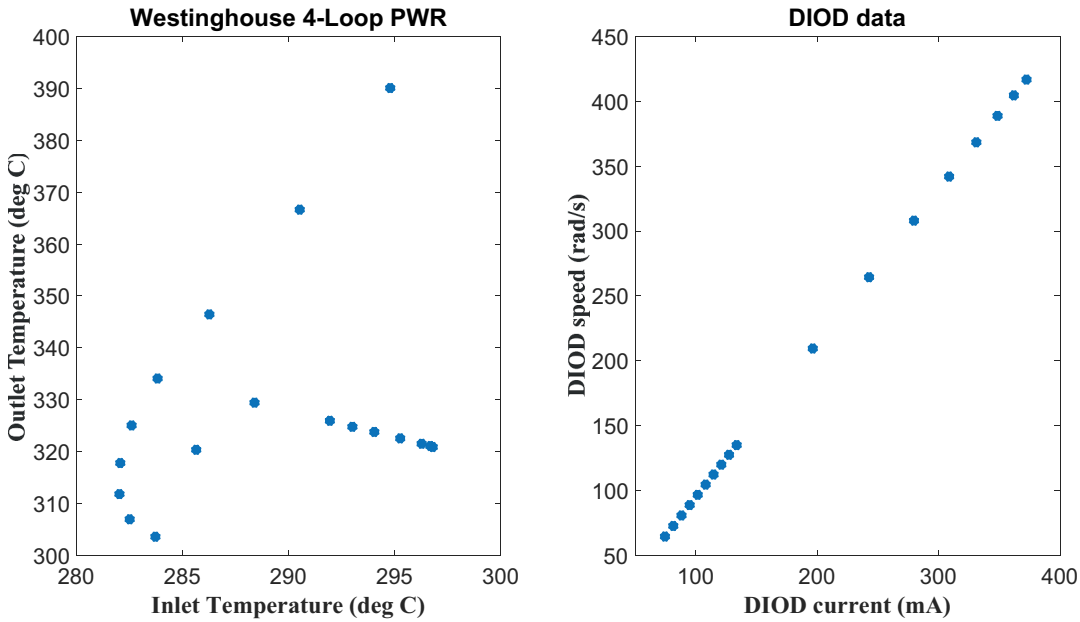

Fig. 6. Correlation among extracted responses.

representing various levels of domain knowledge on the DIOD version of the PWR data. Additionally, the DIOD methodology is validated via feature extraction methods, such as SVD, and response correlations that provide insight into the fundamental and inference metadata of the system. The correlations among the inference metadata of the PWR are preserved while those among the fundamental metadata are masked using the concealment operators developed from the DCPM system.

The presented methodology motivates a few fundamental questions on the DIOD data masking paradigm. For example, is it necessary to mount the inference metadata onto the fundamental metadata of a generic system if the AI/ML application only requires the inference metadata? How can one ensure that the overwriting of the fundamental metadata is sufficient to protect sensitive data? Can knowledgeable adversaries make educated guesses on the fundamental metadata based on the observed inference metadata to solve the ill-posed inverse problem? The concept of mutual information is key to addressing these questions as it is a measure capable of identifying the necessary inference information for AI/ML applications, and all other data from the sensitive dataset may be discarded or overwritten.

The mounting of the inference metadata onto the fundamental metadata of different systems has multiple purposes. For instance, the same inference metadata may be mounted onto different fundamental datasets to create a benchmark for AI/ML applications since in theory, the same inference should be achieved on all the DIOD datasets as long as the mutual information is preserved in the DIOD transformation. Second, from the perspective of masking, the mounting is necessary to obscure the proprietary nature of the dataset and to make the data more convincing and not appear suspicious to unscrupulous third-party services. While the sample implementation in this paper directly mounts the inference metadata, the latter may be further masked and modified via invertible transformations to fit the constraints of the target generic system (e.g., DCPM, spring-mass system) to achieve another layer of masking. This is facilitated by the invariance of mutual information to invertible transformations. For example, sensitive information such as reactivity, delayed neutron fraction, radius of a fuel pin, etc., may be obfuscated to appear like capacitance, resistances, and voltage of an electrical circuit.

In fact, this approach may be extended further by introducing extraneous variables or correlations in the masked DIOD data to render the data more representative of the generic system. These extraneous variables are expected to have no impact on the mutual information of the data and thus preserve the inferential properties necessary for AI/ML applications. Note that in this case, solving the inverse problem is virtually infeasible since the adversary needs to also guess the particular transformation/obfuscation function used on the inference metadata in addition to the fundamental metadata of the proprietary system to find the original sensitive information. In mathematical terms, the problem may be stated as guessing the concealment operator, the obfuscation function, their parameters, and the input data given the output

alone. Future work shall explore these additional layers of masking via the concept of mutual information to ensure that the inference metadata themselves cannot be used to glean information about the proprietary system, further highlighting the flexibility of the DIOD data masking paradigm.

The proposed paradigm also has applications in safeguards monitoring such as the nondestructive assay of nuclear material and may be potentially used in obfuscating and guarding physics model–based defenses against adept adversaries. Additional implications for the nuclear field include widespread collaboration among peers who may be otherwise reluctant to share sensitive nuclear data despite wanting to adapt AI/ML tools for their purposes, e.g., classification and regression, where an equivalent DIOD version of these problems may be found using the same fundamental and inference metadata principles.

Acknowledgements

Idaho National Laboratory is a multiprogram laboratory operated by Battelle Energy Alliance, LLC, for the U.S. Department of Energy under contract DE-AC07-05ID14517.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

ORCID

Ahmad Al Rashdan http://orcid.org/0000-0002-9682-3137

References

1. E. R. GRIFFOR et al., "Framework for Cyber-Physical Systems: Volume 1, Overview," NIST Special Publication 1500-201, National Institute of Standards and Technology (June 2017); https://doi.org/10.6028/NIST.SP.1500-201.

2. M. DOOSTAN and B. H. CHOWDHURY, "Power Distribution System Equipment Failure Identification Using Machine Learning Algorithms," *Proc. 2017 IEEE Power & Energy Society General Mtg.*, Chicago, Illinois, July 16–20, 2017, IEEE (July 2017); https://doi.org/10.1109/PESGM.2017.8274109.

3. J. DE GROOT, "Biggest Manufacturing Data Breaches of the 21st Century," *Digital Guardian* (Aug. 7, 2020); https://digitalguardian.com/blog/biggest-manufacturing-data-breaches-of-the-21-century (current as of Jan. 6, 2022).

4. R. J. SANTOS, J. BERNARDINO, and M. VIEIRA, "A Data Masking Technique for Data Warehouses," *Proc. 15th Symp. International Database Engineering & Applications*, New York, September 2011, p. 61 (Sep. 2011); https://doi.org/10.1145/2076623.2076632.

5. R. RADHAKRISHNAN, M. KHARRAZI, and N. MEMON, "Data Masking: A New Approach for Steganography?" *J. VLSI Sign. Process. Syst. Sign. Image Video Technol.*, **41**, *3*, 293 (Nov. 2005); https://doi.org/10.1007/s11265-005-4153-1.

6. K. MURALIDHAR and R. SARATHY, "Data Shuffling: A New Masking Approach for Numerical Data," *Manag. Sci.*, **52**, *5*, 658 (2006); https://doi.org/10.1287/mnsc.1050.0503.

7. C. BEESLEY, "What's This Data Masking All About Anyway?" *DLT* (Mar. 19, 2015); https://www.dlt.com/blog/2015/03/19/whats-data-masking (current as of Jan. 6, 2022).

8. C. DWORK and A. ROTH, "The Algorithmic Foundations of Differential Privacy," *Found. Trends® Theor. Comput. Sci.*, **9**, *3–4*, 211 (2013); https://doi.org/10.1561/0400000042.

9. B. LIVER and K. TICE, "Privacy Application Infrastructure: Confidential Data Masking," *Proc. 2009 IEEE Conf. Commerce and Enterprise Computing*, Vienna, Austria, July 20–23, 2009, p. 324, IEEE (July 2009); https://doi.org/10.1109/CEC.2009.43.

10. F. AMATO et al., "Artificial Neural Networks in Medical Diagnosis," *J. Appl. Biomed.*, **11**, *2*, 47 (Jan. 2013); https://doi.org/10.2478/v10136-012-0031-x.

11. N. SHLOMO, "Releasing Microdata: Disclosure Risk Estimation, Data Masking and Assessing Utility," *J. Priv. Confidentiality*, **2**, *1* (Sep. 2010); https://doi.org/10.29012/jpc.v2i1.584.

12. C. C. AGGARWAL and P. S. YU, "An Introduction to Privacy-Preserving Data Mining," *Privacy-Preserving Data Mining: Models and Algorithms*, pp. 1–9, Springer US, Boston, Massachusetts (2008); https://doi.org/10.1007/978-0-387-70992-5_1.

13. J. ASENJO, "Data Masking, Encryption, and Their Effect on Classification Performance: Trade-Offs Between Data Security and Utility," Doctoral Dissertation, Nova Southeastern University (Jan. 2017); https://nsuworks.nova.edu/gscis_etd/1010 (current as of Nov. 5, 2021).

14. C. CLIFTON et al., "Tools for Privacy Preserving Distributed Data Mining," *ACM SIGKDD Explor. Newsl.*, **4**, *2*, 28 (Dec. 2002); https://doi.org/10.1145/772862.772867.

15. L. SWEENEY, "k-Anonymity: A Model For Protecting Privacy," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, **10**, *5*, 557 (Oct. 2002); https://doi.org/10.1142/S0218488502001648.

16. A. MACHANAVAJJHALA et al., "L-Diversity: Privacy Beyond k-Anonymity," *Proc. 22nd Int. Conf. Data Engineering (ICDE'06)*, Atlanta, Georgia, April 3–7, 2006, p. 24, IEEE (Apr. 2006); https://doi.org/10.1109/ICDE.2006.1.

17. N. LI, T. LI, and S. VENKATASUBRAMANIAN, "t-Closeness: Privacy beyond k-Anonymity and l-Diversity," *Proc. 2007 IEEE 23rd Int. Conf. Data Engineering*, Istanbul, Turkey, April 15–20, 2007, p. 106, IEEE (Apr. 2007); https://doi.org/10.1109/ICDE.2007.367856.

18. X. XIAO and Y. TAO, "M-Invariance: Towards Privacy Preserving Re-Publication of Dynamic Datasets," *Proc. 2007 ACM SIGMOD Int. Conf. Management of Data (SIGMOD '07)*, Beijing, China, June 2007, p. 689, Association for Computing Machinery Special Interest Group on Management of Data (2007); https://doi.org/10.1145/1247480.1247556.

19. X. SHANG et al., "(k,P)-Anonymity: Towards Pattern-Preserving Anonymity of Time-Series Data," *Proc. 19th ACM Int. Conf. Information and Knowledge Management*, New York, October 2010, p. 1333, Association for Computing Machinery (Oct. 2010); https://doi.org/10.1145/1871437.1871614.

20. Y. BANG, H. S. ABDEL-KHALIK, and J. M. HITE, "Hybrid Reduced Order Modeling Applied to Nonlinear Models," *Int. J. Numer. Meth. Eng.*, **91**, *9*, 929 (2012); https://doi.org/10.1002/nme.4298.

21. M. G. ABDO, C. WANG, and H. S. ABDEL-KHALIK, "Probabilistic Error Bounds for Reduced Order Modeling," *Proc. 7th Int. Conf. Modelling and Simulation in Nuclear Science and Engineering*, Ottawa, Ontario, Canada, October 18–21, 2015, Vol. 49, No. 2, p. 483 (2015); https://inis.iaea.org/search/search.aspx?orig_q=RN:49003487.

22. S. POMROY, "Static Versus Dynamic Data Masking," | Imperva (July 10, 2017); https://www.imperva.com/blog/static-versus-dynamic-data-masking/ (current as of Jan. 6, 2022).

23. C. DWORK, "Differential Privacy: A Survey of Results," *Theory and Applications of Models of Computation*, Vol. 4978, pp. 1–19, Springer Berlin Heidelberg, Heidelberg (2008); https://doi.org/10.1007/978-3-540-79228-4_1.

24. I. DINUR and K. NISSIM, "Revealing Information While Preserving Privacy," *Proc. 22nd ACM SIGMOD-SIGACT-SIGART Symp. Principles of Database Systems (PODS '03)*, San Diego, California, June 2003, p. 202, Association for Computing Machinery Special Interest Group on Management of Data (2003); https://doi.org/10.1145/773153.773173.

25. N. A. QARABASH, S. S. SABRY, and H. A. QARABASH, "Smart Grid in the Context of Industry 4.0: An Overview of Communications Technologies and Challenges," *Indones. J. Electr. Eng. Comp. Sci.*, **18**, *2*, 656 (May 2020); https://doi.org/10.11591/ijeecs.v18.i2.pp656-665.

26. S. DÖBELT et al., "Consumers' Privacy Concerns and Implications for a Privacy Preserving Smart Grid Architecture—Results of an Austrian Study," *Energy Res. Soc. Sci.*, **9**, 137 (Sep. 2015); https://doi.org/10.1016/j.erss.2015.08.022.

27. C. ROTTONDI, G. VERTICALE, and A. CAPONE, "Privacy-Preserving Smart Metering with Multiple Data Consumers," *Comput. Netw.*, **57**, *7*, 1699 (May 2013); https://doi.org/10.1016/j.comnet.2013.02.018.

28. S. ARMOOGUM and V. BASSOO, "Privacy of Energy Consumption Data of a Household in a Smart Grid," *Smart Power Distribution Systems*, pp. 163–177, Academic Press (2019); https://doi.org/10.1016/B978-0-12-812154-2.00008-0.

29. P. BARBOSA, A. BRITO, and H. ALMEIDA, "A Technique to Provide Differential Privacy for Appliance Usage in Smart Metering," *Inf. Sci*, **370–371**, 355 (Nov. 2016); https://doi.org/10.1016/j.ins.2016.08.011.

30. "Guidelines for Smart Grid Cybersecurity," The Smart Grid Interoperability Panel–Smart Grid Cybersecurity Committee, NIST IR 7628r1, National Institute of Standards and Technology (Sep. 2014); https://doi.org/10.6028/NIST.IR.7628r1.

31. T. TUNG et al., "Maximizing Collaboration Through Secure Data Sharing," *Accenture* (Oct. 2019); https://www.accenture.com/us-en/insights/digital/maximize-collaboration-secure-data-sharing (current as of Jan. 6, 2022).

32. O. ABBOSH and K. BISSELL, "Securing the Digital Economy," *Accenture* (2019); https://www.accenture.com/us-en/insights/cybersecurity/reinventing-the-internet-digital-economy (current as of Jan. 6, 2022).

33. F. KERSCHBAUM, "Privacy-Preserving Computation," *Proc. Annual Privacy Forum*, Limassol, Cyprus, October 10–11, 2012, p. 41 (2012).

34. M. M. POTEY, C. A. DHOTE, and D. H. SHARMA, "Homomorphic Encryption for Security of Cloud Data," *Procedia Comput. Sci.*, **79**, 175 (2016); https://doi.org/10.1016/j.procs.2016.03.023.

35. A. OPPERMANN et al., "Secure Cloud Computing: Communication Protocol for Multithreaded Fully Homomorphic Encryption for Remote Data Processing," *Proc. 2017 IEEE Int. Symp. Parallel and Distributed Processing with Applications and 2017 IEEE Int. Conf. Ubiquitous Computing and Communications (ISPA/IUCC)*, Guangzhou, China, December 12–15, 2017, p. 503, IEEE (Dec. 2017); https://doi.org/10.1109/ISPA/IUCC.2017.00084.

36. K. HE et al., "Deep Residual Learning for Image Recognition," *Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, Nevada, June 27–30, 2016, p. 770, IEEE (June 2016); https://doi.org/10.1109/CVPR.2016.90.

37. N. HALKO et al., "An Algorithm for the Principal Component Analysis of Large Data Sets," *SIAM J. Sci. Comp.*, **33**, *5*, 2580 (Jan. 2011); https://doi.org/10.1137/100804139.

38. Y. BANG and H. ABDEL-KHALIK, "Reduced Order Modeling for Multi-Physics Problems," *Trans. Am. Nucl. Soc.*, **107**, 586 (Jan. 2012).

39. K. WILLCOX and J. PERAIRE, "Balanced Model Reduction via the Proper Orthogonal Decomposition," *AIAA J.*, **40**, *11*, 2323 (Nov. 2002); https://doi.org/10.2514/2.1570.

40. D. J. LUCIA, P. S. BERAN, and W. A. SILVA, "Reduced-Order Modeling: New Approaches for Computational Physics," *Prog. Aerospace Sci.*, **40**, *1*, 51 (Feb. 2004); https://doi.org/10.1016/j.paerosci.2003.12.001.

41. M. I. BELGHAZI et al., "Mutual Information Neural Estimation," *Proc. 35th Int. Conf. Machine Learning*, Vol. 80, p. 531 (July 2018); https://proceedings.mlr.press/v80/belghazi18a.html (current as of Jan. 6, 2022).

42. N. TISHBY and N. ZASLAVSKY, "Deep Learning and the Information Bottleneck Principle," *Proc. 2015 IEEE Information Theory Workshop (ITW)*, Jerusalem, Israel, April 26–May 1, 2015, IEEE (2015); https://doi.org/10.1109/ITW.2015.7133169.

43. N. CARRARA and J. A. ERNST, "On the Upper Limit of Separability," *ArXiv High Energy Phys.* (Aug. 2017); http://arxiv.org/abs/1708.09449 (current as of Jan. 6, 2022).

44. M. DEMPSEY, "Dymola for Multi-Engineering Modelling and Simulation," *Proc. 2006 IEEE Vehicle Power and Propulsion Conf.*, Windsor, United Kingdom, September 6–8, 2006, IEEE (Sep. 2006); https://doi.org/10.1109/VPPC.2006.364294.

45. H. ELMQVIST, "DYMOLA—A Structured Model Language for Large Continuous Systems," TFRT-7175, Lund Institute of Technology, Department of Automatic Control (1979).

46. M. S. GREENWOOD, "TRANSFORM—TRANsient Simulation Framework of Reconfigurable Models," Oak Ridge National Laboratory (Sep. 26, 2017); https://doi.org/10.11578/dc.20171025.2022.

47. N. CARRARA and J. ERNST, "On the Estimation of Mutual Information," *Proceedings*, **33**, *1*, 31 (2020); https://doi.org/10.3390/proceedings2019033031.