



International  
System Safety  
Society


[www.systemsafety.com](http://www.systemsafety.com)

Journal of System Safety

Established 1965 Vol. 58 No. 2 (2023)



# Proposing the Use of Hazard Analysis for Machine Learning Data Sets

H. Glenn Carter<sup>b</sup>, Alexander Chan<sup>b</sup>, Chris Vinegar<sup>b</sup>, Jason Rupert<sup>ac</sup> 

<sup>a</sup> Corresponding author email: <mailto:jason.rupert@mtsi-va.com>

<sup>b</sup> U.S. Army Combat Capabilities Development Command Aviation & Missile Center (DEVCOM AvMC); Redstone Arsenal, AL USA

<sup>c</sup> Modern Technology Solutions, Inc.; Huntsville, AL USA

## Keywords

machine learning, data assurance,  
data governance

## Peer-Reviewed

Gold Open Access

Zero APC Fees

[CC-BY-ND 4.0](https://creativecommons.org/licenses/by-nd/4.0/) License

Online: 22-Jun-2023

## Cite As:

Carter, H.G. et al. Proposing the  
Use of Hazard Analysis for  
Machine Learning Data Sets.  
Journal of System Safety.  
2023;58(2):30-39.  
<https://doi.org/10.56094/jss.v58i2.253>

## ABSTRACT

There is no debating the importance of data for artificial intelligence. The behavior of data-driven machine learning models is determined by the data set, or as the old adage states: “garbage in, garbage out (GIGO).” While the machine learning community is still debating which techniques are necessary and sufficient to assess the adequacy of data sets, they agree some techniques are necessary. In general, most of the techniques being considered focus on evaluating the volumes of attributes. Those attributes are evaluated with respect to anticipated counts of attributes without considering the safety concerns associated with those attributes. This paper explores those techniques to identify instances of too little data and incorrect attributes. Those techniques are important; however, for safety critical applications, the assurance analyst also needs to understand the safety impact of not having specific attributes present in the machine learning data sets. To provide that information, this paper proposes a new technique the authors call data hazard analysis. The data hazard analysis provides an approach to qualitatively analyze the training data set to reduce the risk associated with the GIGO.

## INTRODUCTION

This paper focuses on a critical building block on the path to certifying machine learning software items - establishing assurance practices for the data set used to train, validate, and test the machine learning models. Key to addressing data assurance concerns associated with certifying machine learning is conducting the hazard analysis of data sets and assuring the adequacy of the data set. Thus, this paper

works through what makes up data assurance for machine learning and devotes additional time on establishing hazard assessment artifacts for the data set. This paper also presents some techniques the industry is proposing for conducting data set adequacy, completeness, and representativeness, as well as an example of data hazard analysis.

**OUTLINE**

An introduction to highlights of traditional software assurance is provided, which includes a comparison of what type of additional assurance is needed for machine learning, where data assurance plays a key role. After that introduction, what is necessary to successfully accomplish data assurance is covered, where data hazard assessment plays a foundational role. Given that foundational role, additional time is spent in this paper proposing what would be necessary for data hazard assessment. This topic is presented to the safety community to generate discussion and engagement. There are certainly additions that should be made to the approach, and we hope the introduction of this concept generates some of that feedback and recommendations. Also, with this introduction we hope to begin to prepare the safety community for the arrival of data assurance techniques, and their role in the certification of machine learning based software items.

**BACKGROUND**

As indicated by SAE International Aerospace Information Report (AIR) 6988 (Artificial Intelligence in Aeronautical Systems: Statement of Concerns, AIR6988™, 2021) and Aerospace Vehicle Systems Institute (AVSI) AFE-87 (AFE 87 Project Members, 2020), the traditional aviation framework certification guidance is not adequate for the uncertainty added by the probabilistic development techniques used by machine learning:

“Industry standard development assurance processes such as ED-12C/DO-178C, ED-109A/DO-278A, ED-80/DO-254, do not have guidance for AI techniques such as Machine Learning algorithms. For some AI techniques, it may not be possible to meet all ED-12C/DO-178C, ED-109A/DO-278A and ED-80/DO-254 objectives such as those associated with the low-level requirements, implementation,

integration, and verification activities. For artificial neural networks, there may be no meaningful representation of the internal structure of Machine Learning algorithm.” (Artificial Intelligence in Aeronautical Systems: Statement of Concerns, AIR6988™, 2021)

Moreover, AVSI AFE-87 indicates the “fundamentally different nature of data-based systems”, i.e., machine learning. AFE-87 goes on to indicate, “Traditional physical models are explicitly constrained, while data driven models are implicitly constrained by the observed phenomenon in the training data.”

Our approach for the development of airworthiness certification guidance for machine learning considers the recommendations laid out in AIR6988 for establishing a framework for AI/ML, and also that of the AVSI AFE-87, SAE International Aeronautical Standard (AS) AS-6983 (SAE G-34, 2022), EASA Level 1 (Soudain, 2021), and SCSC-153B (The SCSC Safety of Autonomous Systems Working Group (SASWG), 2022).

**TRADITIONAL SOFTWARE ITEM ASSURANCE**

In general, the traditional software item assurance approach can be summarized as shown in Figure 1. Of course, Figure 1 is a bit of an oversimplification for the purposes of this paper. Other critical ML-based system assurance processes are not shown because they are similar to assurance processes for traditional systems. These include planning process, configuration management process, quality assurance process, and certification liaison processes. Processes not shown in Figure 1 but included in requirement assurance are high-level requirements (HLRs) processes, low-level requirements (LLRs) processes and the bi-traceability between HLRs and LLRs. In addition, requirement assurance includes the bi-directional traceability from HLRs to system/sub-

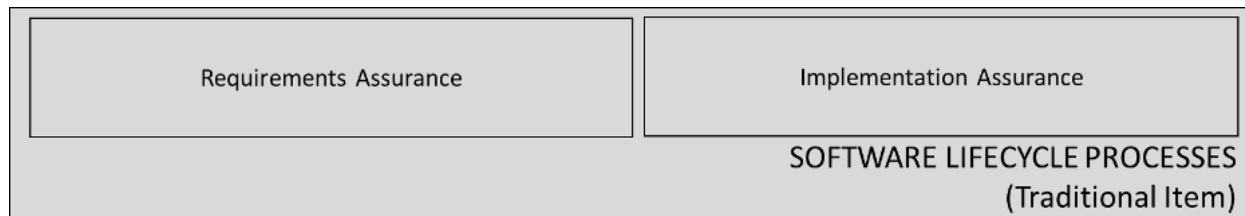


Figure 1: Traditional Software Lifecycle Assurance

Note for Figure 1: Grey fill is used to illustrate existing traditional processes, i.e., processes covered under existing processes and standards. For example, software lifecycle is covered under RTCA DO-178C (SC-205, 2011).

system requirements. Implementation assurance includes design, coding, verification, and implementation, and the appropriate bi-directional traceability between those processes. It is through the execution of the objectives and activities associated with those processes that assurances are provided for the software item to ensure it will “perform [its] intended functions under all foreseeable operating conditions.” (14 CFR 25.1309 Equipment, systems, and installations.) Similar quotes are applicable for CFR Parts 23, 27, and 29.

**MACHINE LEARNING SOFTWARE ITEM ASSURANCE**

As indicated in AIR 6988 and AFE-87, additions are necessary to the traditional software lifecycle process assurance approach to account for the unique aspects of data-driven machine learning software development techniques. Grey fill was used in Figure 1 to indicate traditional software development processes, while in Figure 2 white fill is used to indicate necessary modifications or additions. For machine learning based software item development, data assurance and learning assurance are necessary assurance additions. In addition, enhancements are necessary to the processes that enable requirements and implementation assurance. Enhancements will also be necessary to the planning, configuration management, quality assurance, and certification liaison processes, but those are beyond the scope of this paper.

For machine learning-based software item assurance, the traditional software item assurance of requirement and implementation assurance processes will be augmented by the addition of data and learning assurance. We propose that data assurance consists of ensuring, for example, the training, verification, and test data set correctness, completeness, and representativeness of the operational design domain. Correctness, completeness, and representativeness of the operational design domain are three attributes of

the data set that will determine the accuracy and performance of a machine learning model in the operational design domain.

Learning assurance consists of activities to confirm the intended machine learning model generalization performance is reached, e.g., not underfitting or overfitting, not being susceptible to bias or drift, and appropriate behavior for out of distribution samples. Underfitting occurs when unacceptable error occurs during model validation. This is often a symptom model susceptibility to bias and is an indication of too small of a training data set. Overfitting occurs when validation error is low, but test error is high. Overfitting is an indication that the model has memorized the training and validation set, i.e., is fitting the variance noise in the data, but is not generalizing. Addressing the expectations associated with the machine learning model requirements and learning assurance is out of scope of this paper but will be addressed in follow-on papers.

Such follow-on work will go through the new machine learning development lifecycle (MLDL), which augments the machine learning implementation lifecycle (MLIL). The machine learning development lifecycle includes the processes to ensure the new development assurance expectations for machine learning are met.

Data-driven machine learning software development does not use traditional software development methodologies. That is, data-driven machine learning development does not develop implementation source code and parameter values directly from low-level requirements (LLRs). Instead, machine learning trains a machine learning model, which is a set of hyperparameters, neurons, and layers from a training data set. The data set and machine learning model are based on a set of data and model requirements. The machine learning data requirements are used to drive the data collection process, where the data set must be correct, complete,

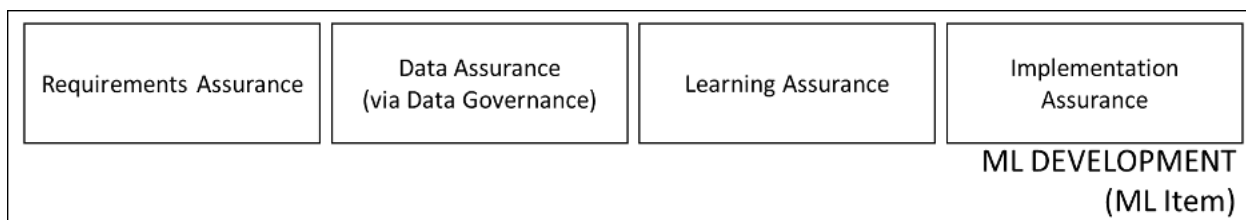


Figure 2: Machine Learning Development Lifecycle Assurance

and representative of the deployment operational design domain. Once collected, processed, and curated, the data set is divided into three independent subsets that are used to train, validate, and test machine learning algorithms. We call the approach to ensure the data set is adequate (i.e., correct, complete, and representative, and collected, processed, and curated correctly, and split appropriately) data assurance, where data governance is the process to ensure data assurance. In this paper, learning assurance is the term used for adequately designing, developing, training, validating, and testing the machine model.

Both new types of assurances (data and learning) are necessary to ensure machine learning is developed in a mature way to consider its use in flight and safety critical applications. Moreover, the addition of these two assurance approaches compensates for the loss of some traditional software development objectives and activities, e.g., loss of low-level requirements, and meaningfulness of traceability and design detail. The loss of the assurance provided by those traditional software development objectives and activities must be accounted for when contemplating the possible use of machine learning in flight and safety critical applications.

## DATA GOVERNANCE

Data is one of the bigger technical debts (D. Sculley, 2015) of the machine learning processing: “Data Dependencies Cost More than Code Dependencies” and “Changing Anything Changes Everything (CACE)”. These technical debts, i.e., resources and risks, are spread across all the processes associated with data governance, i.e., data planning to allocation. Data collection alone spans various types of data acquisition which could involve discovering existing collected data sets or synthetic generation and augmentation of data sets. After the collection process, the preparation and processing begin where labeling and other improvements are necessary. The labeling can be an intensive and technical process involving manual labeling or a semi-supervised labeling technique. Where necessary, improvements to the data may be necessary, which can also be intensive. Through each of these processes, care must be taken to maintain the data set's validity and authenticity. The data set has a large impact on the performance of the model properly reflecting the required generalization behavior in the operational

design domain. Benign and even imperceptible modifications to data can cause unexpected, unanticipated, and undesired behavior of the models when exposed to deployed operational design domain native data sources.

As shown in Figure 3, the data governance process manages the data's sourcing, collection, processing, hazard assessment, and allocation. Data Governance provides the following processes, objectives, and activities to enable data assurance: data integrity, data hazard assessment, data planning, data completeness, data representativeness, data accuracy, correctness, data traceability, data reproducibility (i.e., collection, augmenting, transformation, labelling), dataset independence, data verification, data configuration management (e.g., corruption guards). The Configuration Management Process addresses the data configuration management, and the data verification is addressed by the machine learning verification process.

Notes for Figure 3:

- Note 1: Data set includes the features, attributes, and classes as well as the samples, signals, sources, and collection of those features, attributes, and classes.
- Note 2: Data configuration management, executed in the ML Configuration Management process, will ensure data is only used appropriately, e.g., avoiding data leakage, via methods like blockchain, and data integrity.
- Note 3: Data verification, executed in the ML Verification Process, will ensure the data is adequate, appropriate, representative, and complete as described in the Data Requirements.

Updates to the data set output from ML data governance processes may be driven by the ML Model Development Process or ML Verification Process. Should those updates occur, the ML data requirements, system safety requirements, and operational design domain requirements should be re-examined to determine if updates are necessary to those as well. Because requirements-based testing should be used for flight and safety-critical

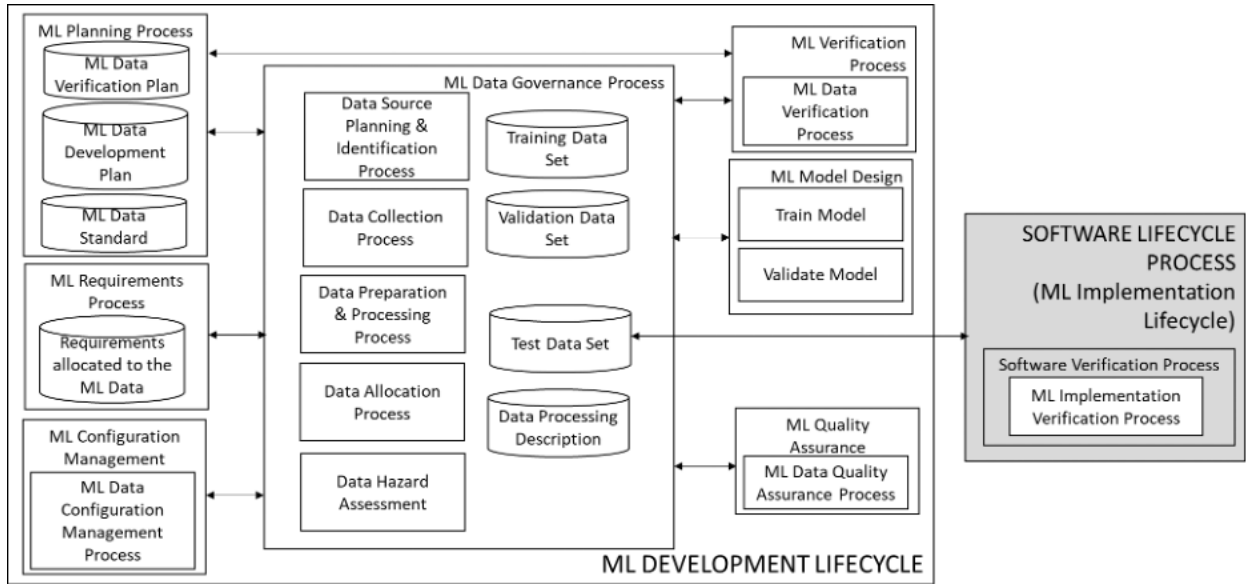


Figure 3: ML Data Governance Process

applications, any updates to those requirements should also be reflected in updates to verification cases. Because data set changes can be costly, all efforts should be made to correctly produce fully representative data set requirements and complete data sets as early in the process as possible.

Except for data hazard assessment, detailed definitions of each of these attributes of data assurance is beyond the scope of this of this paper but will be covered in general. Follow-on work is necessary to fully define the expectation for each of these attributes. The follow-on work will look to luminary guidance like that provided by SCSC-127G Data Safety Guidance (Version 3.4) (Data Safety Initiative Working Group, 2022).

The following sections specifically focus on ensuring data safety and data completeness and representativeness, accuracy, and correctness through the use of data hazard analysis and data verification techniques.

**DATA HAZARD ASSESSMENT PROCESS**

Data hazard assessment is a hazard assessment process that leverages techniques applied to traditional system and software hazard analysis techniques. In addition, the approach introduces novel techniques to assess the hazard impacts associated

with the use of data sets for machine learning model training, validation, and testing.

Figure 4 shows the bi-directional traceability of the data hazard assessment to the traditional system safety hazard assessment process (shown in grey fill), e.g., those associated with and identified in SAE International Aerospace Recommended Practice (ARP) 4754 (S-18, 2010) and ARP 4761 (S-18, 1996). Analysis is on-going to determine if and how traditional hazard assessment processes may need to be augmented for machine learning based systems, e.g., accounting for autonomy level (classification) and methodologies may impact the functional hazard assessment (Copeland, 2019) or development assurance levels.

Notes for Figure 4:

- Note 1: Data includes the features, attributes, and classes present in the data as well as the samples, signals, sources, and collection of those features, attributes, and classes.
- Note 2: Data is most applicable to data-driven ML techniques; however, similar techniques more applicable to reinforcement learning will be specifically covered in the future, e.g., scenario planning, assessment, and verification.

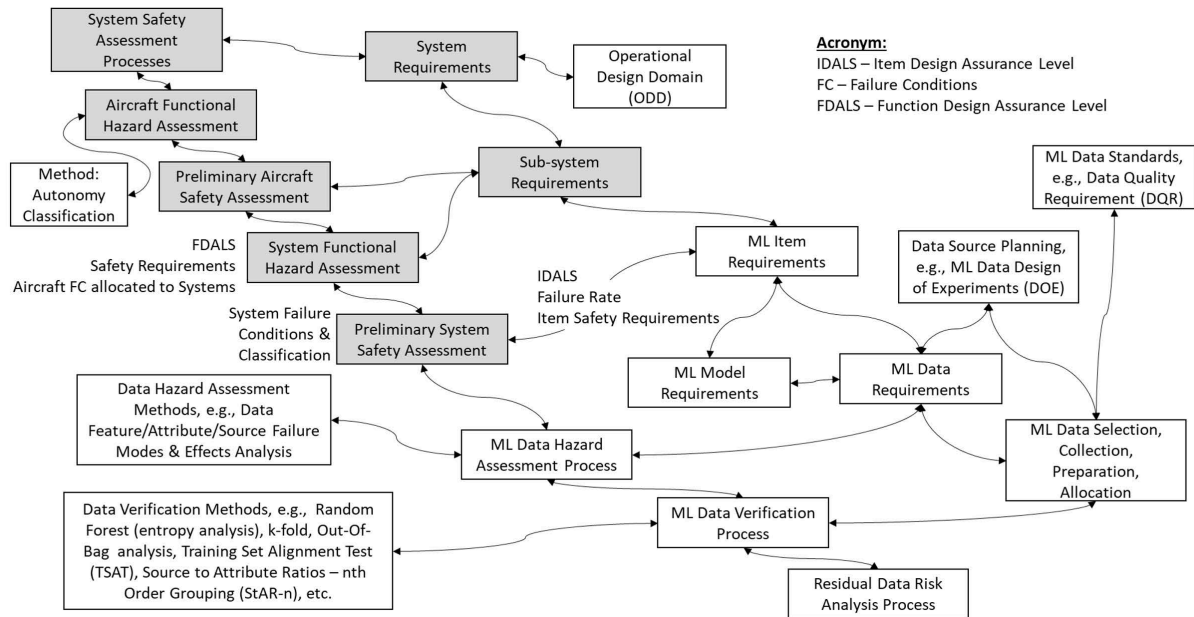


Figure 4: Hazard Assessment Processes

In the discussion that follows, the failure modes and effects analysis (FMEA) is offered up as one type of data hazard analysis technique, but others might be more appropriate, e.g., SCSC-127G proposes the establishment and use of the Data Safety Assurance Level (DSAL).

As indicated in ARP 4761: “An FMEA is a systematic, bottom-up method of identifying the failure modes of a system, function, or item and determining the effects on the next higher level. It may be performed at any level within the system (e.g., piece-part, function, black box). Software can also be analyzed qualitatively using a functional FMEA approach. Typically, an FMEA is used to identify failure effects resulting from single failures.” An additional resource for the application and development of the FMEA is SAE International ARP 5580 (Recommended Failure Modes and Effects Analysis (FMEA) Practices for Non-Automobile Applications, ARP 5580, 2020).

As shown in Figure 5, for the discussion in this paper, the machine learning data FMEA is a data hazard analysis technique that would be used to assess the failure effects of the data, e.g., attributes, features, signals and sources, and the hazard effect on the ML model.

As shown in Figure 5, inputs to the ML Data FMEA include the following:

- As shown in the grey boxes, the systems-level hazard assessment artifacts, including the systems functional hazard assessment which should assess the hazard impacts of the functions within the ML-based system.
- From the ML requirements process, the ML data requirements.
- From the ML planning process, the ML data source planning and identification process artifacts, which may include data design of experiments type artifacts indicating what data is necessary and why.

Other artifacts that influence the machine learning data FMEA are the following:

- ML Data Development Plan
- ML Data Standard
- ML Data Requirements Standard

Data FMEA is a safety assessment of data features, attributes and sources, samples, and signals anticipated to drive the AI/ML model generalization in its operational design domain. Each of those is analyzed similarly to the approach used by a Software Interface FMEA described in ARP5580. Ultimately,

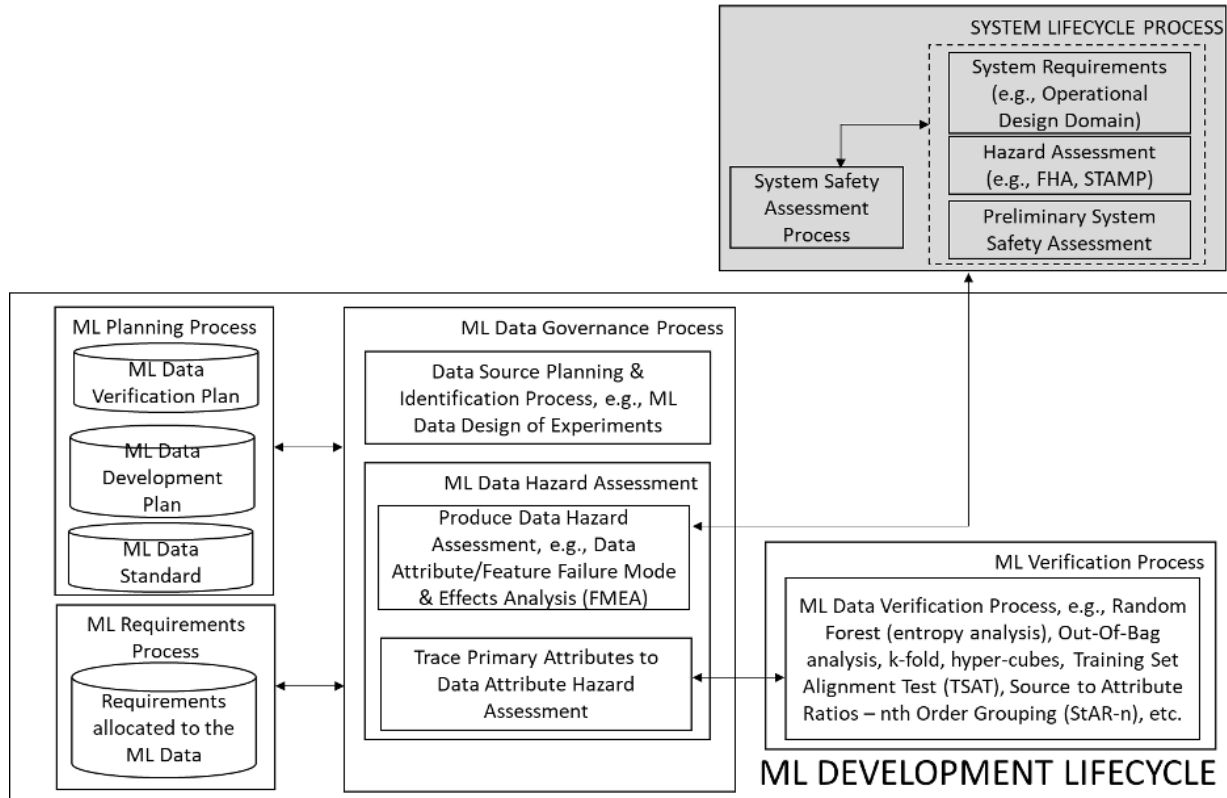


Figure 5: Data Hazard Assessment

the Data FMEAs may be summarized into a Data Failure Modes Effect Summary (FMES) to support the failure modes analysis considerations. Given the sensitivity of the machine learning model to input data sources, developing the data FMEA or its equivalent, is encouraged.

As an example, Table 1, from the Computer Vision- Hazard and Operability Study (CV-HAZOP) (Oliver Zendel, 2017), shows a type of data FMEA establishing traceability from failures (errors) in the data attributes, features, samples, sources, and signals to resultant impact of the ML model. The table identifies the source (location) and feature/attribute (guide word) and the consequences and risk.

Additional steps would be to indicate the expected machine learning effect, and eventually, if the data is available, the probability of each. Currently the CV-HAZOP has 1,469 entries. Such a systematic approach by computer vision experts allows for the machine learning based software item to be appropriately representative and complete data sets to

be collected and used for machine learning training, validation, and testing. Data sets used can be assessed against the data hazard assessment to determine if all negative consequence and high risks effects are covered. Where gaps in the data set exist, appropriate mitigations can be determined, e.g., creation of synthetic data, creation of a derived subsystem requirement to mitigate, or other. Any gaps that remain in the data set should be indicated in the ML data processing and MLDL verification output data item and brought to the attention of the certification authority.

With respect to chronology, a data hazard assessment occurs before data source identification and collection, so as to drive the collection of safety critical features/attributes over those less so.

Similar methodology would be employed for reinforcement learning, but instead of data set/signal attributes/features, the reinforcement training scenario attribute/features would be analyzed.

Table 1: Example Data FMEA

Risk Id	Location	Guide Word	Parameter	Meaning	Consequence	Risk
0	Light Sources	No (not none)	Number	No light sources	No light available	Sensor will receive no light, but thermal noise or black current can cause wrong input
1	Light Sources	More (more of, higher)	Number	Many light sources (more light sources than expected)	Too much light	Overexposure (of whole image)
2	Light Sources	More (more of, higher)	Number	Many light sources (more light sources than expected)	Too few shadows	Algorithms using shadows can be confused
3	Light Sources	Less (less of, lower)	Number	Few light sources (fewer light sources than expected)	Too faint light (in parts of the scene)	Sensor will receive too faint light from some scene regions
4	Light Sources	Less (less of, lower)	Number	Few light sources (fewer light sources than expected)	Too many shadows	Algorithms can be confused by shadows
5	Light Sources	Less (less of, lower)	Number	Few light sources (fewer light sources than expected)	Very sharp shadows	
6	Light Sources	As well as	Number	Mirrors fake additional light sources	Light sources can appear at locations other than where they are	Algorithm confuses position of light sources

### DATA VERIFICATION ASSESSMENT PROCESS

Statistical data verification assessment of machine learning data sets is an emerging field, so various methods are mentioned where their applicability depends on the situation. For this paper, a few different techniques will be mentioned with appropriate references to guide their application:

- Random Forest, e.g., feature importance
- Clustering, e.g., feature redundancy (Tabular Modeling Deep Dive, 2022)
- k-fold cross-validation
- Training Set Alignment Test (TSAT) (Nagy, 2021)
- Source to Attribute Ratios – nth Order Grouping (StAR-n)
- hyper-cubes (focus - data completeness) (Kevin Fuchs, 2016)
- distribution discriminator framework (out-of-distribution)

The machine learning based software item developer may have different techniques they prefer. In such a situation the vendor should indicate their selection. The evaluation and justification of the statistical relevance of the data set should be conducted regardless of the approach for determining such validity. The goal of these approaches is to quantitatively show, through statistical analysis, that the data set selected, e.g., samples, signals, sources, attributes, and features, contains a complete representation of the operational design domain. While the method matters, more important is that the processes is pursued. Approaches to present a valid statistical representation of the data set may involve the following techniques or others unique to the vendor's approach:

- Exploratory Data Analysis (Brillinger, 2011)
- Boxing clever (Rob Ashmore, 2018)
- Datasheets for datasets (Timnit Gebru, 2021)



These machine learning data set verification results should present these statistical examinations of the data sets. The results should explain where the data set does not statistically fulfill the requirements associated with the operational design domain. The goal of these approaches is to ensure the proper data sets were collected, so these processes are complementary to the data hazard assessment processes.

## CONCLUSION

For machine learning the data set is critical and ultimately determines how well the machine learning model generalizes on previously unseen data when deployed in complex operational design domain. Data assurance, specifically data hazard assessment and data verification, is a necessary assurance addition to the certification of machine learning based software items. Data assurance provides the necessary confidence that the data set is adequate, complete, and representative of the operational design domain. The data hazard assessment determines the impact of features, attributes and sources, samples, and signals. Through this process, the data hazard assessment provides guidance for the collections of features, attributes and sources, samples, and signals that should be present in the data set. The data hazard assessment process output will be used to guide the data governance collection and processing processes to help ensure data set adequacy, completeness, and representativeness. The complementary data set verification process ensures those features, attributes and sources, samples, and signals were collected. Through the addition of the data assurance process, and others to be addressed more thoroughly in follow-on work, the assurance community can begin to consider the inclusion of data-driven machine learning based software items in flight and safety critical applications.

## DISCLAIMER

This paper is for information/education purposes only and does not provide the official position of U.S. Army Combat Capabilities Development Command Aviation & Missile Center (DEVCOM AvMC) with respect to establishing the Airworthiness Assurance argument for Artificial Intelligence and Machine Learning. Review of the document and associated

briefing for Public Release was successfully completed (ID 6995).

## AUTHORSHIP CONTRIBUTIONS

H. Glenn Carter: Funding acquisition, Conceptualization, Supervision, Writing - Review & Editing; Alexander Chan: Supervision, Writing - Review & Editing; Chris Vinegar: Supervision, Writing - Review & Editing; Jason Rupert: Writing - Original Draft.

## COMPETING INTERESTS

This work was funded through the U.S. Army Combat Capabilities Development Command Aviation & Missile Center (DEVCOM AvMC). The authors declare they have no potential competing interests.

## ORCID IDS

Jason Rupert  <https://orcid.org/0009-0004-5778-4747>

## REFERENCES

- [1] AFE 87 Project Members. (2020). Machine Learning, AFE-87. College Station: Aerospace Vehicle Systems Institute. Retrieved June 1, 2022, from <https://avsi.aero/projects/current-projects/cert-of-ml-systems/afe-87-machine-learning/>
- [2] Brillinger, D. R. (2011). Data Analysis, Exploratory. Retrieved June 1, 2022, from <https://www.stat.berkeley.edu/~brill/Papers/EDASage.pdf>
- [3] Copeland, R. (2019). An Analysis and Classification Process towards the Qualification of Autonomous Systems in Army Aviation. Vertical Flight Society's 75th Annual Forum & Technology Display. Philadelphia. Retrieved from <https://vtol.org/store/product/an-analysis-and-classification-process-towards-the-qualification-of-autonomous-systems-in-army-aviation-14727.cfm>
- [4] D. Sculley, G. H.-F. (2015). Hidden Technical Debt in Machine Learning Systems. Advances in Neural Information Processing Systems 28.

- [5] Data Safety Initiative Working Group. (2022). Data Safety Guidance (Version 3.4). Safety-Critical Systems Club. Retrieved June 1, 2022, from <https://scsc.uk/scsc-127G>
- [6] Kevin Fuchs, P. A. (2016). INTUITEL and the Hypercube Model - Developing Adaptive Learning. SYSTEMICS, CYBERNETICS AND INFORMATICS, 14(3), 7-11. Retrieved June 1, 2022, from <http://iiisci.org/journal/pdv/sci/pdfs/EA039OY16.pdf>
- [7] Nagy, B. (2021). Increasing Confidence in Machine Learned (ML) Functional Behavior during Artificial Intelligence (AI) Development using Training Data Set Measurements. Acquisition Research Program. Retrieved June 1, 2022, from <https://dair.nps.edu/handle/123456789/4393>
- [8] Oliver Zendel, K. H. (2017). Analyzing Computer Vision Data — The Good, the Bad and the Ugly. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/CVPR.2017.706>
- [9] Recommended Failure Modes and Effects Analysis (FMEA) Practices for Non-Automobile Applications, ARP 5580. (2020). SAE International.
- [10] Rob Ashmore, M. H. (2018). “Boxing Clever”: Practical Techniques for Gaining Insights into Training Data and Monitoring Distribution Shift. SAFECOMP 2018 Workshops, LNCS 11094, 393–405. Retrieved June 1, 2022, from [https://doi.org/10.1007/978-3-319-99229-7\\_33](https://doi.org/10.1007/978-3-319-99229-7_33)
- [11] S-18. (1996). Guidelines for Conducting the Safety Assessment Process on Civil Aircraft, Systems, and Equipment, ARP4761. SAE International.
- [12] S-18. (2010). Guidelines for Development of Civil Aircraft and Systems, ARP4754A. SAE International.
- [13] SAE G-34. (2021). Artificial Intelligence in Aeronautical Systems: Statement of Concerns, AIR6988™. SAE International. Retrieved June 1, 2022, from <https://www.sae.org/standards/content/air6988/>
- [14] SAE G-34. (2022). Process Standard for Development and Certification/Approval of Aeronautical Safety-Related Products Implementing AI, AS6983. SAE International.
- [15] Safety of Autonomous Systems Working Group. (2022). Safety Assurance Objectives for Autonomous Systems V3, SCSC-153B. Safety Critical Systems Club. Retrieved June 1, 2022, from <https://scsc.uk/SCSC-153B>
- [16] SC-205. (2011). Software Considerations in Airborne Systems, DO-178C. Washington: RTCA, Inc.
- [17] Soudain, G. (2021). First usable guidance for Level 1 machine learning applications. European Union Aviation Safety Agency. Retrieved June 1, 2022, from <https://www.easa.europa.eu/newsroom-and-events/news/easa-releases-its-concept-paper-first-usable-guidance-level-1-machine-0>
- [18] Tabular Modeling Deep Dive. (2022, April). Retrieved June 1, 2022, from [https://github.com/fastai/fastbook/blob/master/09\\_tabular.ipynb](https://github.com/fastai/fastbook/blob/master/09_tabular.ipynb)
- [19] Timnit Gebru, J. M. (2021). Datasheets for Datasets. Communications of the ACM, 64(12), 86-92. <https://doi.org/10.1145/3458723>
- [20] United States Code of Federal Regulations. (n.d.). 14 CFR 25.1309 Equipment, systems, and installations. US Government. Retrieved June 1, 2022, from <https://www.ecfr.gov/current/title-14/chapter-I/subchapter-C/part-25/subpart-F/subject-group-ECFR9f24bf451b0d2b1/section-25.1309>