Building Trust in Safeguards Voice User Interfaces: An Experimental Approach

Zoe N. Gastelum, Kristin M. Divis, Sydney Dorawa, Jamie L. Coram, Breannan C. Howell

Sandia National Laboratories | Albuquerque, NM, USA

Abstract: The research community is exploring multiple potential applications of artificial intelligence and its enabling technologies for international nuclear safeguards monitoring and verification. Based on successful implementation in other domains or early safeguards development, some applications hold significant promise. However, those benefits can only be realized if users place appropriate levels of trust – that is, reliance in the system – when delegating tasking or accepting recommendations. Prior work by members of this team has examined user trust in visual-based systems focused on model performance, but voice user interfaces such as those in digital assistants pose unique opportunities. While some safeguards challenges for voice user interfaces – noise in the operational environment, non-native English speakers, etc. – have been addressed in other domains, there remain aspects of voice user interfaces that are unique for trust in safeguards applications. In this work, we use human performance testing to evaluate the factors that impact trust in voice user interfaces within the safeguards context. Our aim is to provide actionable recommendations to the software development community to optimize system performance of voice user interfaces for safeguards. In this paper, we describe our task analysis and safeguards trust factors. We detail our experimental prioritization, focusing on two safeguards tasks: seal checking and nuclear material measurement. We describe early results and our experimental plans moving forward.

1.0 Introduction

The International Atomic Energy Agency (IAEA) Department of Safeguards faces an increasing workload due to the growing quantities of nuclear material and nuclear facilities across the globe. The IAEA's stagnant safeguards budget requires that they continuously seek efficiencies, a recent example of which is the use of artificial intelligence (AI) to assist with or perform safeguards-related tasks. The international safeguards research community and the IAEA have been evaluating the potential AI applications for safeguards analysts and inspectors in their daily tasks. Examples include support for open-source information analysis (Feldman, Arno, Carrano, Ng, & Chen, 2018) (Gastelum, Shead, & Rushdi, 2021), satellite imagery analysis (Rutkowski, Canty, & Nielsen, 2018), surveillance camera data review (Smith, Hamel, Hannasch, Thomas, & Gaiten-Cardenas, 2021) (Thomas, et al., 2021), isotope identification (Rossa & Borella, 2021) (Mishra, Branger, Elter, Grape, & Jansson, 2021), and safeguards digital assistants (Smartt, Gastelum, Rutkowski, Peter-Stein, & Shoman, 2021). Safeguards digital assistants, with which users interact via voice user interface (VUI) are the focus of this research. Voice user interface (VUI) assistants could support inspectors with task tracking, documentation of observations or results, navigation, and other in-field activities. While it is broadly recognized that international safeguards inspectors are highly trained experts, hands-free support for inspectors could make some tasks easier for inspectors, enable tasks to be completed with fewer inspectors (freeing time for higher priority in-field activities), or even enhance performance.

Safeguards VUIs have not yet been fully developed or implemented. In preparation for safeguards VUIs, user adoption and integration are being considered, and user trust is a key consideration in adoption. To be clear, undiscriminating trust in automation is not the objective. If users place too much trust in a safeguards VUI, they may become complacent with the system, and we lose the benefit of their expertise honed over years of experience and training. If users place too little trust in a safeguards VUI, we do not realize the potential for performance enhancement from human-computer interactions nor the reduction of user cognitive load that these systems can provide. This paper describes initial research on the factors that impact user trust in a voice user interface for safeguards-relevant tasks.

In this paper, we describe our human performance testing plans to measure trust in safeguards VUIs. We describe the construction of our experimental plans, including a safeguards VUI task analysis in which we detail the potential interactions between a safeguards inspector and VUI for the various in-field safeguards activities. Building on these interactions, we define the trust factors associated with the VUI tasks against variables identified from the cognition and psychology literature on trust in automation. For each intersection of a trust factor with a VUI task, we described potential manipulations of the VUI to test the trust factors. We present initial results from our foundational studies on seal examination and describe the path forward for this research.

1.1 Safeguards VUI Task Analysis

This work focuses on two distinct types of safeguards inspection tasks: seal checking and nuclear material measurement. These tasks were selected due to their prevalence and relevance to safeguards inspection activities and the opportunities they pose for safeguards VUI manipulations.

The safeguards VUI task analysis builds upon prior work in which we: 1) summarized safeguards activities described in IAEA documents, 2) combined safeguards task information with data from expert elicitation and literature review (Gastelum, et al., 2017), and 3) detailed the tasks a safeguards VUI could perform (Smartt, Gastelum, Rutkowski, Peter-Stein, & Shoman, 2021). For each of the safeguards VUI tasks, we developed vignettes of potential safeguards inspector and VUI interactions. These scenarios were intended to help our team identify potential variants in types of interactions between an inspector and a VUI to support our task analysis. For example, we developed a vignette between an inspector and safeguards VUI (who we named "VAL" for voice assistance laboratory) related to seal checking, as follows:

```
Inspector: Hey VAL, use the camera to verify the seals.
VAL: Initiating camera. Please show me the first seal you would like me
to record.
(Inspector shows the seal)
VAL: Seal 57834 recorded with normal wear and tear.  Ready for the next
seal.
```

We developed hundreds of these scenarios and summarized the VUI tasks into ten high-level categories. For brevity, they are not explained fully here, but those that have been selected for human performance testing are explained in the context of the experimental scenarios below. The VUI task categories include:

1. VUI reads information.
2. VUI provides directions.
3. VUI records inspector observations.
4. VUI confirms input or receipt of information.
5. VUI refers user to another platform.
6. VUI continues task after a pause.
7. VUI alerts inspector.
8. VUI facilitates team communications.
9. VUI exits or switches tasks.
10. VUI provides information on its current state.

Once we summarized the VUI high-level tasks, we cross-analyzed each VUI safeguards task type against a series of trust factors that we identified through our cognitive science, psychology, and human factors literature review. We specifically focused on factors related to trust in automation. Many of these trust factors are not mutually exclusive and include common aspects regarding how to provide information to the user. The factors influencing trust in safeguards VUI that we reviewed for this research include granularity, provenance, privacy, confidence, and explainability. The trust factors that were selected specifically for this research are explained within the context of their relevant experimental task below. Other factors like system performance are also incredibly important to user trust. However, we found the general best practices from the user experience community to be directly applicable, and therefore they were not selected for safeguards-specific experimental testing.

We reviewed the entire cross-section of trust factors and VUI tasks and prioritized the tasks for experimental development. The trust factors are mentioned below but will be more fully explained in situ with their associated task descriptions in Sections 2 and 3. Our highest priorities from this activity are as follows:

- **Communicates decision or analysis outcome from underlying models**, with a focus on explainability and confidence trust factors.
- **Reads information**, such as from source documents or a measurement tool. This task relates most closely to the provenance and granularity trust factors.
- **Requests clarification or additional information**. Aspects of this task that do not rely heavily on unbounded responses from the user are best suited for the scope of our study. This task will be associated with the confidence trust factor.
- **Alerts the inspector**, specifically how to alert an inspector, based on time sensitivity, if action is required, and sensitivity of the alert. This VUI task has emphasis on privacy and state of the system.
- **Confirms input or receipt of information**, specifically related to potential requirements for visual scaffolding for verification of correct capture of complex data. In general domains, this is well covered; however, our high-consequence domain may reveal different findings. This task was considered a mid-level priority, and along with the third priority topics might be considered to re-test or expand upon our highest priority tasks.

Our experimental activities focus on two pervasive inspection activities: seal examination and nuclear material measurements. Our human performance testing includes an experimentally-manipulated version of VAL that provides verbal cues to participants. For this experiment, VAL has a female voice with an American accent.

For both foundational studies, we ask participants a series of questions to understand how much they trust VAL. We measure trust in two ways. First, we measure user compliance with VUI recommendations (VAL will provide incorrect responses 20% of the time, so complete compliance would result in 80% participant accuracy). We can also interpret response times, which indicate the time a user spent with a trial after the initial verbal prompt. This response can imply, for example, how much time a participant spends searching for a tamper if VAL says there are tamper indications. The second measure of trust is collected through direct elicitation from our participants. We survey our participants with a variety of questions regarding their perceived accuracy of VAL, if they would use VAL for future tasks, and if they found VAL useful or trustworthy. These questions are based on the Trust in Automated Systems Test (TOAST) (Wonton, Porter, Lane, Bieber, & Madhavan, 2020), as well as direct queries into participants' trust in VAL and perceived reliability. We also conduct brief assessments on a participant's individual technology preferences to account for individual differences using the Affinity for Technology interaction (ATI) scale (Attig, Wessel, & Franke, 2017).

The intent of the trust questions is to understand changes in those trust measurements as we manipulate factors like explainability and confidence. Given that the Foundational Studies do not manipulate our trust factors, these measures serve as baselines for future activities rather than stand-alone findings.

2.0 Seal Examination Task

2.1 Seal Examination Task Description

Seal examination, for the purposes of this study, refers to the examination of a seal body for its identification number ("seal ID") and tamper status. Seal examination is a frequent, manual, and time-consuming task for safeguards inspectors working in the field. It is also a preliminary task being developed by our sister project *Hey Inspecta* (Smartt, Gastelum, Rutkowski, Peter-Stein, & Shoman, 2021). In our seal examination experiments, participants are presented with illustrations of security seals like those in Figure 1. These seals are not representative of actual IAEA or U.S. Government seals, nor are the markings representative of actual wear or tamper events. In the experiment, VAL identifies the seal identification number and indicates whether the seal has indication of tamper.

Figure 1. Example Tamper Manipulations. From left to right: Normal wear and tear (no tamper), tamper in top left corner (highlighted in yellow), tamper on bottom of the seal, right corner (highlighted in yellow).

2.2 Seal Examination VUI Tasks and Trust Factors

The seal examination task focuses primarily on the VUI tasks **communicates results from other models,** and **requests additional information**. We will describe how each of these tasks maps to trust factors that can be experimentally manipulated in the sections below.

In the task **VUI communicates result from other models**, the VUI provides information that represents a decision, recommendation, or result from decision automation support models. These decisions might include, but are not limited to, identification of isotopes in a nuclear material measurement, recognition of a change in a visual scene, results from visual detection related to seal examination (such as predicting a seal identification number), prioritized frames from safeguards surveillance cameras for inspector review, and communication of a potential anomaly or alerting an inspector to an action that requires follow-up. For our experimental seal checking activity, the underlying models from which VAL is communicating results include an optical character recognition model to identify the seal identity and a tamper indication model to determine if markings on a seal reflect tamper or normal wear and tear.

For the **VUI communicates results from other models** task on tamper status of a seal, we will test two primary trust factors: **explainability** and **confidence**. To simplify our experiments, we did not manipulate trust factors related to object character recognition for the seal identity.

Explainability refers to how the VUI references information that was significant in a secondary model for making a prediction or inference. Our experiments on explainability will focus on how VAL communicates machine learning explainability information regarding tamper detection to a user. This information would presumably be available for an underlying tamper-detection model that VAL is describing verbally. Our explanations will focus on the **granularity** of location information communicated regarding the tamper, in which we will compare users that had no explanations, those with low-detail locational information (e.g., the tamper is on the top of the seal), and those with high-detail locational information (e.g., the tamper is in the top left quadrant of the seal).

Confidence refers to the level of confidence provided with the prediction or inference of a secondary model. For our work, we are interested in how the VUI communicates (or does not communicate) confidence information. While measurement uncertainty is not equivalent to confidence from a machine learning model's inference, we also considered uncertainty of a system or measurement in this category. Our experiments on confidence will focus on how confidence information regarding the identification of tamper is communicated to the user. Recall that this is communication of confidence of the prediction from an underlying model rather than the VUI itself having confidence. We will compare scenarios in which participants are not provided with confidence information with scenarios in which confidence is communicated in verbal terms (e.g., high confidence, low confidence), simple numerical terms (e.g., 95% confidence, 10% confidence) and detailed numerical terms (e.g., 95.45% confidence, 10.36% confidence).

For seal examination, we are also interested in the VUI task of "**requests information**." In the **requests information** task type, the VUI receives data input from the inspector, safeguards equipment, or environment to pass to an underlying model, which requires more information to make a suitably confident decision. Example requests could include repetition of an inspector's verbal input, additional measurements from safeguards equipment, or photographs or other camera-based inputs with additional lighting or from a new angle. In the experimental task, we can manipulate **confidence** or **explainability**. We anticipate instances in the field where a digital inspection assistant would not have sufficient information to make a high-confidence prediction. Alternatively, there may be environmental conditions in the field (e.g., low light, debris on the seal) that inhibit the predictive power. With this experiment, we will compare users who receive requests for additional information with those who do receive additional information about the request.

As we experimentally manipulate explainability, confidence, and other trust factors, we will measure if and how user trust is reported. We will ask our users to report on several measures of trust, including their perceived performance of VAL, their understanding of how VAL is intended to work, and their trust in automated systems more generally.

2.3 Initial Experimental Results

Foundational Study 1 – Tamper Detection

In preparation for manipulating our selected trust factors, we have completed two foundational studies related to seal examination. We ran the foundational studies on the Pavlovia and Prolific behavioral science platforms, selecting for American participants who were fluent in English and had successfully completed previous experiments with the platform. We held VAL's tamper identification accuracy constant at 80%.

In our first foundational study, participants were instructed to remove seals that had indications of tamper and keep those with normal wear and tear in-place. VAL read the seal identification number to participants (which was always correct) and then provided the tamper assessment (with 80% accuracy). VAL's potential response types, with intended participant actions, are provided in Table 1.

Table 1 VAL Response Types and Intended Participant Actions for Foundational Study 1

	VAL indicates "signs of tamper"	Val indicates "normal wear & tear"
Seal has tamper	**True positive** *Remove the seal*	**False negative** *Remove the seal*
Seal does not have tamper	**False positive** *Leave seal in place*	**True negative** *Leave seal in place*

Our results from the first foundational study indicate that that there is a difference in user performance when VAL provides correct versus incorrect prompts to users regarding the status of tamper (see Figure 2). For instance, in cases where the seal does have signs of tamper, there is a 17-point performance difference between when VAL correctly tells the user there are signs of tamper (True Positive) and when VAL incorrectly tells the user there are no signs of tamper (False Negative). We can also see interesting trends in response time (see Figure 2). When there are no signs of tamper on a seal, there is a difference of over two seconds in user response time if VAL correctly tells the user there is no tamper (True Negative) and when VAL incorrectly tells the user there is sign of tamper (False Positive). The two second difference indicates users are performing detailed searches for the tamper that VAL has indicated. Participants' rating for how much they trusted VAL was approximately 6.5 out of 10 (standard deviation = 1.94) for Foundational Study 1.

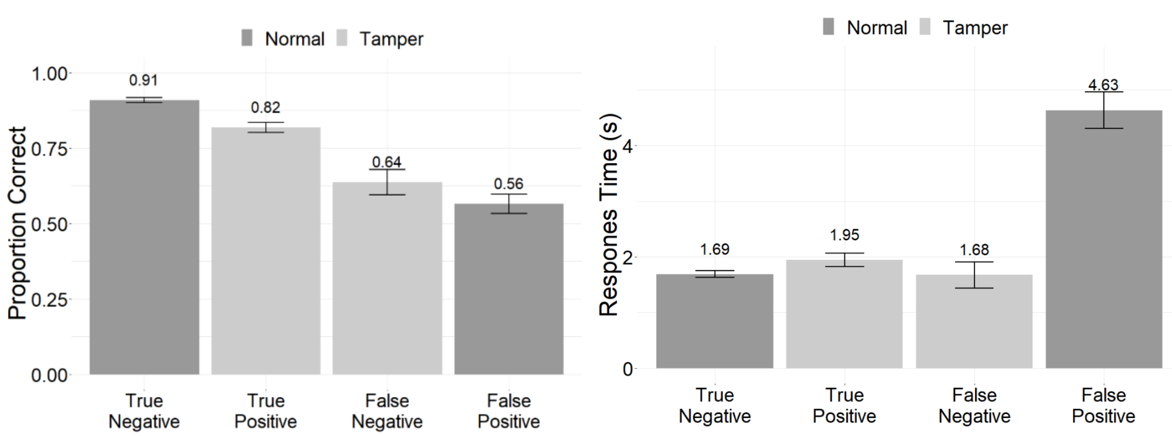

Figure 2 User accuracy (left) and response time (right) for Foundational Study 1.

Foundational Study 2 – Tamper Detection and Seal ID Manipulation

In our second foundational study, we included a secondary task on seal identification. In this study, VAL would occasionally (10% of the time) say the wrong seal ID, and we instructed our participants to remove the seal if VAL said a different ID than was present on the seal or if there were signs of tamper. The secondary task is intended to increase difficulty of the experiment and to provide a task that relies completely on the auditory output (since the pure tamper detection task from Foundational Study 1 could be completed with full accuracy without the prompts from VAL).

To differentiate between seals removed due to the primary tamper detection task or the secondary seal ID task, VAL only said the incorrect seal identification number on seals that would otherwise have no reason to be removed (i.e., they did not show signs of tamper or have a response from VAL indicating signs of tamper). With the seal identification secondary task, we now have five task types as indicated in Table 2.

Table 2 VAL Response Types and Intended Participant Action for Foundational Study 2

<table>
<tr><td></td><td>VAL indicates “signs of tamper”</td><td colspan="2">Val indicates “normal wear & tear”</td></tr>
<tr><td>Seal has tamper</td><td>True positive
Remove the seal</td><td colspan="2">False negative
Remove the seal</td></tr>
<tr><td>Seal does not have tamper</td><td>False positive
Leave seal in place</td><td>True negative correct ID
Leave seal in-place</td><td>True negative incorrect ID
Remove the seal</td></tr>
</table>

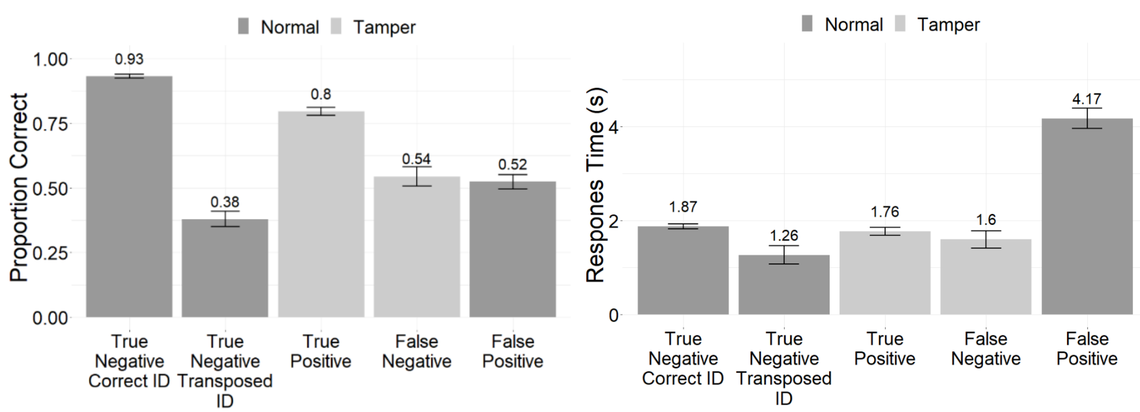

Figure 3 Participant Accuracy and Response Time by Trial Type for Foundational Study 2

Our results from the second foundational study show similar patterns of performance as Foundational Study 1. Performance on the secondary task for seal identification was low, indicating its intended difficulty. Participants rated their trust for VAL as 6.15 out of 10 (standard deviation = 2.09) on Foundational Study 2.

3.0 Nuclear Material Measurements Task

Our future set of experiments will focus on a nuclear material measurements activity. We will use the nuclear material measurements task to measure other trust factors that are less relevant to seal examination. This task will be like the seal examination task in its visual presentation and audio output, but the numerical aspect of the measurement results will allow us to manipulate additional trust factors.

The nuclear material measurement task is centered around determining the level of enrichment of uranium in a 30B storage container. We anticipate using open-source images of the containers from the Limbo dataset (Gastelum, Shead, & Rushdi, 2021) for our visual stimuli. Example container images are in Figure 4.

Figure 4. Example synthetic 30B uranium hexaflouride containers from the Limbo dataset. Similar images, including other container models, that will be the visual prompts for the material measurement task.

In this task, the participants will be presented with a visual representation of a nuclear material container, and VAL will read a quantitative measure related to the material (potentially weight or isotopic enrichment). Participants will then decide either to measure the next container (indicating that the measurement was acceptable) or conduct further sampling on the container. As with the seal examination task, the purpose of this activity is to test how our experimental manipulation of how VAL provides information to users impacts their trust in the system. The nuclear material measurement task consists of two primary VUI tasks: **reads information** and **requests additional information**.

In the **VUI reads information** task, VAL reads multiple types and sources of information for an inspector, such as a sampling list, previous inspection reports, sampling results, state declarations, historical results or current readings from a piece of safeguards equipment, maintenance information about a piece of equipment, or results from an Internet search. This task includes directly reporting on or summarizing information from existing records and does not include provision of information such as predictions or recommendations from other models, covered in a different task. This experimental task allows us to test two trust factors: granularity and provenance.

In the **granularity** experiments, we will manipulate how much detail is provided to participants about enrichment values. For example, participants may be told to expect enrichments between 3% – 5% uranium-235. We will compare users who receive verbal prompts that have minimal granularity (e.g., enrichment level is within expected bounds), low granularity (e.g., enrichment level is 4%), and moderate granularity (e.g., enrichment level is 4.7%). Based on previously established best practices for VUIs, we will avoid higher granularity responses (e.g., 4.6598%).

In the **provenance** experiments, we will manipulate how much information about prior measurements is communicated to participants. For example, participants may be told that a container shows an enrichment of 5% U-235 and that the last measurement taken on that item two months ago was 3% enriched. We will manipulate whether the provenance is provided and how much historical information is provided in the description of prior activities.

The **VUI requests information** task was outlined above in relation to seal checking. In the nuclear material measurement task, participants will be given an opportunity to re-measure a container (or in some cases VAL will compel a re-measurement). The re-measurement activity will improve consistency of measurement values. For example, if participants are told to expect uranium enrichment measures between 3% – 5% and VAL says the equipment reading is 6%, a re-measurement would bring the enrichment down to 5%. This experiment will test either **provenance** or **explainability**, the details of which will be determined by findings from current **requests information** experiments related to seal examination.

Conclusion

A safeguards digital assistant with a VUI is one AI tool that the IAEA could employ in the future to make inspections more efficient. However, low trust in a digital assistant could result in non-use of the system. This work seeks to understand the factors that affect user trust in a VUI for safeguards-related tasks. Over the next two years, our project team will deploy and analyze a series of human performance experiments assessing the impact of various trust factors on user compliance and trust of a VUI. From these experiments, we will document our results with actionable recommendations intended for the development community. While our results will be intended specifically for international nuclear safeguards digital assistants, we anticipate that our results and recommendations will also be relevant to other domains including export controls, physical protection, and homeland security.

Acknowledgements

This work was funded by the National Nuclear Security Administration's Defense Nuclear Nonproliferation office. Additional thanks to Haley Norris at Sandia National Laboratories for her design and production of our seal stimuli. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND2023-02964C.

Works Cited

Attig, C., Wessel, D., & Franke, T. (2017). Assessing personality differences in human-technology interaction: An overview of key self-report scales to predict successful interaction. *HCI International 2017 - posters' extended abstracts part I, CCIS 713*, (pp. 19-29). Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-58750-9_3

Feldman, Y., Arno, M., Carrano, C., Ng, B., & Chen, B. (2018). Toward a Multi-Modal Deep Learning Retrieval System for Monitoring Nuclear Proliferation Activities. *Journal of Nuclear Materials Management, XLVI*(3), 68-80.

Gastelum, Z. N., & Shead, T. M. (2018). Inferring the Operational Status of Nuclear Facilities with Convolutional Neural Networks to Support International Safeguards Verification. *Journal of Nuclear Materials Management, XVLI*(3), 37-47.

Gastelum, Z. N., Matzen, L. E., Divis, K., Stites, M. C., & Howell, B. C. (2021). Not All Errors are Created Equal: Examining Human-Algorithm System Performance for International Safeguards-Informed Visual Search Tasks. *Proceedings of the Institute of Nuclear Materials Management & European Safeguards Research and Development Association Joint Virtual Meeting.* Virtual.

Gastelum, Z. N., Matzen, L. E., Smartt, H. A., Horak, K. E., Solodov, A. A., Moyer, E. M., & St. Pierre, M. E. (2017). Testing Human Performance in Simulated In-Field Safeguards Information Environments. *Proceedings of the Institute of Nuclear Materials Management Annual Meeting.*

Gastelum, Z. N., Shead, T. M., & Rushdi, A. (2021). A Large Safeguards-Informed Hybrid Imagery Dataset for Computer Vision Research & Development. *Proceedings of the Institute of Nuclear Materials Management and European Safeguards Research and Development Association Joint Annual Meeting.* Retrieved from https://esarda.jrc.ec.europa.eu/esarda-43rd-joint-annual-meeting_en

Gastelum, Z. N., Shead, T. M., & Rushdi, A. (2021). A Large Safeguards-Informed Hybrid Imagery Dataset for Computer Vision Research and Development. *Proceedings of the INMM & ESARDA Joint Virtual Annual Meeting.*

International Atomic Energy Agency. (2017). *Emerging Technologies Workshop: Trends and Implications for Safeguards. Workshop report.* Department of Safeguards, Vienna, AUSTRIA. Retrieved from https://www.iaea.org/sites/default/files/18/09/emerging-technologies-130217.pdf

International Atomic Energy Agency. (2020). *Emerging Technologies Workshop: Insights and Actionable Ideas for Key Safeguards Challenges. Workshop Report.* Department of Safeguards. Retrieved from https://www.iaea.org/sites/default/files/20/06/emerging-tehnologies-workshop-290120.pdf

International Atomic Energy Agency. (2022). *Enhancing Capabilities for Nuclear Verification: Resource Mobilization Priorities.* Department of Safeguards.

Mishra, V., Branger, E., Elter, Z., Grape, S., & Jansson, P. (2021). Comparison of Different Supervised Machine Learning Algorithms to Predict PWR Spent Fuel Parameters. *Proceedings of the Joint Institute of Nuclear Materials Management and European Safeguards Research & Development Association Annual Meeting.*

Rossa, R., & Borella, A. (2021). Inferring Initial Enrichment, Burnup, and Cooling Time of Spent Fuel Assemblies Using Artifical Neural Networks. *ESARA Bulletin: The International Journal of Nuclear Safeguards and Non-Proliferation*, December.

Rutkowski, J., Canty, M. J., & Nielsen, A. A. (2018). Site Monitoring with Sentinel-1 Dual Polarization SAR Imagery Using Google Earth Engine. *Journal of Nuclear Materials Managament, XLVI*(3), 48-59.

Smartt, H., Gastelum, Z., Rutkowski, J., Peter-Stein, N., & Shoman, N. (2021). Hey Inspecta! *Proceedings of the INMM & ESARDA Joint Virtual Annual Meeting.*

Smith, M. R., Hamel, M., Hannasch, D., Thomas, M., & Gaiten-Cardenas, C. (2021). A Deep Learning Workflow for Spatio-Temporal Anomaly Detection in NGSS Camera Data. *Proceedings of the Institute of Nuclear Materials Management and European Safeguards Research & Development Association Joint Annual Meeting.* Virtual.

Thomas, M., Passerini, S., Cui, Y., Rutkowski, J., Yoo, S., Lin, Y., . . . Moeslinger, M. (2021). Deep Learning Techniques to Increase Productivity of Safeguards Surveillance Review. *Proceedings of the Institute of Nuclear Materials Management and European Safeguards Research & Development Association Joint Annual Meeting.* Virtual.

Wonton, H. M., Porter, D., Lane, S. T., Bieber, C., & Madhavan, P. (2020). Initial validation o f the trust of automated systems test (TOAST). *The Journal of Social Psychology*. doi:https://doi.org/10.1080/00224545.2020.1749020