SAND2016-7166C

AN IMPACT ASSESSMENT OF ADVANCED INFORMATION TECHNOLOGIES FOR IAEA SAFEGUARDS INSPECTIONS

Zoe N. Gastelum[1], E. Travis Gitau[2], Joel R. Doehle[3], Alexa R. Erickson[2]

1. Sandia National Laboratories, Albuquerque, NM
2. Pacific Northwest National Laboratory, Seattle, WA
3. Pacific Northwest National Laboratory, Richland, WA

Abstract

Since at least the 1990s, researchers in the field of international safeguards have proposed the use of advanced information technologies to support International Atomic Energy Agency (IAEA) safeguards inspectors working in the field. Evolving expectations of safeguards inspectors' investigative responsibilities, combined with the rapidly increasing availability of information that is potentially safeguards-relevant within a State, accentuates the promise that advanced information systems could offer. However, the impact of such systems must be considered across a range of safeguards inspection activities and priorities. In this paper, we describe the development of an impact assessment methodology of advanced information technologies for IAEA safeguards inspections, utilizing the analytical hierarchy process (AHP) as its foundation. We then discuss the application of that methodology to four hypothetical safeguards inspection scenarios which offer various levels of advanced information technology integration. We conclude with the outcome of the safeguards inspection impact assessment and a discussion of its implications for IAEA safeguards.

Introduction

International Atomic Energy Agency (IAEA) safeguards inspectors are now expected to be familiar with, and able to act upon, a wide variety of information related to a State that was previously not considered during in-field verification activities such as satellite imagery, scientific literature, and nuclear trade [1]. Current access to potentially safeguards-relevant information in the field is limited to what inspectors downloaded or printed before leaving the IAEA, or what they could request via email in the field. The information, when available, is not necessarily organized based on the inspector location, activity, or other means to ensure the most effective and efficient use of that data during inspections. The need for secure technologies to support collaborative information sharing and advanced analytics has been formalized in the IAEA Long-Term Research & Development Plan's *Long Term Capability 8*, noting the need for an "ability to use safeguards information in a fully integrated, secure environment, maintained and available to all who need it according to their role" [2]. To facilitate inspector access to all information related to a State, the IAEA could deploy advanced information sharing and analytical tools both at headquarters and in the field.

Since at least the 1990s, researchers have been proposing advanced information and computing systems to support IAEA safeguards. These platforms could enhance inspector situational awareness, increase inspection efficiency, and improve timeliness of information exchange between inspectors in the field and at headquarters potentially requiring less time at facilities, or fewer follow-up information requests from the IAEA. Many of these platforms have been advanced or even "futuristic" for their

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000

times. Indeed, even members of the research team represented in this paper were involved in developing proof-of-concept capabilities [3] and conducting deployment feasibility studies [4, 5] for such advanced information systems. Yet, it is difficult to quantitatively assess how such platforms that could offer additional value for inspectors, or how to compare them.

Before the Agency embarks on costly and time-consuming procurement or development activities for an advanced, mobile safeguards information environment, there is an opportunity to begin a dialogue regarding the full spectrum of capabilities, and the potential costs and benefits of each. By understanding the potential impacts (both positive and negative) of various deployment scenarios, the IAEA Department of Safeguards or other state or regional safeguards authorities could be better informed as they consider advanced safeguards information systems. In response to this opportunity, and building upon several years of experience in developing and analyzing advanced information environments for international safeguards, Pacific Northwest National Laboratory and Sandia National Laboratories collaborated to develop and implement a Safeguards Impact Assessment Methodology to evaluate a series of information technology deployments scenarios for international safeguards.

In this paper, we describe the development of a quantitative safeguards impact assessment methodology based on the Analytical Hierarchy Process (AHP) and utilizing expert elicitation to analyze how mobile information platforms for international safeguards impact the time, accuracy, and other factors of international safeguards inspections compared to current practices. We will describe a trial application of the methodology to four hypothetical safeguards inspection scenarios which offer various levels of advanced information technology integration. We will describe with the outcome of the assessment and its implications for international safeguards and the research community working in this field.

Technology Scenarios

In order to capture the range of potential mobile information technology development options for international safeguards, the research team developed a series of hypothetical safeguards scenarios to represent general categories of technology adoption. The hypothetical safeguards inspection scenarios were developed to facilitate expert evaluation of advanced information technologies for international safeguards, and do not represent actual facilities, inspection techniques, or safeguards equipment or technology. The use of the International Atomic Energy Agency as the inspectorate body was for example purposes only, and the scenarios do not necessarily represent current or anticipated inspection practices.

The development of technology deployment scenarios centered on identification and application of advanced technologies within the context of the DIE/DIV activity of Examination/Verification of Process and/or Containment Design at a gas centrifuge enrichment plant. In order to assess the impact of various types of technologies, four deployment scenarios were defined. These scenarios differ based on their technology development level, and include:

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000

- A control, or baseline, scenario, designed to mirror current IAEA practice with respect to information technology use in the DIE/DIV context.
- A basic deployment scenario focusing on digitization and automation of paperwork.
- A moderately advanced deployment scenario including advanced analytics and collaboration capabilities.
- A highly advanced deployment scenario meant to illustrate the full potential of using collaborative information platforms to support international safeguards verification activities.

The four safeguards scenarios included advanced technologies in five categories: 1) data storage and access; 2) personal computing devices; 3) user interface; 4) sensor platform; and 5) advanced sensors. The sensors included in each of these categories is listed in Table 1.

Table 1 Technologies included for each Safeguards Scenario

Technology Categories	Corresponding Scenario Levels (Base – High Level)
Data Storage and Access	**Base:** Hard copy/written records **Low:** Complete storage on local personal computing device **Moderate:** Backup, store, access some data on the Cloud **High:** No (or minimal) local storage. Completely dependent on Cloud infrastructure
Personal Computing Device(s)	**Base:** Laptop computer **Low:** Mobile devices (such as tablet computers) **Moderate:** Wearable devices (e.g., smartwatches, flexible screens, heads-up display) and smart phones **High:** Holographic heads-up display
Interface	**Base:** Keyboard and mouse and/or trackpad **Low:** 2D gesture control, including multi-touch (e.g., on touch screen) **Moderate:** Speech control capability and Optical Character Recognition (OCR) **High:** 3D gesture control
Sensor Platform	**Base:** Hand-held, and facility process-mounted sensors **Low:** Wearable sensors **Moderate:** User-controlled grounded robotic platform **High:** Unmanned aerial vehicles (single user controlled to autonomous coordinated swarms)
Advanced Sensors	**Base:** Unknown at this time **Low:** GPS; portable, mountable "action" camera (e.g., GoPro) **Moderate:** Portable, mounted laser range finder; simple augmented reality sensors (e.g., mobile device applications using personal computing device cameras) **High:** Advanced augmented reality sensors, depth imaging, computer vision sensors, short range multispectral imaging

Incorporating the advanced technology sets described above and the IAEA's defined objectives and methodologies for DIE/DIV, we developed narratives of hypothetical safeguards scenarios that included the selected technologies into the predefined inspection activities. To avoid biasing participants in the expert evaluation activity used to rank the performance of each technology development scenario, we elected to nominally refer to each scenario by the type of personal computing device used in the scenario (i.e., laptop, mobile, wearable, and holographic display scenarios).

Methodology Development

In developing the Safeguards Impact Assessment methodology, multiple decision theory methodologies were evaluated, as well visualization approaches that are sometimes used to present qualitative and quantitative information to decision-makers. After considerable review, the research team decided that the Analytical Hierarchy Process (AHP) was most suitable for the nature of this research. AHP proscribes both how to mathematically combine information about factors, or measures, which relate to the specific research objective (in our case, safeguards impact), as well as how to quantify those measures, i.e. the metrics.

<u>Analytical Hierarchy Process for Assessing Safeguards Impact</u>

In AHP, users make pair-wise comparisons of alternatives (i.e. scenarios) for each of the defined measures. For our research, this meant pair-wise comparisons of the four safeguards inspection scenarios for each safeguards impact measure (described below), for example: "Does scenario 1 or scenario 2 contribute to better inspection time?" The users/respondents then rank "how much better" on a standard scale. We describe the scale in additional detail in the "Metric Definitions" section.

In AHP, both the scenarios and measures may be ranked in pair-wise comparisons, so that the end result is a prioritization of options based on their performance for each measure, weighted by the importance of each measure on overall objective. While some AHP analyses chose to have all measures weighted equally, we elected to use expert elicitation to determine the pair-wide rankings for both the scenarios and the measures used to evaluate those scenarios.

Analytical Hierarchy Process can be administered in several ways, for example through group consensus when all respondents are discussing in a single location, or by averaging of individual results. For timeliness considerations, the research team elected distribution of electronic surveys to a group of experts who each responded independently. More information on the experts and the survey they completed is provided in the "Expert Survey" section below.

<u>Defining Safeguards Impact Measures</u>

Measures are the factors by which we compare each scenario. They should be independent of each other (to the extent possible), and should be measureable or observable in some way. For this research, we began by brainstorming a list of approximately 25 potential measures, and narrowed potential factors by removing measures were inter-dependent, co-dependent, interesting but not relevant for the purposes of assessing safeguards impact, or relevant but simply not observable within the parameters of this project. The following seven measures were the ultimate result of that process (in alphabetical order):

- Applicability to multiple verification activities: The ability of technologies in the safeguards scenario under evaluation to also support other safeguards verification activities, either in the same manner or potentially in a different way. For example, if the safeguards scenario under evaluation included the use of a flashlight to look for pipe penetrations in a dark room as part of a DIV activity, could that flashlight be used for other DIV activities, inventory verification, or complementary access?
- Inspection time: The time it takes an inspector (or team of inspectors) to complete the safeguards inspection activity under evaluation. This includes not only the on-site inspection time, but any time required to resolve inconsistencies or anomalies. Inspection time also includes any time needed for documentation of inspection activities.
- Inspector accuracy**:** The correctness with which an inspector (or team of inspectors) is able to complete the safeguards activity under evaluation. This includes both performing and recording the task correctly, and should account for the propensity for incorrect measurements or record-keeping due to human errors.
- Operator acceptability**:** The ease in which most facility operators will allow the use of the suite of technologies for the safeguards activity under evaluation. This includes operator concerns regarding safety, as well as information security and transparency of information systems to the operator (for example, so that an operator can have high confidence that commercial proprietary information is not being recorded or transmitted off-site).
- Situational awareness**:** The contribution of a suite of technologies to an inspector's (or team of inspectors') awareness of their environment and surroundings, including geospatial orientation (where the inspector is, and where he/she has been before), personnel behavior or activities, and site awareness (things that were different last time, activities that appear to be taking place). Potential contributors to situational awareness include display of existing and new information, in field analysis support, current location/site tracking, and visualization of previous inspection paths.
- Sustainability**:** The ability of both hardware and associated software to be maintained, repaired, re-purchased, and upgraded into the future. This includes the ability to procure identical or compatible parts or complete replacements for hardware, and inter-compatibility with upgraded operating systems or other software changes for software programs.
- Usability: The ease of use and learnability of technologies involved in the safeguards activity under evaluation. This includes the intuitiveness of the human-machine interface, and the amount of training or expertise required for an inspector to proficiently use the required technologies.

Information security - defined as the assurance of the confidentiality, integrity, and availability of IAEA safeguards-relevant information in an inspection scenario - was also of high interest to the research team. Because the majority of the experts to who the survey would be proctored would not have sufficient experience with information security to make an informed judgement, the topic was included as an optional question on at the end of the survey.

Metrics Definition

The Analytical Hierarchy Process uses a pre-defined scale by which to assess each option's performance on a measure (i.e., the metrics), described in Table 2. A translation of the scale for each of the seven safeguards impact measures defined in this study was developed for use in the expert survey.

Table 2 Intensity Scale for the Analytical Hierarchy Process

Intensity	Definition	Explanation
1	Equal importance	Two elements contribute equally to the objective
3	Moderate importance	Experience and judgement slightly favor one element over another
5	Strong importance	Experience and judgement strongly favor one element over another
7	Very strong importance	One element is favored very strongly over another, its dominance is demonstrated in practice
9	Extreme importance	The evidence favoring one element over another is of the highest possible order of affirmation
2, 4, 6, 8 can be used to express intermediate values		

Expert Survey

The pair-wise evaluation of the safeguards scenarios was completed through an expert survey. Experts were recruited based on their experience with or knowledge of international safeguards inspections or arms control inspections. Additionally, some experts who also had familiarity with the nonproliferation domain were recruited for their experience developing and evaluating advanced information platforms. A recruitment email was sent to a total of 21 experts from Pacific Northwest National Laboratory (and their contractors) and Sandia National Laboratories. Of the 21 recruited, eight experts returned their completed consent forms, and seven returned completed surveys.

The survey consisted of pair-wise comparisons of the four scenarios' performance on each of the seven safeguards impact measures, resulting in a total of 42 questions. The experts were given two weeks to complete the survey. To minimize burden on the expert group, the pair-wise comparison of seven measures' importance for safeguards impact (i.e., the weighting of the measures) was evaluated within the research team using the same AHP approach. Each member of the research team completed a worksheet consisting of 21 pair-wise comparisons.

Results and Analysis

The results of the safeguards impact measure weighting activity are listed below in

Table 3.

Table 3 Safeguards Impact Measure Weighting

Rank	Measure	Weight
1	Inspection accuracy	43.7%
2	Usability	15.1%

3	Operator acceptability	13.6%
4	Situational awareness	8.7%
5	Inspection time	8.1%
6	Sustainability	6.4%
7	Applicability to multiple verification activities	4.4%

Completed expert surveys comparing the four safeguards inspection scenarios were returned electronically to the research team. The results from the expert survey were analyzed two ways – first using the weighted safeguards impact measures above, and second using unweighted scores in which each measure contributes equally to safeguards impact. The weighted results, which take into consideration the relative importance of each measure as determined by the research team, are listed below in Table 4. The unweighted results, assuming equal importance for all safeguards impact measures, are listed in Table 5. While the rank order of the scenarios stayed the same in both the weighted and unweighted results, the absolute scores of each scenario did change, with the Hologram scenario's score actually improving in the weighted analysis, indicating that this scenario excelled in areas that were given high priority by the weighting scheme.

Table 4 Weighted Safeguards Scenario Results

Weighted Rank	Scenario Name	Weighted Score
1	Hologram	0.359959
2	Laptop	0.224569
3	Tablet	0.217529
4	Wearable	0.197943

Table 5 Unweighted Safeguards Scenario Results

Unweighted Rank	Scenario Name	Weighted Score
1	Hologram	0.325571
2	Laptop	0.261286
3	Tablet	0.236000
4	Wearable	0.177143

In addition to these combined, weighted scores for the safeguards inspection scenarios, we analyzed the results for each individual safeguards measure. For the measure-by-measure analysis, we found that scenarios were ranked according to their level of incorporation of advanced technologies. For some measures, for example inspection time, the most advanced scenarios performed the best, followed by the less advanced technology scenarios in descending order. For other measures, the less advanced technologies performed the best, followed by the more advanced technology scenarios in descending order. This indicates that some safeguards measures are more supportive of higher technology solutions, while the others favor lower-technology solutions. The measures which were more favorable to more advanced versus less advanced inspection scenarios are listed in Table 6.

Table 6 Safeguards Measures Favoring More Advanced versus Less Advanced Inspection Scenarios

Measures Favoring More Advanced Scenarios	Measures Favoring Less Advanced Scenario

Applicability to multiple verification scenarios	Operator acceptability
Inspection time	Sustainability
Inspector accuracy	Usability
Situational awareness	

We recognized that the larger number of safeguards impact measures that appear to favor inspection scenarios that incorporate more advanced technologies could skew our results towards scenarios utilizing advanced technologies. In response, we performed an additional in which the least-weighted safeguards measure (applicability to multiple verification scenarios) was removed from the analysis to even the number of measures. Even with the elimination of the seventh measure, the hologram scenario ranked the highest in both the unweighted and weighted scores. This can be attributed to (in the weighted scoring), the very high degree of importance that was given to inspection accuracy, in which the Hologram scenario did very well; and (in both the weighted and unweighted scoring), the high degree to which the Hologram scenario was favored over other, less advanced, technology scenarios across the other measures.

Recognizing the diverse experience and personalities of our seven participating experts, we conducted additional analysis on individual expert responses to better understand how each expert's survey responses compared to the others. We conducted the following comparisons of expert response:

- <u>Most "different"</u>: We examined which expert rated the technology scenarios most differently from the other experts, across the seven safeguards measures. This was measured first by the categorical response (whether A or B was better), followed by degree of the distinction. No expert's responses were found to be considerably different from the other experts for more than three of the seven impact measures. This indicates that experts who either interpreted or scored the inspection scenarios differently from the other experts were outliers for less than one-half of the time, and was determined not to have significant effect on our results.
- <u>Most supportive of advanced technology</u>: We examined to what degree experts were generally in favor of higher levels of advanced technology by determining who selected the higher technology option more frequency in the pair-wise scenarios, and to what degree. To that end, one expert was the supportive of higher technology scenarios for five out of the seven safeguards measures. However, for the two measures in which that expert was not the most supportive of advanced technology, they were one of the two strongest supporters of the lowest technology option, potentially indicating that the expert was more comfortable ranking scenario performance on the extremes of the metric scale than other participants.
- <u>Most supportive of current technology</u>: We examined which experts were generally the most in favor of the less developed technologies, most closely representing current safeguards practices. No expert was the most supportive of current technology for more than three of the seven safeguards measures, indicating that no single expert significantly biased the group towards lower technology options.

- Most extreme scoring: We also examined the general pattern in expert scoring of the degrees of difference in scenarios, regardless of which scenario they thought performed better. This analysis of scoring behavior indicated that one expert tended to score the favorability of one scenario over the other with the highest possible order of affirmation, while one other expert was much more likely to indicate that the two scenarios being compared performed equally to one another, or with minor differences. The remaining five participants tended to score the relative difference of scenarios between three and five on the AHP scale, which meant a moderate to strong favorability of one scenario's performance on a safeguards impact measure over the other.

Conclusions

Deploying new or advanced technologies to international nuclear fuel cycle facilities will almost always be met with operator concerns, especially if those technologies have capabilities that could pose information security, safety, or other operational challenges. Though it may seem that a step-wise progression towards the use of more advanced technologies might be preferred, our analysis indicates that the intermediate level reliance on advanced technologies, such as presented in the Tablet and Wearable scenarios, actually provide limited safeguards benefit in comparison to their obstacles for deployment. Our analysis suggests that a bold move to integrate highly advanced information platforms as associated sensor packages could have the greatest impact on international safeguards activities, even when considering the likely barriers to deploying these solutions.

Despite these conclusions, a more detailed study is required to measure the safeguards impact of individual technologies on safeguards inspections activities, rather than the inspection scenarios incorporating multiple technologies used in this analysis. From participant comments on our survey we noticed that some technologies described in the safeguards scenarios were mentioned much more than others, and seemed to drive how the scenarios performed on each safeguards impact measure. Thus, being able to discriminate the impact of each information technology could provide a more granular level of detailed that would likely be needed for decision makers. Furthermore, a larger group of safeguards and information technology experts, or a set of experts who will be more directly impacted by the deployment of the proposed technical solutions, could provide unique perspective compared to our United States-based survey participant pool.

The outcomes of this study have the potential to inform research and development activities for the use of advanced information technologies for IAEA safeguards. In addition, we hope that the development of this methodology can assist analysts and decision makers who are examining not only information technologies, but any advanced technology solution that could potentially support international safeguards inspections. The methodology was intended to be developed in a transparent and modular manner, which will allow future users to modify the safeguards impact measures that they deem as the most significant, the relative importance of those measures, and the comparison of any proposed solutions in support of safeguards activities.

References

[1] Gayne, Eva. “Information-Driven Safeguards: A country officer perspective.” Presented at the 2010 Safeguards Symposium, Vienna, Austria. November 2010. IAEA-CN-184/41.

[2] International Atomic Energy Agency, “IAEA Department of Safeguards Long-Term R&D Plan, 2012-2023.” January 2013.

[3] Gastelum, Zoe N., Michael J. Henry, E. Russ Burtner, Joel R. Doehle, Dimitri V. Zarzhitsky, Shawn D. Hampton, Ryan R. LaMothe, and Peter L. Nordquist. “The Development of an Example Precision Information Environment for International Safeguards Use Cases.” Proceedings of the Institute of Nuclear Materials Management Annual Meeting, July 2015, Indian Wells, California USA.

[4] Gastelum, Zoe N., Joel R. Doehle, E. Travis Gitau, Christopher M. Toomey. “Feasibility Study of Implementing a Mobile Collaborative Information Platform for International Safeguards Inspections.” September 2014. PNNL-23716.

[5] Gastelum, Z.N., J.R. Doehle, E.T. Gitau. “Computer Security in an Increasingly Mobile World: Recommendations for the Nuclear Community.” Proceedings of the International Conference of Computer Security in a Nuclear World: Expert Discussion and Exchange. June 2015, Vienna, Austria.