LA-UR- 02-3107

Approved for public release; distribution is unlimited.

c. 1

Title: Expert Elicitation And The Problem Of Detecting Undeclared Activities

Author(s): Joseph F. Pilat
Kory Budlong Sylvester
William D. Stanbro

Submitted to: 43rd Annual Institute of Nuclear Materials Management Meeting
Orlando, FL
June 23-27, 2002

Los Alamos
NATIONAL LABORATORY

EXPERT ELICITATION AND THE PROBLEM OF DETECTING UNDECLARED ACTIVITIES

Joseph F. Pilat
Kory Budlong-Sylvester
William D. Stanbor
Los Alamos National Laboratory
Los Alamos

Abstract

Measures applicable to the detection of undeclared activities are not well established, and their effectiveness is uncertain. To detect clandestine paths, the IAEA is still developing processes and procedures. As the Agency gains experience with new measures and with integrated safeguards, dealing with such problems may become more experience-based and perhaps more closely parallel the process with current safeguards where detection probabilities for the measures to be utilized on declared paths are well characterized. Whether or not this point will be reached for undeclared and mixed paths, the only tool that appears suitable at present for the purpose of generating a reasonable detection probability that can over time be tested against reality and, if necessary, adjusted is formal expert judgment, or expert elicitation.

Formal expert elicitation is a structured process that makes use of people knowledgeable in certain areas to make assessments. To provide a "proof of principle" of this methodology for presentation to the Agency, experts in nuclear technology, nonproliferation, safeguards and open source information, as well as in formal expert elicitation processes, engaged in three illustrative expert elicitations on assessing information analysis as a means to detect undeclared activities. These elicitations were successful. This paper will discuss the process of and issues raised by the elicitations.

Introduction

As the International Atomic Energy Agency (IAEA) develops criteria and approaches for implementing integrated safeguards, SAGSI and other bodies have emphasized the need for an evaluation of proposals for these new safeguards, which will directly address undeclared activities for the first time. To meet this requirement, the US Support Program has developed the Integrated Safeguards Evaluation Methodology (ISEM). ISEM has been designed with the ability to undertake evaluations of declared, undeclared and mixed acquisition paths.[1] As a consequence, its full value depended on resolving the issue of how detection probabilities for undeclared and mixed paths could be formulated. This challenging problem has been addressed

[1] A description of the methodology is provided in "Integrated Safeguards Evaluation Methodology (ISEM): Phase 2/Rev. 1: Concept Development," Joseph F. Pilat, Kory W. Budlong Sylvester, George W. Eccleston, William D. Stanbro, Kenneth E. Thomas, Jim Larrimore, and Myron Kratzer, Los Alamos National Laboratory document LA-UR-00-1366, (March 2000). See also, "Illustrative Application of the Integrated Safeguards Evaluation Methodology: ISP-1/Rev. 1 and Current Safeguards," Kory Budlong-Sylvester, Joseph F. Pilat, George W. Eccleston, William D. Stanbro, and Kenneth E. Thomas, Los Alamos National Laboratory document LA-UR-00-1362, March 2000.

in ISEM through the use of expert judgement or elicitation. In the following, this use of expert elicitation is discussed.

Expert Elicitation and Undeclared Activities

Safeguards measures applicable to the detection of undeclared activities are not well established, and their effectiveness is uncertain. For entirely clandestine paths, the IAEA is still developing processes and procedures to resolve inconsistencies or anomalies. As the Agency gains experience with new measures and with integrated safeguards, dealing with such problems may become more experience-based and perhaps more closely parallel the process with current safeguards where detection probabilities for the measures to be utilized on declared paths are well characterized. Whether or not this point will be reached for undeclared and mixed paths, the only tool that appears suitable at present for the purpose of generating a reasonable detection probability that can over time be tested against reality and, if necessary, adjusted is expert judgement, or expert elicitation.

Expert Elicitation: An Overview

Expert elicitation is a process utilizing knowledgeable people in cases, for example, when an assessment is needed but physically based data is absent or open to interpretation. More specifically, it can be used to:

- predict future events;
- provide estimates on new, rare, complex or poorly understood phenomena;
- integrate or interpret existing information;
- learn an expert's problem-solving process or a group's decision-making processes; or
- determine what is currently known, how well it is known or what is worth learning in a field.

Expert elicitation can be informal or formal. The informal use of expert judgement is frequently used. Although it can produce good results, it often provides demonstrably biased or otherwise flawed answers to problems. This along with the absence of transparency can result in a loss of confidence when experts speak on issues.

More formal expert elicitation is a structured process that makes use of people knowledgeable in certain areas to make assessments. The reason for advocating formal use is that the quality and accuracy of expert judgement comes from the completeness of the expert's understanding of the phenomena and the process used to elicit and analyze the data. The use of a more formal process to obtain, understand and analyze expert judgement has led to an improved acceptance of expert judgement because of the rigor and transparency of the results.

Procedures for formal elicitations are employed throughout the process to enhance the technical credibility and consistency of this process, and to enhance its transparency. The process involves:

- identification and selection of issues;
- identification, selection and training of experts;
- development of specific questions and assumptions;
- elicitation of expert judgement;
- analysis, aggregation and resolution of conflicts between experts; and

- documentation and communication of results.

Uses of Expert Elicitation

Formal processes based on the judgement of experts are commonly used in many fields. In the United States, Europe, Australia and elsewhere, governments, industry and academia use expert elicitation for:

- development of nuclear accident risk assessments;
- development of financial institution risk assessment;
- product improvement and new product reliability in the automobile industry;
- development of food inspection procedures;
- maintenance of dams;
- determination of uncertainties in climate change detection and attribution;
- choice of environmental remediation methodologies; and
- other uses.

For the purpose of ISEM's use, expert elicitation can offer:

- specific assessments of detection probabilities for undeclared activities;
- an overall assessment of information analysis;
- comparative assessments of its role with respect to types of facilities and activities; and
- a tool for internal planning purposes.

Conducting Expert Elicitations

To provide a "proof of principle" of this methodology for presentation to the Agency, experts in nuclear technology, nonproliferation, safeguards and open source information, as well as in formal expert elicitation processes, engaged in three illustrative expert elicitations to assess information analysis as a means to detect undeclared activities. The experts were asked to participate in formal expert elicitations to assess the detection capabilities and associate uncertainties of information analysis as a safeguards measure for Acquisition Path (AP) segments for input to ISEM.[2] This meant becoming knowledgeable on the elicitation process itself and completing all required steps in the process.

The elicitations were undertaken between September 2000 and April 2002. In all, sixteen acquisition path segments were selected for analysis. The AP segments chosen were expected to pose different levels of difficulty for information analysis—i.e., they were more and less challenging. Additionally, they were chosen because they would allow us to complete the evaluation of various acquisition paths begun in exercises with the Agency.

For each of the three elicitations, the LANL participants developed an elicitation package that consisted of:

- a cover letter with instructions;
- papers on the expert elicitation process;
- a paper on information analysis;
- a questionnaire with specific guidelines; and
- assumptions to use in answering the questions.

2

The experts were sent this full package and asked to make clear their assumptions and consider the following issues:

- Is the list of indicators complete?
- What is the likelihood that the assumed indicator will appear in the various sets of data gathered by the IAEA (with attention to concealment attempts)?
- What is the likelihood that the Agency will recognize the indicator as such and initiate follow-up activities?
- What is the persistence of the indicator? Will the indicator, by its nature, remain available for detection or will it disappear?
- Is the performance of suggested measures independent of one another?
- Which indicators could be defeated by effective concealment, deception and denial practices?
- If a safeguards measure is not completely defeated, what is its residual effectiveness?
- With concealment practices in effect, what is the effectiveness of detection before material production (i.e., within conversion time) and one year after material production?
- What is the role of human sources? While direct physical evidence may be temporal in nature, participant knowledge of activities may remain and leak out over time. How significant is this prospect? What is the impact on detection capabilities over time?
- What is the likelihood that the follow-up activities themselves will effectively resolve the issue? What is the impact on detection capabilities over time?

The experts were given time to respond in writing to the questions and to the validity of the assumptions. (The later point is important. The experts were specifically asked to assess the working assumptions of the elicitation in their responses.)

Once received, the experts' responses were analyzed and aggregated, and a draft report of the elicitation was prepared and sent to the experts. This report put forward in draft form the results of the elicitation, specified the areas of agreement and disagreement, raised some follow-on questions and requested the experts to explain (in greater detail, in most cases) their rationales for certain elicited values. The experts were given time to respond to this package. The experts then met in Washington, DC, at the close of each elicitation to resolve remaining conflicts and produce a consensus report.

The objectives of the elicitation were achieved. The results of the expert elicitations included

- general observations about the role of information analysis in detecting undeclared APs; and
- specific inputs for segment activities.

A discussion of the general observations and the segment values is not in the scope of this paper, but a discussion of the issues raised in developing those values is considered below.

Issues Concerning Eliciting Detection Probabilities

The experts assessed the detection probabilities, and the uncertainty of their judgements, for the AP segments under review. They reached consensus on nearly all values. Based upon the indicators before them (along with any amendments to the indicator list the experts deemed necessary), the elicited values reflect an assessment of the strength of those indicators, and the

likelihood that the IAEA would be able to see them, view them as raising a question or inconsistency and initiate follow-up activities. [3]

For the purposes of the elicitation, it was assumed the Agency's information collection capabilities would reflect "best practices." Furthermore, if a decision were taken to follow-up a question or inconsistency, it was assumed that such actions would be fully effective. This assumption was questioned by the experts, who believe it should be examined.[4]

It was recognized that the analysis of open source information along with other information available to the Agency could be used as a means of detection. In practice, beyond information collection, the success of information analysis will hinge strongly on capable analysts who are given sufficient time and resources to perform their duties. Even in the best case, however, there were different prospects for information analysis achieving timely detection for different types of facilities and/or activities.

The use of information analysis for the detection of undeclared facilities and activities was recognized to pose challenges. A number of questions were raised:

- How will information, including open source information, be presented?
- Will it be seen to reflect biases of sources?
- Can the Agency process the information it receives, or will it be overwhelmed?
- Will the Agency have adequate and independent analytical capabilities?
- Will analyzed information be seen as a credible "cue" for follow-up action?
- Who will ultimately determine what is "significant" and "credible" information, i.e., actionable information?
- What criteria will be used?

These questions were regarded by the experts as critical, and answers to them were seen by the experts as ultimately defining the prospects for information analysis as a safeguards measure. In this context, determining what information might be available is the Agency is a difficult task. The IAEA will face numerous difficulties/limitations in assessing available information, including incomplete information, media errors and the "echo chamber" effect in which a single piece of open source information is widely reported in multiple sources. The Agency will need to assess information to determine the original source, its credibility and its independence. Regular reviews of scientific-technical literature and import/export information were seen as critical.

In all elicitations, the experts agreed that the possibility of third party information could be decisive, but could not counted on by the Agency. For example, whether or not such information is produced by third parties and brought to the attention of the Agency in a manner in which it is

[3] It is necessary to recall that the results of the experts' assessments are to be seen as initial estimates of detection probabilities for the undeclared segments under consideration. As the Agency gains experience with information analysis and/or there is a decision to bring State-specific information into the analysis that directly bears upon these values, there may be a need for a higher fidelity reevaluation.

[4] It was recognized that IAEA follow-up actions would need to deal with: assessing the credibility of the information; seeking clarifications (which may warn a State and enhance the prospects for concealment); and conducting timely and effective internal deliberations on an appropriate course of action.

potentially useable is subject to a relatively larger number of factors beyond the Agency's control (as compared to other information). It was viewed as taken into account both in the elicited values and via the uncertainties associated with relevant detection probabilities.

The timeliness of detection based on open source information was also seen as an issue. In procedural terms, it has to do, in the first instance, with the frequency with which information is evaluated in the Agency. If it is done regularly at frequent intervals, detection could in principle be timely. If done only sporadically, or rarely, the prospects for this measure decline. This is a matter fully within the Agency's control, and it can be used to increase the prospect of timely detection.[5]

In the elicitations, the reliance of a proliferator on imports was seen as a particularly important factor in the experts' judgements. This raised the question of whether the values were sensitive to State-specific factors. It appeared in the first two elicitations (EEL-1 and EEL-2) that they were largely insensitive, but it was eventually decided that the issue of whether the elicited values were largely sensitive or insensitive to State-specific features should be addressed directly. To more systematically assess whether the values are insensitive to State-specific factors, it was agreed that the third elicitation (EEL-3) would specifically look at a range of three generic States/fuel cycles in order better to gauge the "universality" of the values.

Any elicited differences were not especially wide. Indeed, the values for all the hypothetical states and the reasoning behind them were to a large extent comparable. Moreover, the elicited values were largely (but not fully) consistent with earlier elicitations where the most significant potential differences among states were seen as related to import dependence for higher technology endeavors (i.e., enrichment). On the basis of these results, it appears that meaningful detection probabilities for a "generic" State are possible.

The consideration of this matter resulted in an important finding. To the extent that there are differences, the technical/industrial capabilities of a State appear most significant. The more a State can do indigenously, the less opportunity there will be of detection via information analysis. Therefore, in explaining the differences between States, its technical capabilities and how it operationalizes secrecy appeared most relevant in the results of the elicitation. Although they cannot be dismissed, political factors appeared less significant in the results. Democracies are capable of maintaining secret programs, especially when they involve national security objectives that enjoy broad political support.

---

[5] The persistence of an indicator was also viewed as significant for determining timely detection probabilities. Theoretically, the longer an activity goes on, the more susceptible it is to detection. This factor could, in principle, allow the probability of detection to increase with time. However, if the detection occurs, for example, after a State has already obtained sufficient material for a weapon, it cannot be regarded as timely. Moreover, the possibility of detection probability increasing over time was questions by the experts. In reality, an activity that can survive undetected for a certain period of time is likely to continue to evade detection, assuming that the signature does not change and that there are no dramatic changes in operational security. Furthermore, a persistent indicator may not be sufficient, unless it is clear that the indicator can only have one interpretation, or it is discovered at a location where it is not expected. If neither of these conditions prevail, then persistence may not be sufficient to raise any alarms leading to follow-up actions or otherwise ensure a meaningful detection capability.

A related point is that the more knowledgeable and experienced a State is in technical areas relevant to a given process, the less likelihood there will be of a major accident. However, as with third party information, is not clear an accident would produce meaningful indicators, or whether the Agency would detect those indicators and recognize their implications.

Conclusions

The expert elicitation process achieved its objectives. It looks to be a useful and transparent tool for initial assessments of detection probabilities for undeclared facilities/activities. Further development and testing of this aspect of the ISEM will increase confidence in its effectiveness and practicability. This capability appears promising as a tool to reassess assumptions, allocate resources for information analysis and for other aspects of IAEA internal planning. If there is interest in Agency use of this capability, guidelines/procedures for conduct of elicitations will be necessary. However, unless the process is done formally and transparently, its advantages are lost.