OPEN SOURCE INFORMATION ACQUISITION, ANALYSIS, AND INTEGRATION IN THE IAEA DEPARTMENT OF SAFEGUARDS[1]

Michael Barletta, Nicholas Zarimpas, Ryszard Zarucki

Acquisition and analysis of open source information plays an increasingly important role in the International Atomic Energy Agency's (IAEA) strengthened safeguards system. The agency's focal point for open source information collection and analysis is the Division of Safeguards Information Management (SGIM) within the IAEA's Department of Safeguards. In parallel with the approval of the Model Additional Protocol in 1997, a new center of information acquisition and analysis expertise was created within SGIM. By acquiring software, developing databases, retraining existing staff and hiring new staff with diverse analytical skills, SGIM is proactively contributing to the future implementation of information-driven safeguards in collaboration with other Divisions within the Department of Safeguards. Open source information support is now fully integrated with core safeguards processes and activities, and has become an effective tool in the work of the Department of Safeguards.

This paper provides an overview of progress realized through the acquisition and use of open source information in several thematic areas: evaluation of additional protocol declarations; support to the State Evaluation process; in-depth investigation of safeguards issues, including assisting inspections and complementary access; research on illicit nuclear procurement networks and trafficking; and monitoring nuclear developments.

Demands for open source information have steadily grown and are likely to continue to grow in the future. Coupled with the enormous growth and accessibility in the volume and sources of information, new challenges are presented, both technical and analytical. This paper discusses actions taken and future plans for multi-source and multi-disciplinary analytic integration to strengthen confidence in safeguards conclusions—especially regarding the absence of undeclared nuclear materials and activities.

Following the rapid advances in the digital world, open source information has become a central component of the working environment in both the private and the public sectors.[2] In the course of progressive development of nuclear safeguards, the IAEA has devoted increasing effort to the effective collection of such data, its evaluation and analysis, as well as proper dissemination of acquired knowledge. The IAEA is in the unique position of being able to combine and compare State-declared information and inspection reports and findings, with information collected from a wide variety of open sources.

---

[1] A version of this paper was presented at "Global 2009: The Nuclear Fuel Cycle: Sustainable Options and Industrial Perspectives," Paris, France, September 6-11, 2009, and published in *Proceedings of Global 2009* as paper 9358.

[2] For the purposes of this paper, open source information is defined as information generally available to the public from various sources, such as government agencies, private entities, academic institutions and think tanks, the news media, commercial databases and scientific associations. This information may be available in print form (e.g. newspaper and journal articles, technical and scientific papers, reports and brochures), electronic database entries, and as web pages.

With the advent of the Model Additional Protocol in 1997, comprehensive capabilities for collecting and processing open source information were created in the Division of Safeguards Information Management (SGIM)—known as the Division of Safeguards Information Technology at the time—a division in the Department of Safeguards that traditionally had been responsible for processing nuclear material accountancy data and providing hardware and software support for the Department's computer systems.[3] The development of SGIM's capabilities in this area was significantly aided by voluntary assistance from Member State Support Programs, which, coupled with IAEA resources, have enabled SGIM to acquire specialized tools, train staff and hire individuals with diverse skills and expertise.[4]

Acquisition of Open Source Information

The Department's open source acquisition requirements are met by continual monitoring of major open source information providers and by developing and maintaining specialized information collections. In addition, new sources, collection tools, and methodologies are regularly evaluated and, if found useful and cost-effective, deployed.

Unlike the well-structured nuclear material accounting data provided to the IAEA by states in fulfillment of their reporting obligations under safeguards agreements, the collection and use of unstructured open source information presents several challenges. First, there is an overwhelming and growing quantity of information available on all thematic areas relevant to agency safeguards. Second, multiple languages and information sources, of widely varying reliability and relevance, must be systematically collected, reviewed, and as appropriate, utilized. Third, and most importantly, "raw" information requires significant effort by highly trained collectors and experienced analysts collaborating though systematic procedures to review, correlate, evaluate, and prepare synthesized information products and services.

The principal reason for collecting open source information is to support the Department's State Evaluation process, which was instituted in the late 1990s to bolster the credibility of the safeguards conclusions reached by the agency. This is a regular (typically annual) in-depth review and evaluation of all information available to the agency about a particular state, which is conducted to evaluate that states' compliance with its safeguards agreements. Presently, more than 100 information compilations per year are being prepared for the State Evaluation process. These so-called "State files" comprise a wide spectrum of information—ranging from background political, economic, legislative and regulatory, and security information, to more specialized scientific, technical and other fuel-cycle related data that is directly relevant to the analysis and evaluation of each state's nuclear program. In addition, focused collections of information and analytical reports are frequently needed to address safeguards-related issues and questions, evaluate state-submitted additional protocol declarations and assist with the

---

[3] IAEA, Model Protocol Additional to the Agreement(s) between State(s) and the International Atomic Energy Agency for the Application of Safeguards, INFCIRC/540, September 1997 (Corrected).

[4] J. Lepingwell, M. Nicholas, V. Braguine, "Strengthening safeguards through open source information collection and analysis," INMM Annual Meeting, Phoenix, Arizona, July 16-17, 2003.

planning of complementary access and inspection activities and subsequent follow-up.[5] Since the early 2000s, open source information has also been utilized in analysis of transnational nuclear procurement networks.[6] Finally, satellite imagery analysis often utilizes open source data, as background and to support the corroboration of imagery assessments.

In addition to contributing to State Evaluation and related analytical tasks, open source information is collected for monitoring and "early-warning" purposes. Thousands of information items from a wide variety of sources are filtered through refined search procedures, and hundreds subsequently reviewed by analysts on a daily basis, to monitor new and ongoing nuclear proliferation and safeguards issues. Selected items of highest immediate relevance are widely disseminated within the Department of Safeguards, via an e-mail newsletter, the "SGIM Open Source Daily Highlights." In addition, items of more particular interest are also provided directly to relevant staff in the Safeguards Department via e-mail, according to their specific work requirements. The daily review also feeds information items into the Open Source Information System (OSIS), the Department's main repository of open source information. This is accessible by all Safeguards staff though a free-text search engine that leverages a nuclear fuel cycle "topic tree" organizational structure.

Uses of Open Source Information

Through the State Evaluation process the agency is able to draw credible conclusions about whether a state is fulfilling its safeguards obligations. In essence, the agency's safeguards mission aims to provide assurances that there has been no diversion of declared nuclear material or misuse of declared nuclear facilities, and—where the additional protocol is in force for a state—that there are no undeclared nuclear material and activities. Keeping in mind that "absence of evidence" does not necessarily constitute the "evidence of absence," demonstration and assessment of such a "negative" presents considerable information-analytic challenges. State Evaluation requires the systematic integration, analysis, and cross-corroboration of various types of information: inspection and complementary access reports, nuclear material accounting reports and additional protocol declarations by states, environmental sampling results, satellite imagery analyses, open source information and analysis and, to a limited extent, voluntarily provided third-party information that is subject to independent agency corroboration. While a State Evaluation focuses on activity and capability assessments at each stage of the nuclear fuel cycle, some non-technical parameters, including political, economic, security, legal, and administrative factors are taken into account as background context for the evaluation process.

---

[5] M. Barletta, V. Braguine, F. Claude, E. Gyane, W. Hammond, J. Lepingwell, M. Nicholas, J. Oakberg, R. Wallace, "Information analysis for Additional Protocol evaluation," INMM Annual Meeting, Orlando, Florida, July 18-22, 2004.

[6] M. Tarvainen, M. Derrough, M. Barletta, "Analysis of proliferation networks," (IAEA-CN-148/96), Proceedings of an International Safeguards Symposium, Vienna, October 16-20, 2006 (Vienna: International Atomic Energy Agency, 2007).

Well-researched and reliable open source information can provide a valuable basis for comparison with other sources of information in determining the nuclear-industrial capabilities of a State, evaluating the internal consistency of declared nuclear fuel cycle programs, identifying possible undeclared activities and aiding in understanding historical issues, especially activities conducted prior to the implementation of IAEA safeguards. Closely related are reviews to evaluate the consistency and completeness of additional protocol declarations, some elements of which largely depend on the availability of the following relevant open source information: holdings and locations of source material; fuel cycle-related research not involving nuclear material; the manufacture of specified equipment; imports and exports of specified equipment and materials; and future plans for nuclear infrastructure development.

Focused open source collections and information-analytic reports are prepared to address specific issues in a state or on a safeguards topic. Transnational procurement networks and illicit trafficking of nuclear material may involve activities that cross the borders of several states. For other safeguards evaluation questions, in depth research in the scientific and technical literature, often from non-English-language sources, and close collaboration with technical experts are needed. The extent and complexity of such research and analysis depend on the history and sophistication of the nuclear infrastructure of a state; past and current fuel cycle activities; in some cases to the existence of past nuclear weapon-related activities prior to IAEA safeguards implementation; as well as to safeguards compliance and state transparency. Other focused collections and information analyses are used to prepare for complementary access and inspection activities, and to evaluate findings from in-field verification activities.

Challenges and Responses

With the development of the State Evaluation process and the additional protocol, the extent of information relevant to safeguards and the importance of its role have grown dramatically. The agency has evolved from its traditional role as an auditor of declared nuclear material to an active collector and evaluator of all information available to it. Since the 1990s, revelations of undeclared nuclear activities in several states under agency safeguards, as well as advances in safeguards information-management technologies, are supporting the development of information-driven safeguards as the future of modern nuclear verification.

After 15 years of continual development, the agency now possesses robust open source capabilities, both in acquisition and analysis. However, as outlined above, several challenges still remain.[7] The growing importance of nuclear nonproliferation issues has resulted in a vast increase in the volume of open source information, the number of sources, and the languages and formats in which this information appears. All of which may contain biases and misinformation that must be taken into account in processing and analyzing information for safeguards purposes. The problems are both exacerbated and ameliorated by developments in

---

[7] J. Baute, "Safeguards information challenges," IAEA-CN-148/25, *Proceedings of an International Safeguards Symposium*, Vienna, October 16-20, 2006 (Vienna: IAEA, 2007); J. Baute, "Information management for nuclear verification: how to make it work in a sustainable manner," *Journal of Nuclear Materials Management*, 35 (Summer 2007), pp. 115-23.

the information technology domain, with regard to web-based searching and extraction, and processing and archiving of growing volumes of information. At the same time, internal agency requirements for open source research have been steadily increasing in order to support over 100 State Evaluations currently conducted per year, and for other safeguards needs.

Broader and deeper information retrieval relies not only on the identification of new sources, for example, by developing access to information in non-official agency languages, but also on the recognition of current limitations in available tools. For instance, only a relatively small fraction of information is available through common search engines. Consequently, the agency has assured access to a number of subscription databases, mainly for scientific and technical information. Specialized software has also been acquired to locate and collect information that has not been indexed.

SGIM has pursued diversification of information sources by accessing multilingual open source collections and databases, employing linguistic skills of highly trained collectors and analysts, and utilizing regional open source collection centers set up with the help of Member State Support Programs, among other measures. In addition, while human resources constrain the volume of articles that can currently be translated in-house, advanced machine translation tools are under evaluation.

The establishment of an integrated, service-oriented, analysis-friendly information architecture, as an effective foundation to collect, evaluate, analyze, structure, secure and disseminate safeguards information is an important element of enhancing information management capabilities. The increasing volume of information available makes it essential to continue increasing the efficiency of the collection and distribution process.

Finally, knowledge management has become the overarching objective for information management, if not for the overall management of verification resources. How do analysts manage the information they collect and the knowledge they generate? How can information and knowledge remain easily retrievable? And how is continuity of knowledge propagated? Time is a factor that poses a critical challenge to information and knowledge management. Not only do analysts need to know what they should know today, but there will normally be no credible analysis of proliferation issues without ongoing assessment of the evolution of the "status" of a state over time. An issue that is considered irrelevant or resolved today (given the information available this time), may subsequently become a major oversight if new information becomes available that is not properly assessed against previous information. Maintaining historical memory over time—including maintaining such evaluations in a structured way that enables dynamic reassessment—adds further complexities to the management of information. In the long term, innovative solutions that provide a reliable basis for safeguards knowledge management will be necessary to guarantee that safeguards conclusions remain credible.