NNSA KNOWLEDGE BASE INTEGRATION, VALIDATION, AND DELIVERY OF OPERATIONALLY USEFUL INFORMATION PRODUCTS

Dorthe B. Carr[1], David P. Gallegos[1], Preston B. Herrington[1], D. Craig Pearson[2],
Steven R.Taylor[2], Aaron A. Velasco[2], John J. Zucca[3], David B.Harris[3], Stanley D. Ruppert, [3]
and Leslie A. Casey[4]

Sandia National Laboratories[1]
Los Alamos National Laboratory[2]
Lawrence Livermore National Laboratory[3]
National Nuclear Security Administration[4]

ABSTRACT

The National Nuclear Security Administration (NNSA) Ground-Based Nuclear Explosion Monitoring Research & Engineering (GNEM R&E) program provides research and development in support of U.S. nuclear explosion monitoring. Specifically, the GNEM R&E program provides products that will be utilized by the Air Force Technical Applications Center (AFTAC) in their nuclear explosion monitoring and treaty monitoring missions. To serve this role, NNSA must develop, track, and integrate products as they move from the research level to operations and must ensure that these products are reliable and technically correct.

NNSA's responsibilities include:
1. Development of operationally useful products (scientific data and the tools to manipulate them) that support the U.S. nuclear explosion monitoring mission at AFTAC,
2. Integration of technical products developed by NNSA and others into a comprehensive and cohesive package (Knowledge Base) that can be readily utilized by AFTAC,
3. Technical assessment, validation, and verification of all products prior to delivery to AFTAC, and
4. Delivery and integration support of the operational Knowledge Base to AFTAC.

The practical implementation of the development, integration, assessment and delivery of operational products occurs through the process of creating periodic Knowledge Base releases. Specifically, the Knowledge Base is comprised of a set of Information Products, which generally include pertinent datasets and analytical tools. In addition, each Information Product also includes critical supporting information (metadata) about the datasets and tools.

This paper describes implementation of the Knowledge Base development and integration process, including the utilization of technical working groups to coordinate and evaluate Information Products.

KEY WORDS: Knowledge Base, integration, validation, information products

OBJECTIVE

The National Nuclear Security Administration (NNSA) Ground-Based Nuclear Explosion Monitoring Research & Engineering (GNEM R&E) Program is continuing the process of compiling Information Products (IP) into a cohesive and comprehensive Knowledge Base (KB) for use by the Air Force Technical Applications Center (AFTAC) in fulfilling their nuclear explosion and treaty monitoring missions. A technically sound and traceable process has been defined to ensure the integrity and utility of research results as they migrate into information products that will be used in the Knowledge Base. A process for validating, verifying, and

logically integrating the information products is required to ensure technically reliable and useful Knowledge Base deliveries. This paper provides a general overview of the Knowledge Base evaluation and integration process to facilitate effective participation in the process by all contributors of research products. In doing so, it is intended that this paper also provide transparency into the overall process.

IMPROVEMENTS TO THE INTEGRATION PROCESS

A process for migrating research results into information products for use in the NNSA Knowledge Base has been developed to ensure that the delivery of a NNSA Knowledge Base contains technically accurate and relevant information products that are in a form that is useful for operational monitoring. This process has been implemented through one Knowledge Base release.

If any process is to retain its effectiveness, it must be continuously assessed and improved wherever possible. During the initial implementation of the Knowledge Base Integration Process, a rigorous evaluation was conducted. Two improvements were identified and have been added:
1. The use of Knowledge Base working groups (WGs) to better address validation of data sets and information products and
2. More explicit descriptions of how the integration process works for software products and algorithms.

This paper describes the integration process (Moore *et al.*, 2000), with the refinements included.

The organizational structure of the Knowledge Base is shown in Figure 1. The fundamental building blocks of the Knowledge Base are referred to as *data sets*. A data set is a collection of data usually all of the same kind (e.g., event data, contextual data, parametric grid data). *Metadata* accompany each data set and describe the content, source, quality and condition of the data set. At the next level up, a combined grouping of data sets makes up an *Information Product*, which is organized around a defined logical theme. An information product comprises:
- One or more conceptually related data sets and/or tools
- Metadata corresponding to each data set
- An information product release document, which provides a detailed description of the information product
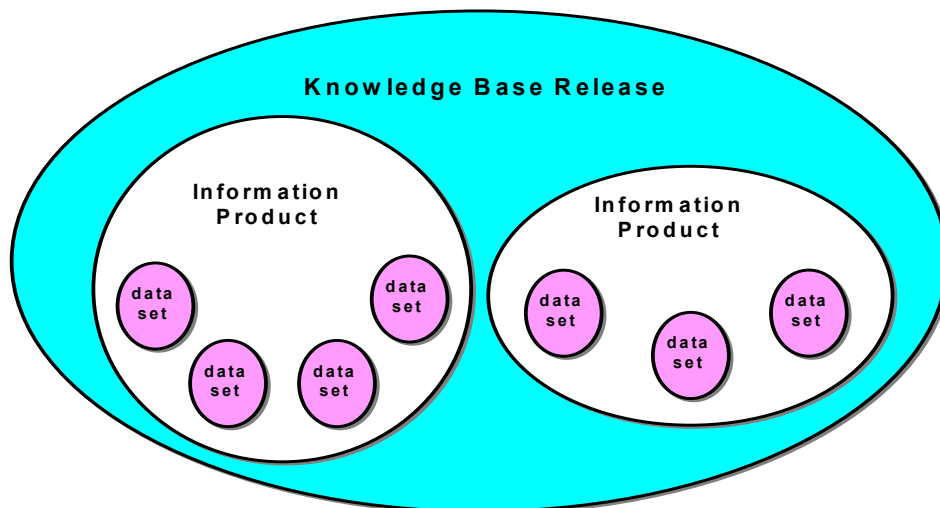Finally, the Knowledge Base is composed of several Information Products.

Figure 1 - Knowledge Base Components

In addition to data sets, the Knowledge Base also contains a number of viewing, analysis, manipulation, and population tools, to facilitate AFTAC's evaluation and use of the Knowledge Base contents. The tools would go through a similar evaluation process as the datasets. In fact, the evaluation of some of the tools (e.g., Knowledge Base population tools) is closely linked to the evaluation and acceptance of many of the datasets.

The participants in the Knowledge Base integration process are shown in Figure 2. The flow of information from Contributors to Scientific Integrators to Knowledge Base (KB) Integrator corresponds to integration of products shown in Figure 1. That is, principal investigators develop products at the level of the data sets depicted in Figure 1. Several principal investigators may contribute data sets to a Product Integrator, who in turn compiles or integrates these into an information product. The theme or concept of a given Information Product is defined beforehand, based on requirements, to ensure that the appropriate data sets go to the correct Product Integrator. The Product Integrator delivers the Information Product to a Knowledge Base Coordinator, who is responsible for verifying that the Information Product is complete and that it has been technically validated prior to integration into the Knowledge Base. The Knowledge Base Coordinators work with the Knowledge Base Integrator to combine several Information Products into a single Knowledge Base. The integrated Knowledge Base is then evaluated and validated by the Integration Board (IB) / Vetting Panel prior to being delivered to the Operational Users.
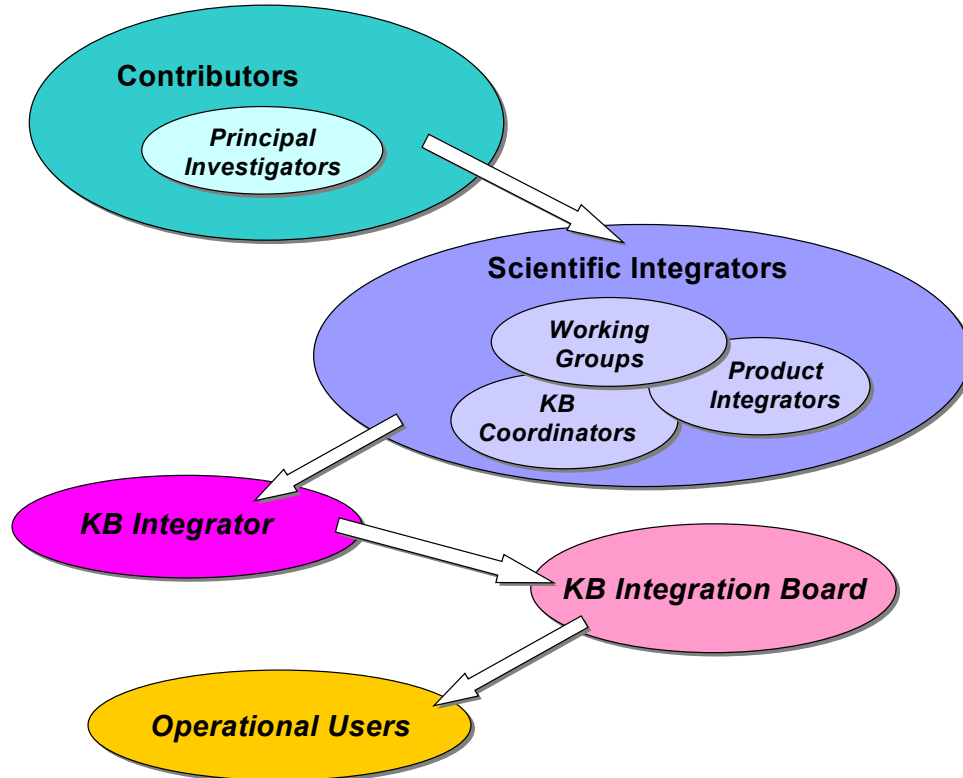

Figure 2 - Knowledge Base Integration Process Participants

The integration and evaluation process is shown in greater detail in Figure 3 and is derived from Moore *et al*. (2000). The single most significant change to the integration process since the initial implementation has been the addition of technical working groups.
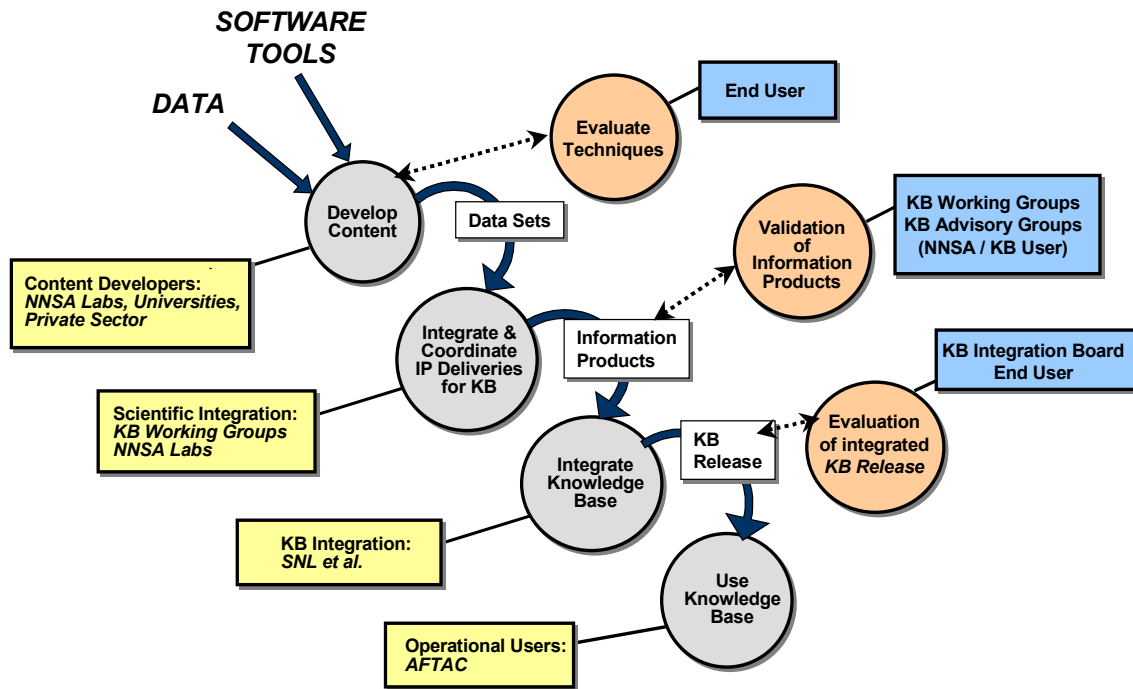
Figure 3 - Knowledge Base Evaluation and Integration Process

In Figure 3, the center series of elements, beginning with "DATA" and "SOFTWARE TOOLS," depicts how data and algorithms are integrated into data sets, then into Information Products, and finally, into a complete Knowledge Base Release. The boxes to the left of this integration series are the specific players who conduct each of the steps. The circular elements to the right of the integration series represent the technical evaluation and subsequent validation that occurs at each step in the process. The boxes to the far right are the specific players that conduct these evaluations. A key characteristic of this process is that the end user (AFTAC) is involved throughout the integration and evaluation process, from the validation of techniques for content development through the evaluation of the integrated Knowledge Base release. This is important because the end user is critical in the early-on determination of relevancy and usefulness of any given data set or Information Product, as well as evaluation of technical accuracy. Another key characteristic of this process is that products (i.e., data sets, Information Products, Knowledge Base) are verified and validated at each step in the process.

The addition of technical working groups has provided not only coordination early on and throughout the process, but has provided an additional component of technical assessment throughout the process. The working groups comprise NNSA technical personnel and representatives from AFTAC. For a given Knowledge Base Release, these working groups will serve the primary functions of verification, validation, and coordination associated with a specific Information Product, which may become part of a Knowledge Base Release. Generally, the working groups would interact with researchers to provide technical guidance and feedback for the research products that form the basis for Information Products. To satisfy this function, the working groups specifically have the following tasks:

- Identify outstanding technical issues, assign some level of priority to those issues, and develop strategies and technically sound methodologies for addressing these issues
- Validate the use of a specific methodology for developing a given Information Product
- Coordinate schedule, formats, and Information Product structure (including metadata)
- Develop and follow a schedule of critical meetings
- Validate site-specific models and the data created through implementation of those models
- Assess and report on the quality of any empirical data to be delivered or used
- Assess and report on the operational usefulness of tools and datasets
- Identify support for researchers, and
- Assist with metadata preparation.

The validation and coordination support, which is provided by the working groups, positions each resulting Information Product for review by the appropriate Knowledge Base Advisory Group. The Knowledge Base Advisory Groups can then verify that the Information Products are technically accurate, cohesive and complete.

Following review of the individual Information Products by the Knowledge Base Advisory Groups, the aggregate set of Information Products is integrated into a proposed Knowledge Base Release and submitted to Knowledge Base Integration Board for review and approval. Once approved, the Knowledge Base is delivered to the end user (AFTAC).

CONCLUSIONS AND RECOMMENDATIONS

The process of validating, verifying, and managing the information products is critical to the successful integration of scientific research to support operational monitoring systems at AFTAC. This process has been modified to provide coordination, integration and validation of products from the onset through the employment of working groups to identify technical issues and to develop methodologies for addressing these issues. The two significant improvements this year are
1. The addition of Knowledge Base Working Groups to better address validation of data sets and information products and
2. More explicit descriptions of how the integration process works for software products and algorithms.

This integration process results in coordinated products from researchers at different organizations. It also involves the researchers at NNSA and the customer at AFTAC in the integration process as the working groups take an active role in defining, creating and validating information products.

Continued success of the integration process relies on the following:
- All contributors, whether DOE-funded or not, are knowledgeable about the process.
- The process continues to be evaluated and refined as new requirements emerge and as various types of research evolve.
- Feedback is received from the end-users concerning the usefulness of the format and functionality.
- Feedback is received from the end-users regarding the validity of the content and functionality.

REFERENCES

Moore, S., H. Armstrong, D. Carr, R. Keyser, E. Shepherd, L. Wilkening, C. Young, M. McCornack, J. Aguilar-Chang, A. Velasco, S. Ruppert, T. Hauk, and C. Schultz (2000), The Integration Process Design for Incorporating *Information Products* into the Department of Energy *Knowledge Base*, Sandia National Laboratories, Report No. SAND2000-0597, Albuquerque, NM (also available from http://www.nemre.nn.doe.gov/shared/papers/KnowledgeBase/sand.pdf).