Open Source Activities

Foreign Aerospace Science & Technology Center

Mr. Thomas R. Pedtke, Dr James C. Canfield and Lt Col Jack O. Sawdy

STIISP: An Introduction

Among the many missions performed by the Foreign Aerospace Science and Technology Center (FASTC), one of "common concern" within DOD intelligence is the Scientific and Technical Intelligence Information Services Program (STIISP). As the largest analytic element of the S&T intelligence community, FASTC is also the intelligence community's largest processor and user of open source S&TI information (unclassified sources of S&TI information). The Defense Intelligence Agency has designated FASTC as the Executive Agent for STIISP for the exploitation of vast amounts of S&TI information available to the intelligence community.

A feature of S&TI analysis is the requirement for information covering long spans of time, especially in examining threats involving technologies. Much of what is developed through research in foreign countries is published in open literature. One of the characteristics of many research efforts is that the people doing the research do not know ahead of time what the application will be for the area being investigated. As a result, numerous research articles are published annually in open literature which eventually contribute to modifications or lead to acquisitions of major weapons systems. Therefore, they publish articles of S&TI interests that ultimately have unique intelligence value. Another factor is the relative importance and prestige placed on being published in world scientific and engineering circles. Unclassified, open source literature provides a vehicle for foreign countries to be recognized for their scientific and technical sophistication.

The importance of published open source S&TI information was recognized at FASTC many years ago - in about 1951. The Air Technical Intelligence Center (one of the earlier incarnations of FASTC) began the process of acquiring and indexing foreign open source publications into what is now referred to as the Foreign Science Library. At the same time, work was started on extracting information to make the foreign literature easier to use by intelligence analysts. Since much of the information was in a foreign language it required development of a translation capability to provide the extracted information to the analysts in English. When computers began to appear, the Foreign Technology Division (the most recent name for FASTC) embraced the technology as a means to get the information to analysts faster. FASTC was also one of the first to make computerized information available to remote users with on-line services. In fact, the first year of an experiment in the 1960's was so successful that the participating organizations continued to use the experimental on-line system in support of intelligence production.

The participants included Army and Navy as well as Air Force agencies. Because of the success of this early experiment the Defense Intelligence Agency

decided to establish a program named the Scientific and Technical Intelligence Information Services Program (STIISP). FASTC (still FTD at that time) was designated the Executive Agent for the program. The program continues to acquire foreign publications for libraries on behalf of the DOD S&TI centers and process the information for inclusion into the data base, the Central Information Reference and Control (CIRC) data base system. STIISP is actually an all-source information services program. Each of the participating organizations provides input of the classified information they need to do their particular mission. FASTC provides all the open source (unclassified) part of the program with the support of two contractors who operate the Foreign Science Library and index the open source S&TI information. FASTC also operates and maintains the integrated, online system. Both unclassified and classified information are hosted on the CIRC system at appropriate security levels.

The CIRC data bases are the best known part of STIISP. The information in CIRC is available to anyone in the government who has a need for its content. As one of the largest data bases ever assembled in the DoD, the 40 plus years that FASTC has been acquiring S&TI literature has resulted in the accumulation of a significant volume of open source literature holdings.

From the beginning recognition of the importance of foreign published literature, FASTC has been a leader in processing open source information for the intelligence community in support of scientific and technical intelligence missions. Because of the foundation developed within FASTC to process and deliver open source information, the FASTC mission continues to evolve and expand within the intelligence community efforts to do more with open source information across the board.

CIRC: Central Information Reference and Control

CIRC is the data base component of the DoD S&T Intelligence Information Support Program (STIISP). The program is defined by DIAM 75-1, managed by DIA/DTI-3C and operated by FASTC as Executive Agent for DIA.

Since 1969, CIRC has been the primary means of storage and retrieval of S&T intelligence information content and references to documents required by the five DoD S&T intelligence analysis centers, the US Army's Foreign Science and Technology Center (FSTC), the AF Foreign Aerospace Science and Technology Center (FASTC), the USN Naval Maritime Intelligence Center (NAVMIC), and DIA's Armed Forces Medical Intelligence Center (AFMIC) and Missile and Space Intelligence Center (MSIC). The data base was adapted from a data base initiated in 1962 to serve USAF needs. CIRC is made available to any organization within security guidelines which can demonstrate to DIA/DTI-3C a need for access to its content.

As the STIISP enters the end of 1992, the CIRC data base contained information from and references to over 10 million documents relevant to DoD S&T intelligence analysis needs. Data base emphasis has been on priority geographic and subject areas; accordingly, the data base is dominated by information about the former USSR, with secondary emphasis on Eastern European countries and the Peoples Republic of China. Virtually all science and technology subject areas

which have military significance are addressed. Worldwide coverage is provided for military equipment and civil equipment with military implications. With the demise of the former Soviet Union, the content of the new CIRC inputs are being adjusted to reflect the new world situation and national priorities being levied on the intelligence community for coverage of the rest of the world. About 85% of the data base is derived from open source publications; the remaining 15% is derived from classified information.

Documents used as source material include foreign books, about 1400 foreign periodicals (including S&T journals, newspapers, patent announcements and other announcement publications), and the entire range of intelligence publications: Intelligence Information Reports, non-recurring finished intelligence and finished intelligence from printed and electrical reporting throughout the intelligence community. All intelligence documents input to CIRC are reproduced on microfiche and distributed without regard to subject to organizations approved by DIA/DTI-3C. Open source material is retained in hard copy by a contractor operating the Foreign Science Library. Growth of the data base is about 200,000 references per year. Data base updates are done weekly.

CIRC is unique among government and commercial data bases in the depth to which S&T intelligence information is handled. Heavy emphasis is placed on personalities, facilities, locations and nomenclature (of equipment). This "PFLN" information provides a robust indexing to the data base unlike any other. These data are extracted from original language documents and recorded along with attributes and relationships among them. Each CIRC entry includes subject content of the source document derived from the title of the document (translated into English if needed) and from the text of the document: Extracts or abstracts from open source documents (translated into English if needed) and extracts or full text of intelligence documents; no keyword assignments are made.

Two customer services are provided from the data base:

1. Distribution of new data entering the data base is done with each update to users who are approved for participation using CIRC profiles. Profiles are standing requests for information that are compared against new data base input to identify references relevant to user needs. Profiles may be as narrow or as broad as desired to serve very specific individual needs or to serve relatively broad institutional needs. The retrieval language used in profiles can address virtually any element of information within data base entries. Organizations approved for participation are generally expected to provide resources for construction and maintenance of their user profiles. (FASTC provides training in profile construction.)

2. Retrospective retrieval through on-line searches of the entire data base (segmented to improve response time) also permits identification of data base content based on virtually any element of information within data base entries. The data base management and retrieval system is interactive, permitting users to browse references on-line for relevance judgements, and to refine retrievals until the desired precision is achieved. In addition to the on-line browsing, off-line output may be requested which is normally distributed to the requester within one working day. An extensive network of secure and unclassified terminals exists (currently limited to the contiguous continental US) that are dedicated to CIRC data base use. Dial-up ports are provided over unsecured commercial telephone lines which permit access to the unclassified portion of the

data base. (FASTC provides training for terminal retrieval operations.)

Contractors are provided access to CIRC through their government contract monitors. As a rule, DIA permits contractor access only via unclassified communications. This limits access to unclassified content of the data base, but permits retrieval against indexes of CIRC content through the secret level. Output of classified content is provided to contractors through their government contract monitors for "need-to-know" determination.

MT: Machine Translations

FASTC, in conjunction with the Latsec Corporation, has been a pioneer in machine translation, particularly Russian to English. The advantages of machine translation are cost, timeliness and consistency. This latter point is particularly advantageous in science and technology. The computer always translates the same way!

Currently, a quality machine translation product is available for Russian, German, French and Spanish. FASTC, with the cooperation and financial assistance of FBIS, has contracted Latsec to develop a machine translation capability for Japanese. The Russian machine translations have been operational since 1970.

Analysts can either submit their requests to FASTC's translation services organization or avail themselves of online, interactive capabilities. Machine translation input is still perform by manual keystroking which simultaneously transliterates. For Russian, four additional keys are identified for the additional parts of the Cyrillic alphabet. In the case of French, special codes are used for the accents. German umlauts are input as two characters. Output from the translations services organization can be provided as raw machine output or with post editing (to include insertion of graphics and tables). Final products are produced in Word Perfect 5.1. The online, interactive system is available through 1400 terminals connected to the FASTC IBM mainframes or through dial-up modems for remote users with proper access to FASTC's ISS computer services. The typical response time is 5-15 seconds.

Near-Term Enhancements

An FY93 "NSR-29" GDIP Initiatives and a companion FY94 GDIP Initiative will provide funding to ensure the survival of the STIISP program and its CIRC data base (in financial jeopardy in recent years) and provide very austere enhancements to the program.

Specifically, the STIISP program will continue to be able to migrate the emphasis of the new subscriptions and indexed inputs to CIRC for the entire world rather than the former Soviet Union and the Eastern European block. It will also afford the acquisition or access to commercial data bases for the STIISP agencies. STIISP envisions a hybrid system relying on the deeply indexed CIRC data base for the "cream" of the world-wide S&T published literature and the commercial data bases purchase or access for the broader spectrum of S&T literature.

A significant and long overdue enhancement will be the automation of the STIISP agency libraries. The system selected is the Scientific & Technical Information Library Automation System (STILAS) developed by SIRSI, Inc. originally for DTIC. This system will facilitate the development of catalog of open source holdings of all the STIISP agency libraries and the formation of a union catalog to facilitate community sharing. The STILAS also features an "intelligent" gateway module that facilitates using a single query language to simultaneously research multiple commercial and government data bases.

Machine translations will also be mildly enhanced in the near-term program. Xerox Corporation has developed a system to integrate scanners, image optimizers, and cyrillic and Roman OCRs with the Systran (Latsec) machine translation software and dictionaries and the Xerox Globalview desktop publishing software. Xerox has entitled the system DOCUTRAN and the near-term program also enables FASTC to acquire this machine translation automation for the STIISP agencies.

Further development of the Latsec dictionaries for enhanced machine translation, migration to C++ open system architecture and incorporation of the Systran dictionaries into a PS\2 environment are also supported.

Long-Term Enhancements

Current dialogue with the Defense Intelligence Agency and the GDIP staff include suggested programs for the migration of the CIRC data base from an IBM mainframe environment to a client-server environment. Included is the evolution of the CIRC data base to an object oriented, multimedia system with an inherent user friendly environment. Very long-term goals are to make available all information in full text with native natural language processing and on-demand machine translation. As an interim step, FASTC will seize the opportunities to incorporate full text where cost effective to do so (e.g. information is already available digitally in full text). Affecting this very long-term program are interpretations of copyright laws and negotiations surrounding fair use policies.

Given the near-term improvements, particularly the library automation; the reality of the very formidable CIRC data base and the possibility of the CIRC object oriented, multimedia data base, the components are in place for the development of a prototype Functional Support Center for the S&T intelligence community (also termed Primary Intermediary or Main Storefront). FASTC has been identified in the open source community architectural strategy to prototype such a functional support center. This is also reflected in the DCI Open Source Coordinator's Strategic Plan for Open Source. The plan is currently being reviewed by the highest levels of the intelligence community and will then have resources attached to it to fund these services of "common concern" that are embraced by organizations such as FASTC.

FASTC points of contact for CIRC:

General information and policy:

HQ AF Foreign Aerospace Science and Technology Center
ATTN: DXLR (Ms Joanne West)
Wright-Patterson AFB, OH 45433-6508

Phone: 513-257-2240 (Commercial)
DSN 787-2240

Access and communications information:

HQ AF Foreign Aerospace Science and Technology Center
ATTN: SCDD (Mr D. E. Leasure)
Wright-Patterson AFB, OH 45433-6508
Phone: 513-257-2242 (Commercial)
DSN 787-2242

FASTC point of contact for MT:

General information:

Hq AF Foreign Aerospace Science & Technology Center
ATTN: DXLL (Ms Mary Folck)
Wright-Patterson AFB, OH 45433-6508
Phone: 513-257-2989 (Commercial)
DSN 787-2989

DIA point of contact for STIISP is:

Defense Intelligence Agency
ATTN: DTI-3C (Ms. Alice Cranor)
Washington, DC 20340-6173
Phone: 202-373-4688 (Commercial)
DSN 243-4688

FIRST INTERNATIONAL SYMPOSIUM: NATIONAL SECURITY & NATIONAL COMPETITIVENESS: OPEN SOURCE SOLUTIONS Proceedings, 1992 Volume II - Link Page

Previous	OSS '92 J. Niles Riddel, Foreign Broadcast Information Service (FBIS) Open Source Intelligence Requirements & Capabilities,
Next	OSS '92 Mr. William Ruh, MITRE, Open Source Processing Research Initiative,

Return to Electronic Index Page