Information Support to Intelligence Production: The Reality and the Dream (U)

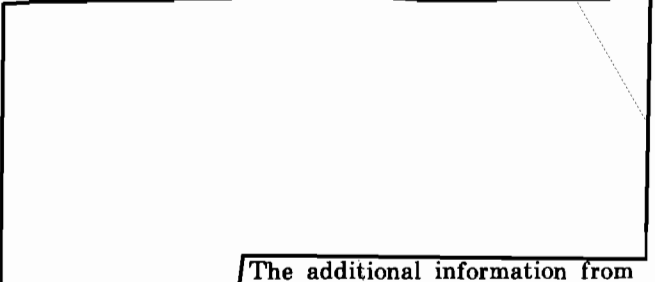*The Role of Information Support (U)*

(U)   Information is the nutrient that sustains the intelligence analysis-and-reporting process. The business of providing information support to the analysis-and-reporting effort plays a significant role in intelligence production and takes a moderate slice of the intelligence budget. But is it enough? Over three quarters of the Intelligence Community's monetary resources are expended to collect, process, and distribute increasing amounts of raw, unevaluated information. By contrast, comparatively few Community dollars are dedicated to the organization, storage and retrieval of this unevaluated information and the partially and fully evaluated intelligence developed from it. This is true even though collection and processing efforts are wasted if the information gathered thereby is not readily available to analysts for their use in intelligence production.

(U)   As well as laboring under budgetary constraints, information support activities throughout the Intelligence Community suffer serious deficiencies because of fragmentation, duplicative efforts, and poor visibility. These are strong statements. The remainder of this article is devoted to explaining why information support is so important to the intelligence production process, why the present level of information support in the Community is inadequate, and what can be done to raise the level and improve the quality of information support in the future. First, let us look at the situation from the perspective of NSA.

Information Support to Cryptologic Production (U)

(S)   It is virtually impossible to produce good intelligence in an informational and experiential vacuum. In a recent book review, Dr. Louis Tordella commented on World War II experiences in cryptology and observed that "a decrypted operational order is often not very illuminating unless read in full context.

Relevant collateral information, operational background, and a thorough understanding of the scene as viewed by the German commanders were always necessary to get more than superficialities from most decrypts."[1] In the summer of 1962, nearly twenty years after the experiences cited by Dr. Tordella,

The additional information from other sources made all the difference in determining the deeper significance of the

Providing these "other" sources to the analyst is the ultimate purpose of information support.

(S)   It is worthwhile to digress here and touch upon a subject that has been the source of considerable misunderstanding and contention within the Intelligence Community. The subject relates to the proper place of intelligence analysis within NSA, which in turn bears upon NSA analysts' need for information support. There are extreme views taken on both sides of the question. Briefly stated, the problem springs from the fact that NSA's principal intelligence mission is to produce "unevaluated" data derived from the intercept of foreign signals; yet NSA must perform intelligence analysis in order to produce this "unevaluated" data. Ergo, the data has been evaluated.

---

[1](U)   Louis W. Tordella, "Review of Ultra Goes to War and Deadly Magic." *Studies in Intelligence*, 23, no. 3 (1979): 55-58.

(U)   Years ago an NSA analyst described something of the difficulty in a colorful little tale entitled "King Eusyb and Queen Deodi."[2] As the analyst described it, NSA (Durensi in the story) must intercept and sift through tons of garbage in order to find the few cocoa beans needed to produce the chocolate cake so dear to King Eusyb's heart. Only NSA has the trained sniffers, sifters, sorters, slicers, ricers and dicers needed to process these mountains of garbage; also, NSA has developed methods of keeping tabs on those garbage shipments most likely to produce choice cocoa beans. Although it wasn't the point of the analyst's story, it naturally follows that if NSA ceased its sifting and analysis and simply started shipping all intercepted garbage to CIA, DIA, et al. (which some Community analysts would apparently like us to do), a great stench would soon arise from the south bank of the Potomac.

(C)   Neither the story of King Eusyb and Queen Deodi nor all of NSA's powers of persuasion have sufficed to convince some Community analysts that NSA is not sitting on a veritable mountain of unexploited information which it covers with hands-off labels such as "technical data" and "raw traffic." Of course NSA misses some cocoa beans, but there is no deliberate intent to conceal. To preserve its sister agencies from inundation, and despite the risk of continued misunderstanding, NSA must persist in sniffing, sifting, sorting, slicing, ricing and dicing to the best of its ability and to the limit of its resources. To do this job properly the NSA analyst will continue to need information support.

(C)   Providing information support at NSA is a multifaceted operation. The Information Services organization is the most visible information support element, but every major organization within the Agency has individuals or even entire units whose principal business could properly be labeled information support. These information support activities at NSA can be generally categorized as follows:

1. Establishing requirements — determining the specific information needs of various NSA/CSS elements.

2. Procuring information — getting the books, periodicals, newspapers, intelligence products and other materials that are likely to contain the required information.

3. Distributing and disseminating information — passing the information as directly and expeditiously as possible to those who need it.

4. Organizing information — scanning, indexing, cataloging, storing, and otherwise arranging or making available information from external sources (books, periodicals, Department of State cables, intelligence reports, press items, commercial services, organized files in other intelligence agencies, etc.), as well as from materials originated by NSA/CSS and other cryptologic producers, so the information can be retrieved and used as needed to support the NSA/CSS mission.

5. Researching information — providing information-retrieval and research services for those who require it, preparing studies and reference aids for intelligence analysts, etc.

(C)   Requirements are determined by a periodic review of information needs by Information Services personnel working in cooperation with various agency elements, by the publications procurement program, and by continuing contacts with managers and analysts at every level. Based on these requirements, an active and successful procurement effort is sustained by arrangements with U.S. intelligence agencies and other government organizations by purchases from private book vendors, and so on. During the calendar year 1978, for example, NSA procured more than 20,000 books, subscribed to some 4,500 newspapers and periodicals, and received more than 300,000 unique documents from other members of the Intelligence Community. These open-source materials and collateral intelligence documents are reviewed by NSA's Information Services organization and, if pertinent, distributed to those individuals who have expressed a need for the type of information contained therein. Computer programs called HIGHMORE and CLIPPERTON

        Multiple copies are usually distributed, and sometimes the same copy is routed to more than one organization. Control and accounting procedures are frequently necessary in order to keep track of the more sensitive materials.

(C)   Incoming materials are often combined with cryptologic products in one or several of the many files, document collections and libraries that have been organized to support the NSA/CSS mission. Some files are restricted to a single type of document or to a single source; others combine open-source materials, collateral intelligence reports, and Sigint products.

---

[2](U)   Robert L. Lowrey, "King Eusyb and Queen Deodi," *Cryptolog* 1, no. 9, (1974): 19-20.

(b) (1)
(b) (3)-50 USC 403
(b) (3)-18 USC 798
(b) (3)-P.L. 86-36

The files are manual, machine (computer-based), or a mixture of manual file and machine index; and, excepting general collections or large files like the main library, the Sigint repository, and SOLIS, most files are structured to satisfy specialized requirements (e.g., the S Branch Library,

(C)   In addition to organizing and maintaining its own files, NSA makes use of a number of external data files and services. Included among these external sources of information are intelligence collections and indexes such as the U.S. Air Force Foreign Technology Division's CIRC II system, the Defense Technical Information Center (formerly Defense Documentation Center) data base, and the many Intelligence Community files that are accessed via COINS. Also, NSA subscribes to commercial systems and services like the *New York Times* Information Bank and Lockheed's DIALOG system and to a national library network (OCLC). Under project MILL-STREAM, NSA staffs five positions in CIA's Office of Central Reference in return for research services on certain world areas. By this arrangement information support is provided to a number of Agency elements without the necessity of NSA maintaining costly and largely duplicative intelligence files.

(C)   During 1979, workers in NSA's Information Services organization researched more than thirty-four thousand information requests from various NSA/CSS elements. That number does not include requests for specific documents or books nor account for the many self-service uses of local information collections and files by researchers from all Agency components.

(C)   Perhaps the most popular of all information-support operations at NSA is the Sigint On-Line Information System (SOLIS), an on-line, interactive system used to index, store and retrieve Sigint products (reports, translations and summaries published by NSA and other cryptologic organizations). System users have access to the complete text of a Sigint product within 30 minutes of the time it is electrically dispatched or received by NSA, and it remains available for interactive access for 14 months thereafter. Overnight retrievals may be batched against products that are more than 14 months old.

(U)   SOLIS employs a sophisticated automatic indexing process that is flexible enough to permit an editor to change or add to the data base. The various parts of a Sigint product are formatted and delimited in accordance with Sigint reporting instructions. SOLIS uses the delimiters to identify the various parts of a product and then indexes them in

one or more of the inverted indices. Among the parts of a product that are automatically indexed are the subject codes (TAGs), the date-time group, the serial number, the terms in the title, and the nontrivial terms in the text.

(U)   System users may enter any number of terms or retrieval items against one or more of the indices during a SOLIS retrieval session. Entry of a query will result in an immediate search of the indices and the return of a count of the number of documents in which the indicated term or combination of terms (or retrieval items) occur. The user may then modify his query and reenter it, invoke a number of display or retrieval options (including a text scan), or simply print the externals or the full texts of pertinent documents. The entire process is conducted in an interactive mode with normal response times in the range of 4 to 8 seconds.

(U)   In addition to interactive queries, interrogations may be stored and run against SOLIS at a later time, frequently on a periodic basis. Many users take advantage of this capability to run retrievals against newly published Sigint products every night. This process, called DAILYDISPATCH, makes it possible for SOLIS to effectively distribute incoming and outgoing Sigint products throughout NSA.

(S)   During 1978, SOLIS users initiated approximately 80,000 retrieval sessions (some 328,000 individual queries) against the system from terminals located within NSA and in external organizations such as CIA, (b)(1) the Department of State, and the Strategic Air Command. As of March 1980, there were 90 SOLIS terminals located within NSA spaces at Fort George G. Meade and at nearby Friendship Annex. Another 33 terminals were located elsewhere, most of them in intelligence agencies in the Washington, D.C. area but some as far distant as Hawaii and London.

Information Support Elsewhere in the Intelligence Community (U)

(U)   The information support situation at NSA is fairly typical of the Intelligence Community as a whole. Every major organization in the Community has an information system to support its analysis-and-reporting effort. Each system is an amalgamation of different elements — specialized computer-based files, manual files, bibliographic systems, and commercial information services. Although not systematized, there are also the ubiquitous analyst files. In addition, each organization taps the information resources of other organizations to some extent, either through formal communication channels and computer networks (e.g.,

(b) (1)
(b) (3)-50 USC 403
(b) (3)-18 USC 798
(b) (3)-P.L. 86-36

MILLSTREAM and COINS) or through informal analyst-to-analyst contacts and "old boy" networks.

(C) Within CIA the Office of Central Reference (OCR) serves as the centralized information-support organization. OCR maintains extensive manual and computer-based files in support of intelligence analysis and reporting efforts.

(b)(1)

Contrasted to information-support operations at NSA, those at CIA appear more structured, more centralized, and more traditional in their approach to information-handling techniques. Also, CIA's OCR has a more direct role in intelligence production than does NSA's Information Services organization. Many of the differences in approach to information support at CIA as compared to NSA are attributable to the fact that CIA spends more of its resources on research and estimative analysis, whereas NSA concentrates more on current intelligence production; the information-support requirements for the two types of intelligence production are somewhat different.

(U) The largest computer-based file currently maintained by CIA's OCR is AEGIS, a thirteen-year bibliographic index to products of intelligence interest. Ten years of AEGIS is kept on-line in a system known as RECON. RECON/AEGIS was designed primarily to support research and estimative analysis. It is a biliographic index in the traditional sense. Neither the full text nor an abstract of an intelligence document is included in a RECON/AEGIS entry. Instead, each entry contains substantive bits of information about the document indexed — the document identification, an expanded title, generalized intelligence subject codes, keywords based on document content, classification and dissemination controls, and publication date. If an analyst requires the document that is described by an index entry, he must use the document identification to extract the document from another file.

(C) CIA analysts use RECON along with a variety of other computer-based and manual files to retrieve the information needed for intelligence analysis and reporting. Most of the source documents consulted are the same ones that are available to NSA analysts; and, as one might expect, there is a great deal of duplication between the coverage of CIA's OCR and NSA's Information Services — except for those areas where, under the MILLSTREAM arrangement, CIA has agreed to cover for NSA.

(C)

(b)(1)

(b)(1)

(C) There are large information files and analyst support systems elsewhere in the Intelligence Community. To name but a few: the Foreign Technology Division (FTD) of the U.S. Air Force in its CIRC II system maintains a massive scientific and technical collection on foreign aerospace developments; NPIC manages a large and valuable collection of Photint materials; and the Department of State has developed a computer-based system for storage and retrieval of State cables and other documents.

(U) The facts given thus far might lead one to conclude that information support at NSA and throughout the Intelligence Community is well organized and functioning smoothly. That picture, unfortunately, is not complete. As stated above, there are many serious problems.

Some Difficulties and Deficiencies (U)

(C) As we examine the analysts' information environment in all parts of the Intelligence Community, it is apparent that each agency does best in organizing its own products for retrieval. NSA takes care of noncompartmented Sigint products through SOLIS, (b)(1) and CIA indexes more than ninety-five percent of its own products in RECON. At FTD, NPIC and the Department of State, the situation is the same — each takes care of its own products first. It is also apparent that each agency attempts to satisfy the information needs of its own analysts with different approaches, varying levels of sophistication, and with a greater or lesser degree of success.

(U) In this environment the intelligence analyst finds that his information support is severely fragmented. A recently completed survey of Community analysts by the Analyst Support Task Group (working on behalf of the DCI's Intelligence Information Handling Committee) revealed that there were relatively few complaints about a paucity of information. Instead, the most common complaint was that

there is no one place, no single data base or facility, from which an analyst can retrieve all types of information pertaining to his or her subjects of interest.[3] Although it may be too much for the analyst to expect that all types of information should be encompassed in one system, surely we can agree that the Community's present information environment is needlessly complex.

(C)   The situation at NSA is typical. The NSA analyst has access to published Sigint, with the exception of compartmented Sigint products, on SOLIS. Photint is available in the NPIC files that can be accessed from NSA via COINS. Index entries on many CIA and DOD intelligence reports are in CIA and (b)(1) files that can also be accessed via COINS. Then, depending on the NSA analyst's area of interest, there are a large number of in-house specialized files that might be consulted. The NSA analyst is thus faced with a multiplicity of files, file formats, systems, and computer retrieval languages as he attempts to pull together references to particular areas or subjects of interest. Many intelligence sources are unavailable to him in an easily accesible form, and there may be other files and sources of which he is completely unaware.

(C)   The NSA analyst's problem with multiple information sources and the complexities of retrieval is mirrored throughout the Intelligence Community.

(b)(1)
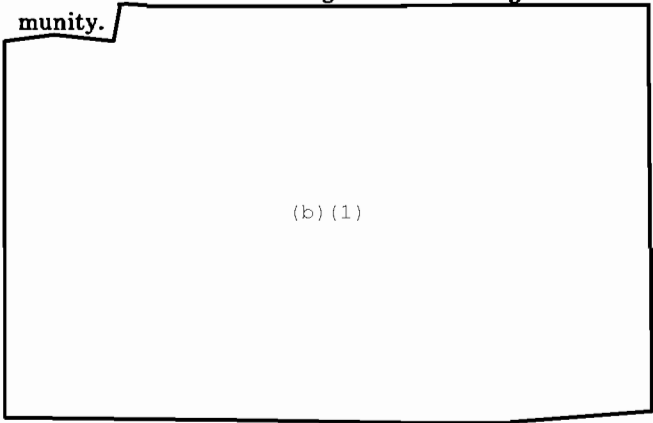
(U)   The fragmentation of information-support activities and files within NSA has accelerated during the past decade or so. Perhaps this has been due in part to a failure of information managers to convince higher management of the importance of information support to the performance of NSA/CSS's mission. There has also been some difficulty in restructuring tradition-oriented information-support organi-

[3](U)   DCI Intelligence Information Handling Committee, Ad Hoc Study Group, *Report of the Analyst Support Task Force*, June 1979.

zations to meet the challenge of the 1960s and 1970s. Whatever the reason, in recent years highly visible support items (such as book purchases and periodical subscriptions) suffered serious attrition during times of even moderate austerity. Indeed, NSA's entire Information Services organization proved extremely vulnerable to the budget slicers, billet hunters and space grabbers. Depletion of the Information Services organization has not, however, saved money or caused functions to cease. In the words of a once-popular song, "The beat goes on." Functions that were formerly centered in the Information Services organization are now performed in many different organizations throughout the Agency. These functions are frequently duplicative and their execution is often less efficient than would be the case if they were centralized; but the support items are masked and relatively invisible to the budget slicers, and the most important tasks get done regardless of the actual costs involved.

To Each His Own (U)

(U)   In the Intelligence Community today, the disjointed, "go-it-alone" approach to information organization, storage and exchange is apparent almost everywhere. The interagency situation is particularly vexing. As one example, for fifteen years the COINS Project Management Office has labored to build communication links that would enable analysts throughout the Community to share data bases; but it has been a struggle every inch of the way. Key managers in some agencies regard COINS with all the warmth and confidence that they would bestow on a registered agent of the KGB, and they are not about to open their files to the uncertain hazards of an interagency network. In other places it is not the security hazards of networking that cause a hesitancy to embrace COINS but a desire to build another network — either to establish control or to satisfy "unique" requirements that COINS is not designed to handle. Nevertheless, despite its problems, COINS (and networking) is one of the bright spots. A greater difficulty is posed by the data bases themselves.

(U)   There are literally hundreds of information files developed to support local analysts in different parts of the Community. Some are manual files; others are computer files capable of being shared with analysts elsewhere in the Community. Of the latter type, there are both interactive and batch files — some well designed and some not. And many different data-base management systems and computer languages are used to organize and interrogate these files. Then there are the different formats, abbreviations, codes, transliteration systems, and so on, that

are apparently designed to obscure the contents of a file to all persons except those who have been properly initiated into the arcane rites of electronic obfuscation.
(U) It would be grand if we could say that the Community has finally learned something after all these years, and that the major information systems and files that are now being built are designed to be shared by intelligence production analysts wherever they might be located. Unfortunately, such is not the case. Even at this moment literally tens of millions of dollars are budgeted for SAFE, a multiphased project for improving the analyst's environment that has been in the planning stage at CIA since the early 1970s. SAFE calls for: (1) automatic distribution of electrically received data on the basis of analysts' profiles, with a retrospective text-search capability against the most recently received data; (2) the capability for analysts to create and search their own personal files; (3) a long-term retrospective search capability by means of a link with RECON (the ten-year bibliographic index) and ADSTAR (a microform-based system for document storage and retrieval; and (4) a capability to assist the intelligence production process by means of text-editing features, reference aids, etc.
(U) SAFE planning is undoubtedly impressive, but one might reasonably ask whether or not SAFE provides for resource-sharing with other members of the Intelligence Community. Unfortunately, the answer at this time must be negative. SAFE involves only CIA and DIA, and it would appear that even the level of sharing between these two agencies will be limited.
(S) As SAFE progresses, other members of the Community seek their own solutions to the problems of information support. Pleased with the success of SOLIS, NSA plans — through Project WEEDER — to apply SOLIS programs and techniques to the electrical materials that it receives from other agencies, thereby giving NSA a capability to automatically distribute and retrospectively retrieve Department of State cables, CIA and (b)(1) intelligence reports, etc. Other efforts to improve the analyst's environment at NSA are going forward internally at the group level. These local systems, which are built around minicomputers, usually include such features as text-editing, storage of reference and reporting aids, a distribution facility, provision for analysts' files, and some degree of retrospective retrieval capability — in other words, the very same capabilities that are being programmed for SAFE. The NSA systems are designed to meet NSA's own needs; concern with Community requirements is secondary. In some instances no attempt has been made to satisfy

the information needs of analysts outside one's immediate office. All over the Community the story is more or less the same.

Toward a Unified Approach to Information Support (U)

(U) In the past each organizational element of the Intelligence Community has dealt individually and independently with problems of information distribution, storage, and retrieval. Indeed, there was no mechanism to deal with such problems on a Community-wide basis. The result, as we have amply illustrated, has been a proliferation of information systems that, for the most part, are narrowly focused on the particular functions and interests of the organization for which they were designed. These systems are also limited by the technology and resources available at the time each was constructed, and, taken as a whole, they are largely duplicative, inefficient and expensive.
(U) The Community can no longer afford a laissez faire approach to information handling. Although it is true that there will always be a requirement for special systems to meet special needs, there is a growing awareness of the commonality of the Community's interests in information distribution, storage, and retrieval. Along with this growing awareness, there has been a rapid advance in the technology of information handling and networking. Indeed, a technological solution to many of our problems is at hand. This answer can be found in a semiautomated storage-and-retrieval system incorporating the best features of SOLIS, RECON and other Community systems.
(U) The basic concept of this Community system should be one of a distributed system with minimum redudancy to include the following specific features:
• All nodes of the distributed system should be linked together via communications services, and the communications facilities should be sufficient to satisfy peak-load requirements of all system users.
• Each of the major participants should establish and maintain its portion of the system in accordance with agreed-upon standards. All files should conform to accepted Community-wide standards regarding structure, indexing procedures, and communications protocols.
• Authorized system users throughout the Community should have the capability to query any combination of the distributed data bases with a single interrogation.

• Responsibility for system development should be centralized. The DCI might choose to establish a team at the DCI level to develop the system, or establish one agency as the executive agent for the development effort.

• All nodes of the system should be configured identically with respect to computer manufacturer and software system employed. (Because of volume requirements, it might be necessary for some nodes to be larger than others in terms of processing or storage capacity, but the concept of identical systems must be paramount.)

• Software maintenance should be centralized.

(U) There are a number of different ways in which the distributed system might be put into operation, but we must not allow ourselves to become trapped in a morass of choice and indecision. The principal ingredients of a plan of action should be fairly clear.

(U) The 1979 study of the Analyst Support Task Force identified the basic requirements of an information system designed to support the intelligence analyst. These requirements include full-text storage-and-retrieval capability, a combination of automatic and manual indexing techniques encompassing concept and subject codes as well as keywords in title and text, and an ability to make distributions of current intelligence materials to analysts based on stored user profiles.[4] Armed with these requirements, we should have no need to launch a massive development effort to build the desired system from scratch. Instead, we should see if there is an existing Community or commercial system with the capability to satisfy the basic requirements. Once we have identified such a system, we should enhance and replicate it as necessary to satisfy the information needs of the Community at large.

(S) The following is a description of one way that a distributed Intelligence Community information storage-and-retrieval system could be put into operation:

1. Order identical processors for CIA, (b)(1) and NSA.

2. Establish a central software maintenance team, operating under the technical direction and control of the DCI, to develop and maintain the distributed system's software.

3. Establish an Intelligence Community standard for the contents and format of the externals of intelligence products.

4. Adopt or develop standard subject codes and apply these codes to all intelligence products.

5. Rewrite SOLIS to operate on the CIA, DIA and NSA processors and to deal with the various types and formats of intelligence products. (As an alternative, an equivalent commercial system might be procured. Another possibility, to avoid a major rewrite of SOLIS, would be to choose Burroughs computers as hosts for the distributed system.)

6. See that each node of the bibliographic/document service system provides a thesaurus and associated software that may be invoked during retrievals. The thesauri could be maintained by a single team, working under the technical direction and control of the DCI.

7. Insofar as is possible, make the entry of data into each bibliographic/document file part of the formal process for preparing intelligence products in the several agencies. When this is not possible, a direct link with the telecommunications circuit should be the mode of entry — unless, of course, the product is not suitable for electrical processing.

8. Provide for entry into the system of those intelligence products not available in electrical form. (A character-recognition device such as the Kurzweil Data Entry Machine might serve as the means of entry.)

9. Program the distributed bibliographic systems to permit automatic dissemination of intelligence products to analysts throughout the Community on the basis of requirement profiles, always with proper regard for clearance and need-to-know policies and regulations.

10. Assign responsibility to CIA for the indexing of CIA products, Department of State products containing substantive information, NPIC reports, and selected open-source material; (b)(1) and to NSA for the indexing of Sigint products produced by the various cryptologic organizations.

11. Ensure that there is a commitment on the part of each Agency to the complete processing, indexing and storage of all intelligence products for which each is responsible. Five years of products should be kept on-line, with a batch access capability against older materials. Each system should contain the full text of intelligence products up to ten pages in length, with abstracts of longer documents prepared in accordance with a to-be-written Community standard.

12. Afford Community access to the distributed bibliographic systems through the COINS network.

---

[4] (U) Ibid.

Conclusion (U)

(U)   An underlying reason for the failure of previous attempts to move toward a Community-wide information-handling system is the basic drive of the several agencies toward self-sufficiency. My intent in making this statement is not to cast stones. The drive for self-sufficiency is the natural result of the varied responsibilities, interests, and experiences of the different agencies and their need to minimize dependence on other organizations over which they have no direct control. Security concerns have also had a significant and continuing impact. This has been true at the Community level, and, to a lesser extent, the situation has been mirrored at the intra-agency level. As a result of this often justified search for self-sufficiency and security, duplication of file building and information support efforts has spread through the Intelligence Community like a creeping plague; and pleas for cooperative Community efforts have received only lip service.

(U)   A certain amount of redundancy in information sytems can be justified. As previously suggested, there will always be a need for special files (or special codes or special procedures) to meet special needs. It is obvious, however, that we have gone far beyond any reasonable level of redundancy — and our definition of special needs has been stretched to the breaking point. Indeed, the situation is out of hand. It is also obvious that in fighting duplication and trying to build an effective Community-wide information-handling system, we find that the technological problems are only part, and the smaller part, of the difficulties that confront us. For us to be successful there must be a DCI-level organization with the clout to make and enforce the hard decisions. There must be a DCI-monitored commitment by the various members of the Intelligence Community to carry their part of the information load in a way that will satisfy the needs of all users. Only by getting such commitment and enforcement can we obtain the broad-based support necessary to a large-scale Community endeavor in this important area; and only by a cooperative Community effort can we give the intelligence analyst the level of information support that he needs and deserves.

(U)   (b)(3)-P.L. 86-36 **has served the Agency in various assignments related to information science and computer systems. He is presently in Information Services (T12) and is involved with the management of SOLIS and other information storage-and-retrieval systems.**

(b)(3)-P.L. 86-36