ROBERT S. HOOPER
Chief, Systems Analysis Staff
Office of Central Reference
Central Intelligence Agency
Washington, D.C.

SUSANNE HENDERSON
Librarian, Systems Analysis Staff
Office of Central Reference
Central Intelligence Agency
Washington, D.C.

The Status of "Paperless" Systems in the Intelligence Community

The Central Intelligence Agency (CIA) and the Defense Intelligence Agency (DIA) have consolidated resources to build a "paperless" information system called SAFE (Support for the Analysts' File Environment). The system will provide intelligence analysts from both agencies with a set of tools to assist them in performing their primary mission: to prepare finished intelligence for national-level policy-makers. This paper covers the evolution of CIA's SAFE System and its current status.

It is estimated that in full operation, SAFE will do away with a minimum of 10 million pieces of paper yearly. While paper reduction has been a goal, it has not been the guiding principle in system development. We recognize that the computer is not always the best way to handle information. Sometimes information is handled most efficiently on paper or microform. In designing SAFE, we have tried to strike a balance and take into consideration the best features of all information processing techniques.

The CIA SAFE System will combine computer and microfilm storage capabilities to support approximately 1350 professional intelligence analysts. These analysts are of the economic, scientific, military, and political persuasions. They lean heavily on their individual or private files and in varying degrees on the services of a central library. Their interests often parallel those of their counterparts in academia, industry, libraries, newspapers, etc. We are not talking about ephemeral information coded on the back of a matchbook cover! We are talking about thousands of pieces of information to be read and processed every day.

Historically, systems built solely by computer personnel who felt they knew what the user wanted have not been successful. We have not made

that mistake here. SAFE, to the greatest extent possible, has been designed by the analysts. In addition to a continuous program of analyst interviews and briefings, the basic SAFE concepts were tested in 1974 in a pilot system which was to have lasted for one year. However, it proved so popular that it was retained and is now known as the Interim System.

After this introduction and explanation, it now seems appropriate that this opus be retitled: "A Plan to Build a Fixed Cost, User-Designed and Pilot-System Tested Computer and Microfilm System, Optimizing the Private and Central File Relationship, Creating a Partially Paperless Environment, in Support of the Political, Military, Economic, and Scientific Analysts of the CIA."

Characteristics of the Users

Currently, we estimate that fully operational, SAFE will support around 2000 users. Of these, approximately 16 percent will be management, 17 percent intelligence assistants and clericals, and 67 percent intelligence analysts. Therefore, the majority of our efforts are directed toward satisfying the needs of the intelligence analyst. There is no typical intelligence analyst. Each one is responsible for the analysis, synthesis, and presentation of intelligence information in his particular area of interest. These ares of interest include: politics, science, economics, military, biography, geography, cartography, photo interpretation, and weaponry. In turn, the topics are in many cases geographically oriented. The analyst is involved with current reporting as well as in-depth research, and is more often a specialist than a generalist. The intelligence reports they write must provide accurate knowledge and estimates of the economic, military, political, and scientific capabilities of foreign countries. They must be timely enough to meet the needs of those concerned with crisis situations (current awareness) as well as those interested in long-term trends.

Analysts receive the majority of their information on a continuing basis through the Agency's dissemination system (see Figure 1). This system routes publications, reports, newspapers, books, etc., in addition to a large number of telegraphic dispatches or cables. These items may have been specially ordered by individual analysts, or may correspond to "profile" or reading requirements that the analysts have on file.

In addition to their own files created from information received through the dissemination system, analysts can call upon the services of the Office of Central Reference. These services include a central library containing 100,000 volumes, 1700 different newspapers and journals, millions of documents, an extensive reference collection, and an experienced staff of professional librarians. There are also professional researchers
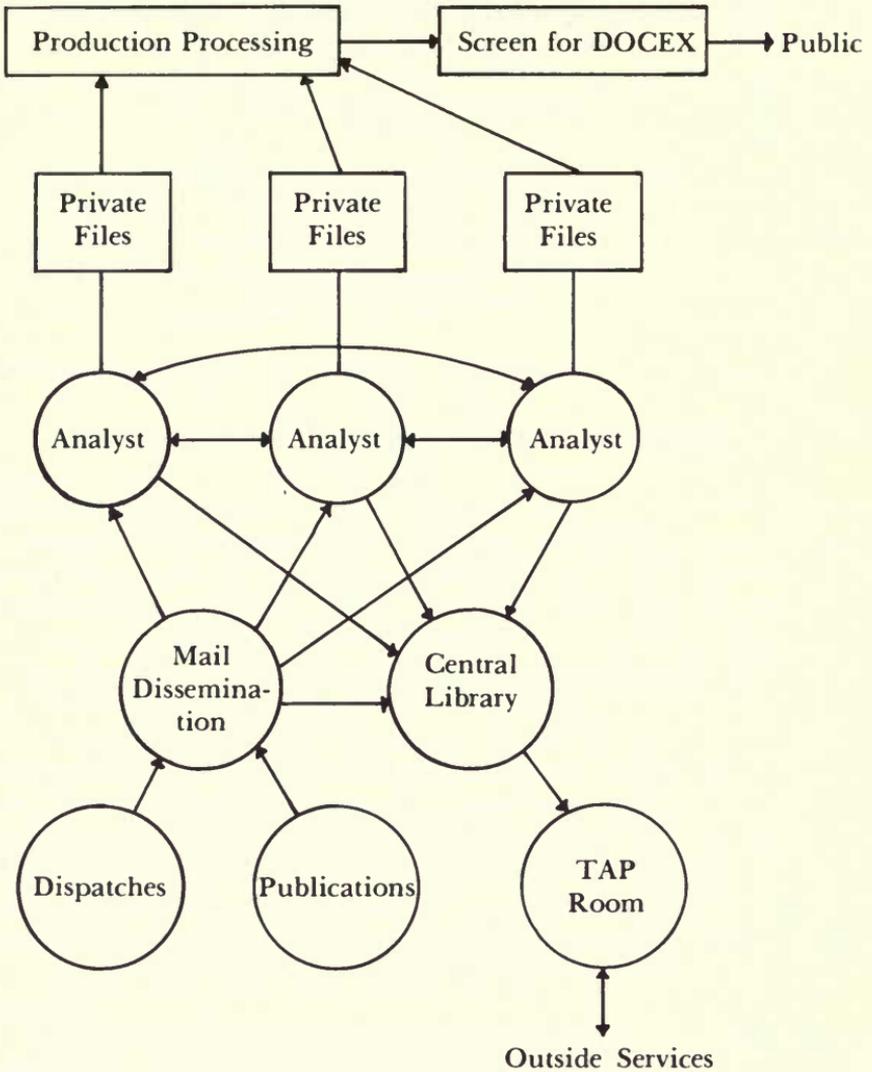
FIGURE 1. CIA Dissemination System

divided into branches of special interest (many of which correspond with those of the analysts). These branches maintain a bibliographic data base which indexes a collection of 3 million intelligence documents. A TAP (Terminal Access Point) room provides access to numerous commercial

and classified data bases, including the New York Times Information Bank, Lockheed Dialog, SDC Orbit, MEDLINE, and several intelligence community data bases.

Analysts read, review, extract/abstract, comment on, and highlight their information, keeping about 20 percent of what they receive. Retained items are eventually filed according to various methods of indexing, subject headings, thesauri, etc. The nature of a particular file organization is solely dependent on the individual mission of an analyst; in other words, the analyst is not forced to conform to a predefined way of entering data into a file.

When an assignment is received, the analyst searches his files and retrieves relevant information. These data are reviewed and interpreted, conclusions are formed, estimates are formulated, predictions are made, and actions are recommended. All of these elements are formed into an intelligence report. The analyst then coordinates and consults with others, rewrites or makes changes as necessary, and publishes a finished product which is disseminated through established channels.

Unclassified documents are reviewed for public interest. If it is felt that there will be interest, the documents are released. CIA publications available to the public may be obtained in one of three ways: (1) USGPO (United States Government Printing Office), (2) NTIS (National Technical Information Service), or (3) DOCEX (Documents Expediting Project) through the Library of Congress. Well-known examples of documents released this way include the famous oil document that President Carter mentioned in his spring 1977 news conference, and the *China Atlas*. Other examples are the *National Basic Intelligence Factbook, Soviet Civil Defense, Chinese Coal Industry*, etc.

Steps Toward Paperlessness

Over the years, the volume of items retained by analysts, as well as the Office of Central Reference, has grown to an unmanageable size. Simply stated, the size of the various document collections was and is a major factor in our looking toward paperlessness. Indeed, it is no joke that the floor loads are being exceeded and eventually, if help is not found, the file collections will have to be housed in the basement (terra firma at last), and of course, as the collection gets larger, the probability of finding anything gets smaller. Thus, by necessity, we have had to look at alternatives, i.e., microform and machine-readable data bases.

The advantages of paper reduction, in addition to increasing the storage space, are several. First, by storing documents in machine-readable form instead of paper we can make use of information retrieval and

text-search capabilities to retrieve those items relevant to an analyst's particular, immediate need rather than having to get involved in a time-consuming manual search. Second, reducing the amount of paper will, of course, remove much of the burden caused by the large amount of paper-handling presently taking place, thus saving processing time and reducing mistakes. Since a considerable amount of coordination is required among the intelligence producers, a paperless communication capability both before and during the production process should speed the process and improve the product considerably. Table 1 summarizes these attributes of paperlessness and associates them with the various functions of the analysts' information world.

TABLE 1. ADVANTAGES OF PAPERLESS COMMUNICATION
RELATED TO VARIOUS INFORMATION FUNCTIONS

Attributes of Paperlessness	Mail Dissemination	Private Files	Central Index Files	Central Document Files	Production Processing
Reduce size of storage		X		X	
Improve search capability	X	X		X	
Improve document retrieval and display		X		X	
Decrease paper-handling time/ errors	X				X
Make remotely located files more available			X	X	
Improve communication among analysts					X

Thus, paperlessness has its desirable attributes, but it is not without its problems. One problem is the cost involved in converting to paperlessness. For example, the cost to have the Office of Central Reference convert all documents to machine-readable form would definitely be prohibitive. Another problem is the readability and "caressability" of paper. Analysts, like most of us, have a love affair with their paper dolls and this is not easy to break. And finally, paperlessness alone will not solve the communica-

tion problem unless an analyst-to-analyst communication mechanism is established.

The SAFE concepts rest on two preliminary efforts. The first was the development of a machine-assisted dissemination system; the second was development of some models of analysts' private files.

Figure 2 is an overview of the dissemination system before the introduction of machine assistance. Note that electrical dispatches (such as cables and receipts) were converted to paper and then read for dissemination. This seemed like an appropriate spot for us to begin our paperless processing operations because the conversion of these items to paper seemed an unnecessary step. We spent two years perfecting a computer text-search dissemination system, which we dubbed MAD (Machine-Assisted Dissemination), which processed the electrical dispatches as they were received.

The MAD operation worked in the following way (see Figure 3). All of the dispatches were formatted to make the text-search operation as effective as possible. Then an initial computer scan of the text compared each word against words in the "Dissemination Dictionary #1." If the proper words were present in a dispatch, and if they appeared in the proper order and with the proper logic, then the dissemination addresses would be automatically appended. Finally, the dispatches were sorted and the proper number of copies were printed. The first dictionary contained, for example, the names of certain kinds of reports which were always sent to the same customers. Hence, a purely clerical function of addressing these kinds of reports was now done by computer.

Those dispatches which were not disseminated by the first process were again scanned by computer against "Dissemination Dictionary #2." In this case, the computer would again compare every word of text with the dictionary words, and proposed dissemination addresses were placed on the message. However, a human disseminator would look at the proposed dissemination, check it against his set of requirements, and add to or change the proposed dissemination as appropriate. All of this was done on a cathode ray tube (CRT) device. The disseminator would then release the item for printing. The proper number of copies in sorted order would then be printed. At this point the paper delivery system took over. Through these actions we had succeeded in reducing the clerical functions and putting the disseminators in a paperless environment.

As time went on, we added the capability of writing special tape files as well as producing paper output. These tape files would contain all dispatches that met a given set of profile conditions. They were produced wholly from the machine's capability to find words and combinations of words in proper order and with proper logic within the text of the dis-
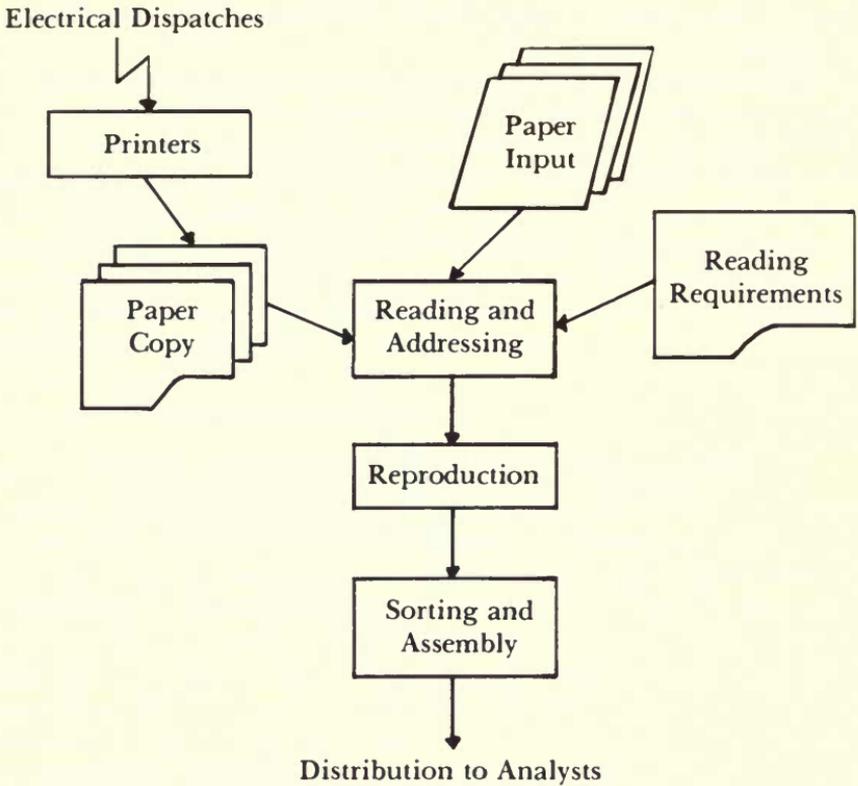
FIGURE 2. Pre-MAD Dissemination System

patches. These files were searchable on the Library's Rapid Search Machine (RSM), a stand-alone computer device developed by General Electric that text-searches computer tapes. These tapes were a step toward improving the analysts' search capability.

Analysts depend heavily on their own or private document collections. In some instances these collections have grown very large and the documents have become increasingly hard to find. To improve the storage and searching of these private collections, we introduced the rather simple notion of microfilming the documents and building computer-based indexes to represent them. The indexes were built on-line and were subject to certain limitations on size and complexity, but the actual design of an index record was up to the individual analyst. One analyst indexed documents according to a "who, what, when, where, how" scheme, while another set up a simplified classification schedule. The key to success was
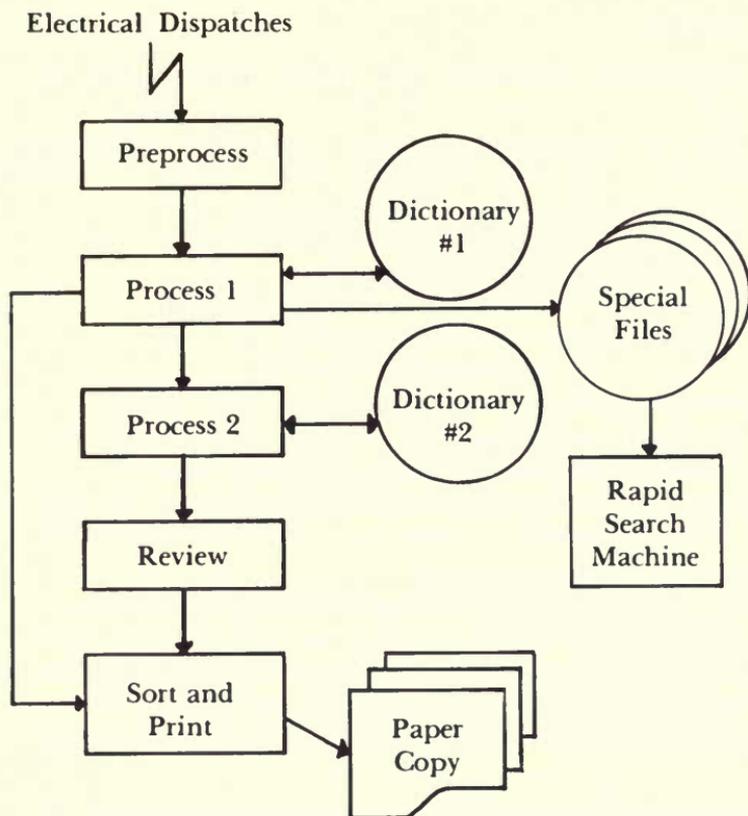
FIGURE 3. MAD System

to make input by the analyst as simple as possible and to let the file be designed by him to meet his individual needs.

From these initial developmental studies we derived a number of significant results:

1. disseminators can make the change from reading paper copy to reading dispatches on-line most readily;
2. the ability to do text searches was a popular innovation;
3. under certain circumstances, analysts will use microfilm in lieu of paper copy; and
4. analysts are enthusiastic about on-line indexes to their private files, even though the actual building of the indexes was considered a chore.

A Pilot System

Armed with this information from our developmental studies and the new emphasis on automation of user files, we began to study the relationship between the analysts' private files and the central library file. The study focused on three questions. How redundant are they? How redundant do they need to be? And the most important question, to what extent could the central library file be used in lieu of or as a complement to the analysts' files? To answer these questions, and because information is only as valuable as it is accessible, we postulated a pilot system that would allow analysts to have as much information at their fingertips as possible.

The pilot system allowed the analysts a continuous spectrum of on-line services, from receipt and disposition of their "mail" (that is, their electrically disseminated dipatches), through the building and searching of their private files and the use of the large central index file, to the use of a large, 4- to 5-week collection of electrical documents. In addition, the analysts had the capability to compose memoranda on-line and to route items from one analyst to another.

There were forty users from four branches representing the major disciplines (political, military, scientific, and economic). We asked them to use the pilot system capabilities as appropriate during the day, and to record basic facts about each use. After nine months of testing, we held a structured interview with each of the users. The most significant findings were:

1. Analysts are not willing to give up the use of paper completely. Paper will still be necessary for cut-and-paste, marginal notes, etc. They are willing, however, to use the processes that store it, search through it, etc.
2. The analysts did, slowly but surely, learn to depend on the pilot system.
3. The best way to introduce office automation is to get an analyst to volunteer to use the capabilities and then let his peers look over his shoulder at the results of his work. A dedicated, convinced user of the system is the best advertisement we have.
4. Analysts would query the central file directly. This previously required a professional researcher to act as an intermediary. The analysts were willing to learn how the big file was organized in order to get the most out of it. However, the occasional user or nonuser of the system indicated that his future preference would be to continue to use the researcher as a go-between.
5. Analysts were very willing to let other analysts have access to their private files.
6. A new category of worker might well appear who would do the "dog work" required of the system; that is, perform data entry, make routine searches, enter compositions into the system, etc.

The analysts' enthusiasm for the pilot system led to a 1974 report that recommended continuing the system as the Interim System and to begin work on an agencywide system—SAFE.

Target 1984

The SAFE System will provide intelligence analysts with a highly reliable, dynamic set of tools. The system, when completed, will be a 24-hour-a-day, 7-days-a-week, dedicated on-line computer and associated microfilm system that provides:

1. faster distribution of incoming intelligence,
2. improved organization of and retrieval from central and personal electrical or hard-copy files,
3. procedures for composing and coordinating information, and
4. indirect access to other intelligence community and commercially available computer systems.

The system will be implemented in a phased approach so that the users will get the assistance they need most at the earliest possible date.

The first phase of SAFE, "Current Awareness," will make electrically received messages or documents (cable traffic) available to the current intelligence analyst as quickly and as completely as possible (within five to fifteen minutes of receipt by the Agency). A SAFE Message Analysis (SMA) module will process and format the incoming messages; then, based on analyst reading requirements contained in profiles, disseminate them. The results of this effort are forwarded to analyst mail files. They are then available for viewing on-line by each analyst at his SAFE terminal. Analysts will be able to process the mail in a variety of ways: file, route, reject, print, and so forth. In addition, all documents that enter SAFE will be retained in a 24-hour text file.

An important adjunct to the mail receipt and disposition will be the capability to "text search" the items within the mail files and those within the 24-hour text file. Analysts will also be provided with a capability to compose articles on-line. This is intended to allow an analyst to capture his thoughts as he reads his "traffic." It is also to be used for forwarding comments to associates, and perhaps for quick text preparation to meet tight deadlines.

The second phase of SAFE, "Private File Management," will provide analysts with an efficient means of creating, searching, and maintaining their own personal files. The analyst will be able to perform detailed filing of material. He may extract data from messages and place them into a file; annotate what he reads and then file it; underscore or otherwise highlight parts of the message significant to him; and/or apply index terms for

retrieval purposes. Information can be retrieved from these files by looking for specific index terms or by searching the full text of messages and/or annotations.

Documents not received as electrical transmissions may be indexed, and the resulting indexes which are maintained in SAFE may be searched by computer. The documents may be filed locally—i.e., in a safe or a desk drawer—or in the ADSTAR System. In the ADSTAR System, the documents are microfilmed and held in a central repository. They can be retrieved and displayed on an ADSTAR terminal located in selected regional areas.

SAFE is designed to allow an analyst to input a single search strategy and get references to both hard copy and electrical items. The electrical documents will be displayed on his CRT; the hard copy documents on his regionally located ADSTAR terminal.

In the third phase of SAFE, "Retrospective Search," several retrospective search services are introduced. Access to the multimillion-record central index file (RECON), maintained by the Office of Central Reference, is made available. This index file points to documents that the office is required to retain for use by analysts throughout the intelligence community. Analysts will be able to view the results of searches against this file on their own CRT or on an ADSTAR terminal. Analysts will be able to perform text searches of electrical documents, referenced either by results of RECON searches or by results of searching a "catalog file" which references all items under SAFE control.

An additional source for retrospective search exists in the Office of Central Reference TAP Room where links to data bases external to the Agency (the New York Times Information Bank, for example) are available. These services will be made available to the analyst by having him contact the TAP Room on his SAFE terminal and provide a statement of his needs to TAP Room personnel, who will then run the query against the appropriate data base and alert the analyst when the results are ready.

In the fourth phase of SAFE, "Intelligence Production," the entire process for producing finished intelligence may be accomplished via SAFE. The simple composing capability of the first phase will be expanded to handle all text-editing and word-processing techniques, and to coordinate the entire production process for finished intelligence. A by-product of this phase will be to direct the finished intelligence back into SMA for input to the Agency mail files and the document file. In this fashion, past intelligence becomes available for future analysis.

The target date for total system implementation is 1984. We anticipate that SAFE will rapidly become an indispensable aid to the intelligence analysts. It will reduce the amount of paper; provide the analyst with

faster, more complete mail receipt; and give him the most complete and accurate retrieval of information possible. The concepts being incorporated into the design are extendible and will be sufficiently flexible to meet changing patterns of analysis that will certainly occur. With SAFE, the intelligence analyst will have the tools necessary to produce the comprehensive and timely intelligence needed by policy-makers to meet the challenges of today's world.