DD/A 75-3888

DDA Registry

File 0-4112-2

18 AUG 1975

MEMORANDUM FOR THE RECORD

SUBJECT: SAFE Briefing for Dr. Albert Hall,
Assistant Secretary of Defense for Intelligence, DOD

1. On 13 August the undersigned was asked by John Slack (OX 5-7072, Grey 2289), OASD (Intel), whether a SAFE briefing could be arranged for Dr. Hall, ASD (Intel). He said that Dr. Hall had attended an IRAC meeting on 11 August and had his interest in SAFE stimulated by the DCI's comments on SAFE. He added that Dr. Hall is very interested in any modernization techniques which would assist the analysts in their work and was also interested in the relationship between SAFE and existing systems, such as COINS.

2. I explained to Mr. Slack that SAFE was in its early stages and we would not be able to talk about any technical system concepts or designs. Our major effort was in the analysis of the analysts working environment and the detailing of requirements that should be satisfied by the SAFE system. We had recently formed a SAFE Project Office that would eventually translate the requirements into a system specification and design. I also mentioned that we had an Interim SAFE system which was being used as a test-bed for trying out some of the ideas for using a computer to facilitate the analysts work. I told Mr. Slack that we probably could put together a briefing for Dr. Hall which would describe the SAFE requirements in some detail followed by a description of the SAFE Project Office and the approach we plan to use to bring the SAFE system into the world. Mr. Slack said this sounded great, and asked if we would include some description of what is being done on the Interim SAFE system and how it has improved the analysts capabilities.

STATINTL

3. After coordination with the Acting D/CRS, Mr. [redacted] and the SAFE PD, [redacted] I called Mr. Slack and suggested a one hour briefing at a time convenient to Dr. Hall between 8 and 12 September. He later confirmed this for 9 September, 0930-1030 hours in Dr. Hall's office, 3E-282, the Pentagon. We agreed that about 45 minutes

STATINTL

would be devoted to the briefing and 15 minutes for dis-- cussion. The majority of the 45 minutes would be on the SAFE requirements and Interim System with about 10 minutes devoted to the implementation approach. I told Mr. Slack that the briefing would probably be given by Mr. [redacted] and [redacted] and they would probably be accompanied by Mr. Eisenbeiss and Mr. Fitzwater. He said "fine" and added that Dr. Hall prefers not to have a large entourage for such briefings. I told him we would confirm our attendees

STATINTL

STATINTL

STATINTL

[redacted]

STAT

[redacted]

STAT

Deputy Director of Joint Computer Support

TRANSMITTAL SLIP	DATE
TO: DDA	
ROOM NO. 7D26	BUILDING Hq

REMARKS:

F.Y.I. 19 AUG 1975

Re: Cook courtesy. Are we getting Bird to Hall on our reports?

C

FROM: D/OJCS		
ROOM NO. 2E29	BUILDING Hq	EXTENSION

FORM NO 241 1 FEB 55 REPLACES FORM 36-8 WHICH MAY BE USED. (47)

FOR : [redacted]

Confidential

25X1A

Project SAFE

Feasibility of an Agency-wide Information System to Support the Analysts File Environment

Confidential

October 1974

Copy No 71

NATIONAL SECURITY INFORMATION
Unauthorized Disclosure Subject to Criminal Sanctions

25X1A

Approved For Release 2002/11/04 : CIA-RDP79-00498A000400050049-9

22 OCT 1974

25X1A

MEMORANDUM FOR THE RECORD

SUBJECT : Facility Proposed for Project SAFE

REFERENCE: Memo dtd 16 October 74 to DDA from D/OJCS
Same Subj

1. A cost estimate to support project SAFE is included in Paragraph 8. The estimate is based on the data presented in the referenced memo and based on the following assumptions:

a. Central utility systems required in support of OJCS proposed expansion will be available. Specifically, a new 2500 kW automatic start generator will be procured and installed, and the Carrier Dunham/Bush chillers will be repiped to a parallel configuration.

b. The installation will be located on the first floor of the South end of the Headquarters Building. The first floor slab to slab height will allow adequate raised floor clearance, and the central utility systems are located in the South end of the building.

c. There will be no unique security requirements. Costs include the provision of a special purpose vault similar to the ORACLE installation.

d. There is no reason to believe that the proposed configuration will exceed the load capacity of the floor. However, a structure analysis cannot be made until an equipment layout is provided.

e. There will be no emination problems, i.e., a screen room will not be required.

f. The provision of standard environmental requirements for this type computer center is included in the cost. There is no provision for uniquely tight tolerances for the control of humidity, dust or temperatures.

Next 3 Page(s) In Document Exempt

Approved For Release 2002/11/04 : CIA-RDP79-00498A000400050049-9

OJCS 1363-74

MEMORANDUM FOR: Director of Central Reference Service

SUBJECT : Comments on Draft SAFE Report

1. General. The SAFE report provides an excellent historical base and a good picture of the kind of system which can and should be developed for the analyst. From our vantage point, the following are important elements of the planned development effort described in the report:

a. The system should be designed as an integrated system, and the development of each sub-system should be well coordinated to fit with the others. A single coordinating group would serve this purpose.

b. The SAFE terminal should be specially selected to meet the SAFE requirements, and this implies that the terminal may be different from the OJCS standard remote terminal.

c. The use of distributed processors for terminal support will reduce the cost and improve the usefulness of the system.

d. The development of a SAFE system should be done in phases with provision to accommodate user feed-back. Our major concerns about the report are in a few technical areas and with the means for achieving your goals. The specific comments below are separated into those which you may want to consider before publication and those which should be the subject of further discussion near the beginning of the next phase of SAFE.

2. Comments relevant to the drafting of the report:

a. The issue of security should be treated in the report. We note that it was emphasized in the vendor reports. Protecting this vast amount of information from both deliberate penetration and unintentional disclosure, and maintaining the "need-to-know" principle, may require a significant effort.

b. The report should clearly identify the critical technical elements of the proposed system, particularly those where some risk is involved. Specifically, the report should caution the reader about the crucial issue of the high-speed search of large volumes of textual information. An additional feasibility and design study must be undertaken to find an appropriate searching technique and understand its response characteristics. We believe it is quite possible that such a study would conclude that response time requirements would have to be relaxed, and this might threaten the viability of the whole project.

c. We believe the report should state clearly the best estimates on the size of the system. Each of the five contractors had different impressions of the size of the problem, as did the OJCS members of the team. Clearly, the major driving forces in choosing a configuration are the number of the various types of records, the amount of expected activity from all of the terminals, and the maximum acceptable access times.

d. The report emphasizes the need for system reliability and mentions the need for processor redundancy. Of equal or greater importance is the need for the file back-up, considering the large volume of data which would be vulnerable to both hardware and software failures. The report should mention this. The greatest vulnerability in the entire system is likely to be the file indexes, which, if lost or subtly modified by malfunctioning software, could result in long periods of file unavailability during reprocessing and restructuring. A scheme will have to be devised to maintain the huge backups without draining the resources of the system.

e. We believe the report should state that the SAFE system should be integrated into the Agency's data processing environment as much as possible. For example, the terminal sub-system should be able to access current OJCS services, such as GIM and CP/CMS. Many SAFE users will also be users of OJCS services such as APL, SCRIPT, GIM and others, and should not be required to have two terminals in order to access all of these systems.

f. Because of our Agency-wide responsibilities in ADP and because of the size of the investment, OJCS representation in the next phase of SAFE should be more significant than the draft report suggests. Specifically, it is suggested that the statement on OJCS participation in the section on the Development Plan be modified to add "one OJCS analyst

responsible for hardware configuration and operating system software." If you agree with this addition, you may wish to reduce the number of CRS people on system configuration analysis from two to one.

g. Given our experience (and yours) on equipment acquisition and installation, the proposed equipment installation schedule is ambitious. If maintaining such a schedule is vital, you should include and underline a statement that extraordinary procedures and priorities are justified and will be needed to achieve your goals.

h. A minor point: the report should not presume that the AMPEX Terrabit memory is the most appropriate mass storage system for the application. It will be available, but other devices exist which might be better for archiving and other purposes (such as the Precision Instruments UNICON).

3. Points for future discussion:

a. The report says that SAFE is only viable if the entire task is implemented ("all or nothing"--page 8). While the design of the system should certainly be broad enough to accommodate all of the SAFE concepts (and more), we believe that each specific sub-system should be evaluated on its individual merits. Many of the services proposed here are valuable; however, unless a cost-benefits analysis is performed on each separate service, it is unclear which services will be worth the cost of implementation. Further, some of the requirements might well be trimmed down to reduce the cost.

b. Must all of the services required in SAFE be written specifically for SAFE? Some of the services mentioned in the design can be supported by existing software, although it is certainly worth the effort to modify outside packages to add consistancy of operation to the whole SAFE system.

c. We believe that backup requirements deserve much more scrutiny. The report implies that all users of the SAFE system require the same level of backup, although a degraded system (all terminals not supported) might be adequate during major system failures if these occur infrequently. This is one example of the need for an important next step; an identification of the critial performance components is necessary to determine the appropriate type and amount of equipment.

d. It is not clear that the system, as outlined in this report, is state-of-the-art or, in fact, implementable.

The endorsements of the vendors are devalued somewhat by the fact that their own implementation ideas are either unworkable or would require significant amounts of untried technology such as special purpose equipment. The use of hardware enhancements, such as associative processors, should be evaluated against the risk implied by esoteric or one-of-a-kind hardware. Other vendors with implementation experience in large scale text searching systems (such as the New York Times Information Bank and Mead Data Central) should be consulted for design ideas and implementation software.

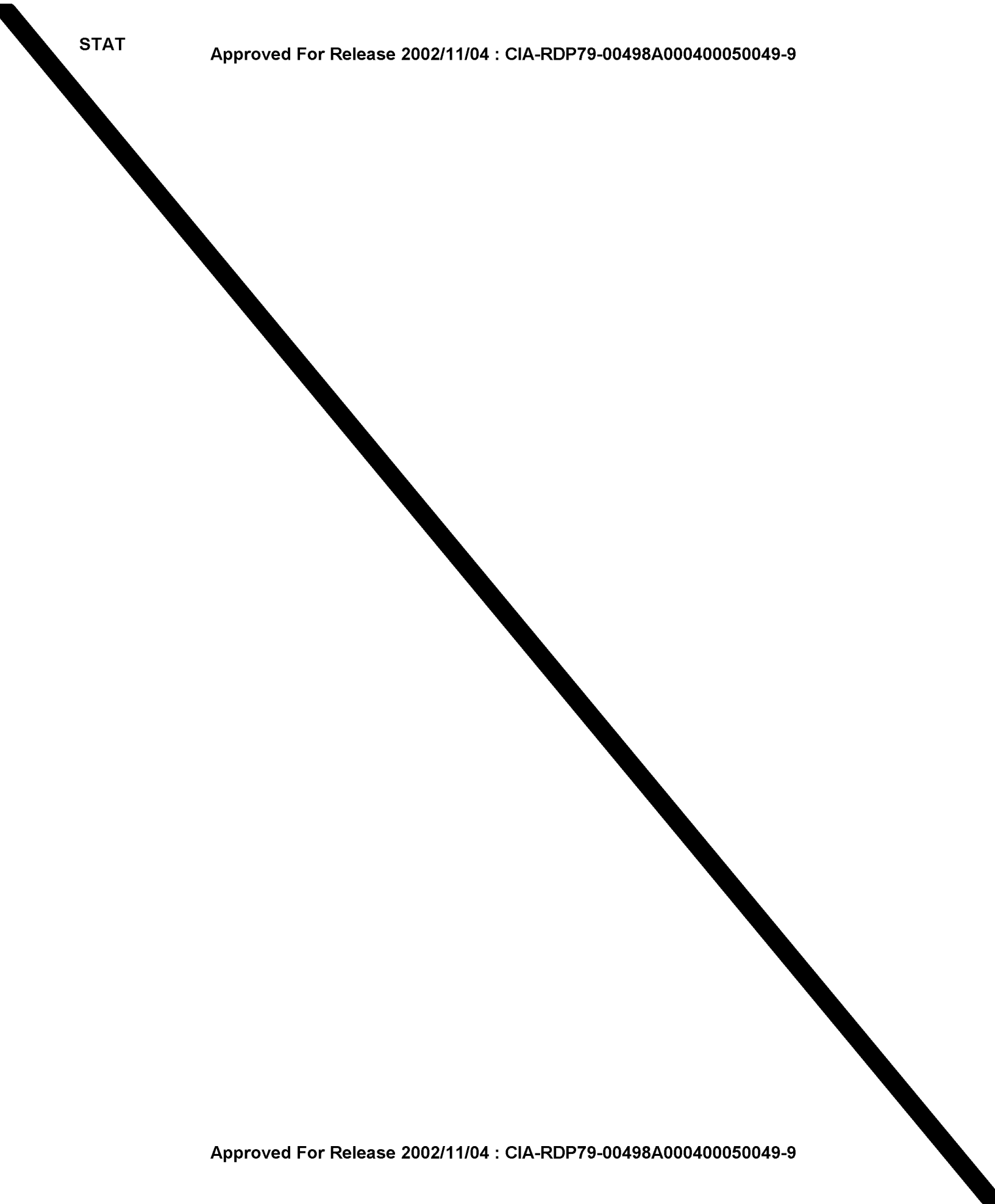
d. Much of the high cost of the system can be related to the requirements for rapid access to huge amounts of data. Would the system concept still be viable if less were retained in the computers, or if the response time requirements were relaxed?

e. The SAFE team should consider the use of a common procedure language facility (such as the CMS EXEC or GIM PROC facilities) as a means of reducing the complexity to the user. This technique could be used in place of a common language like SQUIRL to provide assistance and to restructure input lines for the neophyte user while allowing total access to the full range of query complexities for more sophisticated users.

4. When taken one at a time, the above points can probably be resolved, but I must confess to an uneasy personal feeling about the totality of the problem that we face in building and operating the proposed system. I am advised that the volume of data, the interactions of data elements, the response time and system availability requirements will produce complexity which no other computer system has ever faced. The Agency has no experience in building systems of this size; in fact, no text handling system of this scope has been built anywhere. The risks are considerable, and I caution against letting the analysts' enthusiasm and the absorbing challenge of the job hypnotize us into dismissing them or putting off a review of them to a later phase. In conclusion, it is necessary that we face up to potential problems in the early stages of the proposed program to ensure against nonrecoverable pitfalls that may occur in the future.

STAT

HARRY E. FITZWATER
Director of Joint Computer Support



Approved For Release 2002/11/04 : CIA-RDP79-00498A000400050049-9

Project SAFE

October 1974

CONTENTS

LIST OF FIGURES

LIST OF TABLES

APPENDICES*

I. PROJECT SAFE PAPERS (October 1972, May 1973, November 1973)
II. DATA COLLECTION FORMS (Pilot and Self-Help Branches)
III. PRODUCTION ANALYST FEEDBACK (Reports/Interviews)
IV. HELP LOG
V. PRELIMINARY DESIGN REPORT
VI. CONTRACTOR REPORTS
VII. REDBOOK INDEX
VIII. SAFE INSTRUCTION MANUALS
IX. PRELIMINARY SURVEYS
X. MAIL LOG

*Appendices I – X are located in Room 1E4808, CRS/Systems Analysis Staff, and will be made available for reference.

I. INTRODUCTION AND SUMMARY

INTRODUCTION

Over the years CIA has made a wide array of intelligence resources available to its analysts. Indeed, the Intelligence Community spends a large sum each year to provide these resources and to find new ones. They are made available by such a variety of processing systems and procedures that the individual analyst may have difficulty in finding all the items he needs—particularly if he has a short deadline.

Production offices have continually sought to better exploit intelligence resources by creating their own data bases and files, sharing files of common interest, or introducing new analytical methods or automation. For the most part, these efforts are made at the office level and, at best, answer only office needs.

This report describes CRS efforts to design an Agency-wide, all-source intelligence resource system that would offer all Agency analysts the best support today's technology can provide. It suggests how such a system might be cheaper in the long run than the sum of all the individual systems currently being developed or proposed. The design that emerges is called the SAFE (Support for the Analysts File Environment) Information System.

SUMMARY

CRS began work on Project SAFE in response to a June 1972 directive by Mr. Colby, then Executive Director Comptroller. It said that CRS should "work with the analysts and production offices within the Agency . . . to develop the most effective mix of central bibliographic and document retrieval files and special purpose document retrieval files for individual customer offices, (and) analysts. . . ."

Preliminary development work with the production analysts soon showed what characteristics a SAFE system should have. The concept that emerged was that of a multipurpose Agency-wide information processing system operating through on-line terminals widely distributed among the production offices. SAFE will permit the individual analyst to view his daily mail on-line, route particular items to other analysts, build machine files for himself or his office, and to maintain on-line files. The on-line file building capability will allow the analyst to store a complete text, an extract from it, or an indexed representation of it and to include his own comments on such items. The system will allow the analyst to search the files he creates and, because he has multiple access points to any item, to search them more thoroughly and more specifically than he could normally search a conventional paper copy file. Where document representations are stored in files, SAFE will provide the necessary full text back-up, either by digital storage of text or, more commonly, microforms.

In addition to its role in dissemination and in the support of analyst or office files, SAFE will give the analyst access, through his on-line terminal, to a wide range of resources, including the major CRS data base and several files of the complete texts of intelligence messages. Eventually the analyst may also be able to use the same terminal to reach "external" data bases, including those within the community as well as such commercially available files as the *New York Times* Information Bank. The analyst thus will have, at his fingertips, a wide array of information resources needed in the production of finished intelligence.

CRS implemented a model of a SAFE system and made it available to a small number of production offices over an 8-month period in 1973-74. This was defined as the data-gathering phase of the project. Its objectives were threefold: to determine the general feasibility of SAFE; to learn the user's reaction; and to gather data from which to develop more detailed specifications for an Agency-wide system. The SAFE model was modest in that it used inexpensive and relatively unsophisticated software, existing computer resources, a small number of terminals and a selected sample of users. It nevertheless demonstrated all of the major components of the proposed system.

Close cooperation between CRS and the analysts in the production offices has been an important feature of the data-gathering phase. Those analysts played a key role in the design of the pilot system. Indeed, CRS assumed from the beginning that if an Agency-wide system is to succeed, its real users must be involved in its actual design. The pilot branches cooperated fully, and the large amount of data collected has enabled us to define much more clearly the requirements of an Agency-wide system.

Conclusions

The overall reaction of participants in the SAFE pilot operation has been extremely positive. Our evaluation (described in detail in Chapter V) of the pilot system indicates that SAFE is potentially a very powerful tool, faster and more efficient than the resources we presently have. Most analysts who have used the pilot system are enthusiastic about its present capabilities and its potential. Indeed, there is a strong feeling that this is the direction the Agency must take in information processing. All the proposed features of the system have proven valuable, but the handling of text files and the building of analyst files will probably be the most important.

Two of the most significant values of SAFE will be its ability to get incoming material to analysts rapidly and its ability to provide fast access to a wide array of information. It appears to have great potential utility, therefore, in the handling of crisis situations, as reported by one of the pilot branch users: 25X1A

25X1A

[redacted]. I believe the SAFE system has enormous potential for crisis management.

The SAFE concepts were examined by five companies involved in the design of large computer-operated data systems. They believe most of the concepts, with one major exception, are within the state-of-the-art. The exception refers to the part of the original concept that called for scanning paper copy, digitizing it and entering it into the system. In their opinion this is not currently feasible. Because parts of the SAFE concept are close to the outer limits of the state-of-the-art, implementation of SAFE will present major challenges in systems design, software production, and the coordination of much hardware. A similarly large and complex system is not known to exist elsewhere. The individual parts do exist, however, and the contractors agree that SAFE can be built.

Our experiment has persuaded us that the Agency should move toward the implementation of a system of this kind, having the general configuration described in Chapter VI, and that we should immediately begin work on a detailed system design.

Cost

25X1C

To support [redacted] the proposed system will require a substantial investment over a number of years. Some of this investment will be compensated by a more efficient and integrated use of Agency computer resources; by the assimilation of

certain existing systems and operations; and by a considerable reduction in the generation, movement, storage and disposal of paper copy. The system must be justified on the grounds of benefits to the Intelligence Community not on the grounds of economy. We consider these benefits to be improved intelligence products, generated by analysts who are informed more rapidly, more completely and more precisely than ever before.

The estimated cost of SAFE is abou[redacted] dollars. This sum would cover the software design and development and the purchase of hardware in 1974 dollars. It does not include past costs, personnel costs of CIA employees involved in the project, logistic costs (which may be high), or OJCS costs for continued support of the pilot program. Our estimated cost would be less if the software could be developed in-house (which is highly desirable) and if much of our existing equipment could be used. We have deliberately used the high figure of our cost range to make sure that approval of Project SAFE carries a realistic recognition of the potential financial impact (excluding logistic costs). Development of the SAFE effort is a commitment of up t[redacted] dollars and a development period of at least 5 years. It would also represent a major effort—not yet defined—for logistics as well as an undetermined communications investment. 25X1A 25X1A

These dollar and time costs are as firm as we can determine from current experience. Both could increase, however, during SAFE's development and implementation. Because we have used the higher cost figure, such increases should not have a major impact on the overall cost of the system.

Finally, the SAFE Information System faces three major problems. First, there are important security considerations involved in the development of a computerized file environment which have not been addressed in this report.

Second, it was noted earlier that, although the concepts of SAFE are within the state-of-the-art, there is no system in existence of comparable size and complexity. There is a related risk. SAFE will become an integral part of the analyst's working environment; if it fails him, he is out of business. Therefore reliability and backup are critical. The Agency has limited experience in building and operating applications where the computer is so intimately tied to an Agency function. What experience we do have tells us that, in addition to high equipment reliability, extraordinary developmental and operational discipline is required even for simple applications of this kind. SAFE will represent a challenge different from any that our computer systems people have ever encountered.

Third, the project need not necessarily be completed by FY 1980; but prolonging the work would probably increase both the cost and risk. The funding need not be so heavily concentrated in the first years as we have proposed; but spreading the funds evenly across all the years will delay implementation and probably increase the risk. Most importantly, SAFE must rationally be a complete intelligence processing system. Because of the cost, we expect to hear proposals to create one-half or two-thirds of the system—to handle some sources of information, but not all; or to serve some production offices, but not all; or to perform some of the functions that are technically possible, but not all. We oppose all such proposals.

II. HISTORY OF PROJECT SAFE

CRS INITIATIVES

In December 1971 the Director of CRS created a task team to write a detailed plan for upgrading the 1,300,000-record, computer-based CRS reference file (AEGIS).[1] The general plan was to convert an off-line batch mode of operation to an on-line interactive mode. This would improve service by allowing interactive searches to be made at remote computer terminals as search requests were received. The ability to enter search requests from remote computer terminals would also theoretically allow Agency production analysts to bypass CRS analysts, who presently serve as intermediaries.

The task team was also to consider methods by which production analysts could add keywords, codes, and documents to the basic reference file. It had long been recognized that many of the analysts' special interests could not be adequately handled by the more general indexing performed by CRS.

In March of 1972 the task team began discussions with representatives from OCI, OER, OSI, DDO (then DDP), OSR, and OBGI in order to inform them of the CRS objective, to learn the extent of their interest as potential input or output users of such a system, and to determine whether any of their requirements should be considered in the proposed upgrading of AEGIS.

OCI and OBGI immediately expressed interest in a system that would give them a computer search capability for their manual office files. OCI was especially interested in reducing the size of its paper files by using a computer control system.

As a result of this interest, the task team conducted an OCI/CRS and OBGI/CRS 2-week experiment, which simulated production analyst input to the CRS AEGIS file. The results were encouraging, and in May 1972 OCI asked if CRS could implement interim measures to allow continued OCI input prior to the upgrading of AEGIS.

AGENCY DIRECTIVE

In June of 1972 the Director of CIA, Mr. Richard Helms, approved a series of recommendations by Mr. Colby, then Executive Director Comptroller. The series included a directive that CRS "work with the analysts and production offices within the Agency, and with such other Intelligence Community agencies as may be feasible, to develop the most effective mix of central bibliographic and document retrieval files and special purpose document retrieval files for individual customer offices, analysts, or other requesters."[2]

CRS RESPONSE

Responding to this directive, CRS first critically reviewed its major file building and information processing capabilities:

1. The MAD system, an Agency-wide Machine-Assisted Dissemination system developed by CRS for SI electricals;

---

[1]Already Existing General Information System—this reference file is often referred to by the acronym AEGIS, which is also the name of the computer data management program for this file. Other programs could also "manage" the reference file. In fact, later in this paper the RECON program is introduced as one such alternative.

[2]MEMORANDUM FOR THE DIRECTOR, SUBJECT: Automatic Dissemination, June 1972. (Confidential)

2. The AEGIS system and an on-line version of AEGIS (which, although not considered a candidate for the upgraded AEGIS system as discussed above, allowed for searching from remote computer terminals);
3. The OLDE computer program, an On-Line Data Entry program by which computer files are created and maintained at remote computer terminals (OLDE was developed as part of the task team's AEGIS follow-on activity);
4. The OCI and OBGI experiments, which gave some evidence that the analysts were willing to switch from their manual document files to a computer/microfilm system;
5. The CRS computer center, a center developed to maintain systems like MAD, AEGIS and OLDE.

These five capabilities were the building blocks upon which two related proposals were based:

1. "Proposal for a Demonstration of an On-line System to Provide Production Analysts with Access to Personal, Office and General Bibliographic Files."

This work was written in August 1972 [redacted] 25X1

25X1 [redacted] Its purpose was "to demonstrate a concept, with the object of generating interest and support within the various production offices . . . As the capabilities are demonstrated, user reaction will be observed and gauged . . . We can learn much more about user needs and attitudes from such a working model than we can possibly learn by a paper model and more conventional interviews or questionnaire surveys."

This working model would attempt to simulate the ultimate system ". . . (which) will give the individual production analyst on-line, interactive access to his personal document file, his parent office files, specially prepared extract files, and a wide range of CRS bibliographic files."

2. "Prototype of a CRS Production Analysts File Support System as an Interim Step Toward an Operational CRS On-Line System."

This work was written in August 1972 by [redacted] 25X1A
Analysis Staff in response to OCI's request for an interim capability. It proposed that OCI analysts would mark the terms by which their documents should be indexed; CRS would input the index records for those documents into a special AEGIS file created solely for OCI. CRS would also microfilm the documents for permanent retention and have computer listings printed regularly, to give OCI analysts an index to their microfilm file holdings. The use of microfilm in this remote system would significantly decrease the volume of OCI holdings, and the printed indexes would give OCI analysts improved access to their documents. This experiment with OCI was the origin of the SAFE concept (later called Module 1) that production analysts would create their own document index files.

A Project SAFE paper based on these two proposals was published in October 1972. The paper (See Appendix 1) described a set of concepts that, taken together, postulated and partly defined a new Agency-wide information processing system for intelligence materials.

The paper also proposed a data collection period during which production analysts would evaluate the utility (not the cost-benefits per se) and practicality of the concepts. First, the concepts would be partly implemented through test systems (called "modules" in the SAFE paper) set up with existing or easily developed computer/microfilm techniques; and then a representative sample of analysts would work with and evaluate the test systems.

Next 54 Page(s) In Document Exempt

Approved For Release 2002/11/04 : CIA-RDP79-00498A000400050049-9

VI. PROPOSED SAFE INFORMATION SYSTEM OUTLINE

INTRODUCTION

We interpret the analysts' evaluations of the SAFE modules and SAFE concepts as a general endorsement—with qualifications, or reservations. The qualifications, which relate to system reliability, file contents, user aids, response times, etc., are being studied.

We interpret the contractors' evaluations of the technical feasibility of the SAFE concepts as a general endorsement with qualifications. These qualifications relate to the technical difficulties of digitally converting and storing data obtained on paper copy medium; the problems of response time for large files; and the inherent difficulties in the SQUIRL concept. They are being studied and are taken into consideration in the system proposed in this report.

This chapter outlines a proposed SAFE Information System that will satisfy the analysts' two fundamental needs: computer searching of digitally stored message traffic (Text Files) and maintenance of computer-based analyst files.

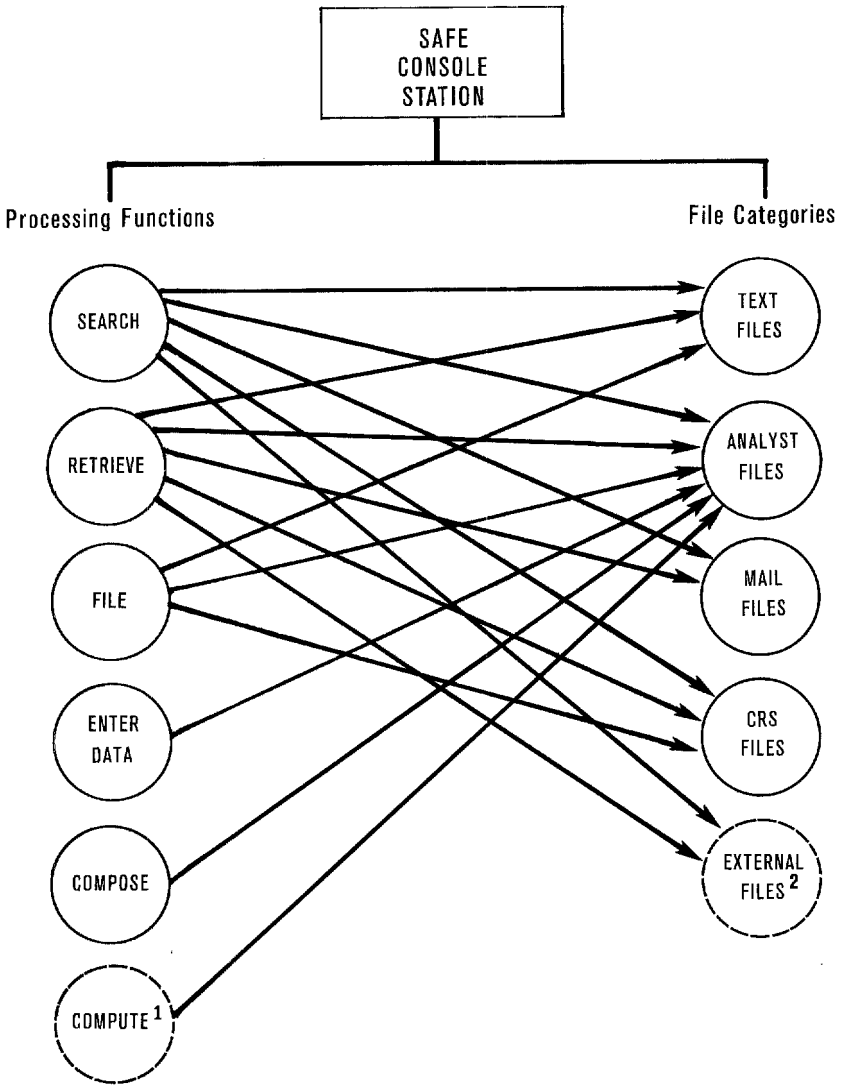
The proposed system resembles that system hypothesized in the SAFE Concepts chapter of this report and described in the Preliminary Design Report (Appendix V). However, because of current technical and cost restrictions, this design differs from the hypothesis in four important aspects:

1. Material received in paper copy form will be stored in microform rather than in digital form. The conversion to digital form is still an objective.
2. An item received by electrical transmission need only be stored once, regardless of the number of analysts who may have "filed" it; but, as a corollary of item 1, material received in paper copy form will have to be stored in as many microform collections as are required.
3. External files, such as the *New York Times* Information Bank, will not be a part of the present system proposal; their inclusion is still an objective.
4. The system response time (time required to complete an analyst's transaction) will vary depending on the size of the files and the "operation" being performed. The original hypothetical response times now appear impractical.

The first step in a system development program would be to design the system in detail; this design would require 4-6 months to complete. The description that follows is in three parts: System Overview of proposed SAFE capabilities; File Operation, which outlines the relationships among the major files; and Preliminary Hardware Design, which includes an estimate of total costs.

SYSTEM OVERVIEW

The system capability can be summarized by describing the SAFE Console Station (SCS), the files it can access and the processes it can perform. (See Figure 19). The SAFE system should, where practical, be integrated into the general Agency data processing environment; a SAFE terminal should be able to access other Agency data bases in addition to SAFE files.



1. Compute functions not to be considered in early SAFE System but is a future objective. Compute would tie the file system outputs into existing (or new) OJCS compute programs.

2. External file not to be considered in early Safe System but is future objective.

Figure 19. Overview of the Proposed SAFE Information System

SAFE Console Station (SCS)

The production analysts will use the SAFE Information System through an SCS. The SCS is more than a simple cathode ray tube (CRT) device. For example, it may consist of a "local" terminal (digital viewing screen and keyboard) stationed at every few desks; a digital printer reasonably close to the terminal; and a "regional" microfilm viewing screen, film storage device and printer. The keyboards will have

function keys that control the file categories to be accessed and the functions to be performed.

The viewing screens must feature readability and general ease of use consistent with today's state-of-the-art. The SCS will be designed with either two screens or a split screen, so that an analyst can view information on one part while entering data on another. The SCS will have an alerting device which will bring a predetermined "priority" message to the analyst's attention. Analysts will be advised automatically of any operating abnormalities.

File Categories

1. Text Files are those electrically received transmissions that may be processed and stored in digital form. They currently include:
   - —FBIS field traffic
   - —SI messages
   - —DoD IRs
   - —State cables
   - —OAKS
   - —CIA/IAS
   - —Military cables
   - —Wire services (Reuters, AP, etc.)
   - —DDO selected information cables

   These items (except for certain sensitive or highly classified items) will be held for 14 days, during which time analysts with the proper clearances can access them for processing and possible inclusion in their own files.
2. Analyst Files are those created and maintained by analysts. They may be document reference files (which contain indexes to specific documents) or information files (which contain data and may or may not refer to the source documents).
3. Mail Files are a subset of the Text Files; each mail file contains a selection of electrically received transmissions that have been processed into it by the Cable Dissemination System. A "distribution index" ties a specific message to a specific set of analysts.
4. CRS Files include the Subject Index File (two million records and growing), a major document reference system. CRS indexers select documents for indexing in this file according to predetermined criteria. Other documents of special merit may be "activated" for the system. SAFE proposes an additional selection criterion, whereby CRS will index any additional document if two or more analysts have "filed" it and if the security classification of the document permits a "public" index record. (The process is described below in the section on Indexing and Filing of Digitally Displayed Items.) CRS files will also probably include certain biographic and installation information files and certain library reference files.

Processing Functions

1. Search—Analysts will be able to perform searches on any of the above files. In the case of Text Files, they may search by specifying any word or combination of words and asking to see the documents in which they appear. The other files will have different search capabilities, but to the extent practical a common language/procedure will guide the analysts through their searching. A search in the Mail File would probably be a simple scan of items

received since the last search. Special aids will be made available to analysts who are unfamiliar with any particular file.

2. Retrieval—Documents or information that match a search parameter can be displayed on the screen and printed at the SCS. The mode of retrieval will vary depending on the file and the file storage medium. Figure 20 shows the retrieval options available.

3. File—Analysts can "file" any document being viewed on the SCS display screen, whether it is a microfilm or digital display. Table 15 shows the file options available. If the document is a paper copy receipt the filing instructions are considered to be in the Data Entry category discussed below.

4. Data Entry—Analysts may create or add to analyst files by calling up the appropriate "form" on the screen and then entering data directly on the displayed form.

5. Compose—Analysts may use the compose function to write and edit. This "document" can then be filed with other intelligence items or in a special project file to which other items can be added.

FILE OPERATION

This section describes briefly how the proposed system will work. For the most part, this description was developed from the outline contained in the more detailed Preliminary Design Report, published in May 1974.

Search and Retrieval—14 Day Temporary Text Files

Figure 21 shows the proposed schema. Digital message traffic is received after being processed through CDS (1) or other OC sources (2). This traffic is processed through the SAFE Automatic Cataloging program (3), which sets up one computer index file record (called the Basic SARDINE record) for each message. The record (4) contains the standard SAFE Number (SANS), classification, date, and file name. Messages in this temporary text file are held for approximately 14 days (5).

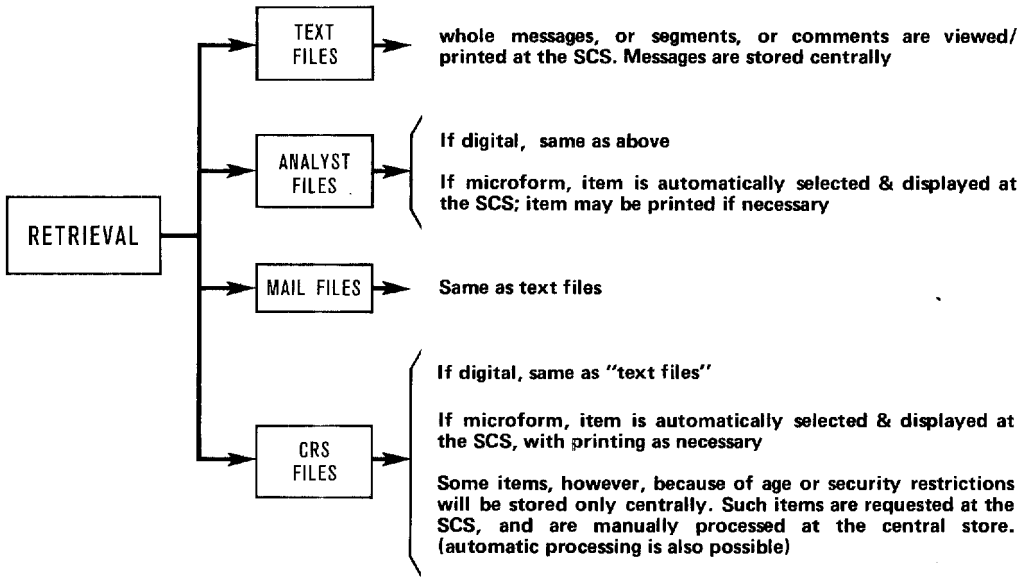


Figure 20. Document Retrieval Options for the Proposed SAFE Information System

Table 15

Filing Options Available to Analysts as They View Documents at the SCS

Filing Option	Description	Applicability
1. Indicate file name.....	Document will appear to be filed under that file name.	Microfilm and digitally displayed documents.
2. Add index terms to the document.	One or more words, or word phrases may be used to further describe the document.	Same as above.
3. Add comments.....	The analyst may add evaluative comments about the document.	Same as above.
4. Extract data.....	Analysts may extract data from the document; whole paragraphs or specific segments.	Digitally displayed data only.

When an analyst searches (6) this file, he may limit his search to any parameter he chooses, e.g., date, post number, security classification, keyword in text, etc. If the number of hits exceeds a certain level, he will have the option of refining his query to reduce the number of hits or having them printed in the OJCS center. Otherwise, he can ask for the whole item to be displayed, or he may ask for only the segment of the item that contains the search terms. He further has the option of printing (7) or filing (8).

Search and Retrieval—Mail File

When a message from CDS is routed into the temporary text file, at the same time (see Figure 22) the list (Distribution Index, DI) of who gets that message is routed to the DI file (2). When an analyst asks to search and retrieve from his mail file, this index determines what messages are sent. The analyst need only ask for "mail" to start scanning the items that have been selected for his office since the last time he viewed his mail file. The analyst can also elect to further route (8) the messages being scanned.

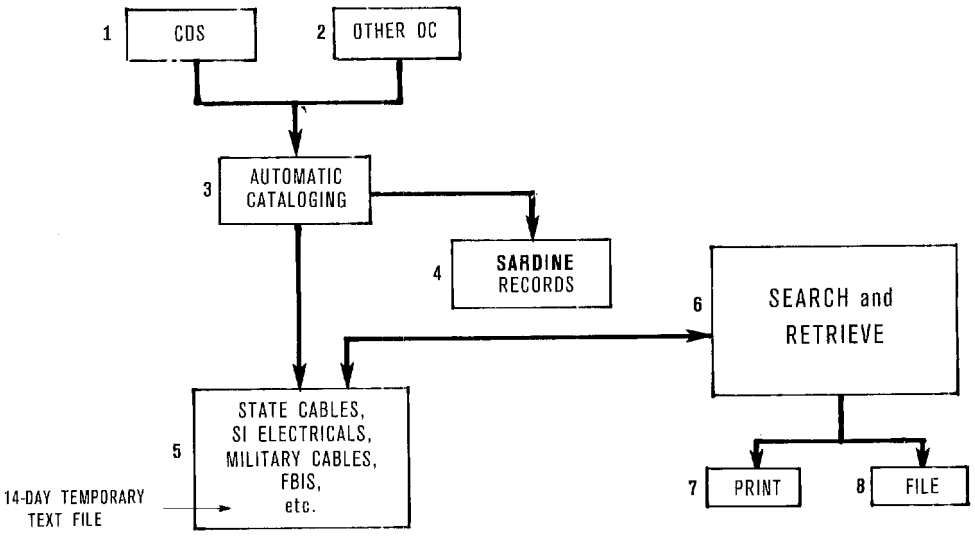


Figure 21. Search and Retrieval From 14-Day Temporary Text Files

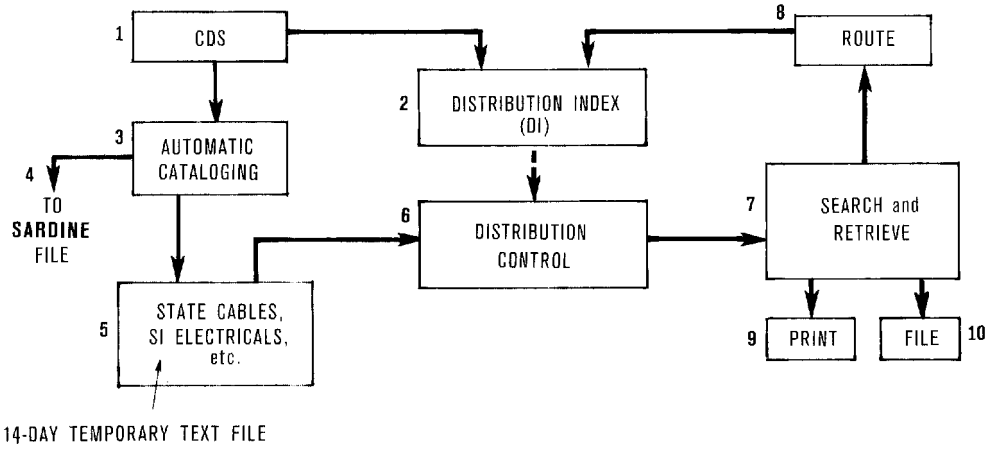


Figure 22. Search and Retrieval - Mail Files

This routing automatically updates the Distribution Index so that it will be available on some other screen—if that analyst has been cleared for the item. Analysts can also print (9) and file (10).

Indexing and Filing of Digitally Displayed Items

The creation (see Figure 23) of temporary text files (4) from OC (1,2) and the creation of the Basic SARDINE record (5) have been discussed above under Search and Retrieval of 14 Day Temporary Text Files. When an analyst chooses to "file" (6) a digitally displayed text item, what he really does is add his file criteria (be they file names, keywords, or whatever) to a record (7) associated with the SARDINE record already created for that item. He may also use a data entry form to create a comments file (8) for the text of comments he wishes to make on the document. When he next retrieves that document, his own comments (but not those of other analysts) will appear with it.

SARDINE relates the proper comment to the proper user and to the proper text document. The above connections are made as the analyst views the document on his SCS screen, and his data entry form is displayed concurrently with the message. If any analyst has added a file sub-record to the Basic SARDINE, it will affect the file reorganization (9), because after 14 days each item in the temporary text file must be moved to another storage area.

If a given item has not been put into any file, even that of CRS, then it is processed via computer output microfilm (10) to a central microform collection (11) or is processed to the lower order digital storage, the Tera-Bit Memory (TBM) (12), which may be an alternative to microform storage. The SARDINE record continues to exist for that item.

If an item has been entered into one or more files, it will be transferred to the primary text file (13). Analysts will be able to do text searching on all items so stored. Items remain stored in primary text until the next reorganization, when the date and activity of each record are automatically reviewed. If an item has not been retrieved for a given period of time, it too will be routed to microform or TBM storage and out of the more expensive digital primary text.

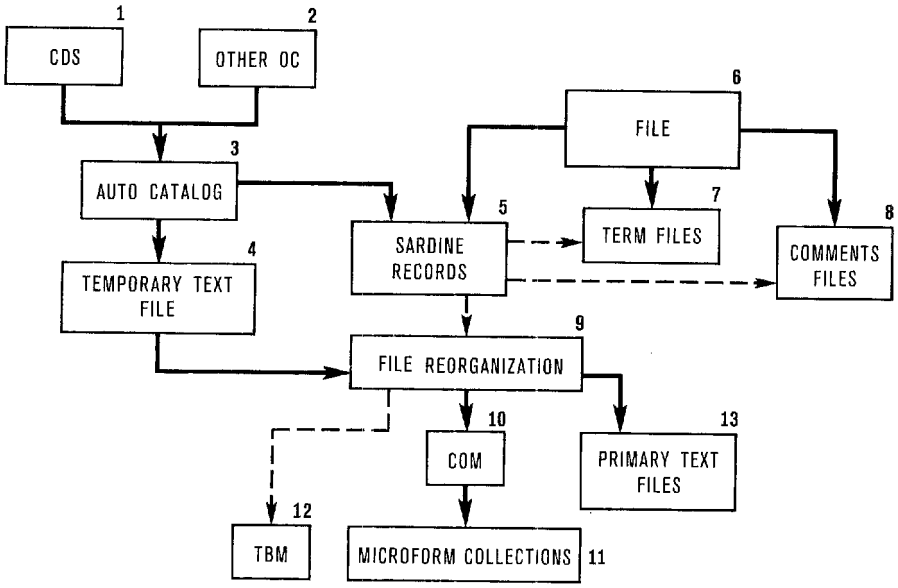


Figure 23. Filing of Digitally Displayed Items From Text Files

Indexing and Filing of Non-Digitally Received Items

In a typical sequence (see Figure 24), an analyst receives a document in paper copy form (1) and reads and marks data (2) that are to be filed. He enters the data on a form that appears as a display on the SCS (3). The particular form is tailored to the kind of file being built. Data so entered goes into term files (4) or comments files (5) as appropriate, and the location is recorded in the SARDINE record (6), which "points" to the CRS microform version of the original document (7). Whenever the SARDINE record is retrieved, it references that document.

An analyst may see only a microform copy of a document. He can still file it by following steps 2-7.

Search and Retrieval—Analyst and CRS Files

When the analyst searches and retrieves from his own or from the CRS files (see Figure 25), he uses various term files (1) and the SARDINE data structure (2) related to them. When the search is complete, he may view the SARDINE records and the term file entries that satisfy his search statement. These may themselves contain the information that answer his question, or the analyst can retrieve the pertinent documents. Documents in digital form are retrieved from a primary text file (3) or the lower-speed TBM (4) device. Once a set of these digital documents (or analyst comments (5) about them) are retrieved, they are available to the analyst in a special computer file called a "work space" (6). Documents thus retrieved can be further searched by text search techniques (7) or refiled (8). Documents in microform are retrieved from the regional storage facility (9) associated with an analyst's SCS.

Some documents will be beyond a given age limitation or will be of a special security category. Such documents must be retrieved from central storage (10). Requests can be made directly from the analyst's SCS; the documents are processed manually.

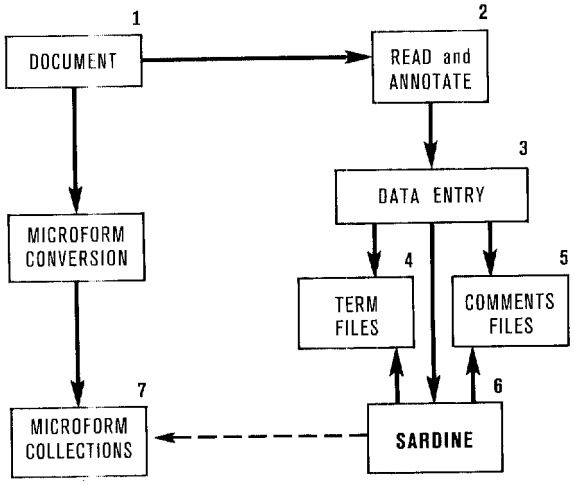


Figure 24. Data Entry

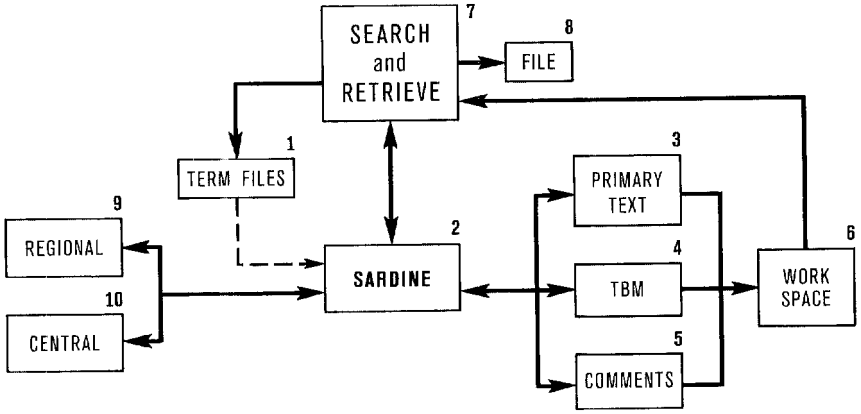


Figure 25. Search and Retrieval of Analyst and CRS Files

PRELIMINARY HARDWARE DESIGN

Introduction

The preliminary concepts of the system design were discussed by a joint CRS/OJCS task team, which had been directed to determine the major parameters for an updated SAFE Information System and to consider how those parameters would influence the system design. Once the parameters were established, the team considered various ways of implementing them and discussed the merits of special versus general purpose computers and of distributed versus central processing. The team decided on a distributed network of minicomputers attached to general purpose computers doing central processing.

The following, more detailed hardware design was made by a team of CRS computer specialists, based upon a consensus of the overall system configuration determined by the joint CRS/OJCS task team. This system design indicates the possible magnitude and cost of a SAFE Information System.

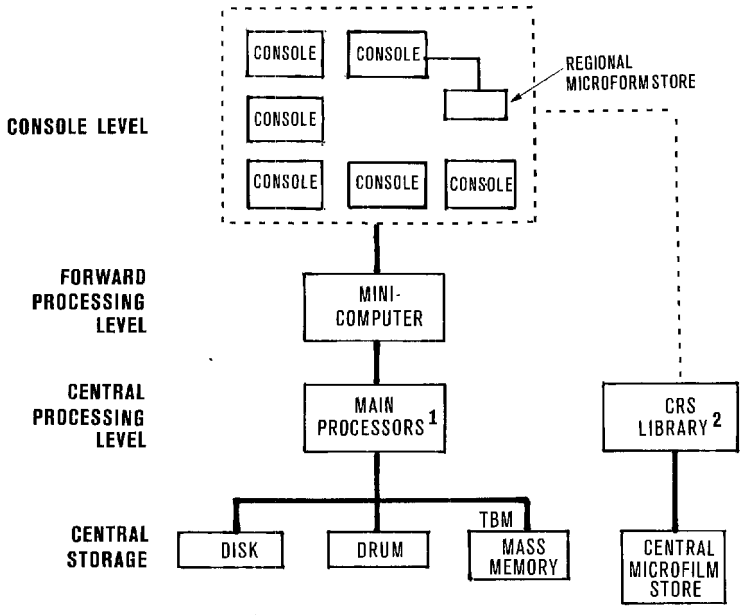
Because many of the SAFE requirements are still approximations, the team considered two possible configurations. The larger and more expensive one might be able to do the job; the smaller and less expensive one probably will not be able to handle peak workloads. Because of the large volume of data that will be vulnerable to both hardware and software failures, file backup and alternate routing procedures will be required at all levels of the system. In addition to backup equipment, SAFE will require processing and electronic file storage equipment to restore service after either an external problem (e.g., fire) or an internal problem (e.g., equipment malfunction) destroys some part of the electronic files in the system.

As exact SAFE requirements are derived, the detailed system design phase of the Development Plan (Chapter VIII) will determine the final system configuration, which will probably lie between the minimum and maximum configurations presented.

SAFE Configuration Description

The proposed system requires hardware for four processing levels: the analyst's console, forward processing, central processing and central microfilm storage (see Figure 26).

- Analyst's Console Level: It is proposed to install some 500 consoles, about one for every two analysts. For every five consoles (approximately) there will be a regional microfilm reader and storage device. This device will contain microfilm images of documents (nonelectrical receipts) that were filed by the analysts and a sub-set of the central (CRS) microfilm storage. The contents of this sub-set will be controlled by security and document age.



1. may consist of two general purpose main frames (small system); or may consist of four special purpose main frames (large system)

2. central processing may remain manual (low-cost system) or may be automated (hi-cost system)

Figure 26. Proposed Hardware Configuration

- Forward Processing Level: It is proposed to station about 50 minicomputers in the Agency, averaging one mini for every 10 consoles. This network of minicomputers allows the SAFE consoles to be less sophisticated and therefore less costly. It also allows the processing of simpler tasks (reading mail, writing and editing reports, and checking syntax of commands for errors) to be accomplished at a level closer to the analyst and relieves some of the work load on the central processors.
- Central Processing Level: The complex computer functions of monitoring the system, text searching, index searching and maintaining the data base will take place at the central processing level. The minimum computer configuration needed is two large (IBM 370/168 size) general purpose computers. All of the functions will be performed in either machine, and each will back up the other. Some members of the task team doubt that this minimum system will have enough computing power to handle the workload, especially during peak periods. The failure of either computer would seriously degrade the entire system. An alternate design uses four large computers (IBM 370/168 size). They are specialized; two maintain the data base and search text files, and the other two search the private and public index files and do text searching of the current 14 day text file. Should any one computer fail, its mate would be able to maintain the function with little or no system degradation. This system is more expensive but guarantees maximum backup and high computing speed.

  In both systems the electrically received data and index files are stored in a two-level storage heirarchy. The primary storage level consists of approximately 75 disk drives (IBM 3330 size) with a couple of fixed head devices used as a buffer. Depending on age and frequency of use, the data will be reassigned to a mass storage TBM system.
- Central Microfilm Storage (CRS): The central storage facility will contain all items processed by CRS as well as some aging items sent back from regional locations because of security restrictions. The minimum system design would continue the present manual system with one additional feature: analysts at their consoles would be able to automatically order those documents not available regionally. The subsequent delivery would be manual. The alternate design calls for automating the central facility so that documents ordered automatically could be delivered automatically. The expense of an automated system might be justified if document requests levied on the central facility were to increase significantly. At present, however, the SAFE plan does not include automating the central microfilm facility.

Hardware Costs

Comparative costs of the two computer systems are shown in Table 16. The price of IBM equipment was used to judge the cost of the main processors and disk/drum storage system. When specifications are better defined, perhaps some other type of equipment of the same computing power could be used. The terminal cost is calculated for 500 terminals. The mini-processor/communication system is based upon 50 mini-processors and the associated computer communication lines. The cost shown for the mass storage (TBM) is not that of a complete system but of an expansion of the system the Agency is currently purchasing. The programming costs include the initial programming of all the software for the system and the maintenance programming needed thereafter. The costs cited do not include the expense of altering existing

facilities to accommodate the new equipment, nor the expense of additional personnel to maintain it. The next chapter will discuss some of the cost savings and benefits associated with the SAFE Information System.

Table 16

System Costs
(In Millions of Dollars)

	2 General Purpose Computers	4 General Purpose Computers
Terminals	5.0	5.0
Mini computers and communication lines	2.5	2.5
Main computers	11.0	18.0
Card reader/punch, printers disk/tape storage	4.0	4.0
TBM—mass storage	1.0	1.0
Microfilm system	1.5	1.5
Software	6.0	6.9
Initial rental for main computer, and total system maintenance cost	2.5	2.5
Total cost	33.5	41.4

VII. COST-BENEFIT CONSIDERATIONS

INTRODUCTION

A cost-benefits analysis of the proposed [redacted] SAFE Information System is not possible at this time. We cannot assign a dollar figure to the potential value of the system to the production analysts for whom it would be built. We can, however, cite the arguments of the analysts that the SAFE system would improve the finished intelligence product by offering new analytic techniques, data bases and data base access. Also we can show that the SAFE system could improve the organization and allocation of Agency computer resources. And we can suggest areas where dollar savings may occur that would at least partly offset the cost of SAFE.

25X1

IMPROVED INTELLIGENCE PRODUCT

The arguments offered here are those made by the analysts in their critiques of the pilot system. They have already been cited in Chapter V but are quoted here, in part, because of their particular relevance.

> "SF/C is a self-help user of the SAFE system. I would venture to say SF/C is more dependent upon SAFE and possibly more convinced of SAFE's indispensability than any other branch . . . SF/C's SAFE system does not merely supplement the branch files; it *is* the branch research file . . . We are striving to establish in SF/C the finest, most comprehensive, most usable repository of all-source information on command and control subjects in the intelligence community. We could not aspire to so ambitious a goal without SAFE . . . Scraps of information of interest to us can be found in all of the file modules being considered for incorporation in SAFE in the future . . . The more files we can dig through, the better chance we have of coming up with meaningful tidbits, and no one can predict where those tidbits will be found. Given the fantastic capabilities of computers, I see no reason to arbitrarily restrict the scope of our search for information by limiting the number of files to which we will have access. We want them all!!! And I promise you that we will learn how to exploit them." (OSR/SF/C comments).
>
> "The most immediately evident one (benefit) is the ability to store and search vastly more information than previously possible . . . A more fundamental consequence is that, with masses of data more easily available, an analyst can bring more evidence to bear on a given problem. Further, the analyst feels more inclined to check his files before writing because he knows it (checking) can be done quickly and comprehensively . . . An interesting effect of having files available on the computer is being able to do searches or use data in ways not previously possible." (OSR/SEC comments.)
>
> "During the Cyprus crisis and more recently in relation to events in the Balkans, I had an opportunity to use the SAFE system in a crisis management mode. The system proved to be an extraordinarily useful device in this respect. The mail distribution system (OLTA) and COLTS were of particular importance . . . SEC was able to receive relevant reports through the OLTA system many hours before the reports were available in hard copy. This capability allowed us to stay well ahead of possible threatening developments

and, in fact, alerted us to potentially interesting developments in the Balkans before reports of this were available through regular channels. I believe that the SAFE system has an enormous potential for crisis management." (Comments of one OSR/SEC analyst.)

"The SAFE system holds the promise of being able to make the ever increasing flow of information a manageable phenomenon, and to help stave off the accumulation of innumerable safes with unmanageable files." (OSR/TF comments.)

"The year-long experiment with Project SAFE has proven that . . . analytic capabilities can be enhanced. The savings of time and space afforded by the system, plus the rapid search capability, represent a highly desirable electronic package." (OER/D/TA comments.)

In summary, we believe that the data collection experiment demonstrated that the proposed system will help Agency analysts provide a better intelligence product. A better product may be a piece of incoming intelligence more thoroughly indexed and annotated for later reference; or information routed to users faster and more efficiently; or a more thoroughly researched piece of finished intelligence.

We believe the SAFE system will offer analysts improved techniques for monitoring and manipulating a large amount of incoming intelligence items, for searching files they could not otherwise use in the time before their deadlines, and for scanning incoming mail minutes after it arrives in the Agency.

In acquiring new technology, the Agency has traditionally emphasized the information collection side of the intelligence problem rather than the information analysis side. As this continues, it resembles building an ever larger cone for a funnel while keeping the same sized neck, and expecting the flow to increase. Agency analysts cannot now digest all the information they receive; they often cannot quickly find yesterday's piece of intelligence when it suddenly becomes relevant today. The task force feels that the development of the SAFE Information System represents the required parallel emphasis on the analysis side of the intelligence problem.

IMPROVING COMPUTER RESOURCES ALLOCATION

Computer and microfilm information systems to support production analysts have often been developed on an essentially individual basis. Each office would set out to meet its particular needs without knowledge of or coordination with other offices with similar problems, and the overall development of the Agency's information system has suffered. Proper development requires a unifying concept that would relate, for example:

—a file building requirement in OSR with one in OBGI,
—a text search and edit requirement in OSI with a text indexing requirement in CRS, and
—a text segment extract requirement in OWI with an automatic cataloging requirement in CRS.

A unifying concept would reveal the relationships between such varying requirements, and enable the task force to derive a common denominator.

Lack of a unifying concept has resulted in unnecessary developmental costs and, probably, unnecessary acquisition of computer equipment.

The task force suggests that the SAFE Information System could be such a unifying concept; that it is wide enough to embrace most of the information processing requirements of the production analysts; and, in short, that SAFE could improve the organization and allocation of Agency computer resources.

POTENTIAL SAVINGS

Savings could follow the adoption of the proposed SAFE methods for handling the Agency's electrical and paper receipts and the proposed SAFE text searching system. SAFE would also change the pattern of CRS use of computers and manpower. These changes are discussed below, but no dollar figures are projected.

Handling of Electrical Receipts

Approximately 20 million copies of electrical messages are disseminated yearly at the Agency. The cost of the existing operation is considerable; the existing equipment, supplies, space and manpower will no longer be needed if they are replaced by more efficient equipment and more efficiently used space and manpower.

Handling of Paper Receipts

The SAFE system plans to continue the current routine microfilming of documents that are received only as paper copy. Instead of keeping them all in a central location, however, SAFE would make a large collection of the microform documents available in regional storage devices and thus lighten this load on the central storage facility. This central facility now manually microfilms documents that were received as electrical messages. SAFE will enable the central facility to receive computer output microfilm (COM) processing, reducing the use of manpower.

Text Searching

During the data collection period analysts used the digitally stored text files to obtain messages that they may or may not have expected to receive through the regular delivery of SI messages, State cables, FBIS field traffic, military cables or DoD IR electricals. Analysts used various parameters in their search of those files and could change the parameters as their requirements changed from day to day. They found these searches valuable:

"I've used COLTS (text searching program) primarily to retrieve messages referred to in other cables but nowhere to be found in our mail."

"COLTS produces messages faster than hand delivery."

The proposed SAFE Information System would regularly update the text files as messages are received from OC. Its improved text search capability will allow analysts to repeat a question without having to reformulate it every time, and to view only titles or segments rather than the whole text, whenever they are scanning many messages for relevant items.

The task force anticipates that text searching will at least partially replace the dissemination of messages to user offices; and, possibly that someday intelligence messages will not have to be read and reread before reaching the ultimate customer. To the extent that shuffling, carrying and reading the mail are reduced, the Agency can save money.

Changes within CRS

If project SAFE becomes an operational reality, it would satisfy most of the present CRS requirements for computer support, as well as some other Agency requirements, and release a significant amount of OJCS resources.

Under SAFE, CRS will continue to analyze documents to create the "public" index record. Some increase in indexing may be required, but we feel money would be saved overall because CRS will be able to use the on-line analysis and automatic cataloging functions. Also, CRS will need fewer specialized analysts for routine reference work, because SAFE will permit production analysts to search many of the CRS files for themselves.

VIII. DEVELOPMENT PLAN

INTRODUCTION

Chapter VI of this report outlined the proposed SAFE Information System, its capabilities, possible hardware configuration, and cost estimates. This chapter describes the development plan of the SAFE Information System and projects the number of developmental phases required through FY 1980 and the expenditure required each fiscal year for the same period. These estimates are tentative and will certainly change as a result of the first phase activity (detailed system design).

DETAILED SYSTEM DESIGN PHASE

In the first phase of the SAFE Information System development, the task force must draw up detailed design specifications. It will have to verify that the system hardware configuration suggested in Chapter VI is correct or spell out the new configuration.

Once the hardware configuration is fixed, the task force must draft detailed specifications on individual components. If the minicomputer/main processor configuration remains the preferred one, studies must be performed to determine the optimum mix of the functions performed by the mini and main computers. The task force must spell out the requirements for the SAFE Console Station and decide whether or not to use existing terminal equipment. The task force must also fix the detailed specifications for the computer software, and determine how the overall project is to be managed.

Task Team

Project SAFE will demand a new task team composed of various specialists. Many are already Agency employees; some must be hired. This team would guide the detailed system design phase and the project management plan mentioned above. It would also maintain the interim SAFE system now in use in the various developmental branches. The analysts who are still working with the pilot system—at their own request—will continue to play an important role as SAFE is developed Agency-wide.

The task team would consist of 13 to 15 full-time analysts from the following organizations:

- CRS/SAS—Six or seven analysts engaged in project management, system design, and interim system management.
- CRS/SSD and OJCS—Two analysts studying hardware configuration.
- OJCS—One analyst, engaged in coordination, would keep OJCS informed of SAFE progress and would seek OJCS expertise as required.
- Contractors—Four or five systems analysts from a major software/system firm to analyze the implications of the expected load and queueing through computer simulation and modeling.

It would also need four part-time personnel as follows:

- OC—One person, familiar with the Cable Dissemination System of the Cable Secretariat, who will coordinate the SAFE requirements with those of the Secretariat.
- ORD—One person who would monitor industrial and academic research developments in areas of interest to SAFE.

Next 2 Page(s) In Document Exempt

Approved For Release 2002/11/04 : CIA-RDP79-00498A000400050049-9