Report of
PROJECT ASPIN

(Automated Systems for the Production of Intelligence)

July 1970

Central Intelligence Agency

FOREWORD

Project ASPIN (Automated Systems for the Production of Intelligence) was conceived as an effort to take a broad look at the employment of automatic data processing to support the production of intelligence in the Central Intelligence Agency. This look was particularly stimulated by the insistence of the President's Foreign Intelligence Advisory Board that the Agency's information activities had not taken advantage of the development of new information technology to support intelligence information handling and the production of intelligence. Thus, although the task was assigned to the Directorate of Intelligence, the task force was made up of professionals drawn from the Directorate of Science and Technology, and from the Office of Communications/Directorate of Support as well as the Directorate of Intelligence. The staff also let a small study contract
25X1A with [redacted] to review the work of the staff and to offer additional suggestions and comment on its findings and recommendations.

The basic objective of ASPIN outlined more than a year ago in an effort to expand the ASPIN concept and terms of reference for approval of the Deputy Director of Intelligence seems equally apropos now. "The basic objective of ASPIN is to develop a broad conceptual design for automatic data processing support to intelligence production. This design should indicate:

(1) the types of ADP applications which may be profitably undertaken,

(2) the relationships among these applications which ought to be preserved in their design, or modification, and implementation,

(3) specifications for the general system(s) which might bring together these processing activities,

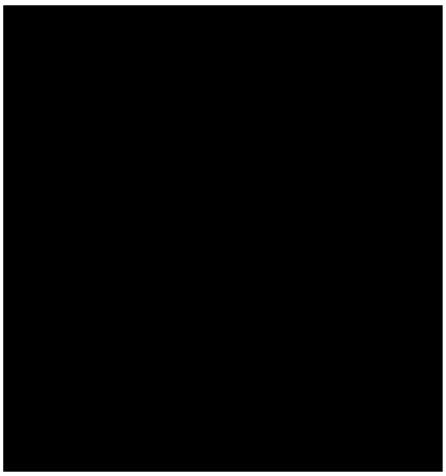
(4) procedures for approval and development of component elements of these systems, and

(5) organizational arrangements for the development and operation of this system."

Project ASPIN was undertaken in early July 1969 and completed on 17 July 1970.

The members of the ASPIN staff were

25X1A



OER/DDI
O/DDI
NPIC/DDI
OSR/DDI
ORD/DD/S&T
FMSAC/DD/S&T
OCI/DDI
OSI/DD/S&T
OCS/DD/S&T
O/DDI
CRS/DDI
OCS/DD/S&T
OC/DDS
O/DDI

CONTENTS

PRECIS

PRECIS

1. The general condition of automation systems to support intelligence production in the Agency is excellent. It is large in terms of total outlay, comparable in general to total outlays for automated support for intelligence collection and for the support of business or administrative activity. The number of applications, their variety in terms of size, scope, and elegance were impressive to the staff, substantially exceeding our expectation of general performance. This is not to say that there are no unsatisfactory applications, that is applications which have failed to realize their objectives. The number, however, is far less than the average percentage of failures in this business; and surprisingly few of the applications which have not been able to amortize the total outlay on the project are thought to produce an output that is not well above the current cost of the application. Moreover the failure rate in terms of inability to develop at least an operational application, once feasibility of an application has been determined, is almost nil. These applications are discussed an enumerated in Part II of the report.

2. The Agency has had its most serious difficulties in trying to implement large systems, a familiar complaint elsewhere in the business as well. Each of the three processing centers serving tne intelligence production components has at least one successful large system in operation, and each of these systems was designed to provide broad support to functional intelligence analysts, two of these centers are presently supporting large scale on-line real-time activities. But each has also experienced the frustration of protracted system development delays in the course of which system objectives are redefined and software shcedules melt away.

3. In general we thought that the overall achievement of automated and manual elements of information processing to assist intelligence production was excellent in quality and well balanced in terms of the allocation of activities among manual and automated systems. For example, we can identify no major information processing activity over the spectrum of offices interviewed that we would claim unhesitatingly should be converted from manual to automatic processing to assure greater economy of operation. The contractor-consultant team from [redacted], whom we employed to review our work and to assist us in identifying important problem

25X1A

areas, concurs in the general findings with respect both to the scope and to the balance or allocation of activity among manual and automatic systems.

4. Data Processing support for existing intelligence production applications is probably the most responsive that it has been for several years. Processing runs are turned around more rapidly. The number and the skill of computer analyst-programmers has increased as has analyst awareness and understanding of automatic data processing. A large, interactive, time-sharing service has been introduced and connections provided to almost all major ADP users engaged in intelligence production. This service provides the analyst direct access to a computer to perform a wide range of computations, to create and manipulate small files, and to provide a wide range of computer programming support. We have presented a brief discussion of the role of OCS in Part III of the report.

5. In analyzing the information processing requirements for effective intelligence production, both in the general section of the report and in Annex I which deals with the intelligence production components in greater detail, we encounter a problem seen before, but it seems to increase as well as persist. The problem is a centrifugal movement of information which appears to accompany increasing acquisition of understanding and increasing specialization or research. The problem manifests itself in the canalization of certain data and information flows in such a manner that they are never received in the conventional information center except in some combined form in which the source data or information is unrecognizable. Thus we believe that the concept of an information or reference center must change appreciably to accommodate this phenomenon. Part IV in the report has been developed to deal with this concept.

6. There will tend to be a considerable number of information centers. The data or information contained in these centers will be a great deal more specialized than what has been sought in the traditional reference center. This in turn will mean that the user must be a specialist or he will require the service of the specialist to search and to understand the resources of such an organization. Moreover, it will mean that many of these collections of information will not be known to a great majority of analysts. The immediate need is to know what exists, where it is located, and how or through whom it may be accessed. The subsequent need will be to structure these collections in such a way as to optimize the productivity of these data among those who must use them.

7. We believe that the most likely approach to operational large integrated systems will be via generalized data managements systems. These systems provide a common framework, yet one with great flexibility, within which to define ones data or files and a common set of programs which may be used to manipulate such files. Most such systems provide the ability to augment such a basic framework by user supplied functions to permit greater specificity of processing. Generalized data management systems provide some other attractive functions as well. They provide rapid access to a computer processing capability against a set of files, a relatively simple file-handling language, and the ability to operate against a wider information base with somewhat greater confidence in the manner in which the data base has been deployed.

8. Having said this, one must add that the development of general data base management systems to date has not produced either the broad flexibility in data manipulation or great facility in information maintenance and report generation expected. A large number of these systems has been built, including one developed by the Agency with IBM contract support. None offers a system that can clearly be identified as superior to the others, nor does any seem to offer a clear promise of flexibility and power that would offer substantial assurance that it could be developed into a system that would provide acceptable large information system management. Nevertheless, we believe there is some value in adopting one or more of these systems as standard for the Agency and getting on with deploying those data which we think may be used rather generally under such control. In the meantime we could use a part of our programming and design resources to push toward the sort of ideal system which one would like to achieve.

9. Even as our most enduring positive impression from the ASPIN study was the breadth of scope and the general excellence of the applications, our most enduring negative impression was that automatic data processing activity was poorly managed in spite of its accomplishments. This impression was conveyed to us independently by the individuals and organizations to whom we talked and by our consultant-contractor team. The nature of the complaint also was much the same, that there was not enough central planning, central coordination and central control of ADP activity. None of the critics had any brief for central control per se but essentially all were agreed that there was too much atomization of authority in this area. The general problem of management control is discussed in Parts V and VI, and in Annex II which contains the working papers of the contractor-consultant.

10. A number of forces tend to atomize control even as a number of others tend to concentrate it. ADP embodies a new technology; there is always a scramble to control a new technology by existing activities as well as by its practitioners. Problem solution tends to drive control to the organization or to the individual with the problem, but this is tempered by the need for an interpreter who can cast the problem in terms of the machine's quite different logic and function. There are still other claimants for control in the data or information collection and recording business, and in the data processing activity itself. The mucilage that keeps reintroducing an element of centralization of control is

1) the desire to have our information as compatible or mutually interactive as possible, and

2) the very considerable economies of scale in data processing.

11. Information is frequently developed by more than one organization and it is almost universally used for more than one purpose. Thus, there is a vested central interest in serving as many different users of a given data set from a single source as feasible. This is particularly true of data which are received in a machine readable format because once data have been released from that form, the cost of reconstituting them in machine form is usually as great as it was to create them initially. Lacking some central means of coordinating information processing activity, there will be widespread development of essentially the same data and similar interpretive computer programs to capture, store, retrieve, update, and publish them.

12. Economies of large scale operation in computer processing have been a factor of lesser importance in some periods. But with the rise of essentially automated computer systems to which the user is offered easy access by remote control, large scale operations have again become a major factor. Large hardware cost to provide the speed and redundancy required for reliability in on-line processing must be combined with large expenditures for programming experts who can assure that the programs under which the machines and these remote communications systems function are equally reliable. With costs of $2 million to $4 million per year for highly reliable systems likely to be commonplace, there will be a tendency to want to assure that these kinds of speed and reliability are really necessary and that these capabilities are effectively used once they are obtained.

13. More effective control of Agency ADP activities seems to need some relatively simple management instruments. One critical instrument is a more complete and more articulate and formal documentation of computer processing projects. We have seen explicit project proposals prepared in the Agency that are models of specificity and clarity. They are unusual. Given better, more specific presentation of automation projects, there needs to be a continuing systematic review of these activities by responsible managers, including centralized review of large projects and projects which impact on more than one directorate. Finally, data processing activity is sufficiently interchangeable with money and people in terms of operational results that an effort ought to be made to allocate these costs to users through the traditional budget and program process.

14. The presentation of the ASPIN report separates a series of general findings and recommendations into the main element of the report. Detailed statements of the role of automated systems in the information processing and intelligence production components of the Agency are presented in Annex I. This annex also contains the specific recommendations and conclusions that may be appropriate to the office level activity. Annex II contains a series of consultant-contractor working papers which we commissioned be done to assist the staff in understanding certain problem areas that emerged in the contractor review of the ASPIN documentation. These papers should be read in the context of problem identification rather than a balanced discussion or presentation on each of the subjects.

15. The recommendations and conclusions of the report have been summarized at the close of each of the segments of the report (and each subdivision of the Annexes) in an effort to keep them as close to the relevant discussion that led to their development as possible. We hold no brief that the recommendations of this report are the only means which may be used to attack the problems identified. They suffer the disadvantage, or enjoy the advantage of some distance from the problem and the immediate responsibilities for these activities as they were prepared, and of preparation by a very catholic group representing a wide range of the activities with which the report is concerned. A desire to restrict the volume of the summary report led us to exclude most alternative solutions inasmuch as we felt a need to lay at least as much groundwork in discussion in the report for such alternatives as had been given to the principal recommendations. We were less concerned about space in the annexes and where we thought viable alternative solutions existed, they were included.

I. BACKGROUND

I. BACKGROUND

1. The use of automatic data processing (ADP) in the Agency has been both imaginative and pedantic. Some of the computer applications developed in the Agency were first efforts in the areas in which they were applied. Others drew upon ADP experience from analagous research activity for an automatic processing concept which was then adapted to the intelligence problem. Still other processing has drawn directly upon standard programs used by industry and government, adapting those programs to the internal data processing requirements.

2. On the whole we believe that the Agency has broadly encouraged the use of ADP techniques in its analytical and its collection role as well as in administrative activities. The Agency was not a pioneer in the development of automatic data processing equipment. Some of its production and data processing analysts were using computers at other installations in the mid-1950's to supplement indigenous mechanical calculation equipment which was incapable of providing timely solution to the complex mathematical problems which they were beginning to formulate. With the advent of the use of computers by Agency analysts, the Agency moved quickly to acquire an internal computer processing capability. Agency management sought to encourage its operational managers to explore the potential for use of automata to support all levels and all types of Agency activity.

3. A wide variety of ADP applications emerged from these missionary efforts. Some were initially successful and have been systematically increased in scope and analytical power. Others were initially only partially successful, some growing fitfully to present day utility, others disappearing because the successful elements were of too little help to defend the time spent on them. Few computer applications disappear, however. Many applications, only partially successful at the outset, hang on because the user has a heavy investment of his time in them and the user -- not required to account for anything but his own input to the program -- finds less disutility from continuing to exercise the program occasionally than he does from admitting that the application is not productive.

4. This sort of history of ADP activity is quite representative of other governmental, industrial and commercial use. Indeed, with pardonable pride, we believe that the Agency can probably show a higher percentage of outright successes and a lower percentage of clear failures than most organizations with a comparable ADP commitment. Our observation indicates that the Agency management has been generous in its support of the development of ADP activity in the Agency.

5. The business of the Agency requires that it be aware of and actively engaged in a wide range of intellectual and operational processes. In only a few of these, such as the development of para-military capability, personal biographic files, the development of targetting or search strategies, and the development of long-range orbital predictions has the Agency a major interest. In none of these activities is the Agency unique. Thus our ADP activity tends to be a microcosm of many different large scale processing activities and individual applications reflecting or drawing upon the state of investigation in a wide range of fields. Data processing innovation here is confined to the development of algorithms which can exploit an exceedingly limited set of data to provide an acceptable reflection of reality or which can reduce a host of only partially relevant information to the same end.

6. Be it said that we have talked to few people -- no one engaged in the production of intelligence -- who have been denied funds for the exploration, development or active implementation of computer processing support whether or not the individual had any real evidence of the economic viability of his proposal. Nor do we find fault with this generosity. Faced with a new technology for processing information and being engaged in the information business, Agency management would have been derelict to do otherwise. We find that Agency use of ADP to assist analysts in the production of intelligence has been limited thus far only by the difficulty of identifying and translating elements of that process to automatic systems. There has been no visible effort to restrict the use of automata to assist this process. Indeed, many proposals for incorporation of automatic assistance to this process which were of highly questionable viability were carefully reviewed before they were discarded.

7. On balance we would characterize the analyst evaluation of the support provided by ADP systems as being generous rather than restrictive. Few analysts believe that their analytical capabilities have been greatly enhanced by ADP assistance even

where this assistance is very substantial. Many analysts profess significant ADP enhancement of their analysis. They are quick to add, however, that "all" ADP has given them is a capability that they simply could not physically exercise in an acceptable time frame. They convey the impression that ADP has not had an important impact on their intellectual or conceptual process. (Perhaps this is a reason why we do not use it better than we do.) Rather it has affected their physical ability to compute (to handle) large, complex, mathematical or logical formulations. Where ADP applications support an analyst by performing functions he has hitherto literally performed himself, few if any of the analysts seem impressed with such support. Apparently the time saving in these circumstances is not great nor does the additional accuracy achieved by machine processing seem to have impressed them appreciably.

8. An analyst will tell you that the greatest contribution ADP has made to the analytical process is to increase his confidence in the reliability of his final estimate or judgment. Large-scale, high-speed computation means that he can test a wide range of possible hypotheses before presenting a judgment where heretofore the best he could do was to partially test the most likely hypothesis which occurred to him. Analysts, however, are inclined to discount the value of this change because their bosses and the final consumers of the intelligence product seem no more -- nor any less -- willing to accept the present product than they were the former one.

II. COMPUTER APPLICATIONS IN SUPPORT OF INTELLIGENCE PRODUCTION

II. Computer Applications in Support of Intelligence Production

1. Computer applications designed to support the production of intelligence may be characterized in several different ways. First, they may be distinguished as general or special applications; and, second, they may be distinguished in terms of the function performed by the application -- whether information storage and retrieval, calculation, data reduction, or modelling and simulation. The initial distinction is the most discrete, although some general data processing applications are so flexible that they can support essentially private or special applications within them. Functional distinctions between systems tend to be the most blurred. Both private and general applications often contain more than one type of functional processing.

2. We have attached, as an Appendix to this chapter, descriptions of the computer applications developed to support intelligence production which are operational or in an advanced stage of development. With few exceptions, these applications are special or private applications. Fewer than 6 qualify as general systems. The list of present applications has been developed as a brief to present each of the activities, by office of origin, and give some measure of its relevance and of its cost. Such an effort is necessarily imperfect but we have for the first time collected and characterized these activities in one place. Moreover we have permuted the list to present the projects by functional array and by relevance and cost. A more detailed statement of the methodology is presented with the Appendix.

Functional Applications.

3. We have generalized a wide range of computer applications by attaching certain functional titles to them. In broad terms we believe that the layman can get both a simpler view of the total computer processing activity in support of intelligence production and some general criteria about the difficulty and the utility of existing or proposed applications. We again warn that these are generalizations, that they depend heavily on an "other things being equal" sort of assumption. Few applications would meet the functional type in pure form. Some applications contain such a balanced mix of two or more of these functions that it would be feckless to

characterize them as "a" functional type. Distribution of the existing applications by function in the Appendix to this chapter produced only a handful that were considered so mixed as to be listed under such a heading.

Calculation Applications

4. Calculation programs tend to be application specific and except where they deal with closed systems tend to be unique to the analytical requirements of one or a small group of users. We believe that the observed Agency policy of adapting where available programs written by other analytical centers dealing with analogous problems is both substantively satisfactory and economic. It capitalizes on existing intellectual effort yet is responsive to the specific data characteristics of security environment and to the accuracy required for intelligence analysis.

5. The most productive applications have been those which execute large calculations. Application design and programming time may consume many man months or years, but they seldom consume more time than would be required to hand-process the data for one full solution. Computer solutions to these same problems tend to require seconds or minutes, providing the power and time for several iterations of the problem using different values for critical variables. Calculation programs require high analytical skills to design and to understand and implement in terms of data inputs and outputs. Thus, an analyst may spend almost as much time trying to take an existing program and get it ready for operation in his environment as it would take him to write the program were he sufficiently qualified. Once such understanding has been developed in an organization, implementation is simple and straightforward and data can be extracted from observations and fed to the program for solution by junior professional or often even semi-professional personnel.

Information Storage and Retrieval Applications

6. Large information storage and retrieval systems which require routine computation for processing are probably the second most productive set of applications. They differ from calculation problems principally in their requirement for large scale data preparation and storage, and greater complexity of design. Thus, these applications tend to be considerably more expensive to develop and to require long lead times. Once in operation, they continue to require substantial maintenance expenditure although their expenditure for computer time may be less than for strictly calculation type

applications. Large information storage and retrieval systems with routine calculation are directed at tasks which simply would not be undertaken without a computer.

7. These applications have usually been developed with a unique set of programs. Some of the recent efforts to develop generalized data base management techniques for large management information systems (MIS) have begun to display the sort of file definition, computation routines, and file handling logic which may produce more generalized solutions to this class of application. It is particularly necessary that such generalized systems be carefully analyzed because both our largest ADP expenditures and our highest application risk rate are met in the development of these complex systems.

Data Reduction Applications

8. Data reduction or pattern recognition applications tend to be present in any general processing activity. They tend to be heavily concentrated in the data collection and data processing sectors, however. They consist basically of comparing a large flow of binary data which may have been converted from text, accounting or statistical information, acoustic or electrical signals, etc., to a mask or stencil which represents the interest profile(s) of the particular operation. Data reduction extracts those elements of a data stream relevant to a particular activity and passes them to that activity where they may or may not involve further application of automata. The productivity of data reduction systems varies widely. Productivity in these systems is an inverse function of how elaborate the mask must be to reproduce the manual - mechanical process of recognition which it replaces. The human brain aided by its sensory system is a particularly effective processor in this situation if not always an attentive one.

9. Data reduction schemes which depend essentially on sampling a particular environment and producing a series of average values for various characteristics of the data are highly productive and require relatively moderate periods for design and programming. Incidentally much of the design work may best be devoted to formatting and conditioning the data which is to be presented to the computer. Data reduction schemes which seek to process natural language have been largely wasteful and have had to be rigidly confined to a limited range of functions to show any economic advantage over human processing. The principal advantage a computer can demonstrate in data reduction is an unusually low error rate which may make the

computer more attractive for those systems which must be essentially failure-free. We must emphasize again, however, the relative advantage of signal processing versus natural language. Data reduction schemes are often the life-blood of intelligence processing. The necessity to resort to clandestine means to collect particular information often leads to "broad-band" collection of information by a necessarily indirect avenue of approach and to the acquisition of a great deal of irrelevant data along with that which was sought. Economic extraction of the target information then becomes a sine qua non if we are to achieve the objectives of these collection systems.

10. We are impressed with the need to emphasize the early and complete immersion of the individuals who do the data reduction and analysis in the design of new collection systems. We have seen the beginning of recognition of this proposition in the coordination of the [redacted] collection guidance and collection format programs with the data processing people who would process the input to and output from these systems. Early access to the proposed format for collection and reporting makes possible changes in that format to accommodate both data processing systems and analytical techniques which would otherwise be overlooked. Perhaps more important, early coordination makes possible the orderly design and development of the automatic and manual capabilities required to process the output of these systems once they become available. Lead time will be long for large systems and is as important to effective processing as it is to effective collection. 25X1A

Models or Simulations

11. Related to the information storage and retrieval application, yet basically different, is the mathematical model or simulation. The objective most often sought in the simulation or model is to develop a forecast of future events or to try to develop an estimate of a past occurrence from data which would be related to the phenomenon but would not describe it specifically. Thus, this class of applications seeks to provide or to generate information which we do not presently possess. Its relevance to intelligence analysis, which must seek to understand existing deployment of activities in detail and to anticipate which action may develop, is apparent. The generally high ratings accorded these applications in the Appendix attest to their importance.

12. The model resembles an information system in the sense that it contains a considerable body of information on the phenomenon which it has been designed to represent. The model tends, however,

to contain as little information as is possible to provide an operational representation. It is a discrete system, although it may seek to represent a complex phenomenon such as a nuclear exchange, or an entire economy, or a missile trajectory to name a few operational models presently being exploited. It might represent a political interaction between 2 or more states, or blocs of states. The model seeks to simplify and reduce the data required to reproduce the process. One then exercises the model, varying the values of those constituents which we opine may be varied in the real system, in an effort to estimate the value of certain unknowns or to forecast future values of certain critical elements of the system.

13. The computer has been instrumental in permitting the operational solution of increasingly complex models, under an increasing range of constraints. Its large memory and its high speed computation make such solutions possible. It is necessary for man to provide the understanding of the process and the data that are necessary to arm the computer for this process.

Special Computer Applications

14. Most of the established Agency applications devoted to assisting intelligence production have been implemented as special or private applications. Each such application creates one or a series of files which are processed by a set of programs written specifically for this set of files and this application. Some of these programs have been developed by other organizations to solve similar problems. They must nevertheless be adapted (sometimes extensively redesigned) to meet the unique data or processing environment of the Agency. Individuals engaged in these applications have specific analytical objectives which they must achieve. They seek the most effective and accurate solution that they can develop and implement. The result is usually basic differences in data formats and programs to process similar data. Such differences distress many observers of this situation. Those who have redesigned programs and data are distressed that others cannot recognize the unique situation they must face and the superiority and validity of their unique solution. Those who observe these events at a distance see only the essential similarity of data, and the waste of scarce systems designers and programmers on almost identical operations.

15. We can offer no pat solutions to this sort of occurrence. Some of these apparent duplicative activities are essential to the quality of performance that the analyst and his supervisor believe are required. Others of these activities are clearly what they seem, duplicative activities; some are occasioned by failure to coordinate

diverse and dispersed activities, others are occasioned by a decision to begin from scratch because only then will they have effective indigenous control of the application. A responsible, technically qualified staff officer who has the confidence of the highest levels in the Agency can create a climate in which this type of dispute can be resolved in a manner which makes optimum use of computational resources. We have proposed such an approach in the section on Organization and Management.

16. The intelligence production analyst has developed a large number and a wide range of computer applications to assist him. Most of these computer projects provide the analyst with an effective and often highly personalized information system. Some are no more than detailed bibliographies of an analyst's or a group's collection of literature on a particular subject area. Others may provide detailed time series or detailed data observations from a particular event. The only opportunity for development of greater interchange among applications of this sort rests with the computer processing organization. The systems analyst-programmer may use a single program with variants to process a number of applications. He may participate in the development or acquisition of a general data management system, a collection of programs which will permit a wide range of computer processing operations on data which can be expressed to the system within specified constraints.

17. The analyst, whether he has an ADP application to support him or not, will accept and use an additional ADP service if it provides better and more responsive support than he has built for himself. Only with the direct participation of representatives of the best analytical talent will it be possible to furnish a system that has any chance of providing a useful ADP system for analyst support. We are faced with a dilemma. In the area of general data management systems, one needs to provide software to which as wide a range of users as possible can conform. To build a system that performs well in such a variety of applications unfortunately is to build a system which performs each additional set of jobs less well. Thus unless an application is a very common one, it is not likely to be well served by such a system. Similarly, in the situation in which an information support function wishes to employ ADP to improve and expand its customer services, the only way to build a more effective and more attractive system is to involve the user in its design and eventually in some feedback loop. The user, however, is usually only interested in assessing a finished product in which he has invested little but his evaluation.

General or Large Computer Applications.

18. One central theme which emerges from our analysis and evaluation of analyst sentiment toward large, general computer applications (e.g. CHIVE, COINS, IIS, etc.) is that these systems must provide the functions which he wants to perform if he is to be expected to use them.

Analyst specification for the System.

19. The analyst wants to be able to specify what functions the system will perform for him. Whether or not he achieves any considerable role in influencing the system, he will attempt to use it or to get a colleague's opinion of its performance to determine whether or not the system will do what he wants to do. The designer for such systems listens to the customer descriptions of what is desired. These descriptions are diverse and contradictory. Under the best circumstances, the designer attempts to build a highly flexible system which satisfies a broad area of the demand expressed. More often he seeks to build an elegant system that will respond to what he thinks the customer would need if the customer were more rational and better understood how to do his job.

20. Note is taken above that the computer analyst and the research analyst often have difficulty communicating with each other. A system built to do what the programmer thinks the research analyst wants to do may be of little help to the analyst. The systems analyst-programmer looks upon a query language (for user interaction with the computer via remote console) built of Boolean expressions as the quintessence of clarity and facility. The research analyst often considers it a nuisance which leads him to frequent errors of expression which impede his problem solving.

21. The user analyst is occasionally, perhaps frequently willing to identify the elements which he requires from an automated system or applications. He is more often reluctant, if not adamantly opposed, to participate in the detailed development of the system. This opposition stems from several factors. He already has an information system. He supports it with whatever resources he can bring to bear. His system usually requires considerably greater depth in indexing and discrimination than a "general" system. He has an enormous job trying to produce the various reports expected of him and at the same time assuring the maintenance of his information base. Any effort to reorganize the system immediately imperils what he has achieved. The general system provides the analyst no assurance that it will furnish him with one equal to what he has, yet takes away the time he has to support that system.

Analyst Interest in Fast Response Systems.

22. The research analyst or the analyst engaged in data reduction has been more interested in an on-line, time-sharing system than he has been in the traditional batch processing activities. This may be simply the culmination of his effort to secure faster trun-around on his jobs; many observers we talked to believe that analyst interest in the time-sharing phenomenon may be transitory. We believe, however, that there are some basic appeals to the analyst in these systems which, if they can be technically supported at an acceptable cost, will provide a sustained, widespread use of this system. There already is a clear demonstration from observation of NPIC/AID as well as OCS Interactive Services that established users of ADP service can and do make wide use of time-sharing. Many analysts who wanted powerful computational assistance on an ad hoc basis constitute one body of new users. Many research and data reduction analysts have wanted a facility to create, change and rearrange files on-line in an effort to escape what they regard as tedious and unnecessary steps in present file creation and file manipulation techniques. Some have achieved this capability to a limited degree; most of this use is in the offing. (computer programmers have made particularly effective use of this capability both in developing and maintaining the program instruction on which they are working and in developing files of test data for checking out their programs.)

Analyst Exposure to Large Systems.

23. The analysts engaged in intelligence production in CIA have been exposed to at least 3 big system concepts eliciting their support: CHIVE (a large information storage and retrieval system to support what is now the Central Reference System), COINS (an experimental information storage and retrieval system to explore the use of time-sharing and file-sharing among the components of the Intelligence Community) and the Office of Computer Services (OCS) Interactive Services (an on-line, time-sharing computer system to support the general set of users served by OCS).

24. Each of these systems promised a good deal more than was delivered but there has been a tendency through time for the performance to get closer to the promises. The last of these, the OCS interactive system, succeeded on 2 out of 3 of its major objectives, and the system has recently been expanded in a further effort.

CHIVE

25. CHIVE, the first of these systems, began life as the design of an ultimate information system which would extract relevant

information from all incoming documents, capture these data in machine formats, store it and retrieve it on demand in response to analysts requests. These objectives, although they were probably technically attainable, exceeded the human resources which the Agency was prepared to spend to index and abstract the documents. In retrospect, they also exceeded the processing costs which the Agency would have been willing to spend to process the records as well. The CHIVE project was reduced in size and scope until it was consummated in a system called AEGIS after the general data management programs purchased to process the system.

26. AEGIS was inaugurated with little fanfare, at small cost for equipment or programs, with a sharply curtailed staff, and it worked. It unquestionably owes some of its success to the enormous development effort which went into the CHIVE project. Analysts familiar with the system have come to depend on it to provide a reference service they formerly received from a more expensive accounting machine system. Most consider the present service superior to the old, a better product in less time. Still, many are unfamiliar with the system and short response time (less than 24 hours) to secure both the computer index and documents is only available for the most urgent request or for requests which generate an exceedingly small system response.

COINS

27. COINS has been of almost no use to the production analyst although it represents something of an achievement in concert of community action on a collective data processing experiment. We believe that the procedures used to support and perpetuate COINS will seriously delay rather than hasten the advent of an inter-agency system. The emphasis on the development of automated files and their processing in a large computer network as a goal in itself is a highly questionable procedure. But it is exceeded by the notion that this process should continue until it is successful. Certain of the premises of COINS with respect to the technical achievement of time-sharing systems and of the identity and duplication of intelligence community files appear to have been seriously in error. The COINS effort seems to have generated a life of its own. It has been an exceedingly expensive effort for the Agency, and it promises to become even more expensive should the Agency have to dedicate an entire computer system to it as additional evidence of good faith.

OCS Interactive Services

28. The OCS Interactive Service System represents a large scale effort to provide a general, on-line, time-sharing service to

the Agency. The system was developed from 2 lines, one an internal experiment in developing a time-sharing monitor (called TSMON) to create, maintain and process information files. To this was added a commercial time-sharing system referred to as CP-CMS designed to provide an interactive programming, text-editing and program test (debug) capability. This system has been augmented by a highly versatile mathematical programming language (APL) which effectively converts the computer into a powerful desk calculator able to perform in a few minutes, including time taken to enter the identification and the values for the data processed, a calculation which might require months of time on a mechanical calculator. The combination was installed on a large IBM System 360/67 computer.

29. The system has been primarily used to support programmer activity. Information storage and retrieval activity, much of it on small files, ranks second in importance and calculation activity third. Large files, using the TSMON programs because CP-CMS has no facilities satisfactory for economically reconstituting and querying these files, have suffered serious degradation in response time over their previous performance on smaller, slower equipment in an experimental system. This degradation has become sufficiently serious that some of these files have been removed from the 360/67 processor system and installed under TSMON on a 360/65 OCS machine previously devoted wholly to batch processing.

30. We concur in the OCS judgment that it is necessary to provide a satisfactory on-line, time-sharing service for large information storage and retrieval files. The large, on-line, missile and space file (MISTAC) is essentially unavailable at present because of development work undertaken on it in an effort to get faster response from the on-line system. A large ground force file (QUIKTRAK) is being operated experimentally with considerable assurance that part if not all of this system will require an interactive environment in the near future. The AEGIS document index system seems to us an excellent candidate for operation in an interactive environment in the next two to three years.

Problems in Computer Applications.

31. The principal problem of computer applications in the Agency and elsewhere is the delay between the conception of the application or processing system and its implementation or delivery. Increasing maturity in ADP has in recent years made some reduction in the gap between announced plans for the availability of computer equipment and programs and their delivery, but there remains a serious gap which erodes the credibility of would be ADP customers. Interrogation of the recipients of the ADP applications outlined in the

Appendix to the chapter identify this lag as their principal complaint or warning to the unwary. Some would not have undertaken the application they presently use had they been aware of the difficulty of this delay. Most regret the delays but consider the application which has been built as well worth the effort and are now anxious to improve its operation.

Delay in Project Implementation.

32. Delay in project implementation stems from a number of circumstances, the most important of which are:

1) lack of an unambiguous definition of the objectives of goals of the application,

2) lack of appreciation of the difficulty and complexity of producing data inputs for the computer,

3) difficulty in communication of the characteristics of the data and the procedures for data manipulation and data output between the substantive analyst and the computer analyst, and

4) lack of a central point of project control; the user has control over indigenous office resources but no control over OCS resources.

33. We believe that the lack of clarity in definition of the objectives and control of ADP applications, projects or systems is a management responsibility. This is dealt with in the Appendix to the Organization and Management section.

Problems with Data.

34. Until an analyst has worked his way through to -- and begun to support -- a sustained operating computer application, he can seldom be convinced of the complexity of data preparation, data formatting, and data testing. Every user attests to this. It seems to us that anyone seeking to develop an ADP application ought to be carefully apprised of this by the systems analyst-programmer with whom he works. We are confident that this is done in most cases, but each time the systems analyst-programmer is either optimistic or finally intimidated by the customer's certainty that his data, unlike that of other people, is straight-forward and unambiguous.

Reality is usually the contrary. ASPIN has data output from established, operating computer applications where the failure to apply sufficiently rigid review over data entry on security controls has produced the very security violations that the data and software search programs were designed to prevent.

Problems in Analyst Communications.

35. One aspect of the difficulty in communications between the data processing and the substantive analyst which bedevils most of the applications cited below is illustrated by the data preparation and data description problems outlined in the preceding paragraph. The data problem tends to be dwarfed by the failures which characterize communications on procedures for manipulation of data and data output. The frequent failures of the substantive analyst to specify fully the procedures required for processing is often accompanied by the data processing analyst assuming that he understands the activity sufficiently to substitute for such an omission. The good data processing analyst can usually bring off an operational (or at least partially operational) application with such incomplete specification. Subsequent extension or modification of the application within the framework which the substantive analyst assumed was included in the program precipitates the communications gap if one or more of the major objectives have not previously raised it.

36. The talking past each other and mutual recrimination between two individuals which arises from the failure to develop a computer application which meets the objectives of the user -- whether expressed or not -- creates little data processing support. It produces, rather, damage and distrust. No matter how delicately stated the different positions are, they amount to mutual denigration of the skills of the participants. These proceedings reinforce previous delays and enlarge the area of dissatisfaction and distrust.

37. An active effort to reduce these communications problem is long overdue. There are doable things to ameliorate this problem discussed in the sections dealing with OCS and Organization and Management. We believe that the identification of the problem, and recognition of the common goals of the participants is an additional avenue of attack on the problem. Assignment of applications programmers to intelligence production components and hence to the organizations which must ultimately be responsible for system performance or output will provide a more acceptable climate in which to address this problem.

An Effort at Reducing Communications Problems.

38. Our observation of the general effectiveness of ADP applications indicates that there is a close relationship between the amount of indigenous effort at understanding and direct assistance to the application and its success. We note that small staff efforts such as that undertaken in the Office of Economic Research (OER), initially by one person, may be highly effective. The OER effort concentrated at first on training a small group of interested indivaduls in the rudiments of computer programming and in the application of known quantitative techniques via the computer to specific analytical problems in their individual components. As the staff was expanded by one or two, additional time was spent on advising analysts who had operating applications and helping them improve the scope and efficiency of these applications. The trainees were often used in the latter activity as well.

39. After about three years of intensive effort at improvement in the use of data processing, OER has witnessed a sharp improvement in the output of its established ADP applications and the development of several new fully operational applications. In addition, there exists now an excellent ad hoc capability to take a complex quantitative problem, program it (many times using established routines developed for general use or for other special applications), test it on the computer and solve it all in a few hours or a few days. This Office is one of the large and successful users of the OCS Interactive Services System as well.

40. Individual leadership is important to the success of a venture of this sort. Office leadership must be satisfied that ADP offers a substantial contribution to production. Staff leadership is equally critical. An individual who has both a high degree of professional research and production experience and a sustained operational experience with ADP activity or applications has the highest probability of success. Individuals with this combination of talents are still rare and expensive, but they are worth the time and cost it takes to find them. The specific case history is not presented for invidious comparison. This type of staff assignment has been filled with obvious skill by senior substantive analysts in some components and by senior systems analyst-programmers in others.

Time Constraints on Intelligence Production Principals.

41. Another difficulty in the development of seriously needed computer applications is the limited amount of time the analyst

has who must both continue his reading, research and writing and take on the description of the data, its organization, and its manipulation for the computer analyst-programmer who is to design said application. People who have the greatest need for computer assistance usually have the least time to seek it out and provide the close interaction needed for effective applications. Often when an individual or organization is willing, indeed anxious to have an activity automated, the time pressure on the analyst and/or supervisor who can most appropriately describe the data or informatin inputs and outputs desired is such that a relatively junior person is given the job of working with the computer programmer-analyst. The organization of data, the specific procedures for computation, and a clear notion of how the project may evolve are all critical decisions which must be made by those directly responsible for the project. Made by a junior analyst or semi-professional, a junior application is apt to be the result.

Application or Project Definition.

42. Discussion of recently developed computer applications with their sponsors indicates that OCS has become somewhat more patient of the difficulty of outlining a definitive (as in definitive forever) application objective. The data on which intelligence analysts must base their work are constantly changing not only with respect to value (which the computer can handle with comparative ease) but also with respect to the factors and the relationships which may be observed and thus become part of the analytical system employed.

43. It must be as important for the systems analyst-programmer to recognize the inherent uncertainty in what intelligence information may be available and prepare, with the research analyst, to accommodate programs and data structures to this prospect as it is to assert that the research analyst must develop the best characterization of the data and their relationships or process that he can achieve. We believe that the increasing experience with and use of a generalized data management system, CAPRI, has provided a flexibility in file handling and processing which was heretofore unavailable.

A Continuing Need for Imaginative Development

44. Observation of the existing set of applications and of new applications under development persuades us that there is a great tendency to repeat types of activity with which we have had operational experience. Few of these activities seem stodgy, but

there is a tendency for new applications to be a relatively safe replication or extension of existing programs. We believe that a greater effort should be supported to explore the use of quantitative techniques in political and social analysis, particularly in the analysis of the decision process and the evaluation of alternative courses of action.

45. Our consultant team observed a considerable repetition or overlap in program development among the several data processing centers in the Agency. One avenue of concurrent effort, however, is in time-sharing activity which probably represents the most promising recent capability in computer processing. This represents a major effort in an area where parallel effort is warranted. In spite of the obstacles that large system designs have encountered to date, we believe that the Agency should be engaged in an effort initially to integrate headquarters' time-sharing activity into a single system, then incorporating with this system the background processing of jobs handed off from the time-sharing and, finally, remote entry of regular processing jobs. OCS is engaged in all of these activities in at least partial conjunction. Inasmuch as this realm of activities can be largely effected by the computer processing facility, it offers greater prospect for concert of action than a meld of activities which would require wider coordination. Such a system will nevertheless require a major effort to develop an adequate set of data management systems and a system management package that can effectively meet these demands.

46. We have stressed the need for a single large time-sharing system principally because of the exceedingly large investment that appears to be involved in an effective system. Users of the present system are insistent on the need for uninterrupted access to the system both because any considerable delay in system response interrupts their attention and slows their work, and because any machine failure which requires the computer to be initialized again means a loss of work in progress which has not been written to user storage. Essentially continuous operation of a time-sharing system which serves a substantial number of users will require a considerable amount of redundancy in expensive equipment, and the large requirement for control programs in this environment demands a large systems programmer staff for its support which further inflates the cost.

File and Program Maintenance.

47. File and program maintenance pose a serious problem in data processing, particularly in little exercised applications. Little used files tend to be little maintained and to offer a

serious threat of misinformation or at least incomplete information. Often programs which support these files are not kept up to data and abort or require reruns to provide a valid response of any sort.

48. File maintenance has been jeopardized even in well established and frequently used systems by the delegation of data preparation and data entry to components which do not have immediate responsibility for preparation of this data and to people who cannot distinguish erroneous data. This problem has been exaggerated in many active applications by failure of the designers and those who input the files to establish and to enforce a rigorous set of data standards which can assure an unambiguous reference to all critical elements of the system. Bad file and program maintenance are like the weather. Everyone talks about it, but little is done to correct those records which subsequently emerge to embarrass the file owners.

49. File maintenance is more than just correcting errors of inclusion or of recording. It also includes assurance that satisfactory conventions are established for extracting or developing data for the file, or processing data within the file. These conventions must be known to the file user(s) so that they may guide their use and their evaluation of output. Nothing so frustrates the user as a negative answer to a query when he knows that the information is in the file. How much more frustrating it is for the uninitiated to seek information in the file, to find there is none, and to act in confidence only to discover later that such information existed. He has simply been duped by unfamiliarity with the recording convention used by the system managers to register it.

Restriction of File Access and File Use.

50. As a result of our discussions with substantive analysts who build file systems, with computer analysts who develop programs to handle automated files, and with individuals outside of these sets who have sought to use automated files, we wish to insert a word of caution about general use of other people's data bases or files -- "Don't!" Few existing automated file systems have been built to support a large number of users. Those systems, that have been built, have been prepared for professional specialty groups and have been carefully tended by journeyman members of these groups in close collaboration with ADP specialists. System builders, maintainers, and users, all speak a common tongue in their activity which permits them to reduce sharply the amount of documentation required to understand and use the system. Moreover, knowing also the major sources of uncertainty and ambiguity in their activity,

these are subjected to precise definition and are prominently identified for system users. With all this care, mistakes or misinterpretation still occur. They can in general be sufficeintly contained that individuals are willing to pay for service.

51. "Publicly" available files ought only to be offered in circumstances in which the information structure and its characteristics are known to the user and he has certified this understanding, or the files contain a level of guidance and documentation which relieve the user of the need for previous familiarity. In the intelligence community, the file will also be classified and otherwise administratively restricted and the system must be satisfied that the person seeking to use the file qualifies on these counts as well.

52. Development of a file for general use in the Agency must anticipate a need to develop and present for general consumption sufficiently detailed documentation on the file to permit the user to access it on the premise that any misinformation he receives will be the result of errors in recording the information in the system. The documentation for CAPRI, CAPRI File Management System Volumes 1 & 2, is a satisfactory example of our estimate of the documentation required for de novo access. For batch processing systems, the user can still be partially kept at a distance and his mistakes forgiven by the operations element which enters his request. In an on-line system he must be controlled by the information and options offered to him at the remote terminal. Thus, we endorse the creation of general or joint files but we wish to underline that their development and operation is a complex matter which must be carefully controlled if they are not to misinform rather than inform the user.

Security

53. The problem of security is treated exhaustively and intelligently in a large number of publications. We think that the ARPA report Security Controls for Computers and the documentation of the USIB Computer Security Subcommittee document Identification of Computer Security Problem Areas IBSEC - CSS-R-2, present the most relevant treatment.

54. Our interviews both with analysts using the OCS Interactive Services System and those who have developed applications of their own reveal no difficulty whatever in meeting the security requirements which have been imposed on their personal activities.

Time-sharing users have expressed dissatisfaction with being denied access to the OCS system during the hours in which the Agency was participating in COINS, but they "understood" that this was an obligation which was to be temporary and which must be discharged to meet the Agency's role in a Community system.

55. The single security requirement which is urgently needed is for security personnel to become broadly aware of and personally involved in the data processing environment. This is being done by training, by direct assignment of people to OCS, and by review of computer security problems with other security components in the Intelligence Community. In the interim, security measures which have been taken have been identified as temporary. These measures have generally sought to permit and encourage the development and use of ADP rather than to restrict it. Restraints have been imposed only where there appeared to be a clear parallel to conventional physical and personal security and then on an essentially analogous base.

56. Automated data processing security however will often have to be treated on its own merits, divorced from traditional concepts of physical security and from myths of machine infallibility. These security decisions will require the same analysis and evaluation of risk and operational necessity as those which have applied to personal and physical security.

Conclusions and Recommendations

We recommend that:

57. The Agency establish as standard procedure in the development of new collection systems the coordination of the data collection and data forwarding formats with the individuals who must perform the data reduction and analysis of the data should the collection system become operational.

58. The Information Processing Board assure the acquisition, development, and use of one or more general data management systems which are sufficiently close to the general design requirements for Agency data processing applications to permit their adaptation and use for a wide range of data processing applications and data processing centers. We believe that general systems to incorporate such applications may best be acquired from commercial vendors in the interest of economy of maintenance, and simplicity and generality of system operation. Experience with indigenous development of such systems seems to indicate that we tend to build overly elaborate systems which provide better for certain internal requirements but tend to violate the other canons outlined above.

59. The Information Processing Board assure that the present capability for development of a unique program to process an application is maintained, so that any application whose objective is clearly unattainable by incorporation into a general data management system may continue to be developed. We would make it clear that increased use of general data management systems should not displace other applications programming activity except where its speed, economy, and prospects for ease of data exchange make it more attractive. It seems ideal, however, for small, infrequently used programs which are subject to change over time.

60. The Information Processing Board and its Technical Panel be charged with creating the means for the development of meaningful communications between the systems analyst-programmer and the substantive analyst who may be engaged in the common development of a computer application. Have them assure that the requisite direction and training is given to accelerate the reduction of communication barriers which still exist.

61. The Agency provide time and professional and clerical assistance to a few talented individuals each year to explore, develop, and test essentially new techniques or new concepts in the use of ADP to support intelligence analysis and production. These applications may be developed under the leadership of either the substantive component, or OCS depending upon the nature of the application and the resources required to do the job.

62. The Information Processing Board assure that the present effort to provide a general time-sharing capability in OCS to serve the interest of the Agency as a whole be strengthened to provide not only on-line but also remote batch processing and remote job entry via terminals distributed so as to make them convenient to users throughout the intelligence production components.

63. The Information Processing Board,in consultation with the interested parties, assure that the OCS Interactive Services System provides a general data management system capable of providing an on-line, quick response capability for large information storage and retrieval activities of the type characterized by the MISTAC, AEGIS, QUIKTRAK data bases. We believe that the present and foreseeable rates of use for these files in an on-line environment are not high enough to warrant economic use of individual processors to support them.

64. The Agency seek to secure an evaluation of the present COINS experiment at the earliest possible moment in an effort to provide clearer guidance for future Agency planning for participation in Intelligence Community ADP activities.

65. The Information Processing Board define minimum standards for control over data entry, data base documentation and file maintenance for any ADP application serving more than one component (defined as a unit under the first-line supervisor).

APPENDIX: Descriptions of Computer Applications

This Appendix contains descriptions of each of the computer applications used in the Agency to support intellignece production and information processing, arranged by Office. Each description includes its name and acronym, if any, an evaluation* of the quality and performance of the application together with information on its initial development cost, who developed it, (contractor or Agency personnel) and operations cost for Calendar year 1969 and the first quarter of 1970.

Following the detailed enumeration of the applications is an abbreviated list titled "Computer Applications to Support Intelligence Production and Information Processing" which lists these applications by evaluation classification* and within the category of function performed i.e., Information Storage and Retrieval, Calculation, Model or Simulation, Data Reduction and Mixed.

*Definitions for the 5 classifications used to evaluate the applications are shown at the end of the Appendix.

1. FOREIGN MISSILE AND SPACE ANALYSIS CENTER

Next 19 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

2. OFFICE OF ELINT

Next 2 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

3. OFFICE OF SCIENTIFIC INTELLIGENCE

Next 11 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

4. CENTRAL REFERENCE SERVICE

Next 13 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

[redacted]

STATSPEC

(No Applications)

6. OFFICE OF BASIC AND GEOGRAPHIC INTELLIGENCE

Next 2 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

7. OFFICE OF CURRENT INTELLIGENCE

Next 5 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

8. OFFICE OF ECONOMIC RESEARCH

Next 21 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

9. OFFICE OF STRATEGIC RESEARCH

Next 9 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

10. NATIONAL PHOTOGRAPHIC INTERPRETATION CENTER

Next 8 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

11. INFORMATION REQUIREMENTS STAFF

Next 7 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

Definitions of Computer Application Classifications

1. Application is directed at a critical need of the office. It performs well enough to make it worth several times its cost in terms of accomplishment.

2. Application is directed at a critical need of the office. It functions at acceptable but less than expected levels. or The application improves both the quality and the quantity of office output by an amount well in excess of it's cost, including amortized investment.

3. Application is directed at a critical need of the office, it functions acceptably but its cost was far more than can be amortized, yet it's output is worth more than our current expenditure on it. or Application produces a part of the original objective and with further work we believe that the original design objective can be achieved which would raise it to a class 2 level application.

4. Application is directed at a useful function the office must perform. It produces only a part of the objective. It cost more than will ever be recovered in performance but present performance is worth more than present expenditure on the project. or Application is directed at exploring our ability to use ADP. It functions satisfactorily and performs a useful function about as well as can be done manually but at a higher total cost than manual processing.

5. Application is directed at a critical or useful function. It has not achieved design objectives although it works in part. Cost of maintenance and operation of the present application exceeds its product. There is no immediate or medium range prospect that the application can be made more productive.

III. OFFICE OF COMPUTER SERVICES (OCS) ACTIVITIES

III. Office of Computer Service (OCS) Activities

1. Three computer processing centers provide most of the data processing used to support intelligence production activity in the Agency: the Office of Computer Services in the DD/S&T, the Automated Information Division of NPIC and the Electronic Data Processing Systems Division of the Central Reference Service. Each of the latter two organizations was developed initially to process large scale computer applications which had been developed in the organization, the NPIC computer facility being the first developed in the Agency. The Office of Computer Services developed out of an amalgamation of the business or administrative data processing activities of the Agency and the rapidly growing need for additional scientific computation facilities with the formation of the Directorate of Science and Technology.

2. The Office of Computer Services was organized to operate: 1) a general computer processing center for the Agency as a whole and 2) to provide personnel with computer programming and computer applications design experience to assist analytical and operational components in the development of computer assisted solutions to their work. These objectives have implied that OCS would process programs and data developed by the customer alone, or in conjunction with some third party, as well as programs developed between the customer and OCS. Providing such a wide range of processing services imposes a considerable burden on OCS to maintain large processing systems, a wide range of computer system programs, and system programming personnel.

3. Such an effort eases the difficulty to the customer of trying to take advantage of programs which have been developed in other facilities and offers a wide range of approach to problem solution. It is however expensive to maintain such a diverse operating environment. Moreover, it tends to encourage the customer to be expansive in permitting contract work for development of his application to be done with bizarre local languages or dialects of the contractor on the premise that OCS can make anything work. Indeed, they usually can. Most of the using components have the opinion that they could operate as efficiently and more economically if the variety of language and system options were more restricted, this despite the fact that such limitation would inevitably limit the scope and flexibility of computer applications.

4. Customers who have operational computer applications tend to be pleased with the service they receive from OCS. Everyone would like his work done more quickly but there were few customers who thought the attention their processing requests received was less than completely satisfactory. The most frequent complaint

expressed by customers of the processing center was their inability to maintain the job control language and the reference calls in their programs sufficiently current to assure that they would run on any occasion without intervention by the production control or technical programming staff. The most general concern of a substantive nature was on the uncertainty about either near-term or longer range ADP plans.

Rate of Technical Change

5. The rate of technical change in OCS has made it difficult for either users or, one suspects, OCS to know what to expect in the near term. Some of this uncertainty is the result of rapid technological change in the industry. Some too seems to be the result of a tendency to premature announcement of changes in OCS plans based upon the availability of some new equipment or program which will remedy one or another urgent problem facing the office. The data processing activity must be dynamic. But the customer, led from one potential solution to another with little apparent coordination of these programs and little opportunity to participate in the decision process, often receives the impression that he must pay a great deal in day to day reworking of his computer applications to support changes which provide no benefit to his work or to support changes which may never occur. Most users would be happier with a rate of change which was no faster than could be supported by a careful review of the advantages and disadvantages of any system change (in which they could be represented) and a decision announced sufficiently in advance of carrying it into effect to permit them to adjust to its impact.

Expansion of the OCS Procedures Manual

6. The present OCS Procedures Manual does not present a complete set of procedures for the programmer in the substantive organization to assure satisfactory preparation and operational readiness of his program. This guidance should provide more complete information on the Job Control Language (JCL) needed, or someone should be furnished full time in OCS/Operations who can prepare JCL for anything that may be brought in to run. The analyst who is compelled to run his program in an emergency at off-hours often finds that it is not producing valid results and must delay completion of his work until someone can be called in or until the next prime work shift.

7. One solution to this problem might be to create critical parts of, or perhaps the entire procedures manual in storage on the Inter-active Services System. There the procedures could be available almost continuously, and they would be amended more easily. Although there would be some need for a print version, we believe that this could be managed with a high degree of mechanization and at no greater cost than the present manual.

Assignment of Applications Programmers to Customer Offices

8. The identification of the problem of communication between computer analyst-programmer and the functional research analyst in II para. 34ff is also susceptible of attack via OCS. Many complex technical processing centers approach this problem by assigning a major portion of their application programmers (sometimes all of them) directly to the customer for whom they are working. In most of these activities the programmer reports to supervisor of the functional research unit, but he is paid by and his professional advancement is controlled by the Computer Center. The entire programming design staff, whether OCS or functional personnel would look to OCS for programming and design standards or procedures. Although several intelligence production components have been hiring and developing their own computer programmers, few of these individuals have the design and programming skills necessary to plan and execute a major computer application. We favor development of programming talents among production analysts to provide better conceptual design for applications. We nevertheless believe that professional programmer-analysts, recruited by and a part of the computer center and assigned to the production components, should undertake the detailed design and programming of computer applications. They should also be used to assist the production component in planning the development and evolution of computer applications.

9. Problem solution is and must be separate from computer operations, viz the very organization of OCS. We believe that problem solution is the responsibility of the data processing user and that it should only be undertaken by the data processing center under the detailed direction of the customer component. Indeed, we would urge that the basic and detailed design of the computer application, its data structures, even a considerable portion of its programming would best be done in the customer's area as well as under his control. He can then terminate any excursions which seem more elegant than the job demands or increase effort on segments which seem to fall short of expectations. He is the programmer-analyst's boss, the person who must ultimately be satisfied with the job that is done.

10. The need to assure professional review of the programmer-analyst's work is recognized. We believe that this is easier done on occasional visits to his home base in OCS than is substantive review accomplished at present by occasional visits to the customer office for whom he is working. Moreover, the influence that the professional programmer can have on the work of functional specialists who are doing design and programming work in the office in which he is cited should assure greater conformance with OCS standards as well as greater productivity on the part of these individuals.

11. There will be an inevitable loss in day to day flexibility in computer programming assignment in OCS, although we assume a total assignment of certainly no more than 75% of the various application division personnel under an arrangement of this sort. This flexibility presently tends to concentrate on ad hoc efforts to shore up schedules on existing projects rather than undertake new initiatives. If these schedules could be set and be met more realistically by direct assignment of some programmer-analysts, the net result would be to reduce the need for some of this day to day flexibility. In any event we are inclined to believe that flexibility which sacrifices work on major application development activity is suspect.

Computer Graphics

12. Computer graphics, which we would define as the development of programs to transform stored digital data into the meaningful spatial relationships contained in the data, have been given little attention in OCS Aside from AUTOMAP which was developed to facilitate cartographic presentation (See Appendix to Part II), there has been little use of graphic presentation or display. We note this omission because CIA analytical reports are distinguished by their profuse and excellent use of graphic presentation to augment text or data. One of the areas in which the Information Processing Research and Development Laboratory in ORD was most active was the development of graphics applications, probably reflecting the limited ability to exploit this technique in OCS.

13. The employment of the machine to aid man's perception of his data through graphic display offers an achievement which is almost as important as speed in raw computation. Man comprehends much better with his eyes than any of his other senses. Indeed, he often finds it necessary to transform data into some sort of graphic before he can appreciate its relevance. Moreover, most individuals have much less facility for presenting data in graphic form than they do for computing. Thus the development of an excellent general repertoire of techniques for transformation from data to graphics would add considerably to the analytical strength of Agency data processing users.

14. The high productivity of AUTOMAP and the great interest in the development of QUIKTRAK (See Appendix to Part II) testify both to the achievement of graphic presentation and the interest it evokes in analysts. We see a wide range of use of graphic displays for existing batch processing applications, graphic presentation of data summaries such as frequency distributions, ballistic or orbital characteristics, penetration routes, etc. The last two applications

have been implemented quite effectively on the plotters employed for AUTOMAP. The ability to access graphic display in on-line processing seems even more urgent, particularly if the computer is to be an analytical aid rather than simply a fetcher of forgotten minutiae.

Generalized Data Management System

15. The present plans of OCS to seek a proprietary set of generalized data management programs to support both time-sharing and batch processing and to abandon internal effort to produce such a system represent excellent judgment in our opinion. Agency experience with both avenues of approach to data management systems in recent years indicates that the cost of acquisition and the cost of maintenance of a commercial system are far less than for an indigenous system. The commercial system while often less elegant than the local product, can provide the functions necessary to get our job done. Moreover, it appears to require less modification and maintenance than the internal product to continue to meet our requirements.

16. We do not think, however, that the efficacy of any available generalized data management system is so clearly established that the Agency should eliminate all development effort in this direction. A small, perhaps experimental program, should be implemented in OCS directed at:

1) the definition of specification of Agency-wide requirements that must be met with such a system,

2) the investigation of and conceptual or even detailed design of the critical elements required in a generalized data management system, and

3) a continuing review of the development of and evolution of existing proprietary systems in terms of their ability to meet Agency requirements.

17. The long-range need of the Agency to provide as much compatibility among the various files and to provide as much uniformity of access as possible demands that the system or systems chosen be established as an Agency Standard. There is already an effort on the part of other processing centers to develop or acquire such systems separately. NPIC has already had a system built and we would expect that system to be implemented. CRS has been investigating the acquisition of an on-line bibliographic system, the processing objective of which would be essentially similar to both the OCS and NPIC systems. A common effort should be made to develop specifications for any proprietary system.

18. The emphasis on the specification, acquisition and development of a single general data management system should not be construed to mean that we believe that such a system will now, or can ever, provide the only programming required to produce an effective set of computer applications. We believe that a system can be built which will go a long way toward provision of inter-file compatibility and control of data for both on-line and batch processing of a wide range of applications. It may be necessary to acquire two or three or even more such systems to cover general system requirements. Our [redacted] consultants thought we might need three such systems: one for large general information storage and retrieval applications such as document indexes or information abstracts, another for processing large economic time series or other scientific or political quantitative arrays, and a third for processing large name searching or name finding files. Even a mix such as this would require that certain applications be processed with unique programs. It would, however, reduce the number of unique programs prepared to execute similar processing requirements. What is important is that acquisition and effective implementation of a general data base management system or systems would reduce the total amount of programming performed to achieve a given level of performance. It would both produce greater processing compatibility and ease of exchange or common use of files over the Agency.

25X1A

OCS Interactive Services System

19. The user concept of an effective ADP system is in fact one in which he is connected to the system from his regular work station or nearby that station. He may use any information he has in the data processing system by simple identification of himself to the system and indicating what he wants with a few key strokes or light-pen. Equally he may exercise any programs he has stored in the system by a similar procedure, adding only the required data by key stroking (when it is negligible), by card reader, or by call from some other part of the system's store where it was put on receipt.

20. Although the user prefers instant return of the results he seeks, he can often abide delays in receipt of his reply if he has not been delayed in introducing his problem and if he knows its status in the system. The user has done all he can do, he is then "Waiting for Godot." In every application we have observed, the user clearly prefers an interactive, hands-on capability to any other association with the machine. The "hands-on" environment to which the researcher or engineer or accountant brought his program and often his data or files in the early period of computing is still the most preferred. Technology now provides for its approximate return.

21. A desire for ease and simplicity in the generation of results has replaced an earlier Spartan attitude of direct coding the program in machine language. The user wants a menu of files available to him and activity he can undertake on the system displayed by a single key stroke. By successive keystrokes, exercising alternatives, he wants to direct the search or processing to the result sought. This ease of access can be programmed although often at great expense and with considerable restraint upon the manner in which the file is organized and manipulated. We believe that this sort of expenditure is warranted 1) where real-time applications are required to control field operations, including sensor systems on or over unfriendly territory, and 2) where the functions required are executed so frequently by the users of the application that their programming as a named function for the system is warranted.

22. The original objectives of the time-sharing system in CIA were to put one foot in front of the other in moving toward these general user objectives. More fundamentally the system sought 1) to provide improved on-line access to large file systems, 2) to provide quick computational capability to assist production analysts, 3) to continue the support for program development, and 4) to continue support for creation and maintenance of small private files. It was to provide capacity for additional growth of file processing applications, to let us understand better how to employ on-line processing to support intelligence analysis and intelligence production, and to provide a few years' growth in this sort of capability unhindered by equipment change.

23. The system has proved particularly effective in support of program development. No engineering measures of the improvement in either program quality or the number of instructions produced have been collected. A survey of users recently conducted by OCS shows that a small group thought their productivity had increased by 300% but the majority of those polled were about evenly divided between no increase and a 20% increase. With the incorporation of APL and other simplified sets of standard programming languages, the opportunity for a quick shot on the computer to process a complex statistical or engineering routine which might require days to complete heretofore has become a reality. The widely heralded capability to provide rapid, essentially interactive access to large files for retrieving and manipulating information has not been achieved. Indeed, service to large file users has been less satisfactory (both less reliable and more difficult to maintain) on the large IBM Model 360/67 than it had been on the 360/50. With the full implementation of the control programs which were to be used to operate the 360/67, service to file users whose files had been created under the earlier time-sharing system was seriously degraded and has never been restored to the level obtained before the installation of the 360/67.

24. The reliability of the time-sharing activity has been less than satisfactory. It is subject to frequent outages, most of which are corrected within a few minutes time. There have continued to be occasional outages in which the system may be unavailable for more than an hour. System reliability to the time-sharing user is a wholly different concept from the batch processing application. Even a short outage may not only destroy his train of thought but also his data inputs for the last hour or more. Unreliability and delay in response has been sufficiently serious that the SANCA application in OS, which demands quick response, has been unable to abandon hand files -- over 2,000,000 hard copy records which were scheduled for destruction and had to be moved to another processor to support the required level of service. The need for greater system reliability to support critical programs which use time sharing may add an entire new dimension to this type of processing. The need for a back-up processor would substantially increase the cost and thus alter the comparative advantage of on-line processing.

25. Agency use of the expanded time-sharing system available on the IBM 360/67 has grown even more rapidly than the preceding year. Rapid expansion had been forecast at the time of acquisition of the 360/67. It had been thought, however, that the machine might see us through a two year transition rather than be fully loaded in a little more than a year. The probability of full loading within a year had been recognized however. Again the increase in use had been absorbed in large part to support program test and development -- that is to improve the performance of the time-sharing system itself and to provide support for programmers rather than to provide more rapid access to intelligence information or to provide a given intelligence product more rapidly. The program test work performed on the time-sharing system, however, makes possible a reduction in the use of the large batch processing systems (360/65's) for this work because it must be done irrespective of what system is employed.

26. Tables 1 and 2 present a partial picture of the present use of the 360/67 with greater emphasis on the employment of the system for intelligence production. The limitations of the Tables are presented in the notes. Specifically, the dedication of the 360/67 system to COINS during a three hour period produces almost zero benefit for Agency analysts and little more than that for the Community. Few Agency analysts use either the internal files loaded for COINS or the COINS files lodged at other Agencies. There is in fact a good deal of use of the 360/67 system between 2000 hours and 0700 hours the following morning. There is little reason to believe that this use significantly alters the ratio of program development work versus intelligence production found in Tables 1 and 2. Its effect is probably to enlarge each of these activities at the expense of activity for business or administrative applications and those devoted to collection which are almost wholly confined to prime shift time, i.e. 0800-1800 hours.

Table 1

Summary Terminal Time -- OCS Time Sharing System

Summary is arranged generally by activity. It includes only system hours between 1000 and 2000 hours. Thus it excludes three hours daily, 0700 - 0900 and 1200 - 1300, of block time handed to COINS. It also excludes a large volume of work, still much of it developmental, performed between 2000 hours and 0700.

	Hours Minutes		
	December '69	February '70	May '70
Intelligence Production Total	366.52	519.25	813.11
CRS	8.36	0.10	22.10
FMSAC	127.31	189.58	399.07
OBGI	2.49	-0-	1.13
OEL	57.55	118.59	96.18
OER	67.52	70.56	187.08
OSI	60.05	23.55	107.15
UNAL 1/	42.04	115.27	219.37
OCS Programmer Time 2/	581.37	724.54	2066.11
Admin Time	53.50	96.02	89.49 3/
Training Time	39.57	65.35	-0-
Collection Time	87.57	53.43	257.57 3/

1/ Unallocated time is the APL time for which there is no basis for division although we believe the major APL use is in intelligence production.
2/ Recognize that a large share of programmer time is spent on development and maintenance of applications programs for other elements in this breakdown. The major share of production office use of the system is for indigenous programming.
3/ No data on TSMON use which was divided between administrative and collection activity was available in the May '70 report.

Table 2

Summary CPU Time -- OCS Time Sharing System

Summary is arranged generally by activity. It includes only system hours between 1000 and 2000 hours. Thus it excludes three hours daily, 0700 - 0900 and 1200 - 1300, of block time handed to COINS. It also excludes a large volume of work, still much of it developmental, performed between 2000 hours and 0700.

	Hours Minutes		
	December '69	February '70	May '70
Intelligence Production Total	10.20	12.10	23.10
CRS	0.18	-0-	.24
FMSAC	5.30	5.56	8.48
OBGI	0.06	-0-	.01
OEL	1.32	2.27	8.45
OER	1.06	1.01	3.00
OSI	1.16	1.07	2.12
UNAL 1/	.32	1.39	2.49
OCS Programmer Time 2/	18.51	21.23	67.43
Admin Time	4.15	7.16	3.32 3/
Training Time	.29	1.46	-0-
Collection Time	4.18	4.03	3.10 3/

1/ Unallocated time is the APL time for which there is no basis for division although we believe the major APL use is in intelligence production.

2/ Recognize that a large share of programmer time is spent on development and maintenance of applications programs for other elements in this breakdown. The major share of production office use of the system is for indigenous programming.

3/ No data on TSMON use which was divided between administrative and collection activity was available in the May '70 report.

27. The amount of computer time-sharing activity devoted to support of intelligence production has increased both relatively and absolutely during the past 15 months. Intelligence production usage has increased from about 10-15% to about 25% of the system. The number of private files in the system has grown considerably more rapidly, from less than 2% of those in the system to more than 10%. There has been even greater growth in the use of the computer as a powerful and versatile desk calculator using APL, although this use remains relatively small. These changes have not altered the basic character of the Agency time-sharing activity. It is still predominately used for computer program development and test. Even production organization use of the system is predominately devoted to program test and development. Why? There has been little apparent need for on-line processing of most of the large production oriented files which are presently in machine form; the COINS files are an excellent example. These files were designed for serial processing as were the applications they support and the needs to which their manipulation is addressed. The MISTAC files which FMSAC had developed initially for sequential processing were immediately made available for on-line processing when it became available because of the time requirement for response imposed on the FMSAC Control Center. They were difficult to operate in that environment, however, and have been withdrawn for redesign and reprogramming. The more we seek to process batch files in a direct-access, time-sharing environment, the more we believe it becomes apparent that this is not a feasible activity given the present structure of the data. We see no reason to urge the development of on-line processing of large-files except where there is a genuine need for such speed or until there exists a clear economy for such a transition.

28. The present status of the OCS Interactive Services System which divides the processing activity for large file systems from that performed to support other analytical activity is acceptable as an interim measure to try to improve the ultimate performance of 360/67 system as a whole, but it would create a serious separation of function in analytical use of the system were this division a necessary condition of on-line processing. We believe that even the prospect of the latter event is a sufficient basis to suggest a more detailed analysis and evaluation of the requirements for and the performance of the time sharing system. It may be that there is a sufficient demand for on-line programming and program test to support an independent or dedicated time-sharing system for that purpose. There is a well defined need, however, to develop an integrated time-sharing system for support of intelligence production which provides for:

1) access to analyst files and general files created for analyst use,

2) ability to manipulate both quantitative and textual data in these files to produce new results or new files, and

3) ability to exercise these functions on-line or in the background.

The tendency to divide time-sharing activities by function may make for a neat processing solution to the problem but it tends to create a difficult, if not unacceptable, solution for the production analyst.

29. Further, the development of multiple time-sharing systems violates one of the basic tenets which both ASPIN and OCS had suggested be observed. The remote user of the time-sharing computer should be able to communicate in a single language from a single terminal to the processor(s) which contained the data he needed. Now, in fact, analysts who have both an interest in any of the large files and in quick computational capability must have 2 consoles, speak two languages, etc. To illustrate, OS officers who must deal with both OS time-sharing applications must go from the 3rd floor to the basement or vice versa to move from transactions on one application to the other.

Conclusions and Recommendations

30. We recommend that: OCS develop in consultation with the Information Processing Board a mechanism for communicating plans for major computer system changes to user components and of eliciting and reviewing user input to these plans before they are ready for submission to the Information Processing Board for review antecedent to approval by the Executive Director-Comptroller.

31. A complete set of procedures be published and maintained which provides enough information to assure that a job can be written (including JCL) and run without intervention from OCS programmers.

32. Applications programmers (this would presumably include a major share of the applications divisions' personnel) from OCS should be assigned to and, where feasible, colocated with analysts in the production organization for whom they are designing and programming. Their work during their period of assignment should be controlled by the host production organization except that their rotation back to internal OCS assignment should be negotiated with OCS.

33. Increased attention should be given by OCS, in close cooperation with NPIC/AID and ORD/An, to the development of a strong computer graphics capability for support of analyst use of the time-sharing system. The AUTOMAP data base should be developed toward an on-line utility which can be summoned as an outline map or chart for superimposing other data for analyzing spatial relationships. These data in combination with the map should be susceptible to linkages to computational routines to further test intuitive visual observations.

34. Present planning for OCS to acquire and test a proprietary general data management system should be encouraged. Plans for this acquisition should move forward as rapidly as a careful coordination of the proposal can be concluded. We believe that this movement should be coordinated with the major users of OCS and with each of the intelligence production components who have their own data processing center, i.e. NPIC and CRS. The objectives in acquiring general data management software ought to be 1) to move toward as wide a coverage of our major processing activities within any given system as is intellectually and operationally acceptable, 2) to establish each system selected as an Agency standard for the type(s) of application identified, and 3) to recognize that there will still exist computer applications which will require unique programs.

35. A single, integrated, interactive services system to provide on-line service for intelligence production components at headquarters should be the Agency near-term objective.

IV. A CENTRAL REFERENCE SYSTEM

IV. A Central Reference System

1. We began our investigation of the notion of a central reference system with Genesis. Is there a need for a central reference system? If there is no need for the system, how is the Agency structured to provide the services that are normally provided by such a system? If there is a need for such a system, how is the need met by existing arrangements? In either event, how may new information processing technology be used to improve the fit between information needs and the requirements levied for collection of information over the next few years?

A Reference Center Concept

2. The existence of an organization identified as the Central Reference Service (CRS) is neither a necessary nor a sufficient condition for the existence of a central reference system to support intelligence production. Many analysts are inclined to be equivocal on the need for certain services that CRS performs, particularly the creation of a large index to intelligence publications and the preparation of biographic reports, but there was unanimity of opinion among the people interviewed that the document acquisition and dissemination functions, the biographic reference files and the library were critical utilities which no one wished to be without. Usually the same claim would be made for the document and photo indexes and collections for certain recurring requirements, particularly for the necessity to build a collection for an activity in which the analyst was not regularly engaged. But the latter is one of the primary reasons for which a document index exists. We found the analyst generally inclined to indict the document index system for its exhaustive nature rather than to employ it to verify the scope of his own frame of reference. This may be more an indictment of analyst technique than a deficiency of a central reference system. Almost all individuals engaged in intelligence production were inclined to emphasize that the analyst files and the special information systems created by functional organizations are an even more critical reference facility for them and for their customers than are the formal reference holdings.

3. There is now and will continue to be a need for a central reference facility in CIA. The emphasis in this facility will tend increasingly to be on "reference" and what that concept implies, a directory, rather than the sole source or repository of the information used by and generated by the Agency. Thus we see a central reference facility as one which can route a request for information to the appropriate source or sources in the Agency and see that the request is answered. Part of the information available may rest in the central reference facility itself. More often the facility will

simply provide a connection to another facility or center in the Agency.

4. The notion that there should be a single repository of information disappeared with increasing specialization, with the very specialization which created the rapid growth in information which we refer to as the information explosion or the information revolution. Highly specialized information tends to reside with those who have the ability to interpret it and to relate it to other activities. No amount of pressure to create "total information" systems has altered this situation. The problem of the policy maker or the analyst is to recognize and cope with the complexity of information arrangements rather than seek pseudo-simplistic solutions in computerized information processing systems. Thus a central reference system needs to be built around:

1) an information center which serves as a directory to assure that anyone unfamiliar with Agency information resources gets to the right point, and

2) an information system that provides both for rapid dissemination of intelligence information and for discrete analysis and evaluation and controls of this information employing the substantive analyst to the greatest degree feasible in this operation.

Agency Effort Toward a Central Reference System

5. The Agency has achieved an excellent beginning of a modern central reference system and is currently planning a number of major programs to extend and strengthen this system in the direction outlined above. Specific identification of a "Central Reference Service" has been achieved. This office has been charged with the dissemination of positive intelligence receipts and with their indexing, storage and retrospective retrieval. A retrospective retrieval system has been carefully designed with a substantial element of automation assistance. A major design effort is underway to provide partial automation of the dissemination function. Extension of the present information storage and retrieval system to an on-line system is a near-term objective of the office.

6. The Central Reference System, however, clings tenaciously to the notion that reference analysts are the only ones who can (or will) index documents. The present machine record input for the document index system provides no formal or regularized procedure for permitting the substantive analyst to index or tag particular records, nor does it provide any means to link the document index to other (non-CRS) document collections. At best, the substantive

analyst may be able to influence document indexing by sheer rhetoric and/or by repeated visits to his friendly reference analyst.

7. Nevertheless, there has been an evident growth in the number, the scope, and the structure of special information collections developed by intelligence production and information processing organizations throughout the Agency. Some of these collections are of such a magnitude that their annual total operating costs rival individual elements of the Central Reference System's collection. Most of these functional information systems are indexed in considerably greater depth and contain annotation of an analytical or evaluative nature as well. Many have been designed initially, or recently redesigned, to accommodate ADP systems. The majority of these holdings, however, are indexed collections of "hard-copy" documents and files. The extent of knowledge of the existence, quality, scope, size and structure of these collections varies widely. Although intelligence officers confidently claim that all of these sources are known widely enough that one may discover them through the "informal organization," we found too many cases of surprise among the ASPIN Staff and individuals whom we interviewed to be satisfied with this soothing allegation.

8. Although the need to identify formally the existence and location of other collections of information in the Central Reference Service has been widely recognized and CRS sought to develop such a system under CHIVE, no such directory or file has yet been created. The principal objections to the creation of a "file directory" seem to rest on the resistance of authority to:

1) the existence of detailed lists of personnel associated with special holdings, and

2) the implied responsibility to service a query which might be forwarded by CRS.

Neither of these objections seem persuasive in themselves. Referral could as well be routed over a few people. Both the file owner and CRS should be equally able to verify the need to know of the inquirer or the limit of their capability which would be the only premises for denial of service. It should be said that reference analysts have personal knowledge of these collections -- they know of more of them than any other individual at least. They are, nevertheless, limited by personal knowledge or recall and the circumstances of any search they may be assisting in recommending other information resources or systems. Indeed, one has the impression that if there are a number of references turned up by the CRS system, other systems will not be identified unless they are specifically asked for by the customer. This tends to protect the special system from attention. May it not protect it too much?

9. Substantive analysts are much more isolated in terms of information sources, however. The individual substantive analyst tends to create his own information resource by filing documents or extracts of documents relevant to his research responsibility from the flow of documents he receives in response to his reading requirements. Although analysts know of the existence of CRS, we found many analysts who were unaware of certain of the major reference functions performed, and few analysts used CRS as a reference center in the context outlined above. The existence of a specialized collection of information held by a production office is usually known to its members. Nevertheless, few members can identify every major file of such a collection.

10. We believe that an urgent present need in our information activities is to recognize formally the existence of many specialized intelligence files or information systems and to provide a means for individual analyst access to these collections on a need to know basis. In each of these reference systems a skilled analyst is required to collect, evaluate and analyze the data and to build the information system. Some aspects of his work may be passed on to automata but even then only those he or his colleagues can specify. Moreover, only the builder of the information system and certain comparably trained associates are likely to be fully competent to use it to respond to a question.

11. Nothing rankles the person in search of information quite as much as this. People want "to form their own opinions" on the basis of the "facts". Where the facts are a myriad of electric pulses, or imagery taken from a distance of tens of miles, or thousands of references in books and periodicals, how is this possible? One can get an answer and have the facts delivered to a professional who can interpret them for the user (interrogator). Or one can simply have the facts delivered to the user without analysis. The latter verges on folly. We have interviewed some moderately complex operations whose managers simply refuse to comply with a request of the latter type. Any system user, however senior in authority, unless he is part of a professional group with the information builders, exposes himself to a better-than-even chance that he will misunderstand and thus be misinformed by the "direct" output of either a machine or manual information system which has not been interpreted by a person familiar with the system.

12. Why then should we not accept the notion of an information center as readily now as we have in the past? What is critically needed in each of these information systems is direct access between the person who has the question and the individual who has an authoritative answer to the question. We must build systems to facilitate this linkage rather than build electronic systems which

dispense "answers." Electronic computers have not lessened the need for such assistance one iota. If anything, they have increased it. They have added greatly to the ability to explore or manipulate data and by this process to the total volume of information. Where we could once develop only a theoretical construct but could not reduce it to precise values, we can now produce such values quickly with an accompanying estimate of the probability that they are correct. Thus we have even more information thrust upon us.

13. We can produce such solutions, however, only where we have a well defined concept of relationship and process. Thus, the computer produces no new theoretical systems to help us understand an increasingly complex society. It only helps us to calculate and record those relationships which we have understood for some time but could not find sufficient resources to quantify.

14. Faced with a greater volume of information, its content more specialized than in the past, and with a more rapid rate of technological change, we must build more rather than fewer information centers. These centers must be more rather than less specialized. Specialist access to these collections can be controlled through counterparts. General access tends to be easiest if it is through an able interpreter, often a contributor to and a user of the center or general information center.

Document, Dissemination and Indexing -- Information Storage and Retrieval

15. The Agency clearly needs reference control over the information and intelligence products it generates and those of the "intelligence community" (those agencies engaged in the collection, processing, production and use of information from covert sources or for covert purposes). It also needs control over the output of information from non-intelligence agencies which may have particular relevance for intelligence analysis, for example ocean shipping information furnished by Lloyds or national and regional economic statistics prepared by central statistical bureaus. Control over the latter sources needs only to be highly selective. It must permit rapid response to questions on a wide range of national security issues. Some of these can be foreseen rather easily. Others may emerge with little or no warning.

16. A central reference system is the only means for providing effective and economic control over the first requirement. Once established to carry out the first function, such a system has a structure, training, and at least a part of the overhead to perform a major role in providing the second (non-intelligence product control) function as well. Such a system by creation of a controlled,

standard notation and by training individuals in the application of simple numerical and mnemonic codes can permit the development of an index record which will provide a wide range of handles with which to grasp the document. Control of this sort may have little direct value to the intelligence analyst -- inasmuch as it catches all documents without respect for the perishability or relevance of their content. What one might hope is that this rather superficial control could be created in a way that documents under such control could be easily excluded from or included in retrospective search of the documentary base. Minimal control of this sort is of paramount value in looking at the output of a particular intelligence series or station -- a little used capability but one which strikes us as important.

17. Beyond such simplistic indexing as is suggested above, finished intelligence documents and information reports of considerable present and/or future value should be provided more detailed content indexing which will make it possible for them to be recovered by any analyst who has a current or future need for information on the subject.

18. We believe that indexing to the level outlined in the above two paragraphs, accompanied by appropriate storage of the individual documents for rapid recovery provides the minimum service required by the Agency to perform its mission. No present techniques for automated indexing or automated search of full text can compete with a human indexing system for handling large volumes of documents in mixed media. Nor do we see any probability that this situation will be altered in the next 3-5 years.

19. CRS is presently conducting an experiment with a General Electric processor (GESCAN -- earlier called RSM) which will execute high speed search on machine-readable text that requires little formatting. There has been a great deal of analyst interest in this processor for searching large bodies of machine readable text. The device is attractive initially because it can be used essentially in an interactive mode. While it is slow in comparison with interactive search of direct access files, it is the "only game in town" for interactive search of unformatted text files.

20. Most used data bases to date have been large structured files which have large variable length fields in which the analyst has recorded free text. Such files are easy to search. There is great economy of language by the analyst who creates the file. Moreover the user analyst knows the "language" which the analyst who creates the file employs. This type of search is productive but may be of use for a rather limited time. On-line computer systems as they are employed to process this kind of file will tend to be

used to develop inverted indexes of terms in these fields which will mean that this type of application can be processed with greater speed on a standard on-line system.

21. On less structured files the GESCAN system is questionable value. Textual information even where it is highly specialized is ambiguous to automata in a way that it is not ambiguous to a human. The former looks only for characters in a single, particular configuration, where the latter is able to bring a wide range of techniques to bear such as context, structure and meaning to all combinations of characters in the text. Machine methods thus miss virtually all misspelled words and all synonyms, yet produce records (false drop) for all homonyms which are spelled alike and are able to provide no contextual basis other than frequency or position of the term(s).

22. Automata can be programmed to cope with some of these problems but the GESCAN which incorporates logic which is embedded (hard-wired) in the processor will suffer a progressively increasing disadvantage to the programmable general purpose computer as time goes on. We believe that exploitation of the GESCAN processor the Agency has acquired should continue both because it is producing a useful interim product and because it permits observation of an alternative approach to text processing which may be partially employed to improve the general purpose text processing systems of the future.

23. CRS is presently introducing a limited system for automatic dissemination of documents received in machine readable form. We believe that the principle contribution that this achievement may provide for indexing is an improved opportunity for the analyst to evaluate the document and to participate in indexing it if he chooses. This improvement will be a product of the ability of the analyst to communicate with the document system while the document is in the system in machine language. He may indicate that the document has sufficient value to warrant incorporation in the system and index it himself or yield its indexing to the indexing analyst in the central reference system. Failure to note an item would be prima facie evidence that no indexing in depth was arranged. In the longer run, the techniques used to control the dissemination system might be extended to provide for the minimal index control described above in this section.

24. At the present state of automatic text processing, a substantial assist from human indexing is required to do an effective job of automatic dissemination. This assistance is presently only offered by NSA and is introduced by application, at the point where the document is originated, of a simple set of codes to identify

content of a document by subject, geographic area, and certain broad functional indicators. We believe that should the initial implementation of this system by CRS show promise of feasibility, the Agency should immediately begin to develop a comparable coding effort on Agency information reports. This coding provides sufficient indexing that only a negligible amount of content indexing would subsequently be required to support or, indeed, to extend the present Central Reference Service control over content of the document. Additional formatting and computer programs would have to be built to provide for the translation or creation of data for system management information. Much of the latter may be done by a simple table look-up technique in the machine which would be essentially identical to the manual look-up done by the Head of the Line (HOL) indexer in CRS.

25. Computer handling of text is enormously difficult because of the ambiguity of the language. This deficiency, paired with the notion that probably no one is quite as able to identify the content of the document as its originator, has led to the concepts involved in the USIB Content Control Code concept discussed above. Intelligence activity tends to generate a large volume of information which has a very limited life as well as a considerable volume of data which is misinformation rather than information. An automated system which operates without human screening of the input is wholly unable to distinguish these products although an index analyst or a substantive research analyst can identify most such material on sight.

26. Thus, an important objective is to get incoming information under machine control and under analyst scrutiny as quickly as possible, preferably at the point of origin of the information. We want the machine control so that we may generate management information needed for administration of the collection as well as the content information to operate information storage and retrieval systems. We want machine control so that the research or operational analyst can quickly and inexpensively tag information according to its need for retention in the system and for further indexing.

27. The most common analyst indictment of existing large information systems is the volume of information which they generate. With any sort of large scale system, irrelevant information begins to play an important role in deterring analyst use. The analyst minimizes the problem in his personal information files by being much more selective of documents chosen for the file. Such selectivity can be obtained for the central files only if the substantive analyst can play a more active role in determining what is held in the system and over what period it may be useful. Until the analyst can tag a machine record directly, that is until he has the direct access key to the record, and until he can do this

essentially at or near the time of receipt, and essentially prior to the moment of its processing, the analyst will seldom seek to exercise direct influence over the content of a central file system.

28. We believe a concerted effort should be made to develop the present CRS automated index storage and retrieval system called AEGIS into an on-line system. Initially the only component with direct access to the system should be CRS document analysts. CRS should work closely with OCS on the design of a file management system which would provide eventually for the creation of a data base design and a query language which can be used throughout the Agency for the exploitation of large, content oriented files. Indeed, we would see very persuasive reasons for including representation for the intelligence collection and Agency administrative or business applications people in this activity both because they have large files of this type and because there is a direct interaction between the content files and the effectiveness of, or need for, collection.

29. An interactive central file system (one in which the user or an intermediary can communicate rapidly and directly with the file) becomes exceedingly important if the analyst is to play a larger role in determining the scope of central files and if we are to provide for greater substitution of central files for local files. Ability to ascertain quickly what the file contains on a particular subject or area is of critical importance in determining whether or not the reference system contains the information an analyst needs. It may also help the reference or functional analyst determine whether or not to include a marginal document in the file. It is probably equally important to file maintenance when a rather definitive document becomes available. For such a document may present an opportunity to eliminate a number of related documents from the current file to some more remote storage. Finally, interaction with the index file provides a means of quickly reducing the number of reports an analyst may wish to inspect by permitting him an interactive search in the event his initial request against the system produces too much or too little response.

Development of an On-Line Information Storage and Retrieval System

30. We believe that a concerted effort to develop a series of on-line information storage and retrieval applications should be undertaken over the next few years. A serious effort in this direction is already underway both in CRS and in FMSAC. These efforts should be encouraged. These systems should be so constituted that the individual user may address any aspect of any system for which he is licensed to receive information and obtain that information. Such an information structure will require a number of

basic changes from existing processing arrangements if it is to be both useful to the customer and economic for the Agency.

31. Most of the present information storage and retrieval processing is directed at searching large bodies of information serially -- one item at a time -- to find a few items with the characteristics we seek. This process is necessarily slow although relatively inexpensive. Techniques are already available (on-line access to a random access data base with time-sharing on a large computer) to increase the speed of such searches dramatically, but initially such techniques will be more expensive than present methods. Moreover, they demand not only a change in the media of information storage but also a change in the organization and structure of the information which is stored. Anticipated decline in the cost of storage on these new devices alone should permit movement of many of these files to an on-line, time-sharing computer environment within 2-3 years.

32. The critical element to hastening the advent of high-speed response to direct interrogation of large files is the development of an effective and economical time-sharing software -- a reliable set of programs which will create and store the files economically, search them rapidly and provide the response to the user all in a secure environment. It is this capability that the PFIAB assured us was here in 1965, that the Agency sought to assure the IHC would not be here before 1970, that in fact is still not here. There has been considerable CIA effort spent on achievement of a time-sharing capability, unfortunately much of it redundant, and we presently have only a partially operational system, a system which will distinctly not handle large files effectively and economically. We believe that the Agency should forego additional system design and programming effort in this direction for the present and concentrate our effort on a clear definition of our needs in such a computer environment. Armed with these, we can better enter a market in which many large and responsible software producers will be offering such systems and obtain one which offers us the best general fit we can achieve for our need.

33. There is no software system available at this time which will perform to the general specifications suggested above, nor may one be expected to be available in the next 2 or 3 years at the earliest. What seems to us advisable is to choose among available systems those that most closely approximate Agency requirements and concentrate our scarce internal resources on adapting the selected system(s) to indigenous processing problems and on better defining our future requirements. We will in the final analysis have spent less money to get a better system which we shall feel freer to amend than if we were to continue our present course.

Not consistent
with page IV. 5

34. Information systems of the sort contemplated inevitably mean a need to provide service to many users, a large portion of whom will seek access to information from more than one system. None of the systems on the horizon are capable of permitting use of natural language for interrogation. Thus we must either provide a common framework for the user to access these systems or deny him much of the benefit from their creation. We prefer the former. We believe that information systems which support intelligence production, and which are to provide on-line access to analysts should be concentrated in a single processing center. The large expenditures for back-up hardware and for essentially error free software to provide the necessary reliability for a system of this sort are such that no effort should be made to reproduce them in more than one center except where security considerations are paramount. The only viable alternative to this within foreseeable technology would demand that separate systems would provide a common data management system, either a common query language or a meta-translator, and a data communications switch. We believe that such a prospect is so small that we should not anticipate it. Moreover, should such a prospect eventually obtain, we could probably make more effective use of it in separate systems for having moved initially through the path suggested above.

Conclusions and Recommendations

35. We recommend that: the Central Reference Service be established as the point of contact for any general request for intelligence information from outside the Agency or from within where there is no immediate known point where the needed information is available.

36. Only those data which are generated and accessioned by the reference center be provided as a direct response by the center and that all other data are sought first from another center in the Agency which may have resources to respond.

37. Work under way on an automated dissemination system should be maintained and each distribution point to be employed in the initial system test should be directed to cooperate with CRS in providing carefully constructed "dictionary" terms to try to guide this system. The work should be recognized as experimental at this stage, but it should be widely encouraged for its long-term prospects.

38. Planning for undertaking an extension of the automated dissemination system from SI input to all State, Defense and Agency positive intelligence information received in machine language should be undertaken coincident with the beginning of feasibility testing.

39. The present concept of CRS indexing should be continued, and a systematic effort undertaken to facilitate indexing input from the substantive analyst and to encourage such input to the system.

40. The Central Reference Service should seek as a general objective a standardized document reference number which can be put on the intelligence information document before it is disseminated. This reference should be capable of being generated and included in the format of any automated document dissemination system, and should become a part of that system as quickly as possible. It should be made an Agency standard immediately and expanded into a Community Standard eventually.

41. The concept and scope of document indexing by a reference center should be developed by a top management decision. Established at a lower level, it results either in extensive duplication of effort or in abandonment of control over the use of intelligence documentation. Document index processing has, however, been customized by each processing organization which supports an individual or organization reference activity.

42. The Central Reference Service should create a personnel-area-subject index to other organized collections of information in the Agency. This index should include both personal and organizational collections of information and authorization points for control of access to the respective collections. This index is an important and complex system which must be carefully defined, coordinated and implemented. CRS should be assigned responsibility for design and development of the system but they must have the full cooperation of all the other offices and directorates. Development of such a system would pose an excellent test of the Information Processing Board.

43. The present method of document storage and retrieval is acceptable and should be maintained. It provides speed when it is genuinely needed and is far more economical than any system of electronic storage or video storage that we have encountered. We believe that the Agency should continue to experiment with a limited number of applications in which documents are created, stored, and searched retrospectively in an electronic format, because development of an on-line document index will almost certainly require a simultaneous improvement in the speed of delivery of documents.

44. An extensive interactive (man - machine - data base) capability with the Central Reference Services intelligence document index should be developed and tested as quickly as feasible. This is one of the few large data bases in which we think there is both wide interest and frequent use. Indeed we are told by analysts that the principal limitation on their use of the system is its slow response time.

V. RESEARCH AND DEVELOPMENT (R&D) IN INFORMATION PROCESSING

V. RESEARCH AND DEVELOPMENT (R&D) IN INFORMATION PROCESSING

Its Origins and Objectives

1. Almost every computer application undertaken in support of intelligence production is in some measure an experimental or development activity. It was the rapid transition from feast to famine in terms of OCS support for applications development in the mid 1960s which gave rise to an effort on the part of the DD/S&T to explore the prospects for creation of an information processing research and development facility. This facility was in general supported by intelligence production components on the premise that much of the straightforward computer assistance had been undertaken with OCS. OCS had reached a point at which its principal preoccupation was production work. The new work to be undertaken in the production organizations was less well defined, and less quantitative than that which had been developed earlier. It was thought that it might find a happier home initially in a genuine research environment.

2. Review of the interaction between ORD/An and intelligence production organizations indicates that this environment has not produced the results expected by the latter. There are certain areas of considerable achievement in the exploration of text processing activities (the GESCAN processor), in the development of an interactive graphic capability (QUIKTRAK and its antecedents), and in the acceleration of the development of an on-line or interactive computer processing capability for the Agency (TSMON). The development of computer processing capabilities which can be transferred in an orderly way to an operational status has been fraught with great difficulty. Two of the three solid achievements are at sort of halfway house in CRS; and one, the TSMON time-sharing programs, has been operational in OCS for more than 2 years.

3. Definition of a "user" in the R&D context is necessary both because of the wide variety of users and because the reader tends to have a single stereotype. Our user ranges from an individual who has an operating data processing application and who thinks that it is poorly designed and wants a more responsive system, to an individual who has an activity which processes a great deal of information both textual and numerical and thinks that he could

get help from a machine processing application, to another who has no idea of data processing technology or its potential application to his activity -- he's simply been told by his boss to look at how computers might be used in their organization.

4. Selection of ORD to create an information processing R&D facility in spite of the substantial R&D interest of OCS may have been a reaction to two lines of development. Experimental development in OCS was concentrated on monster information processing applications: foreign language translation (ALP), large scale information storage and retrieval (CAPRI), and a large scale, and high speed analog to digital converter (ANDI) which held little prospect for providing a significant improvement in the quality or reliability of intelligence production. Moreover, OCS had a mounting list of going computer applications and an imminent threat of delivery of the new IBM System 360 computer which would require both a sustained, major programming effort to convert existing programs and a major operational and systems programming effort to achieve the smooth introduction of this new system. ORD evinced an interest in research and development on a set of problems much closer to the interests of research analysts.

New R&D Problems

5. Indeed many potential computer users did get help from ORD/An. Much of that assistance tended to match user problems to pieces of hardware or software in which ORD/An analysts were interested rather than a critical analysis of the user's problem de novo. Moreover, most of the assistance was generated from contractor developed concepts rather than indigenous effort. ORD/An served primarily as a broker in these matters -- with one important difference -- they usually furnished the funds for the contracts. Two problems arose from this arrangement. First, OCS, which would eventually have responsibility for processing any application which grew out of such arrangements, had no means of access to this work and was not represented by ORD/An in spite of the fact that they were in the same directorate. Second, the users could do little to alter the course of experimentation because they had no control over the funds which were financing them.

6. We continue to believe that there is a class of problems on which long-range research and development are required to enhance the Agency's data processing performance. These requirements are in general so similar to those of the military, commercial and industrial communities that the existing competitive pressures for

hardware or software advantage in these fields will assure more attention to the problem than our individual effort could engender. A more critical problem in our view is to provide an opportunity for Agency staff people in these fields to keep in touch with the research and development which is in progress, particularly those developments which might provide more responsive computer systems for the Agency. More regular attendance at major professional meetings and visits to research installations on the part of the regular staff will probably accomplish more for awareness of ADP developments than a professional research staff, particularly one which is dissociated from the mainstream of Agency data processing activity. We may emphasize here, as we have elsewhere in several office appendices, that simple physical separation of a research facility from its principal customers contributes to this problem.

7. Our critical research and development need is understanding more fully how to use the equipment available to us at present rather than for new equipment. Indeed, we know of few users of modern computing equipment who would aver they make genuinely effective use of it. Unless new equipment emerges with whole new computing concepts, we should be in a struggle with ourselves for the next 5 years just to use effectively what we now have.

8. Then too in observing R&D work we are often faced with more solutions than we have problems but with no fit between the solutions and the problems. Developers have systems which are a sure cure for problems that they know must exist but which they haven't quite found yet. And users can conceive of systems which would answer their problems in a trice, it's just that they can't articulate them well enough that anyone can implement them.

A General Statement of our R&D Problem

9. Thus, we believe that the major research and development requirement of the Agency is for problem definition, solution, system design and programming support for new computer applications. We are engaged in the production of intelligence rather than the production of computers. We need better definition of that process and its potential interface with automata if we are to use the latter more effectively. Only the user and the processor can give a general estimate of the value of such an application which will serve as a guide to whether or not work on the application may be undertaken. ORD has provided support for this type of activity. Indeed its major achievements are in this area. Nevertheless,

development of computer applications is the primary preoccupation of the functional divisions of OCS. Moreover, the Advance Projects Staff continued to work on the development of and experimentation with major general software systems throughout the period ORD/An has been in existence. Indeed, this staff was one of the major users of ORD equipment during their early efforts to develop a time-sharing monitor.

10. OCS development work at ORD proceeded quite differently than that of other components. It consisted of OCS personnel on the premises doing the design, programming and program test work on a system specified by OCS. Other users conventionally were consultants who sought to describe general requirements they might seek to satisfy on systems which ORD had obtained or intended to obtain. These consultants were later employed to collect and organize data bases used on the test system and to serve as test personnel on interactive systems, those requiring some form of man-machine interface. Where problem definition was essentially complete in the case of OCS, development work went forward rather easily. Where problem definition was negligible, development was slow and the problem tended to be defined by the research facility rather than by the user or by joint effort.

11. Thus, OCS experimentation produced systems ready to run in OCS. ORD experimentation with other parties often produced systems which were incompatible with OCS operational standards and operational philosophy, even where there was general agreement that OCS would provide the processing environment for the application. Such systems could only be absorbed by OCS by redesign of the application system or the host system or both. Such a practice should not occur, but it has been difficult to eliminate. We make no brief that all new computer applications should be processed in OCS. Patently where a particular processing activity is highly organization specific, and equipment specific, or where the organization involved is remote from the headquarters location, there is an immediate presumption that a stand alone computer may be a better solution than a link to the OCS computer center. Finally some judgment other than OCS or the target data processing center is desirable. We believe that this might better be provided by an independent review facility than an independent application design activity.

12. An experimental or development facility to be successful must be free, indeed ordered, to take on all comers who have problems which they expect may be reduced by the incorporation of automata

into the processing or production activity. More than that such a facility must be fully conversant with the present and planned processing systems available to OCS and to other computer processing facilities in the Agency. Finally, it must be able to visualize the fit of any given application to these systems or its susceptibility to economic processing on a stand-alone computer.

Conclusions and Recommendations

13. Both ORD and OCS have long experience in general contact and service throughout the Agency. Beyond that point most of the activities which seem critical to us for information processing research and development have always been performed in OCS. A genuine effort was made early in the development of ORD/An to coordinate its activity with OCS, but this effort disintegrated from a halting start to a general awareness of each other's existence although some effort to restore an interaction has begun this year. Individuals located in OCS have both the technical expertise and the awareness of processing activity throughout the Agency required to provide an optimum service to the user. What must be created in OCS is confidence that Agency management will support the separation of general development activity from the press of production activity. Having seen the Agency willing to support development work with money and positions elsewhere and having seen the diseconomies of the present system, we believe that OCS would be willing as well as able to undertake control over this activity.

14. We recommend that the DD/S&T review the division of effort between ORD and OCS in the area of information processing research and development against an alternative allocation of function and effort which would:

a. Provide for the subsequent problem definition, and computer application design and development effort to be moved from ORD to OCS.

b. Provide for the transfer of essentially standard computer processing equipment from ORD to OCS and for OCS to provide a level of experimental or developmental computer processing time necessary to support the expanded experimental function outlined above. We would for example urge the processing time might be made available on machines appropriate to the work involved rather than on a single machine which is used only for experimental work.

c. Provide for the transfer of other equipment from the IPRD laboratory to those surviving or anticipated development programs which may use it most effectively, the rest to be transferred to surplus.

d. Provide for a review of existing ORD contracts through the Information Processing Board and selected prospective users to determine which of these contracts should be continued and under whose leadership they should proceed.

e. Provide that subsequent ADP equipment or software test and analysis be conducted by OCS except where the items are a direct adjunct of a special processing center such as NPIC. The special unit would procure and test the latter products.

f. Provide for OCS to issue a current awareness publication similar to its present Tech Notes to announce new activities, new products, and new developments which its research and development component consider of general interest for Agency components engaged in information processing.

15. In addition we recommend that the DD/S&T and the Information Processing Board reject the proposal of the R&D Subcommittee of the USIB Information Handling Committee which proposes a community wide R&D Center on the basis of the recent experience with COINS and the IPRD which we believe demonstrate both the difficulty of an integrated community activity and the impracticality of performing research and development on non-existent or badly defined requirements.

16. Finally we recommend that research and development projects or programs in the area of information processing be submitted to the same scrutiny as that proposed for ADP projects in the section below dealing with management.

VI. ORGANIZATIONAL AND MANAGEMENT ELEMENTS OF AUTOMATIC DATA PROCESSING

VI. Organizational and Management Elements of Automatic Data Processing

Introduction

1. Our remarks on organization and management derive from our observation of the intelligence production milieu, from the presentation of organization activity and the discussion of organization problems. In the context of the ASPIN study and findings, they seem to us better made than omitted.

2. The most manifest complaint voiced by individuals and groups in informal briefings was dissatisfaction with the management of ADP activity. Perhaps this is a convenient outlet for the frustration at lack of more effective ADP performance on the part of the individuals voicing these complaints. The nature of this complaint was less a criticism of too much management than it was an indictment that there did not seem to be enough integration and coordination of ADP activities, particularly between offices and directorates. True, there were mechanisms established to treat such coordination, but they took too long to act. And their solutions were too often a series of partial solutions rather than a well integrated total solution to the problem. We believe that some of these problems are sufficiently complex and diverse that they are difficult to attack operationally in any other manner than by successive approximation. Other problems are quite amenable to solution.

3. The frequency and force of concern about management problems were such that we asked our consultant-contractor to review these matters and to prepare an informal paper on their reactions to this problem. It appears in Annex II. It contains essentially the same complaints that we have had except that we would tend to characterize the extent of duplication of program activity as somewhat less objectionable than they. Our contractor, unlike our indigenous critics, is inclined to fault our failure to be more sweeping and more far reaching in our observations and recommendations. We are unquestionably somewhat constrained by what we think operationally possible, but we harbor no ideal solution which blushes unseen because we think it cannot be executed.

Centralization vs. Decentralization

4. The dichotomy of centralization vs. decentralization is perhaps unfortunate. We use this apposition here to characterize a general position on computer processing as opposed to computer problem solution which we have discussed above. We believe that the responsibility for the definition of the functional problem and the

computer solution to it must be developed in a highly decentralized environment, by the analysts and their supervisors. Indeed certain computer analysts who are responsible to OCS for technical standards and personal career development might best be assigned to production organizations for assistance to this problem definition, problem solution process. In contradistinction, we believe that the computer processing activity is best treated by a high degree of centralization of processors. There are great economies of scale in central processing both in equipment (hardware) components and in the concentration of skilled operational and programming personnel (software). We believe that the evolution to greater use of on-line or interactive processing activity will markedly increase these economies of scale in large processing units.

5. The present de facto management of Agency ADP resources is rather highly decentralized despite the concept of the Executive Director-Comptroller as the responsible ADP manager. Such decentralization offers the advantages of being more responsive to user needs. It is able to satisfy both short-term and long-term requirements from individual units for data processing support. It suffers, however, from overlapping and conflicting long-term objectives on the part of different Agency components, from inability of one unit to profit from the experience of another and from insufficient planning and inadequate control.

6. [redacted] The earlier decision not to centralize all ADP activity has become accepted as a decision to localize such activity by any organization whose ADP activity has been questioned by any central authority. Such an interpretation tends to magnify the difficulty of obtaining increasingly complex systems. Unless central review can produce a route to coordinate design of such systems, we shall spend a great deal more money getting a distinctly inferior product. That product (really a series of organization products) will become increasingly difficult to support in terms of communications, in terms of analyst time devoted to learning additional systems conventions and languages, or conversely, in multiplying copies of individual machine files among many different computer systems. Our specific comments on each of the elements which we believe need attention indicate a position vis-a-vis centralization. We believe this position may be rather near what Agency management thought its direction was; we think that this position is in fact quite different from what obtains at present.

7. We realize that the Agency has previously struggled with the question of centralization vs. decentralization of control over data processing. The earlier decision to avoid the complete centralization

of data processing in an effort to increase the ADP alternatives available to the individual with a problem to solve seems as valid to us today as it seemed to those who announced it in 1965. We believe, however, that the present permissiveness in the extent of acquisition of ADP equipment and programming and processing personnel by CRS and ORD, for example, is considerably greater than was implied or intended in the 1965 decision of its subsequent implementation. We believe that a small stand alone processing capability for CRS to maintain its contact with its punch card processing records and reduce the future processing costs of maintaining these records was economical and essential, even the expansion of this activity to provide the initial capability to test the AEGIS software. Subsequent expansion of the system to full scale AEGIS operation, and to taking on the OSR experimental programs being moved from ORD involves a potential long range expansion that must be weighed more carefully than it has been to date. CRS now makes the clear statement, and we may expect others to make it if CRS is successful, that it does not intend to be dependent on OCS for any data processing service required for the performance of its basic mission. Generally adhered to, this would provide the basis for creating more than 20 ADP centers to perform the vital data processing required to execute the mission of each of as many like offices (each of which is engaged in performing a service of common concern).

8. Such decentralization will unquestionably give the individual office better control over when its work gets done, but it gives this control at the expense of increased requirements for (1) highly skilled personnel for whom the offices find it difficult to offer a genuine career opportunity, (2) space for processing centers and their attendant logistics, (3) larger long-run increases in equipment and personnel expenditures to provide the redundancy needed to back up on-line or real-time systems, and (4) a Babel of systems and languages which will make the long-run communication between analysts and these systems, or among the systems themselves, exceedingly difficult. It potentially puts a large number of offices squarely in a business with which they have had little experience, a business which has the highest rate of technological change in our society. It creates a need for office level management to perform functions they have never performed before. Finally, it puts a great pressure on top management to assure the economy and the easy interchange of information among these systems. Either of the latter responsibilities would imply the need for a much larger and more complex staff review function at the Agency level than has existed in the past or is proposed in this paper for the future.

9. A centralized processing system on the other hand could be permitted considerable head-room in terms of its capacity to accommodate processing requests and still permit great economies

over the proliferation of general purpose processing activity over many functional organizations. Indeed, this economy could be obtained and still permit certain functional organizations to acquire their own computers for special and compelling technical processing reasons such as the remote location from headquarters, a unique processing activity which can only be economically supported (programmed and operated) by the functional office, or the need for a small computer to provide economy of interaction between a remote area in the headquarters area and the large computer center.

Problem Definition and Project Control

10. The most critical element in the control of ADP activity is to seek to control what may be characterized as the definition of the problem and the approach to its solution. Our strong preference is to concentrate problem definition and solution in the hands of the functional components. Thus, functional offices would be responsible for developing and administering a set of projects to provide solutions to their problems. Computer processing centers incidentally would be responsible for developing and administering a set of projects to provide a processing milieu capable of dealing economically with the processing load generated by the functional organizations. The weakest link in the present control chain is that of project control. Organizations often embark on highly complex programs with little identification of what is to be done and even less indication of how it is to be controlled. To attempt to meet a minimum statement of the need of project definition or statement we have developed Appendix A. Appendix B which accompanies it was directed at a brief statement of the problem of dealing with the contractor in project control for those who are unfamiliar with the DD/S&T Project Officers Handbook which should serve as a more definitive statement of this problem.

11. A lack of coordinated Agency-wide planning and review adds to the difficulty of problem definition and control. Dispersion of computer and manpower resources among a number of processing centers working on similar applications, and with little central information on the specific characteristics of these applications or their performance, makes for inefficient use of scarce, skilled computer programmers and operations personnel. Development of systems designed at the office level but exercising an impact over other Agency components tends to produce an eclectic Agency plan which is seldom better than the lowest common denominator of office objectives. With efforts and costs thus dispersed, total expenditures for common activities are often difficult to bring together and analyze at the Agency level and they are never a matter of concern below that level. An Agency Five Year ADP Plan published two years ago constituted an

initial effort at identifying the level and distribution of ADP activity in the Agency and set some forecasts for growth. It has not been revised to bring it up to date or extended to provide greater definition in Agency ADP objectives.

Project Review

12. Obviously, the first step in project review rests with the manager of the functional organization to determine that the project will accomplish an objective which he thinks important to have the Agency achieve and that resources to process the project would be available at the time the project would be ready for processing. Such a review should suffice for all small applications or systems software projects.

13. Large projects which we would define as those requiring a total development expenditure of $50,000 or more for in-house and contract expenditures should, in our opinion, require a centrally controlled technical review in addition. This review might best be performed by a technical panel composed of Agency personnel who are familiar with the class of application proposed in the project. The objective of technical review of the project would be to assure that the project was technically feasible (or for R&D efforts was within the prospective range of the technology whence it sought support), and that it was taking optimum advantage of existing developments in this area that had been undertaken by the Agency.

14. Project review should be a continuing process. Thus, threshold projects of the sort outlined above might well be reviewed annually until implemented, given a general performance review some 12 to 18 months after implementation, and be reviewed at 24 month intervals thereafter. Projects which were never afforded technical review in their development process should be submitted for a general performance review and continuing review in the same manner as threshold projects after implementation.

15. The lack of adequate documentation on project initiatives, or their goals, accomplishments and particularly on their costs makes control at any level difficult. The manager of the functional component has only recently begun to get a rough measure of the resources spent in his data processing. Even now he usually gets programmer or computer time figures rather than dollar figures, although all of the major processing centers can now give quite precise dollar figures as well. The functional manager has no control over the programmers who have allegedly worked several hundred hours for his account during the last quarter with no apparent change in the status of his project. Thus what potential control data exist are sufficiently removed from the functional manager's responsibility

that he cannot exercise any choice save the choice to proceed or to stop. The data processing manager is equally free of responsibility. He is simply rendering a service, doing what he was told to do. He may be partially responsible if the ADP job or its objectives are not met, to the extent that the data processing system would not perform the functions his analysts had indicated it would.

16. Automatic data processing activity meanwhile increases in cost and potential impact on Agency operations. Independent development of processing systems by separate Agency components which was little more than nettlesome or aesthetically displeasing a few years ago has started to become both wasteful and a threat to integrated operations over the organization as a whole. We have entered a series of developments which promise a "Babel" unless greater concentration or compatibility of processing systems can be achieved. A more aggressive approach to central planning and control is required, but we believe that it must be tempered by permitting user initiative in seeking ADP support. The process of project review and its documentation are so limited and localized that there is a substantial overlap in applications development. Again this leads to less than optimal use of ADP personnel. These technical personnel are too immobile because of their isolation in production and processing units to provide the communication needed to avoid such repetition. One avenue which the ASPIN Staff thinks may make a considerable contribution both to data processing personnel and to the units for which they work would be the development of an Agency ADP Career Service.

ADP Career Service

17. Personnel engaged in the analysis, design, programming, and processing of ADP systems are scattered over a set of Agency career services which existed prior to the introduction of computers to the Agency. The training, experience, and skills required for these people to perform their duties, while not necessarily identical in every respect, are basically the same in each of the components. That is, a computer systems analyst in OCS requires essentially the same background and training to perform his duties as a computer systems analyst in AID/PSG/NPIC. Although the equipment these people work with is not the same e.g. IBM in OCS vs. UNIVAC in NPIC, a relatively short period of time (2 to 3 weeks) is required to train such personnel on new or different equipment whereas the basic skills of the job require years to acquire. The same thing is true to a great degree of some of the other types of personnel required i.e., programmers, machine operators, etc., in an ADP support component. A Career Service for ADP personnel could take full advantage of this situation to improve Agency ADP support and also to expand the career development possibilities for the ADP professionals.

18. The more important benefits of such a Career Service to the Agency and to the personnel involved would be:

1) more flexible use of ADP personnel by Agency management would be possible, the assigning/detailing of personnel from one ADP component to another would be simplified. When high priority projects and/or unusually heavy workloads existed in one of the components, the detailing of key, scarce personnel could be made with a minimum of paper work and in a minimum of paper work and in a minimum amount of time.

2) greater variety of tasks and career development possibilities would be available for ADP personnel which would make their work more interesting and challenging, help reduce the loss of such personnel to private industry and to other government agencies, and increase their value to the Agency.

3) the possibility for cross fertilization of ADP experience and developments among the various ADP components would be greatly enhanced.

4) a more uniform set of position standards, qualifications, and responsibilities for each job level could be developed. This then would result in a more uniform grade structure for like positions throughout all of the ADP components.

5) unfair competition between ADP professionals and professionals in other disciplines for promotion to higher grade levels would be greatly reduced if not entirely eliminated. This competition has often worked to the disadvantage of ADP personnel in Offices where the primary responsibility of that Office is something other than that of providing ADP support. As a result morale has suffered with consequent effect on the quality of work.

19. We recognize that there are certain weaknesses in a career service of this type, i.e. one which extends across several directorates and offices and several categories of professional personnel. But we believe that the potential for improvement both in the range of professional work alternatives and in the opportunity for professional development and career advancement offered by an arrangement of this sort more than compensate for the disadvantages.

Training

20. Education of Agency personnel in the fundamentals of automatic data processing and its application to information processing and intelligence production was delayed too long. Early training in

this field was highly specialized in terms of the range of data processing activities and was confined to a small corps of individuals who were to establish and operate a data processing center. As a result the early computer processing activity in the Agency was almost wholly devoted to a procession of information storage and retrieval applications and a few modified punch-card accounting routines. The only substantial computation activity afoot was carried on by the Agency's photo interpretation activity wholly outside the prime Agency computer center. It was designed to reduce the data associated with the position of the camera platform vis-a-vis photographic images so as to make possible accurate measurement and interpretation of the objects displayed on photographs.

21. By the mid-1960's, the need of the data processing user for better data processing support and the disappearance of the computer center as a source of magic created an interest in more general training for intelligence analysts and for better technical training of computer programmer-analysts. The series of courses for Agency personnel prepared in 1966-67 and the associated external training have provided an excellent base for the subsequent education of individuals associated with data processing. Essentially this training has been made up of five constituents, each of which we think has an important place in staff training.

1) A general ADP orientation course which deals with computer fundamentals and the character and scope of Agency data processing.

2) Specific programmer trainee (for beginning, intermediate, and advanced professionals) courses to prepare programmers for operation on Agency ADP equipment in support of the full range of applications processed.

3) Individual office ADP training programs which combine specialized programming instruction with analytical training in computation and/or logical techniques required to improve the adaptation of computer processing to support office analytical objectives.

4) Highly specialized courses at external training sites that provide preparation for specific ADP techniques which are too limited in their application or too specific in their time requirement to warrant the expense of internal preparation.

22. On-the-job training is a regular part -- and usually the major element -- in each individual's preparation for an ADP or ADP related activity in the Agency. Even established journeyman computer programmers or computer operators must spend a considerable

number of hours on the job being instructed in local systems and local operating conventions. Novice professionals or trainees will find most of their time during their first 6 months of activity devoted to training on the job and a regular effort to sustain this sort of training over their professional career in the Agency.

23. We believe that this training program is well balanced, offering the full range of educational opportunity required, appropriate attention to local activities and standards, and both of these at the lowest cost consistent with the quality of training desired. We would encourage two trends in training over the next few years:

a) More widespread development of office level training in intelligence production and information processing offices. Such training provides not only more effective use of ADP resources available but also a better coordination or interaction of office analytical activity employing ADP than is possible by any other means.

b) More widespread assignment of OCS programmers for on-site support of specific offices for extended periods of time. We would suggest a minimum of 12 months and a maximum of 18 months. This provides excellent training for both programmer and functional analyst in their mutual problems in the production milieu, for better control over problem solving and programming support, and for the programmer's escape back to his processing center base quickly enough that he has not lost his computer center discipline or his touch with technical developments. (We believe present use of programmer trainee course for beginners provides the opposite flow required for production analysts).

24. One other training requirement is on the horizon, the preparation of individuals who will interact with large time-sharing systems. The question of what constitutes use and what abuse of a time sharing system may be critical to system performance and should be carefully understood by all who are licensed to use the system. The OCS Interactive Services system was explained in some detail to the computer programmers who were to use it before it was installed. Anyone else interested in the system was required to prepare himself with system and facility manuals on a catch as catch can basis. The result has been less efficient use of the system, and an expensive trial and error learning procedure in which no one is quite certain as to whether or not he makes optimum use of the system in his present activity. The advent of the Integrated Information System (IIS) at NPIC this autumn has been prepared for much more intensively via a series of formal training courses. Although the two systems have

different general objectives, the difference in approach to preparation of analysts for system operation may give us another measure of the appropriate level of training to accompany introduction of on-line systems.

25. A training program must be developed as an integral part of any large on-line or time-sharing system. This program should acquaint management with the general content and capabilities of the system and specify an operational evaluation period (about 14 months). It should provide for training users of the system in techniques of operation and results they should expect from the system. Finally it should provide operators of the system with a detailed explanation of handling input to the system and of assuring continuous, reliable performance. Even with careful preparation, there will be many surprises, but these can be dealt with more quickly and their impact upon the system evaluated more reliably if the individuals dealing with the system have an appropriate understanding of it at the outset.

Conclusions and Recommendations

26. We recommend that: The Agency reassert a policy of providing a high degree of centralization in data processing activity in the Office of Computer Services, that this policy be tempered by permitting the acquisition of small or medium computer processors by functional organizations where there is a demonstrable technical-computational economy in using a stand-alone computer system, and that this policy continue the present emphasis on the functional component (user) responsibility for problem definition and problem solution. In short we recommend that computer organizations develop the systems necessary to run the computers and run them, and that functional production people prepare the data and the processing steps required for its transformation by computer.

27. A central technical management review of major ADP projects be created under the present umbrella of Executive Director-Comptroller responsibility for Agency ADP management, that a full-time position of ADP Advisor to the Executive Director-Comptroller be created for an experienced ADP professional whose responsibility it would be to:

1) advise the Executive Director-Comptroller on all professional/technical matters relating to ADP;

2) be chairman of the IPB and the director of its permanent staff;

3) review the various local plans, provide technical input to IPB and, periodically, develop a statement of long term ADP objectives for the Agency;

4) assign computer application design proposals to the suitable functional/technical review components;

5) prepare Agency-wide ADP technical standards;

6) serve as chairman of the Agency-wide ADP Career Service Board;

7) serve as focal point for internal leadership and for external relations in ADP technical/professional matters.

28. Existing central ADP planning be strengthened to provide:

1) for a more definitive outline of Agency objectives to be achieved in related or overlapping office plans and for regular revision and publication of the Agency ADP Plan,

2) for the definition and publication of Agency-wide ADP technical standards beyond the present work on nationwide (USASI) standards, and

3) for a standard format and procedure for the proposal and review of major requests for the acquisition of computer systems or of computer processing applications.

29. A means of pricing data processing services performed by computer centers be developed, and that each user component be required to budget for its data processing services and transfer the funds to pay for these services in essentially the same way that property funds are handled.

30. An Agency ADP Career Service be created.

31. Existing ADP training programs introduce additional emphasis on the changing responsibility or role of the user in an on-line and/or real-time computer environment, and that functional organizations review the need for unit training of personnel in the use of quantitative and/or logical techniques in indigenous analytical problems.

32. The Director, OCS be an ex officio participant on the Information Processing Board and that the DD/S&T should be represented on the Board by an individual who reflects the computer user population of the whole Directorate. The presence of the Director, OCS on the Board is imperative, but we believe he should participate in his capacity as director of computer processing rather than as the representative of a Directorate with large processing requirements.

Appendix A

DEVELOPMENT AND MANAGEMENT GUIDELINES FOR REVIEW OF ADP PROPOSALS

General Objectives.

1. ADP project proposals may be expected to achieve one of the three following general objectives.

a) Develop a system which produces at lower total cost an activity which we need and produce now in an acceptable time frame.

b) Develop a system which replaces existing activity but improves the accuracy, speed and reliability of the product by an amount sufficient to exceed its cost of development.

c) Develop a system which provides for activity which cannot be undertaken at present other than in experimental of exemplary form, i.e. provides for a level of complexity, speed, and accuracy which cannot be achieved economically through existing man-machine techniques.

2. We believe that the greatest intelligence gains may be expected through encouraging the development of class c. applications. Such applications should be encouraged through systematic support of at least 3-5 experimental programs a year looking toward increasing the role of this type of product in the Agency ADP mix. In general class a. applications should be avoided unless there is an immediate or near-term reduction in the cost of processing of an amount which will recover the initial cost of the application within 5-7 years. Class b. applications should be undertaken readily where they enhance processing speed on established and enduring activities which are labor intensive, or where the need for accurate, reliable results is critical to national security. They should, however, be avoided where they replace mechanized applications unless they provide an immediate or near term reduction in cost that will permit recovery of the cost of the application in 5 to 7 years.

Management Guidelines for Review of ADP Proposals.

3. The most critical need in automatic data processing control is a formal means for review of ADP activities. Small ad hoc processing projects should obviously be excluded from such control except insofar as they are collectively reflected in significant additional requirements for computer time or for additional remote connections to a central time-sharing computer. Although stated in terms of guidelines for initial project approval, these controls are equally essential for a continuing review of established projects. Criteria for establishing the merit and the viability of a potential ADP application are readily identified. They consist first and foremost of a definitive statement of the objectives of the application and its relevance to the sponsor's mission. Such a statement should provide both management and operating personnel with specific goals in terms that each can recognize and evaluate, a condition which is all too often absent.

Project Objectives.

4. An ADP objective such as "automation of a particular activity," otherwise unspecified, contains no means to evaluate the objective. It may be considered satisfied by some if only a small fragment of automated assistance is developed for the activity. Such a procedure becomes a vehicle for spending considerable sums to do very little. On other occasions it may become a vehicle for continuing development, frequently without any implementation, on the argument that all elements of the activity have not yet been designed and integrated into the application or system. Thus large sums may be spent for no tangible result, or such sums may be spent to achieve much less than was anticipated.

5. And an additional note. Specificity of objectives should not be confused with the notion that applications or systems once implemented are fixed for all time, a concept which "systems analysts" have been prone to encourage. An effort ought to be made to be reasonably definitive when developing an automated activity because changing such systems may be rather expensive. It is enough that individuals approach any given objective with a notion of how the system may evolve if it is effective. This will make possible initial design that facilitates such changes as they may be required. Changes in applications must be expected, because both people and their understanding (technology in the general sense) change. Moreover, we must expect the rate of this change to accelerate over time.

Statement of Work.

6. A statement of work or tasks should accompany the objectives. The work statement should identify what needs to be done to achieve these objectives, who needs to be engaged in this work, and what interdependencies may exist among the tasks and the individuals involved in the project. Deadlines for discrete segments of the proposal as well as for the whole project must be established. A need for realism in schedules pervades the intelligence business but is particularly acute in ADP projects. The statement of work should also contain an identification of important alternative means of accomplishing the project objectives with "trade-offs" among these alternatives clearly set out.

Expenditure Estimates.

7. A statement of the estimated expenditure for the application is automatic. Estimates of the costs of identified alternatives should be included. Where internal manpower and machine power is to be used on the project, an estimate of these expenditures at actual or average cost should be included in the total cost.

Statement of Risk.

8. An important factor, particularly in the development of large systems, is the risk involved in undertaking the project and the risk in its pursuit once it is undertaken. Certainly many projects are nearly risk-free or at most threaten the loss of the resources devoted to a particular segment of the project. Occasionally -- indeed often on large projects -- the risk to the organization and to the Agency may significantly exceed the expenditure for the project during a particular time period or for the entire period over which the project has been active. The manager should keep himself continuously apprised of this element of risk even as he should assure himself on the objectives of the project.

Project Review.

9. Finally, a project proposal should contain an indication of the procedure for internal review of progress on the project and a series of check-points or decision points at which the project may submit a status report which will permit executive review of progress and decision as to how or whether to proceed. If this is an intermediate point in the period for which funding or other authorization is sought, the statements of cost and risk ought each to present expected level of cost and risk at that point.

10. We have incorporated these concepts into a suggested outline to accompany an ADP proposal or to serve as a means for periodic review of major operational ADP Applications.

An Outline for ADP Project Proposals.

(1) What is the objective of this proposal?

A. End products
B. Programs or tasks supported
C. Existing activities to be affected (how?)
D. Intermediate or longer range objectives that are relevant.

(2) A statement of work.

A. What tasks must be done to accomplish these objectives?
B. Internal resources men and equipment
C. External (contract) resources
D. Substance of these assignments and the extent of interdependence.
E. Who will manage the project?
F. Alternative implementation patterns.
G. Major "trade-offs" involved.
H. Completion date(s) for review and further decision.

(3) Additional equipment required?

A. Type
B. Processing organization
C. System (existing or new)

(4) What will be the estimated total expenditure for the project?

A. Amount by each major substantive element?
B. Amount by each major type of work or product provided to the project?
C. In-house vs contract expenditures.

(5) What risks are presented by this proposal?

A. Can a partial success be operational?

B. An all or nothing activity?
C. Any secondary risks, i.e., extend beyond the expenditure commitment to the project?

(6) Management level review points or decision points.

A. Intermediate review and evaluation.
B. Final review and evaluation.
C. Operational Review.

Appendix B

MANAGEMENT GUIDELINES FOR ADP CONTRACTOR RELATIONSHIP

1. All of the information presented here is contained elsewhere in the Project Officer's Handbook which should serve as guidance for the individual who must perform the porject monitor function. It is abbreviated in this presentation to characterize the process for the management consumer of this document. It is important, first of all, that the problem to be addressed by a contractor be written in clear, understandable language and reflect the needs and efforts of both the user office and the computer facility. See Annex A. This paper should include a detailed description of the problem together with the objectives to be achieved by the project and the role of the contract in achieving these objectives. In addition, where appropriate, detailed programming specifications and documentation requirements should be furnished the contractor and followed by him. The time frame for the completion of the contract and the specific products expected from the contractor should also be spelled out clearly. A detailed check list for the manager is set out in paragraph 5 below.

Contract Monitor Function.

2. Emerging from the preparation of the paper discussed in the preceding paragraph should be a monitoring team consisting of a representative from the user office and one from the computer facility to oversee the work of the contractor and periodically review his progress. It is important that this team establish the kind of relationship and understanding with each other and with the contractor that ensures contractor responsiveness to team guidance and direction.

3. At the very outset of a contract, a meeting should be held between the monitoring team and the contractor at which the contractor would spell out, in detail, his understanding of tne problem and his plans for attacking it. Any misconceptions on the part of the contractor should be corrected by the monitoring team and any differences in approach should be discussed and resolved at this meeting. Periodic meetings, perhaps weekly at first, tapering off to monthly or even quarterly depending on the size and complexity of the

contract, should be held to ensure that the work is progressing satisfactorily and in line with the desires of the monitoring team. Alternations or changes in the scope of the contract should also be discussed thoroughly, with the monitoring team exercising approval or disapproval authority over any and all suggested changes. Approved changes should of course be fully documented and where a change in the cost of the contract results, the contracting officer from the Office of Logistics should be informed immediately so that he can take whatever action is necessary.

4. Monthly reports should be prepared by the contractor which set forth the percentage of contract completion, the funds expended during the month and to date, milestones reached in the contract and any problems encountered or anticipated which need resolution.

5. We have summarized the process outlined above in a contractor relationship checklist. This has been designed as an illustrative rather than an authoritative document.

Contractor Relationship Check List

A. Clear Statement of the Problem Should Include:

(1) Detailed description of problem
(2) User needs
(3) Computer facility needs
(4) Objectives to be achieved by the project
(5) Role of contract in achieving project objectives
(6) Detailed programming specifications
(7) Documentation requirements

B. Composition and Responsibilities of Contract Monitoring Team.

(1) Representation required from the user office and from computer facility

(2) Team provides guidance and direction to contractor

C. Monitoring Team and Contractor Relationship

(1) Initial Meeting

(a) Contractor's understanding of the problem reviewed and any misconceptions corrected

(b) Contractor plans for attacking problem reviewed and any differences in approach resolved

(2) Regular Meetings

(a) Frequency of meetings

(b) Progress of work reviewed

(c) Changes, alterations, changes in scope discussed

(d) Monitoring team exercises approval authority over any and all suggested changes

(e) Approved changes thoroughly documented

(f) Office of Logistics contracting officer informed when changes require alterations of basic contract

D. Monthly Reports from Contractor Should Include:

(1) Percentage of contract completed

(2) Percentage of funds expended during month and to date

(3) Milestones reached

(4) Problems encountered or anticipated requiring resolution.

1. FOREIGN MISSILE AND SPACE ANALYSIS CENTER

TABLE OF CONTENTS

1.1 Introduction.

The Foreign Missile and Space Analysis Center (FMSAC) is responsible for production of all-source intelligence on foreign missile and space operations; production of all-source intelligence on foreign offensive missile and space systems; for maintaining a 24-hour Control Center to monitor foreign missile and space events; and for disseminating, as appropriate, results of analyses and evaluations on foreign missile and space activities. In addition, FMSAC provides substantive contributions to National Intelligence Estimates, and both administrative and substantive support to the Guided Missile and Astronautics Intelligence Committee (GMAIC), a committee, of the United States Intelligence Board.

25X1A

Next 10 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

[redacted]

1.4 Conclusions and Recommendations.

The comprehensive ADP needs of FMSAC for its intelligence production requires daily consultation and support from the Office of Computer Services(OCS). The established excellent working relationship between OCS and FMSAC personnel has been instrumental in developing and maintaining ADP applications to support FMSAC's intelligence production. The following recommendations are suggested to continue and expand the OCS-FMSAC ADP efforts:

(1) Expand coordination between OCS and FMSAC when changes in OCS digital computer equipment, systems, and/or operating procedures are to be implemented.

(2) Extend OCS efforts to obtain and/or evaluate a general purpose data management system for Agency use and to support FMSAC's Information System requirements.

(3) Continue experimentation with use of ADP for intelligence production, (e.g., generate the FMSAC quarterly reports of analyses on foreign missile and space events from the new comprehensive Information System).

(4) Provide expansion capabilities to improve data processing of current information receipts, to economically process peak loadings of information, and to provide for processing a steady growth of foreign missile and space information. Information refers to both measured or intercepted data and textual messages.

2. OFFICE OF ELINT

TABLE OF CONTENTS

2.1 Background.

The Director of ELINT is charged with the primary responsibility for Agency ELINT activities, with establishing and coordinating the Agency ELINT program, with providing technical support and guidance required for and analyzing and reporting the product of Agency ELINT projects, and with supervising or conducting all research and development required for Agency ELINT and related COMINT activities. (OEL does not produce intelligence, although its intelligence information output may be incorporated verbatim in intelligence reports published by other components). It is the responsibility of the Director of ELINT to advise the CIA SIGINT Officer in matters of ELINT policy and to maintain liaison on technical matters pertinent to the Office of ELINT with NSA and other Government Agencies.

The Director of ELINT shall:

(1) Provide management and support for tasking and technical guidance for field ELINT activities.

(2) Provide for optimum analysis of the ELINT product of Agency operated and supported activities for the purpose of the reporting of intelligence information.

(3) Originate, develop and control special ELINT projects in support of DD/S&T needs and operation.

[redacted]

25X1A

Next 3 Page(s) In Document Exempt

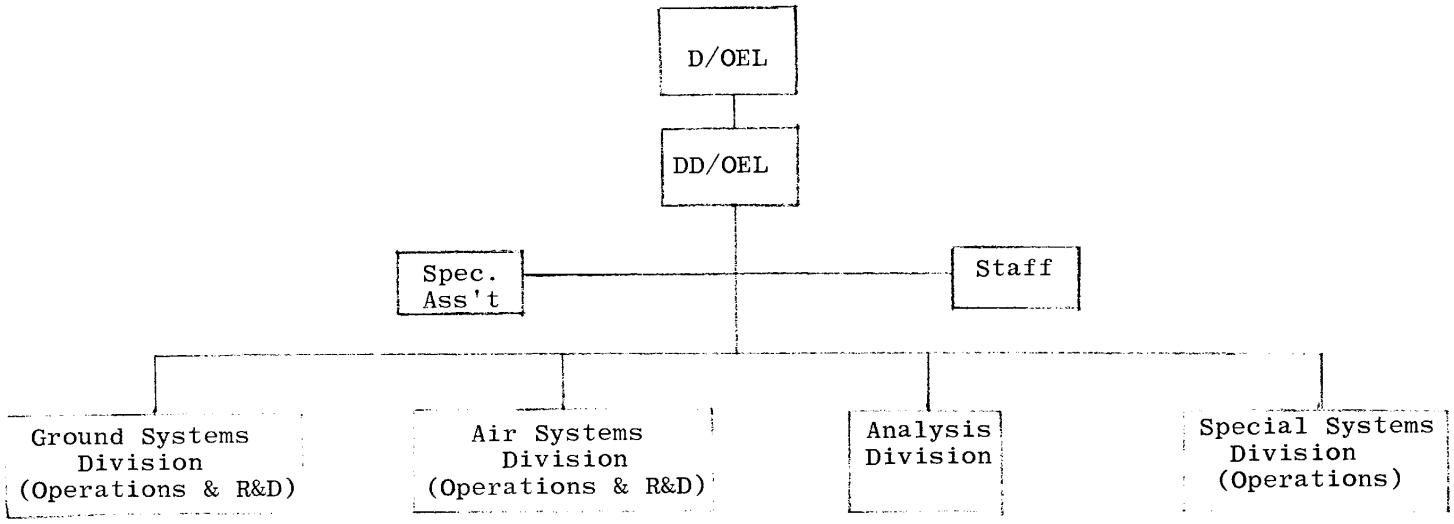
Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

2.4 Recommendations.

We believe that the development of advanced ELINT processing systems, capable of fully exploiting the signal environment in which the Agency has undertaken a collection and production responsibility is essential. Better coordination of OEL-contractor interaction with OCS systems analysis efforts to support both the OEL processing and the OEL collection activities seems to be needed. Indeed, we observe that there is a large and diverse effort in signal analysis spread over a wide range of organizations including OEL, FMSAC, OCS,

Table 2.1

OFFICE OF ELINT



3. OFFICE OF SCIENTIFIC INTELLIGENCE

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

3.1 Background

The Office of Scientific Intelligence is responsible for the production of intelligence in the following fields:

(a) all foreign atomic energy activities, with particular emphasis on weapons development;

(b) foreign R&D on ABM systems, antisatellite systems, air defense systems, antisubmarine warfare systems, and offensive and defensive aircraft, cruise missile, and naval weapons systems;

(c) foreign R&D in the life sciences, including biomedical aspects of space-flight systems, biological and chemical warfare, medical and public health practices, [redacted] 25X1B

(d) foreign R&D in physical and engineering sciences and long-range military threats based on scientific and engineering advances. 25X1A

[redacted]

Next 9 Page(s) In Document Exempt

Approved For Release 2001/11/19 : CIA-RDP78-04723A000400050002-6

[redacted]

3.6 Conclusions 25X1A

OSI's "walk-before-you-leap" approach to ADP seems sensible to us. It is easy to inaugurate projects with outside "help" but more difficult to assure the direction of these projects to show improvement in operations, particularly a cost saving. In the "number-crunching" field -- solving calculations that would be impossible because of their complexity or number of interactions, as for the ABM problem -- OSI has been successful and these efforts seem well in hand and evolving satisfactorily. The recent addition of remote terminals in an on-line system will particularly facilitate program development and test as well as interactive running of computational programs. In general OSI's computer usage is growing steadily as more uses are found, more personnel become familiar with it, and more equipment becomes available.

The area that seems to hold promise for the greatest improvement is in data indexing, storage, and retrieval, (ISR), but there appears to be no single high-priority OSI intelligence problem that would justify an ISR research effort by itself. OSI should support any Agency effort in this area, such as that by CRS to extend FMSAC's automated dissemination system.

What are the characteristics of ADP that make it unique with respect to manual systems? For years we were led to expect savings because of the savings ADP produced in the performance of clerical functions. In the development of assistance to intellectual

or analytical activity, neither money nor manpower is likely to be saved. Speed is great but there's what has been called a "5 millisecond and 48 hours" effect - the machine operates in fractions of a second, but delivery of input and output may take days. To some extent development of real-time sensors hitched to reliable communication systems has begun to reduce this problem. Remote terminals have helped the analyst in his design or computational problem, but OSI, by the nature of its tasks, has few problems requiring instantaneous responses.

The factor of most importance to OSI would seem to be what I term comprehensiveness. As files grow the analyst tends to remember best the most recently acquired reports, whereas an older report could be of more significance for his immediate problem. One expert in the field has said that nowadays most decisions are made following consideration of no more than 20% of the facts that bear on the question; an automated data retrieval system should be able to raise that percentage. This might well require reeducating analysts in how to use large volumes of material. Most today "intuitively" feel their decisions are correct, based on their experience and on a small percentage of the available facts. With more facts, if they know how to separate wheat from chaff, more solidly based decisions should be forthcoming.

There is also the case of the 20- or 30-year analyst who retires and his personal files, whether well organized or not, are kept for awhile and eventually discarded. This is a tremendous waste of our resources. The information in such files, if they could be in an automated file, would then remain available for use by anyone.

The other principal factor is ability to handle large volumes. Most analysts have reasonably efficient files, i.e., they are reasonable in size and cost, furnish speedy access to desired information and are reasonably easy to maintain. The average OSI analyst receives 82 documents per day and spends 5% of his time maintaining his file; less time and he would not know what was in his file. But the charts of CRS indicate an ever growing influx of data for the 1970's and most analysts are going to need the help of ADP in managing their data bases efficiently. Even now some files are large and threatening to become unmanageable. Manual files should be set up so that they can easily be converted to ADP if desired. For instance typefaces should be standardized with respect to optical character readers (OCR). Research in the OCR field should be pushed as part of the input problem. Such developments as that at Compuscan should be

[redacted]

3.7 Recommendations

Where do these conclusions lead us? There seem to be two prime areas of consideration for OSI: support of the R&D effort and education and training.

1. OSI should support strongly, with money and manpower if needed, promising research and development efforts in the following areas:

(a) Sort/dissemination/storage system along lines of Hooper proposal in CRS.

(b) Personal file management, along lines of Lancaster's work for CRS.

(c) Input/output research of ORD and other Agency components looking into OCR, computer output to microfilm (COM), or other methods.

2. OSI should push training of its own personnel in ADP so that they may better adopt this technology to the analytical and operational problems of the office. This user education should cover both interactive systems and batch processing. The users cannot sit back and await developments, but must jump in and learn enough about ADP to influence development. There exists a continuing need for training of analysts in the development and manipulation of files with or without ADP to improve our research effort.

3.8 Original ASPIN Survey Questionnaire for OSI (attached).

3.9 Original ASPIN Questionnaire for Established ADP Applications (attached).

SECRET