July 8, 1966

Dr. William T. Knox
Office of Science and Technology
Executive Office of the President
Washington, D. C.

Dear Bill:

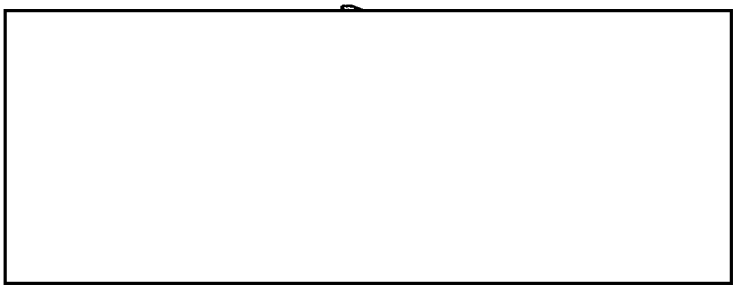
We take great pleasure in transmitting to you the enclosed draft of the report of the subpanel on biographies of the Guidance and Evaluation Panel.

This report cannot claim to be final. Indeed, as we see it, it is only a bare beginning.

We wish to take this opportunity to thank those members of your staff, and of the staffs of various community agencies whose help was invaluable, not only in terms of excellent technical presentations but also in guiding our way through the labyrinthine administrative obstacle course the community presents to the innocent outsider.

25X1A

Sincerely yours,



7/7/66

SECRET

DRAFT

SECRET/NOFORN/CONTROLLED DISSEM

President's Foreign Intelligence Advisory Board

Special Assistant to the President for Science and Technology

Guidance and Evaluation Panel

William T. Knox, Chairman

DRAFT

REPORT OF THE SUBPANEL ON BIOGRAPHICS

25X1A

SECRET

SECRET/NOFORN/CONTROLLED DISSEM

BIOGRAPHICS SUBPANEL REPORT

Table of Contents

Table of Contents (continued)

INTRODUCTION

The subpanel on biographics, like the whole Guidance and Evaluation Panel, faced the question of how to provide increased technical support for the information processing activities of the intelligence community. The technical problems presented themselves as inextricably woven into the pattern of information use and therefore into the structure of operations, management and management policy.

The rapidly increasing growth rate of information in the national data base obviously generates a need for additional resources to aid in processing and using this information. However, since the potential information pool is practically infinite, the rate of improvement in information processing is directly dependent not only on the rate at which new technical approaches are implemented, but also on the growth of management recognition of the fact that information processing policy must be one of its basic concerns and that this policy must take into account the need for selectivity created by the evident hopelessness of ever marshalling "all" relevant information.

The intelligence community is fortunate in having extensive resources of modern science and technology available for use now. The rate of advance in information handling technology is itself increasing. Moreover, the community already has a strong capability in applying automatic data processing techniques to cryptanalytic and communication

problems. The community is therefore in a unique position to apply these resources before the magnitude of the data base outstrips human capability to process it effectively.

One area that seems amenable to the application of modern data-processing technology and methodology is that of biographics. This subpanel has studied the suitability of automatic data processing for solving some of the problems found in the biographics area. In the process, we have examined biographic files both in and out of the intelligence community. We conclude that given strong management policy guidance, support and evaluation, biographics could benefit greatly.

I. SUMMARY

A. Background

This document is the report of the subpanel on biographics of a Guidance and Evaluation Panel created under the Special Assistant to the President for Science and Technology and the President's Foreign Intelligence Advisory Board, in response to a memorandum to the President by the Chairman of the PFIAB (CLIFFORD)*.

The Clifford memorandum charged the panel with responsibility for:

A. "Providing guidance to the intelligence community in the forwarding of methods and facilities for information handling and access;

B. Evaluating in technical terms the true meaning of the enormous and somewhat heterogeneous growth of the intelligence community's information pool."

The memorandum then goes on to say "It is emphasized that the proposed panel of technical experts would not be tasked with the too-obvious assignment of simply applying modern machine methods to the existing, specialized, and rigidly-maintained activities of processing and distributing information within the intelligence community. The panel would have the overall task of guiding the necessarily large, and

*Capitalized expressions refer to the corresponding entry in the bibliography.

presently ignored, planning for the realistic and long-term development of mechanized facilities for the processing of information in the manifold forms in which it is encountered within the intelligence community."

Adhering to this spirit, the subpanel has not hesitated to seek information outside the confines of the intelligence community narrowly interpreted as participants in the United States Intelligence Board.

The panel was given an extensive series of briefings and site visits in the major intelligence installations in the Washington area and elsewhere, and access to some documents produced by earlier studies. [LISTING TO BE INCLUDED]

It quickly became clear to this subpanel that the whole subject area, and biographics especially, has an excellent history written not only by innocent outsiders, but mainly by very able, knowledgeable, conscientious and thorough members of the community itself. Since the Kirkpatrick Report (KIRKPATRICK, 1960), more time has been spent, more detailed data has been gathered, more debate and interpretation has taken place than any part-time "experts," however well intentioned, could possibly equal.

Without apology, this report therefore makes very little claim to originality. In the words of Tom Lehrer, "Everything, I stole from somewhere else," including of course information gathered from briefings and some more informal on-site spot checks of critical items.

B. Major Findings

The major finding of this subpanel is that the main thing preventing an immediate intelligent community-wide attack on information problems in the biographics area is a combination of inability and unwillingness of the intelligence community (as presently structured and operated) to pay more than lip service to the idea of concerted effort. It is the so-called "realities" of organizational politics, not any basic technical obstacles, that spawn apparent impossibilities at every step of the way. No serious technical obstacle is visible on the way toward automation in the biographics area. We do not doubt that technical problems will appear further on, but the way is clear enough to set out on the journey at once.

C. Major Recommendations

Our major recommendation is that a clear-cut locus of management responsibility and authority for effective decision and concerted action be created to:

1. dispel the fog that sweet talk of unity casts over a fundamental parochialism that precludes intelligent collective action.

2. focus attention on basic problems of research and development, technical feasibility, systems analysis and design and management policy and control instead of frittering it away on arguments over such profundities as the matter of how many columns of a punched card should be allocated to a name.

3. establish technically competent lines of management to be supplied with rational system performance and cost data -- now almost totally lacking -- in order to guide the selection, collection, storage and retrieval of information in an environment where the ideas of sampling and statistical confidence must overtly displace the vain ideal of completeness already abandoned in apparently random practice.

4. insure that the sound, valuable but fragmentary ideas generated -- but not always presented -- in earlier internal community studies be given a genuine hearing leading to the vigorous exercise of management decision-making responsibility in place of the current committee practice of making recommendations reduced to the lowest common denominator.

5. that the results of isolated experiments be broadcast within and outside the community, evaluated critically and impartially and, when found valuable, applied in response to vigorous management direction.

6. ensure that the needs of users are taken into account by drawing users actively into the earliest and subsequent stages of systems analysis and design.

The remainder of this report aims to make more specific, to detail and amplify, and to document these basic opinions.

7. To ensure that a plan for developing a Community-wide biographic capability with the highest professional standards of performance be submitted to --- by 2 January 1967. Action to carry out such a plan should follow with the highest priority

SECRET

II. THE BIOGRAPHICS PROBLEM

A. Information Sources for This Report

Throughout this report we shall have occasion to refer to the report of CODIB Task Team 5 in a draft form dated January 17, 1966 (CODIB V). Although that draft report does not have the blessing of CODIB as of the time of this writing, it nevertheless reflects the concentrated effort of a group of twelve members of all components of the intelligence community in gathering and evaluating data concerning the scope of biographic files, the problems associated with creating, maintaining and using them, and, in the words of the statement of objectives of the team "to identify means for improving the storage, retrieval and exchange* of information from the major name files and related data files in the intelligence community." As of 31 December 1965, members of Task Team 5 "have reported a total of 3,993 hours devoted to this effort." (CODIB 1966) The subpanel is most grateful for the opportunity to use the fruits of such extensive labor as well as other related documents detailed in the bibliography and several supplemental briefings

There is a major gap in the work of the CODIB Task Team 5 which, unfortunately, this subpanel has lacked the time and resources to fill. Although one of the strengths claimed for "the CODIB approach" is that "participation by officials

*Our underline.

with operating responsibilities gives greater assurance of utility and acceptance of results" (BOREL 1966, p. 6), the report of Task Team 5 is restricted to technical analyses that give no evidence of participation by any official with responsibility for seeing to it that the needs of file users are indeed met. The notion of "operating responsibility" thus seems to have been narrowly construed to mean responsibility for operating the files. Since it is now generally accepted that sound systems analysis requires paying attention to the users (and users paying attention to systems analysis), we see this limitation as a major shortcoming of the CODIB approach.

The impression that this limitation is not mere accidental oversight is reinforced by the following quotation:

> "The SCIPS study group and the panel of consultants recommended the establishment of a permanent inter agency body of sufficient size to cover the problem areas. CODIB however has rejected this approach in favor of a number of ad hoc panels to be supported by a small permanent executive secretariat (group of documentalists). CODIB not only found excessive* the time necessary to build up a sizable competent permanent group (unless this were done at the expense of promising intra-departmental programs), but considers counterproductive any plan which calls for solutions developed by specialists not themselves involved in departmental information operations" (BRIGGS 1964).

Since the directive creating this panel requires it to "supplement the longer range Task Force projects being pursued by the USIB" (CLIFFORD page 2, paragraph 5) the availability of a

*Sic!

report of the relevant Task Force even in draft form is a major asset. It is exceedingly rare for a panel of "outside experts" to have access to such timely data on its subject of concern, produced by those most intimately acquainted with the problem. The findings of the subpanel are based in addition on visits to the major files [ADD LIST HERE] and on discussions with those concerned with their maintenance and their use, as well as with the CODIB support staff. We have also specially requested presentations of certain specific technical facts and opinions (Appendix 2) and these are attached as Appendices 3, 4, and 5.

We are sensitive to the qualms of the community about the intrusion of outsiders. One CODIB Task Team urges the acceptance of its recommendations on the quaint ground that, otherwise, the intelligence community "will continue to be vulnerable to external investigative and evaluative groups without having any recognized negotiating position" (CODIB VI, p. 11), further warning that such groups, however well-intentioned, will tend to gather data that are fragmentary, without interrelation and will therefore tend to make recommendations that are even more fragmentary and less related to real problems. The warning concludes that, in the process, the outsiders "will remove a thorn and, by so doing, implant a tumor" (CODIB VI, p. 11). We do not feel immune to such propensities but, by relying extensively on the community's own documents, we hope to present it with a mirror in which it will see itself.

B. Subpanel Qualifications

Although the members of this subpanel are clearly outsiders to the intelligence community, their qualifications have nonetheless been acknowledged by the community.

The Fifth Annual CODIB Report (CODIB 1963) quotes from a report by an ad hoc study panel of PSAC on non-numerical information processing completed in mid 1962, [redacted] The SCIPS report includes the following recommendation (CODIB 1964, p. 31, Recommendation 3): "That a technical review panel such as the PSAC ad hoc study panel on non-numerical information processing be assembled to review the detailed findings of the study and provide comments thereon." This subpanel consists of two-thirds of the aforementioned PSAC ad hoc study panel on non-numerical information processing. One of us was in the technical review panel for the SCIPS report and the submission of this report further implements at least this one ~~CODIB~~ SCIPS recommendation.

The March, 1964, "Progress Report on Foreign Intelligence Objective No. 3" [redacted] in mentioning the establishment by the National Academy of Sciences of an Advisory Committee on Automatic Language Processing comments that "This disinterested group is to advise the Department of Defense, the Central Intelligence Agency and the National Science Foundation on Research and Development in the general field of mechanical handling of languages ..." One member of

25X1A

this subpanel was an active participant in the National Academy Advisory Committee.

Since, as will be made clear in the remainder of this report, we have found that inaction and, indeed, deliberate inhibition of action, is a major factor blocking progress in improving information handling activities of the intelligence community, these expressions of confidence in the subpanel's previous activities are gratefully acknowledged.

Moreover, the portions of the report by the PSAC ad hoc panel on non-numerical information processing quoted in CODIB 1963 ascribe to that panel a certain conservatism:

"The two fundamental difficulties faced in present applications are that, even though large-scale commitments to the development of operational hardware systems have been made, there is (a) a lack of precise definition and specification both of the objectives to be reached and of the relevant non-numerical techniques and (b) a shortage of first-rate research workers in the field.

"Symptomatic of these difficulties are the present and growing external pressures towards large-scale hardware procurement in cases where the operational purpose to be served is vague and where, in any event, the present state of the art in hardware and software is inadequate to support the grandiose 'intellectual processing' that is so ardently desired."

The initial requirements of biographics are much more modest. Our major finding (Section I, B) that is is the so-called

"realities" of organizational politics, not any basic technical obstacles, that spawn impossibilities in the biographics area and elsewhere should not, therefore, be ascribed to any starry-eyed belief by this subpanel in the miracles of automation. We think, rather, that no degree of automation can compensate for bad management.

C. Definition of Biographics

Various intelligence community and related agencies of the government maintain extensive biographical files to serve the needs of Positive Intelligence, Security and Counter-intelligence as the case may be.

Our definition of these three major areas is taken from CODIB V, Annex 1:

> "COUNTERINTELLIGENCE BIOGRAPHIC AREA: That activity which deals with information on personalities who constitute a known or possible threat of national security. These normally include members and agents of foreign intelligence services, Communist Party officials, and others engaged in organized subversive activities.
>
> "POSITIVE INTELLIGENCE BIOGRAPHIC AREA: That activity which deals with information on personalities, usually foreign, who are of general interest to the intelligence community. These include leaders in the scientific, political, governmental, economic, military, and other professional/governmental fields.
>
> "SECURITY BIOGRAPHIC AREA: That activity which deals with information held by those organizations which have the normal function of investigating and granting clearances on individuals or counterintelligence interest in respect to the internal operations of the holding organization."

Biographic files within the Intelligence Community include the CIA's Special Register and Biographic Register, the National Security Agency's files in the Office of Central Reference, and files in the Defense Intelligence Agency and the State Department. Major biographic files are kept also by the FBI, the Immigration and Naturalization Service, the Civil Service Commission, the National Security Agency's Office of Security, the Air Force Office of Special Investigation, the Navy, the Army, and the CIA's Office of Security and RID. Some consolidation of the indexes, at least, to certain of the Armed Services files is evident at Fort Holabird. Additional details on the nature and use of biographic files are given in Appendices 1 and 3. Still greater detail is available in CODIB V and SCIPS 1963.

D. Significance of Biographics

The importance of biographic files in Counter-intelligence and Security is obvious since these activities are concerned primarily with people. In the Positive Intelligence areas, the biographic files are said to play a critical role in support of the evaluation and production of foreign intelligence but we have not completely verified this nor, apparently, has the CODIB Task Team 5 (see Section II, A). In either area, the collection, storage and dissemination of biographic material is presumably not an end in itself.

At present, the different locations and the real or imagined differences in the uses of biographic files in various branches of the community lend existing files a considerable measure of apparent variety.

A case can be made, however, that similarities on the whole outweigh differences (Appendix 5). The members of Task Team 5, hobbled by their terms of reference and apparently hypnotized by the rigidities of the 80-column punched card officially see incompatibilities and impossibilities wherever two members disagree over the number of card columns that should be allocated to such data as names or birthdates. In the opinion of this subpanel, the technical problems of biographics are sufficiently simple in contemporary terms to suggest biographic files as a primary candidate for competent and honest experiments designed to explore the relative advantages and disadvantages of various positions on the spectrum from complete isolation to total integration of files (Appendix 5a, p. 1). The criteria for evaluating these experiments should be based on the satisfaction of explicitly formulated needs of the community of users, as to both substance and security, not on catering to the crotchets of keepers of the files.

E. Technical Simplicity of Biographics

The biographics area lends itself particularly well to a study that cuts across the whole of the intelligence community, in relatively simple yet non-trivial terms. While the problems of creating, maintaining and using biographic files are by no means negligible, in comparison with most other files used by the intelligence community biographic files are a great deal cleaner, simpler and better understood (Appendices 3, 5). The biographic files are a very good example of files that are widespread throughout the community, used throughout the community, serviced by the whole community and which, while relatively simple in structure, present all of the problems of file maintenance in a non-trivial way. Simpler files such as order of battle or installation files exist in various areas but, in general, these are based on a narrower input spectrum than biographic files and serve a narrower area of the community. Because they can usually be rigidly formatted and organized and accessed through clear-cut keys such as geographic coordinates, their mechanization is well within the state of the art and therefore ~~generally already~~ easily accomplished.

We do not wish to understate the ultimate complexity of biographic files. We emphasize, however, that major advances over the present state of affairs can be made without fear of technical barriers.

The report of the Biographics Task Team states that "no major name index in the intelligence community has yet been fully automated. Therefore proof of success has not yet been conclusively demonstrated" (CODIB V, p. 2). The parallelism between the restrictive phrase "in the intelligence community" occurring in the preceding quotation and in the Task Team's statement of purpose (CODIB V, p. 1) may account for the absence of reference (see also Section III, E) to such files as that operated by the National Driver Register in the Bureau of Public Roads of the Department of Commerce, which this subpanel visited (Appendix 3). This file, now accounting for 860,000 driver's licenses revoked in each of the 50 states, the District of Columbia, Puerto Rico, the Canal Zone, the Virgin Islands and Guam, for offenses involving drunkenness or fatalities got no overt consideration by the Task Team under the headings "major," "fully automated" and "proof of success," or any other heading.

We regard it as most significant that such a central service can deal effectively and efficiently with fifty sovereign states plus other jurisdictions. The degree of trust in the operation and of understanding of its non-deterministic aspects (see the " may be's" in the Matched Driver Record Reports of Appendix 3) is apparently such that the practice of returning no response whatever to the state if a name is not found is generally accepted, with consequent savings in transmission and paper-shuffling costs. Registrars of Motor Vehicles apparently

understand and believe that if there is a probably match in the Register they will hear about it within 24 hours.

Admittedly, the National Driver Register Files must deal only with intentional camouflage by drivers in danger of license revocation, while the community's agencies must sometimes deal not only with that, but with illegible, fragmentary data, aliases, names overheard on noisy channels, etc. In the panel's view, these problems are real and serious, but in no way preclude immediate automation of a substantial portion of the community's biographic files.

In commenting on the problem of variations in "categories of identifying data recorded if available in the reporting," the Task Team states that it is "impossible to develop rigid rules on what constitutes the minimum identifying data required" and that "each agency, in recognizing these problems and the nature of its own index, forms its own rules regarding minimum identifying data for recording and the depth of search according to the nature of the request" (CODIB V, p. 8). Our site visits and conversations with members of the Task Team suggest that the following exegesis of these comments merits consideration: "Impossible" means politically, not technically impossible; the "realities" get in the way. "Impossible" perhaps means impossible with punched cards. There is no technical reason/why only "rigid rules" must be considered, when flexible rules might do and modern storage and logic technology permit their use (Appendix 5). In short, "impossibility" is a CODIB thing.

known to this subpanel

The foregoing should not be taken to imply that this subpanel thinks that all the problems of the intelligence community have been solved by the Bureau of Public Roads or that "not impossible" is synonymous with "easy." It is, however, clearly implied that there is a need for candid and unfettered reporting of facts and for scientific experimentation.

F. Biographics as a Laboratory for Technology and Management

The Clifford memorandum of June 15, 1965, (CLIFFORD) characterizes the problems of the Intelligence Community in the following terms:

> "The problems of the intelligence community in connection with information access and retrieval include, but are not restricted to, those common to all who must maintain very large bodies of information in accessible form. This is even true in the handling of information from unclassified sources. The importance of negative information, and of patterns of information, requires that access to intelligence information produce a completeness of response beyond that which is expected from many large files of stored information. Like statistics, intelligence cannot be satisfied with the highly anecdotal, but requires that all available items of information are allowed to contribute their part of the final summary or other intelligence product. (p. 3, No. 7)

> "As a consequence of intelligence community requirements for high recall, the mechanized and automated means of access to many sorts of intelligence files cannot be required to meet simultaneously rigid requirements as to relevance. Accordingly for some time to come the mode of gaining access to intelligence information will be through combined machine-human systems that will seek the machine retrieval of stored intelligence information in order that its relevance may be established by human examination. It is this combined machine-human factor which generates systems problems of great difficulty and dimensions. (p. 3, No. 8)

> "In the area of experimental approaches to the adaptation of machine processing to the storage and retrieval of intelligence information, an encouraging beginning has been made within the National Security Agency where the Technical Information Processing System (TIPS) study is presently under way. This experiment, although on a limited scale and confined to a selected number of organizational units and information files within the National Security Agency, is producing important lessons for the achievement of a realistic system for the interrogation of a computer by remote users requiring access to a common information base." (p. 4, No. 11)

In addition, Recommendation No. 2 concerning the Technical Information System (TIPS) project, (CLIFFORD, Page 5) specifically calls attention to the fact that "the capability for extensive handling of the Russian biography problem should be available in the community-wide system by the summer of 1966."

The successful operation of some quasi mechanized biographic systems in CIA, NSA and the Bureau of Public Roads demonstrates that many technical and security problems arising with biographic files can be solved at least to some degree. Our observations in Section III suggest that no insuperable technical problems blook coordinated action by the community in the biographics area once management problems are resolved.

In particular, biographic files are free of most of the serious linguistic problems which bedevil attempts to handle the finding and retrieval of the extensive textual materials that form the bulk of the intelligence community's files. The CODIB biographics task force implies in one of its

recommendations (CODIB V, Page 5, Recommendation 8) that a successful biographic system could form a good basis for approaching these more difficult general retrieval problems through "a coordinated program (which) should be developed using EDP methods to provide machine indexes of the bibliographic data processed by any organization in this field, so that the personality information is accessible to a recipient in machine form, with quick followup to the translated source." The converse would be equally valid, that biographic indexes could lead back to bibliographic material containing information about installations, organizations, etc., related to the personalities. This subpanel agrees that coordinated biographic files together, perhaps, with installation and organization files, may well form one excellent basis for indexing other holdings of the intelligence community since curiosity about an individual, an organization or an installation is so frequently a source of curiosity concerning the more detailed and extensive material that might be in the files.

The subpanel takes exception only to the mode of implementation proposed by the CODIB task team, namely that another task team or its successor be tasked to follow up on this recommendation. The record shows that CODIB, as presently constituted, has been singularly ineffectual in getting action on difficult problems. Task teams have operated within severely constrained terms of reference conducive to a bizarre

frame of mind in which members felt constrained to "go at their task as CODIB, deliberately avoiding the use of information available to them as individuals" (Statement to full Panel at Briefing of 4/19/66). Serious problems are recognized even within these limitations, statements that strong efforts are required appear, but the process seems inevitably to conclude with bland ineffectual statements about cooperation and consensus devoid of concrete suggestions for action and enforcement of agreements.

The history of CODIB Task Team 5 - Biographics yields an illuminating example of that subtle magic. While CODIB itself boldly recommended in February 1964 to "Develop a Biographic Intelligence Processing Plan," (CODIB 1964, p. 21, Recommendation 4f), the Task Team's terms of reference as of January 1965 (CODIB V, Annex 6) and the January 1966 Draft Report contain only a limp eviscerated shadow of this goal in the statement that "the objective of this Team was to 'identify means for improving the storage, retrieval and exchange of information from the major name files and related data files in the intelligence community'." (See also Section III, D)

It is equally curious that while it was recognized in 1963 that "the most significant problem in the present period is that of organization," that "collective efforts within the community will become increasingly important" (BOREL, 1963; Attachment, p. 2) and that "steps can be taken within the

present state of the art to ease some of the current problems" CODIB, 1963, p. 10), such pieties are transmogrified by 1965 into the idea that "attention (and management support)" be given "to improving the individual systems in each USIB agency, with projected compatibility monitored by CODIB and the PFIAB/OST Joint Guidance and Evaluation Panel through briefings, demonstrations and discussions."

The findings of the Task Team VI on Research and Development that "security barriers are used quite effectively to barricade against attempts by other agency personnel to acquaint themselves with ongoing R & D efforts" (CODIB VI, p. 13) and that "very few intelligence staffs permit realistic evaluation of their practices or products and results of any evaluation are usually tightly held" (CODIB VI, p. 31) coupled with the absence of observable experiments that would support the change of attitude between 1963 and 1965 suggest that this change is not as down-to-earth and commonsensical as might appear at first glance. As of this writing, the report of Task Team VI has yet to be approved by CODIB.

The biographics task team observed that, in the positive intelligence area, "every organization has its own standards for selection (from collected material) based on the mission it is supporting and budgetary limitations" and that "the same source document is frequently processed by different PI organizations." They further remark that "there is an

overlap of information in PI files because the different file systems support the same requirements, or because the personality mentioned in the source meets the selection criteria for two different requirements. (CODIB V, pp. 13-14). Given, on the one hand, that repeated handling of the same document is prima facie wasteful while, on the other hand, interpretation of the same material from different and possibly conflicting points of view is likely to be helpful and perhaps vital to reliable intelligence production, it is sad but not surprising that the task team felt unmoved or unable to recommend any approach to this vital problem of management control, especially since the same problems had already been recognized earlier by the SCIPS Study (SCIPS, 1963).

Indeed, a disturbing proportion of the technical people this subpanel came into contact with seemed conscious only of details of card formats, while the extent of some management personnel's appreciation of information technology seems typified by such a curious anachronistic statement as "our problems have largely to do with the processing of language, while automatic equipment is essentially designed to handle numbers" (BOREL, 1963; Attachment 1, p. 2). There is a vacuum of competent consideration of system design problems combining the understanding of needs and of possibilities.

This problem is not altogether unrecognized, as shown by the following expression of the outlook for CODIB:

> "We will need to guard against the ADP tail wagging the information processing policy dog. But we do need to give more emphasis to technical problems than we have in the past." (BOREL, 1966, p. 6).

The problem is, perhaps, that the tail and the dog have not yet met each other. As for wagging, while we have repeatedly emphasized the need for keeping the user in mind, we also believe that in a community where information is the most important product, the processing of this information should be far more than a managerial afterthought. ADP, as we conceive it, has more to do with the head than will the tail but, if pressed, we would insist at least on thinking kangaroo rather than dog.

Where information is concerned, process and substance assume nearly equal weights and management cannot safely assume that if it pays attention exclusively to substance the technicalities of processing will take care of themselves. The top and middle managements of most organizations, not only in the intelligence community, were trained in days when information processing was for accountants or scholars, but not for them. The fact that modern information technology affects the very fabric of organizations has achieved some notoriety in The New Yorker cartoons, but has yet to be translated into realistic organization charts. (See also Section III, E).

It is the opinion of this subpanel that the biographics area provides a unique and promising laboratory in which to explore the general technical and management control problems raised by the Clifford

memorandum and to find out which of CODIB's several minds reflects reality. The technical problems of biographics, while by no means trivial, are not so severe as to becloud the primary organizational and management control problems. The COINS effort may be a beginning.

Any experiment, however well-intentioned, will be worthless if set up in the CODIB pattern. The reasons why are evident from the following expression of the CODIB credo:

"We recognize and accept certain constraints upon what can be accomplished by joint action because of:

> the primary responsibility of USIB members to their command channels for carrying out basic departmental and service missions; and,
>
> the impact of the NSC allocation of intelligence collection and production responsibilities among agencies upon supporting information processing programs." (BOREL, 1966, p. 3).

The deadening effect of the imposition of such administrative constraints is made explicit in the recognition of the following weakness of CODIB and its approach:

> "Common denominator of agreement may be so large as to negate utility of the solution." (BOREL, 1966, p. 5).

That these constraints bring out the worst of ubiquitous bureaucratic tendencies is made plain by the less elegant advice one of our informants assured us applies under the circumstances: "Cover your ass." We can't help thinking that this motto is as inappropriate a slogan for a healthy

intelligence community as "Expose your rear" would be for a crack military unit.

We have, in addition, fair evidence that "the system" not the people is inherently responsible for this sad state. Individuals whose wits seem hopelessly dim in an official light visibly brighten in the sunshine of informality. This phenomenon is common to all organizations, but seems pathologically pronounced in the Community. While, once again, we recognize the need for users and operators to exert the strongest influence on systems design, we recommend that members of any group constituted for this purpose be instructed to act according to their individual wisdom and conscience, and that they be protected by appropriate administrative devices against any wrath they may draw from home offices as a consequence. Someone, presented with conflicting opinions instead of bland, lowest common denominator consensus, may then have to make painful decisions, but we submit that the Community urgently needs management able and willing to act in this way.

We do not think this recommendation is naive and inapplicable. We found it refreshing to learn, at Fort Holabird, that the very same people who saw only impossibilities, parochial interests and lowest common denominators when acting in a CODIB group, stopped dragging their feet and leaped into action when a directive issued by the Secretary of Defense on May 27, 1965, ordered them to provide a centralized

(See also Section III, L)

index to the Armed Services' investigative files by May 27, 1966. Indeed, the job apparently got done. Critics do point out that excessive haste has, in their opinion, produced a product lacking elegance, generality, or power. Nevertheless, the illustration shows that management has a real choice. Under the constraints described in its credo, CODIB had none.

III. OBSERVATIONS

A. Relative Characteristics of Positive Intelligence, Counterintelligence and Security Files

The Positive Intelligence files are in a sense the least sensitive of all. Since they deal largely with foreign personalities and a substantial proportion of their source material is from the open literature, the problems of indexing these files are subject only to the normal technical constraints common in the Intelligence Community, without touching more delicate areas such as might be raised in the evaluation of uninterpreted data on U. S. citizens - where the very mention of a name in the file can create a problem. On the other hand, the files are among the more problematic in terms of acquisition, indexing and completeness.

In the Security and Counterintelligence files a goodly portion of the inputs is on controlled forms, often supplied by the individual himself and positive identification in tags, such as Social Security number, etc., are readily possible. The PI files, on the other hand, are constituted from a much greater variety of sources, they tend to be far more fragmentary and their use is more difficult.

It therefore seems reasonable to infer that general solutions of Positive Intelligence biographic file problems could be readily specialized and adapted to the needs of Security and Counterintelligence file but that the reverse

very likely would be far more difficult. These statements are supported by evidence to the effect that PI requests tend to include 20% name finding* and 80% name searching*, while in the Counterintelligence and Security files the proportions are, respectively, 5% and 95%.

The study of PI files thus presents the greater challenge to research and development as well as a hope of finding general solutions. Such study should be encouraged. In the near future, however, useful practical results are much more likely to be obtained by concentrating on the automation of CI and Security files.

B. File Overlap and the Single File Problem

CODIB V points out (page 2, paragraph 6) that "there are several thousand people involved in biographic activity in the Intelligence Community. Approximately 1000 of these, at an annual salary-only cost of $5,000,000 are directly involved in work at the index level, in the preparation, maintenance, and searching of the major biographic indexes." The report then goes on to point out (page 13, 2.b) that

*The terms "name finding" and "name searching" are defined in the CODIB Task Team V Report as follows:

"NAME FINDING: The searching for name information about one or a group of individuals by looking for data elements other than the name, such as date of birth, position, location, organizational affiliation, occupation, military rank, nationality, including a combination of such factors."

"NAME SEARCHING: Search of indexes or files organized by the names of persons to determine if information exists on the individual, or to validate basic information."

CODIB V, Annex 1, p. 1, Glossary.

> "The basic criterion of any agency for selecting an item for a PI file is whether or not the item supports the foreign intelligence effort on a particular country or area. Every organization has its own standards for selection based on the mission it is supporting and budgetary limitations. The same source document is frequently processed by different PI organizations ... There is an overlap of information in PI files because the different file systems support the same requirements, or because the personality mentioned in the source meets the selection criteria for two different requirements: e.g., CIA and State have an interest in military personalities who are prominent in other fields such as politics, science, space, etc., whereas DIA and NSA are interested in the same person because he is in the military field. There is no assurance, however, that because a personality is mentioned in a source document that he will eventually be processed into a PI file."

Further on, the same report points out (p. 15, 4.c) that

> "Some files are restricted by security classification as to what can be processed. Research in such a limited source file often gives incomplete or outdated information."

The CODIB report further states that

> "It is doubtful that any single file, whether it be computerized or manual, can ever be considered a complete or sole source for biographic information."

This subpanel takes strong exception to this last statement, largely on the ground that it has not been able to uncover the existence of any data, produced as a result of a comprehensive systems study concerning the advantages and disadvantages of a single community-wide biographic file. (See also Section III, I) It should also be noted that the Clifford memorandum (CLIFFORD, Recommendation #2) specifically suggests the Russian biographic problem as an experimental vehicle for the expanded TIPS system (COINS).

The evidence available to us suggests that consolidation or at least common indexing and easy mutual accessibility of CI and Security files is technically feasible. The issue of whether or not this should be done was left in limbo by CODIB's departure from its resolution "to develop a biographic intelligence processing plan" (see Section II, F).

In the PI area the technical issues are more confused and therefore all the more in need of candid intelligent investigation. For example, the prevalent practice of scanning PI input materials for names, and resolving name variants and name variations before entering the material in files has led not only to a substantial backlog of unentered and therefore unavailable material but also to the repeated resolution of the same problem by every recipient of the same material.

The alternative of entering materials only once for the whole community, using as a key whatever name variant or variation appears in the raw material, and applying name-grouping techniques to draw together scattered relevant material at the time an actual request is made for information has received only scant attention. Such a task is clearly impractical with manual or older punched-card techniques, but not at all ruled out with modern equipment. The feasibility, potential dollar savings and increased information availability that such techniques might provide must be investigated in spite of the specter of centralization (Appendix 5).

The SCIPS report points out that, although "DCID 1/9 allocates responsibility for production of biographic intelligence and the collection and maintenance of biographic data on foreign personalities," (SCIPS 1963, p. 57), various agencies (NSA is singled out), maintain substantial biographic files without specific allocation in DCID. Also, as an aside, it points to an "incidence of many files that are no longer 'input to'."

The report then comments that "a set of huge files with their implied amount of processing effort is being maintained in a manner not identifiable in terms of DCID functional responsibilities. This is not to say that NSA does not have a need for this information, for it manifestly does. Therefore the inference is that those responsible under DCID are not performing satisfactorily to NSA needs or that subject category is not a good way to allocate this responsibility or both" (SCIPS 1963, p. 68).

Had CODIB undertaken "to develop a biographic intelligence processing plan" some light might now shine on this obscurity. They did not and therefore it does not.

C. Evidence of Current Automation Efforts

There is encouraging evidence of widespread beginnings of modest automation efforts throughout the Intelligence Community. Within the CIA the Special Register, the Biographic Register and the Walnut system with its associated indexes are

mechanized to varying degrees and with a variety of different technical and file structure approaches. However, in some of these cases, punched cards were introduced 10 years ago and nothing new has happened since. Wholly or partially mechanized biographic files exist also in the NSA, DIA, and at Ft. Holabird.

However, as discussed above (Section II, F) community efforts to mechanize biographic files are fragmented and generally uncoordinated between agencies. In view of the recognized overlap of biographic files (CODIB V, p. 14, para 2c), it is difficult to understand why a community-wide effort to evaluate the possible support of ADP has not been sponsored, except on the hypothesis that management has been unwilling to sponsor it. Why CODIB V does not recommend such action is clear from its impotent charter and the pressure toward the lowest common denominator. The irony of this situation is made plain by Recommendation 8 of CODIB V which states (CODIB V, p.9)

> "The Task Team III (or its successor) be tasked to study those various programs exploiting open source scientific and technical information which generate personality information of positive intelligence value as a by-product. A coordinated program should be developed using EDP methods to provide machine indexes of the bibliographic data processed by any organization in this field, so that the personality information is accessible to a recipient in machine form, with quick follow-up to the translated source." (See also Section II, F.)

Apparently, community-wide mechanization is good in this bibliographic area but not good in the derived biographic area. This subpanel has not perceived any substantive

reason for this distinction, and our recommendations reflect this fact.

D. Need for Research and Development

There is some encouraging evidence of recognition of the need for research and development activities in the biographics area. Notably, the CODIB V Report stresses the critical technical problem of dealing with name variants and name variations in name searching through biographic files, especially those with a preponderance of foreign names or even of U. S. citizens in cases where source format control is not possible (see also Section III, B) and Appendix 5.

Problems created by the high proportion of common names in various files (CODIB V, p. 8, item 8) are also recognized, as is the more difficult technical matter of providing usefully formatted identifying data on named or unnamed persons to permit searches of the name-finding type or identification in the absence of names.

Nevertheless, the Panel is encouraged by the finding of the CODIB Task Group (CODIB V, p. 11, item 5) that "in observing some of these typical and widely used forms, the Team found that certain basic data, such as name, place and date of birth, Service Serial number, Social Security number, sex, etc., were included on each form." While this finding applies primarily to the CI/Security files, perusal of the Chart in Annex 4 of the CODIB report indicates that an effort

is being made to provide similar fields, even in the Positive Intelligence files, although the probability of no information appearing in these fields is obviously greater. The fact that certain standard elements, blanks in the forms as it were, may occasionally be left empty in some individual forms does not negate the clear possibility of defining standard elements even in PI files. The specter of empty fields haunts those accustomed only to punched cards, but is readily exorcised by more modern techniques.

As pointed out in the Introduction to this report, the principal technical issue faced by the intelligence community is that of making maximum use of the technology and methodology available in the nation. In some problem areas, there will be a need for directing resources toward the development of technical solutions designed especially for biographic information processing problems. In order to guide such efforts, the community will need a coordinated R & D program and strong in-house capabilities in specific areas which are not covered in other government or nongovernment agencies.

(The original CODIB Terms of Reference (DCID 1/4) charge CODIB with the responsibility and authority to develop a long range plan for R & D in information processing. HOLD FOR CHECK OF 4/23/65 TERMS AGAINST 6/26/59.) CODIB has never

TRY BACKUP TO 4/21/58 - MAY NEED

concerted action in developing such a plan; the efforts of Task Team VI represent an admirable start, but its tartly critical report has been referred for further drafting by [redacted] usual CODIB technique. The effect of CODIB action in the R & D area has been to prevent other groups from initiating a community-wide program by keeping the R & D responsibility [redacted] in its charter and, by inaction, preventing itself from initiating such a program.

The biographics area is merely one of many that are hampered by this policy. It is therefore understandable that CODIB V did not recommend any R & D activity as an aid in solving biographics problems; the Task Team objective appears carefully emasculated in order to preclude any consideration of fundamental methods for solving biographic problems.

The Task Team was barely able even to state the problem under its modified objective. Finding #1 (CODIB V, p. 1) states that:

> "1. Improvements in the speed and quality of biographic information processing involving interagency exchange on U. S. citizens and foreign nationals are necessary to further improve security, and to afford policy makers and analysts better response from biographic intelligence files on foreign nationals of interest from a variety of angles - military, subversive, political and scientific. The Team finds that use of computer techniques and interagency telecommunications links may provide significant improvements."

Finding #2 then remarks:

> "2. There are, however, profound, complex problems and significant costs in making major changes in the large biographic holdings of community concern, particularly if the changes involve conversion to computer systems."

The apparent fact that intelligence research and development funds under the control of the Deputy Director of Defense Research and Engineering (DD R&E) is not considered a part of the Community deserves a deeper analysis which I would welcome.

The inclusion of policy makers in Finding #1 implicitly argues for the importance of the files and the associated problems. The recognition of the profound nature of the problems in Finding #2 implicitly argues for the need for very strong and powerful recommendations for action in the R & D area as well as other areas.

The actual recommendations made by the Task Team are pathetically inadequate when held up against the problem as stated in Findings #1 and #2. However, when the recommendations are compared with the Task Team objective, the fit is exceedingly precise. This subpanel can only conclude that the original objective was modified (see Section II, F) in order to place strict control on the recommendations of the Task Team. The Subpanel suggests that, in the future, control be placed on the findings rather than on the recommendations; this would seem to be a more effective technique for the apparent purpose.

Given the context, this subpanel understands why the Task Team could not make recommendations in the R & D area; therefore, this report includes R & D recommendations the Task Team should have made (Sections I, C; III, B). Faced with Findings #1 and #2, this subpanel would be remiss to do less.

E. The Nature of the Biographics R & D Problem

There are many biographics problems that can be solved with available technology. However, there is a wide

gap between available technology and the community's capability to apply it. It is apparent to the subpanel that a new qualification, namely computer systems analysis, must be added to the community's wide range of expertise. This qualification must be distributed broadly and at every level including the highest. Once a major fraction of community personnel have some training and experience in this discipline, it will become much easier for the community to avail itself immediately and rationally of the benefits of off-the-shelf technology and methodology.

Ancillary benefits will also come in the R & D area. Specific and important benefits should accrue directly in biographic problem areas. The most immediate and important benefit would be an appreciation by senior management personnel responsible for biographic file installations and their use of the need for and the nature of R & D work. With appreciation and understanding would come stronger leadership and better control.

Of the entire listed membership of Task Team V, no single person has the experience or qualifications of a professional EDP systems or applications analyst. A few have medium heavy experience, but the rest have very little, if any, experience in applying EDP equipment to problems. This casts no aspersions on the personnel involved, because there has not been a need in the past for such experience.

Most biographic files were started as manual systems; the need for mechanization is relatively recent. Thus, the lack of highly qualified personnel in biographic file systems may be seen as the root of the lack of appreciation for R & D work.

But, even granted the existence of a cadre of highly qualified EDP systems analysts in biographics work, there remain the "profound, complex problems" of Finding #2, and these have to be solved. Since these problems, combining the technical and the administrative, they are therefore problems for which methodologies or technologies cannot merely be picked off the shelf.

In order to shed further light on these problems it is instructive to examine the Task Team Findings (Section III, D) and their statement of the problems from the viewpoint of a professional EDP systems analyst.

Finding #1 begins with the clause "Improvements in the speed and quality of biographic information processing ... are necessary ..." This statement is very good because it emphasizes speed and quality, two of the three basic parameters by which any information processing installation can be measured (the third parameter is cost). The Team further found "that use of computer techniques and interagency telecommunications links may* provide significant improvements." At this

*Our emphasis.

point, the Team reveals its level of expertise; there is no real doubt that computers and telecommunications links can improve biographic information processing because of the high capacities (bandwidths) of both modern computers and telecommunications equipment!

Evidence, if needed, is found in the fact that some members of the community have already installed both types of equipment (e.g., CIA-Walnut and LDX) precisely on the ground that both quality and speed are thereby improved. The best evidence for the feasibility of improvement via EDP is therefore found within the community itself. A professional systems analyst would also be able to deduce this trivial conclusion from an examination of the files themselves. For the most part they are formatted, and even the unformatted files (such as in CIA-BR) offer no technical impossibilities (Appendix 5). Cost is the most important determinant, barring politics, of course.

Finding #2 is somewhat more instructive. Here the emphasis is on "making major changes ... particularly if the changes involve conversion to computer systems." This statement distorts the realities of conversion (Appendix 5c). Usually, the cost of the job is the only hindrance, and it is hard to believe that any other parameter could be seriously considered. The "profound, complex problems" have been solved in several community installations visited by the ?

subpanel; usually the profound complex problem turned out to be changing the attitude of management. This change, however, can hardly be charged as a technical component of the systems problem.

More to the point, it should be noticed that Finding #2 sees the problems as consequences of making changes in the file installation. In other words, the new file will be a modified version of the old file, and the greater the changes, the greater will be the problems introduced. The inevitable conclusion is that the best way to solve the problems is to make no changes at all. This gratuitous finding thus gives the coup-de-grace to Finding #1 which weakly states that changes may improve the speed and quality of biographic information processing. The professional systems analyst would not be intimidated by Finding #2, thus leaving the subpanel with the conclusion that this Finding was intended for readers who are not systems analysts.

The remaining Findings offer substantial data for the systems analyst.

Findings #3 through #10 present pertinent numerical data that is very useful in deriving a picture of the biographic file problems. Indeed, the problems are great, but by no means beyond the limits of technical possibility.

Finding #8 is worth a modicum of attention. Despite the implied threat of Finding #2, Finding #8 says that for name searching, EDP may provide improvements, but this statement was apparently considered rash and puerile, for it was immediately slugged in the following sentences with the pompous and only tenuously true pronouncement that "no major name index ... has yet been fully automated," (see also Section I, E) thus invoking the threat of Finding #2.

In Finding #10, the body of Finding #1 is still twitching. We are told here that name finding may possibly be improved, but only by exchanging information about "EDP techniques for improving speed and flexibility of response." Notice how the golden "quality" of Finding #1 has been transmuted to leaden "flexibility."

Finding #11 then concludes with a wide-eyed view of the wonderful world of professional interchange. The whole set of Findings reminds one of Franck's Symphony in D where the good theme jousts with the bad theme. In CODIB V, the bad theme wins, and Finding #11 depicts the rainbow after the storm.

The foregoing considerations, while necessary to achieve catharsis, do not take the most constructive viewpoint of the problems. The standard approach of systems analysis is to create something new out of something old: change is not something to fear, but something to use to advantage.

Preoccupation with format standards can prevent one from seeing a better solution to the overall problem. Annex 4 (CODIB V) illustrates the reality of the problem in different agencies. There seems to be little doubt that each agency will need more or less data in its records than another agency. There is no technical reason why each agency cannot have information entered, stored, and retrieved in any format of its choice. Annex 4 merely tabulates the differences which the system must accommodate.

From a systems design viewpoint, it would be a mistake to try to force standard formats on every agency when there are more significant and difficult problems to be solved. As discussed earlier in this report (Section II, F), biographics offers unique and promising opportunities as a vehicle for R & D. Sample suggestions are as follows:

1. What would a single community-wide biographic file installation be like? Specify its size in storage capacity, response speed, request rate, etc. What administrative problems would it create? How serious are these?

2. Develop methods for recognizing proper names in unformatted text and for extracting text which includes references to or data about people.

3. Develop a specification for a very low cost technology for entering data into files, to be installed 1972-75. Compare with cost projections of optical character recognition technology. Will the latter be the final answer? If not, a different approach must be developed and work should be started now.

4. ~~Develop specifications for a [illegible]~~ telecommunications technology that is hybridized ~~with~~ digital data processing technology for use in 1970's. What type of ~~processing~~ will each perform best?

F. The Invisible Colleges

There is a strong need for positive interagency cooperation, information exchange and joint enterprise in the biographics realm. This need is clearly demonstrated by our observation of efforts in several agencies without evidence of mutual knowledge. This impression, which might otherwise be attributed to the superficial character of an inquiry by outsiders, is corroborated by Finding #11 of the CODIB Task Force, which states:

> "The Team agreed that the professional interchange derived from the Task Force effort was highly valuable to each member in providing new insights in manual and machine techniques, interagency channels, sources of information and policies of other agencies."

A similar reaction was observed by the Panel in the course of a meeting on March 29th 1966 which confronted representatives of the Electronic Data Processing arms of NSA, DIA and CIA,

with representatives of the counterintelligence offices in a discussion of security and compartmentation problems. The invisible colleges seem to be responsible for the work that does get done but they are not adequate enough in their precarious existence.

G. Compartmentation and Security

There is convincing evidence that compartmentation exists not only for security reasons but, as the R & D report points out, (CODIB VI, p. 13) also for administrative reasons masquerading as

> "security barriers that are used quite effectively to barricade against attempts by other agencies' personnel to acquaint themselves with ongoing R & D effort."

The problem of file security fails to be adequately distinguished from administrative and in-fighting problems. We acknowledge the genuine needs for security and compartmentation but we believe that these have been grossly exaggerated with adverse effects on Community cooperation.

The Community clearly recognizes the possible damaging effects of compartmentation, witness the statement

> "Research in such a limited source file often gives incomplete or out-dated information." (CODIB V, p. 16).

Anecdotal evidence of the effects of delayed or impossible collation of relevant facts abounds, but we notice a total absence of either a mechanism for studying this problem or even a genuine concern for it. The CODIB Biographics Task Force report nimbly skirts around the problem by making such innocuous recommendations (CODIB V, p. 4, item 5) as that

"the CODIB Support Staff be directed to prepare and maintain current publications to inform users of biographic information ..." with the hedge that this effort be limited

> "within the limits of security classification and need-to-know prescribed by each agency.*"

This recommendation firmly endorses motherhood with zero real effect; in fact, it flatly contradicts the one previously quoted from page 16 of CODIB V.

H. Conceptual and Managerial Problems

The R & D report rightfully points out that

> "Conceptual and managerial problems are more crucial than purely technical." (CODIB VI, p. 3).

The R & D report further points out (CODIB VI, p. 5) that

> "There is neither an organized set of R & D objectives, a policy for establishing R & D objectives, nor a mechanism for accomplishing either."

It is clear to the Panel that the product of one agency or compartment is the input of another; that there are many loops, cycles, interconnections, etc., but that the prevalent concern for local optimization of resources and performance is unmarred by any visible concern for total systems design.

There is an almost total absence of planning data or statistics in a form that would enable management to determine appropriate courses of action. This Panel would

*Our emphasis.

welcome a refutation of this statement in the form of appropriate data. For instance, although the question of security of shared files is agreed on by everyone to be of primary importance, no one can supply useful data or criteria concerning such elementary matters as the relationship between the level of security precautions and expense and the safety acquired at such a level. It is striking, in fact, that the Chairman of the USIB Security Committee regarded this question as a novel one when it was asked during the Security Briefing of the Panel on March 20, 1966. The data may exist somewhere, but the fact that this Panel was unable to uncover it is itself symptomatic. Elementary facts concerning management techniques should be made available if a decent system is to be developed. (Section II, F)

We have dealt briefly elsewhere in this report with our concern that process become less of a stepchild (or tail) of substance, with the need for users to participate in systems design, with the problems of enforced consensus and lowest common denominator, with the need for a probabilistic view of collection and retrieval, and with the need for introducing competence and understanding of systems design problems at the highest level of management.

At the root of much of the uneasiness over the introduction of automatic information processing technique there is a fear, not only about the matter of security, but also about agency or management authority. After all, folk wisdom tells us that knowledge is power.

This is a valid concern, but we suggest that it can be exaggerated, and that the fear is, to a large extent, fear of the unknown. While there are legitimate reasons to guard information, at least part of this concern arises from a mistaken confusion of information gathering with the exercise of authority. Clearly, the opening of information lines up, down and across would legitimize a leaping over organizational boundaries that, while essential for real accomplishment, is done nowadays only at official risk and peril. Organization lines reflect lines of authority, but while knowledge is power, the gathering of information is not the exercise of authority. It seems, therefore, perfectly proper for a manager to leap several levels down in search of answers, for a subordinate to leap across organization lines and occasionally over his boss's head, so long as decisions and orders travel by normal channels and care is taken to protect legitimate privacy.

This vital question must be dealt with if real progress is to be made in Community-wide information processing.

I. Manpower

The Subpanel believes that there is a serious shortage of skilled manpower applicable to biographic problems. The skills in short supply include both computer systems specialists and biographic specialists. There appears to be a wide variation in competence and experience of these people in the installations visited, and the absence of positive communication channels precludes a sufficient interchange of experience which could help upgrade weaker personnel.

The Subpanel is impressed with the quality of expertise found among senior biographic personnel. However, these specialists appear to be made, not born, and their skill is the result of extensive on-the-job experience. At the same time, this pool of talent is fragmented along the same lines as are the files. This strikes the Subpanel as an unwarranted division of a critical capability. To have experienced indexers review identical source materials in several agencies and yet have analysts be unable to have view of the total picture of an individual strikes the Subpanel as a serious deficiency. We applaud once again the intent of Recommendation #8 of the Biographics Task Team Report which recommends in part,

> "A coordinated program should be developed using EDP methods to provide machine indexes of the bibliographic data processed by any organization in this field, so that the personality information is accessible to a recipient in machine form, with quick followup to the translated source."

(CoDIB V, p.13, p.15; See also Section III, B)

[redacted] Considering the strategic necessity of a pre-eminent capability in biographics, the Subpanel feels that more than good intentions are required.

Therefore, it is recommended that a plan for developing a community-wide biographic capability with the highest professional standards of performance be submitted by 2 Jan 1967. This plan must include means [redacted] for upgrading the skills of individuals to a much higher average level and it must also include plans for adding EDP expertise to the important biographic capability. Appendix 5 of this Report presents some specific suggestions which should have been included in the Biographic Task Team Report.

J. Communications

Recommendation #2 of the Biographics Task Team Report says that each Agency should

> "Study the feasibility of telecommunication links within the national agency check complex to facilitate the exchange of requests and replies."

In this connection, it should be noted that the extent of this problem and its security implications may have been exaggerated. The figures compiled by the Biographics Task Team reveal that an extraordinarily high percentage of requests addressed to biographic files are returned with an indication that no record could be found. (See also Section II, E. Appendices 3, 4) This fraction varies from 80% for Civil Service Commission files to a low of 2% for DIA files. For positive

intelligence requests, three of the five agencies cited stated that 50% of the responses were no hits. Since this response can be communicated by one bit (or none) of information, it strikes this Subpanel that neither traffic density nor security plays a major role and that current time delays are outlandish. (See also Appendix 5a, p. 2 #4) There is no evidence concerning how many of the no-hit responses are due to the diversification and multiplicity of the files.

This Subpanel believes that telecommunication links are obviously feasible and that there is no need to study anything except the costs and the configuration desired. This is routine engineering work and should be implemented within 6 months.

K. An Approach to Security Problems

We note with pleasure a tradition and an experiment which suggest that the security problem for biographic files may be tractable. The Biographic Task Team Report points out that

> "Information about individuals comes from a great diversity of sources, through a large number of channels, and has been stored in a variety of retrieval systems in diverse formats." (page 6, item 1)

The Report then goes on to point out that

> "These have traditionally taken the form of index references, either self-contained or leading to dossier files or individual documents."

Here, the tradition of the Community seems in complete accord with the needs of a readily accessible, mechanized or unmechanized, biographic file system. The CIA experience with Walnut and its associated systems reinforces this impression. The Walnut File contains at least two major types of biographic documents; one

of the counterintelligence type, the other of the Russian biographic type - to which access is controlled by different compartments.

There is, as we understand, a unified index to these collections but security compartmentation is still maintained within the index as a safeguard. The separation of index and dossier leads naturally into a stratified retrieval technique which could minimize the restraint of security problems as the pace of mechanization. Index file volumes are considerably less than dossier file volumes and the one can be mechanized while the other is not. Together with the use of such techniques as LDX transmission, this may lead to reasonable solutions.

The separation of index and dossier leads naturally to the concept of free search on the index with a filter at the output. This is, in fact, the way in which the Walnut File operates. Anyone who is authorized to use this file has access to the index, but when a hit is made in the file, only items which the requester is authorized to see are returned to him - the others are sent to a control point and the requester is notified that he must seek permission of the controller to have access to the data if he is qualified to have it. The fact that this kind of system can work in such a sensitive area is most encouraging. The fact that this technique seems to be unique illustrates how slowly ideas travel, even on innocuous matters of technique. COINS is presently trying to grapple with some of the problems of adding biographic files to

At least the requestor knows there is something to be after. If he decides to insist on seeing the material, he has a position.

the system. It should be instructive to study their selection of files for inclusion in the system and of files that will not become part of the system. Except for large dossier-type files, there should be no particular bias except that of utility. Security should not present any serious barriers.

The implication of the Walnut approach is quite serious. It suggests that, in principle, the notion of free index search can be adopted throughout the community. The material in dossiers may then be kept at any level of classification, subject to appropriate output filtering. The requester can be notified that material is immediately available; that there are more hits, but that he should consult the controller for permission to see these.

The importance of this technique lies in that, without it, ~~this technique~~ access to otherwise denied files is at best a passive matter. An analyst may send out a request for information and may hit a very sensitive file. Those who control the file may, if they feel so inclined, take his request seriously, sanitize the material, and make it available to him. This is in principle the role that resident controllers in groups that keep sensitive files play in substantive production areas. Given the general climate in the community, however, we suspect ~~one would suspect~~ that the tendency to give no response in the face of great trouble is indeed be very great. A detailed analysis of how such requests are handled is not available, but the Subpanel would be delighted ~~happy~~ to have this point refuted.

The value of positive information about hits is that without a need for creating a central authority (~~with~~ hence out the disadvantages inherent in ~~such a service~~ centralization), one could insure much greater pressure for a request to be filled. An analyst who knows that someone else holds material that may be interesting to him, and that he has the right and the duty to request access to such material unless positively denied by security rules, is likely to be far more alert and diligent ~~to~~ in digging out this information and exercising pressure via his superiors than if he is at the mercy of passive responses to a broadcast request. The labor of justifying a denial under these circumstances may also counterbalance the labor of sanitizing material to make it available.

The Walnut File approach thus offers potential advantages to both those who must maintain security of the file and to the user who needs access to all information which he is justified in having.

L. File Conversion

The enormous size of present biographic files creates a major problem in automation. ~~That problem grows with every delay and so does the waste inherent in filing~~ Conversion from manual, or partially manual, files to machine systems is hindered by the effort required. It is interesting, in this connection, to observe that the excellent principle of conversion of minimal usable files has, either consciously or unconsciously, been followed in some of the experimental situations we have seen. The example here is the Walnut/Russian biographics scheme where a name search system has been developed experimentally for application to the totality of a relatively small Russian biographics scheme. One of the

material that will later need to be converted

problems that must be considered in doing conversion is that of maintaining usable files at all times during the conversion process. The Subpanel noticed that some installations have managed to convert their files, or are in the process of doing so, while others are wringing their hands and pointing to the differences which ostensibly prevent them from starting a conversion effort. (Appendix C)

While it is by no means true that the totality of all files should be converted, it is nevertheless true that no fundamental technical problems prevent conversion. The Subpanel was also encouraged by finding that stiff directives from high authority (e.g. at Fort Holabird, See Section IV, B) has great effect in overcoming non-technical obstacles of all types. The implication is that the Community is not nearly as helpless as appearances are, as we see it, intended to suggest.

For instance, on page 13 of the CODIB Report on the SCIPS Study (CODIB 1964), the following statement is quoted from the SCIPS Report:

> "The most pressing problems of systems integration or interface appear to be between components within agencies rather than between agencies."

The CODIB comment concurs, but with a hedged statement:

> "This statement is probably true and deserves careful consideration. This is not to say that information processing does not warrant community consideration to a considerably greater degree than it has had to date -- it does; but this finding reflects a logical first-things-first philosophy."

In the light of our Analysis of CODIB history, we see this as clearly killing with kindness. The biographics area is such that it is by no means clear-cut that more can be done within an agency than across agency lines. This Subpanel recommends this question be experimentally. However, what we see as a remarkably refined CODIB ability to relegate reports to limbo, water down findings, and pull the teeth out of recommendations must be overcome during the course of such an experiment.

In conclusion, this Subpanel recommends that file conversion studies be integral to any interagency biographic file automation experiment. The cost and effort required for file conversion will not decrease - it will increase; the resulting strain on manpower resources will grow. Hence a major effort should be initiated to convert manual biographic files on a large scale in the community. The experience and data obtained will be very helpful in other file conversion efforts outside of biographics.

DRAFT

SECRET

SECRET/NOFORN/CONTROLLED DISSEM

25X1A

BIBLIOGRAPHY

[redacted] "Processing of Printed Intelligence Information" This paper was prepared for the ODDR&E during June 1961, by
25X1A Mr. [redacted] while he was serving a two-week tour of military duty with the Defense Department. (S&T Cont.No. 0553-b4)

BOREL 1963: United States Intelligence Board, Committee on Documentation "Foreign Intelligence Objectives" CODIB-D-107/4, 16 September 1963, Paul Borel, Chairman (CODIB). Attachment: Excerpt from DCI Report on Intelligence Objectives (SC 08732-63/1) (S&T Cont.No. 0229)

BOREL 1965: Memorandum for Mr. William T. Knox on "Problem Areas of Possible Interest to Panel", 22 December 1965, and Attachment "Selected Problem Areas in the Intelligence Community", 22 December 1965, by Paul Borel, Ass't Deputy Director for Intelligence, CIA. (S&T Cont.No. 1861)

BOREL 1966: Memorandum for Mr. William T. Knox on CODIB Presenta-
25X1A tion to PFIAB/OST Joint Guidance and Evaluation Panel, 25 April 1966. (S&T Cont.No. 0691)

[redacted] United States Intelligence Board, Committee on Documentation "Progress Report on Foreign Intelligence Objective No.3", CODIB-D-107/8, 9 March 1964, [redacted] Secretary (CODIB) 25X1A
Attachment: Memorandum for Deputy to DCI/National Intelligence Programs Evaluation (6 March 1964) by Paul Borel, Chairman (CODIB) with attachment: (Proposed Text for) "Progress Report on United States Foreign Intelligence Objective No. 3: Automatic Data Processing" (S&T Cont.No. 0229)

CLIFFORD: President's Foreign Intelligence Advisory Board, Clark M. Clifford, Chairman, Memorandum for the President on U.S. Intelligence Community Capabilities for the Handling of Intelligence Information, June 15, 1965

CODIB V: Report, T/V/R-1, dated 17 January 1966 of Task Team V (Biographics) of the Committee on Documentation of the United States Intelligence Board (CODIB) (S&T Cont.No. 0142)(Stamped "DRAFT") [Revise to latest version sent by Landau/Borel 6/66]

CODIB VI: Report T/VI/R-1, dated 28 September 1965 of Task Team VI (Intelligence Data Handling Research and Development) of the Committee on Documentation of the United States Intelligence Board (CODIB) (S&T Cont. No. 1586)

CODIB 1961: CODIB Third Annual Report, 30 June 1961, Serial No. CODIB-AR-3

CODIB 1962a: CODIB Fourth Annual Report, 31 July 1962, Serial No. CODIB-AR-4

SECRET

SECRET/NOFORN/CONTROLLED DISSEM

CODIB 1962b: Information Processing Programs in USIB Member Agencies (Appendix C of CODIB 1962a)

CODIB 1963: CODIB Fifth Annual Report, 5 September 1963, Serial No. CODIB-AR-5

CODIB 1964: United States Intelligence Board Committee on Documentation "Stage 1 Report of the Staff for the Community Information Processing Study (SCIPS)" USIB-D-39.7/5 COBID-D-82/28, 26 February 1964

CODIB 1965a: CODIB Draft Task Team Quarterly Report for the period 1 July to 30 September 1965, Serial No. CODIB-D-112/5, 1 November 1965, Limited Distribution

CODIB 1965b: CODIB Seventh Annual Report, 3 December 1965, Serial No. CODIB-AR-7

CODIB 1966: CODIB Draft Task Team Quarterly Report for the period 1 October to 31 December 1965. Serial No. CODIB-D-112/6, 18 January 1966, Limited Distribution

KIRKPATRICK 1960: The Joint Study Group Report on Foreign Intelligence Activities of the United States Government, 15 December 1960, (TS Control No. 172800) (S&T TS Cont.No. 598)

SCIPS 1963: Staff for the Community Information Processing Study (SCIPS) Stage 1 Report Vol. II, Oct. 1963 (SC No. 12481/63, 15 January 1964) pp. 57-70

DRAFT

SECRET/NOFORN/CONTROLLED DISSEM

Appendix 1

Nature of the Biographics Problem

DRAFT

SECRET

SECRET/NOFORN/CONTROLLED DISSEM

Appendix 2

Request for Information and CODIB Support Staff Response

SECRET

SECRET/NOFORN/CONTROLLED DISSEM

NATURE OF THE BIOGRAPHICS PROBLEM*

"The Intelligence Community has for many years collected an ever-increasing amount of information about individuals from a great diversity of sources through a large number of channels, and has stored these data in a variety of retrieval systems in diverse formats. These have traditionally taken the form of index references, either self-contained or leading to dossier files or individual documents. The Team decided, as a point of departure, that the relative pay-off in system improvement would be higher in respect to the larger biographic files in which there is a high degree of activity and interagency communication. Thus, many of the smaller files studied by SCIPS (the Staff for the Community Information Processing Study) were not included.

"There are three types of major biographic indexes and files now in operation. They are the Positive Intelligence, Counterintelligence and Security holdings. There is relatively little exchange of requests between the PI biographic files and the Security files, moderate exchange between the CI and PI communities and frequent exchange between Security and CI. The Counterintelligence (CI) biographic system centers around the foreign counterintelligence repository of CIA and the domestic counterintelligence holdings of the FBI. The Security and PI holdings of the agencies referred to in this report also lead to CI data in some degree. The interagency exchange of Security data centers around the name search type operations performed by CIA, State, Army, Navy, Air Force, FBI, Secret Service, Immigration and Naturalization Service (INS), and Civil Service Commission (CSC). The major PI biographic records are contained in the files of the CIA Biographic and Special Registers, DIA, NSA/Office of Central Reference, Department of State and Air Force/Foreign Technology Division (FTD).

"There are important and fundamental differences between, and some similarities in, the basic operating procedures and kinds of searches that are made in the PI systems versus the CI/Security systems. The PI biographic systems are deeply intertwined with, and in many cases actually part of, larger intelligence collection and storage systems which are mission, subject or area oriented. In contrast, the CI/Security systems are clearly oriented to the heavy use of name searching among alphabetically ordered biographic indexes which, in most cases, lead to dossier files. The Team determined that there is name searching and name finding going on in both the Positive as well as the CI/Security activity. However, the bulk of the requests in both areas involve name searching (above 95% in the CI/Security area and about 80% in the PI area)."

- - - - - - - - - - - - - - -

*(CODIBV, p.6, Par. 1,2,3)

SECRET