SECRET

NATIONAL SECURITY AGENCY
FORT GEORGE G. MEADE, MARYLAND

CRYPTOLOG

2nd Issue, 1989

P.L. 86-36

THIS DOCUMENT CONTAINS CODEWORD MATERIAL

SECRET

NOT RELEASABLE TO CONTRACTORS

To submit articles or letters by mail, send to:
Editor, CRYPTOLOG, P1, NORTH 2N018

If you used a word processor, please include the mag
card, floppy or diskette along with your hard copy, with
a notation as to what equipment, operating system,
and software you used.

via PLATFORM mail, send to:
cryptlg@bar1c05
(bar-one-c-zero-five)
(note: no 'o')

via ALLIANCE, send to:
PLBROWN [note: all caps]
attn: CRYPTOLOG

Always include your full name, organization,and secure
phone; also building and room numbers.

For Change of Address
mail name and old and new organizations to:
Editor, CRYPTOLOG, P1, NORTH 2N018
Please do not phone.

The first issue of CRYPTOLOG came out fifteen
years ago. It was designed to be a monthly,
"written by technicians for technicians,
informal, newsy, controversial, lively, and
timely." The budget crunch did away with the
"monthly." But thanks to you, the reader-
writers, the other attributes still apply.

"To be successful, CRYPTOLOG must reflect
current operational topics in a way that
interests you and others." So wrote General
Wolff, the then DDO, in a letter of introduction
to that first issue, August 1974. He went on to
say, "I hope that you will want to read it and
will help to write it."

We are repeating his words for the benefit of
newcomers to CRYPTOLOG. You, the readers,
are also the writers. The vitality of the
publication is due to the animated exchanges of
views among readers.

And now for two high-tech events to mark this
anniversary. (We are state-of-the-art ourselves,
as befits the high-tech sponsoring organization.)
CRYPTOLOG can now receive contributions via
ALLIANCE as well as over PLATFORM.
Instructions appear elsewhere on this page.
Also, beginning with this issue, we are moving
over to desk-top publishing on the Macintosh as
conversion software permits.

CRYPTOLOG
has moved to OPS-1, 2N018.
New Mailing address: P1 NORTH

FOR OFFICIAL USE ONLY

In Honor of Women in Science and Engineering
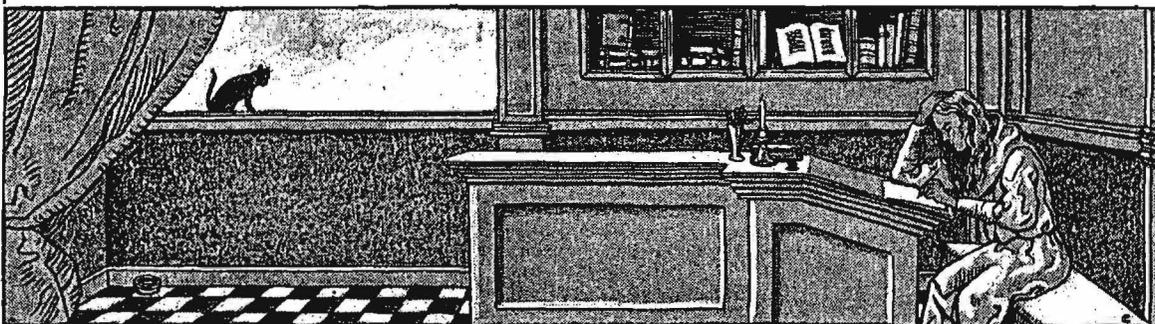
Elizabeth Rindskopf, GC

This article is based on an address at a luncheon of the Federal Women's Program honoring NSA's women in science and engineering.

It is an honor to be asked to speak to you today but rather daunting too—I have read your bios and know the talent that I have before me. Indeed, such professional distinction and accomplishment — and in such great quantity—makes me wonder just what I might have to say today which could be directly relevant to you in your busy lives. In short, I asked myself, as I sat down to prepare these brief remarks, what message might I want to deliver this morning and what could I hope to accomplish in speaking to you.

The answer that came to me is simple and intensely personal. It may be that what I most want to say will have greater meaning for me than for you. And so I ask your indulgence and permission to express my own personal feelings on the meaning and importance of women in Science and Engineering today. For what I want to say today is first, thank you, and second, keep it up, and third, spread the word. Let me spend a minute telling you a little more about each of these messages.

Thank you. Some of you may recall an interesting program produced and directed by Peter Ustinov several years ago on the dangers and likely impact of a nuclear holocaust. It combined multiple interviews with those involved day-to-day in "manning" — and I use that word advisedly — our nuclear defense. Interspersed with these interviews was a dramatic portrayal of the impact of a nuclear attack on a small group of people—a mini "day after", if you will. In all of the many interviews not one woman appeared, never mind spoke. The only woman, in fact, was one actress in the dramatic inserts. I was struck by this total absence of women. Weren't we half of the population? Where were we when decisions about matters so central to the future of the nation— our civilization—were being made? Shouldn't we be involved with our unique feminine perspective? Of course we should.

Well, why weren't we there? Was Don Regan really right when he commented that women simply weren't interested in disarmament talks and related issues? No, I felt in my heart that that view was wrong.

The correct answer was a different one: Women as a group have traditionally lacked the technical training to be able to participate effectively in many of the technical discussions surrounding issues so critical to the nation's security, issues such as nuclear war, disarmament, and so on. And lacking technically informed voices, they have not been heard at all.

To my mind, this is not only wrong, it is wasteful and foolish as well. Wasteful because we need all the good minds we have to solve the problems that confront us as a nation and as an agency. Let's be frank. This Agency and this nation are on the brink of a serious shortage of mathematicians and technically trained people. To survive in the next century the nation must muster every bit of its technical and scientific strength. And that means

NSA Women Honored for Contributions in the Fields of Science and Engineering, Women's History Month, March 1989

women must be a part of the solution; without them there may not be a solution. Foolish because I believe that women are in fact different than men— neither better, nor worse, but with an important and often unique perspective and wisdom all their own. We waste a precious asset when we organize ourselves to make important decisions without the presence feminine character and point of view. And, as a lawyer, I might also add that we undercut the very genius of our democracy—a system based on the strength in diversity and the healthy tension in the competition of different ideas, different strengths and different perspectives. In short, we need the participation of women in all avenues of public life. But they cannot participate unless they are properly prepared to do so. And so today I thank you for the hard work you have done to obtain the training that will allow you to excel in the technical leadership of this very high tech and very important agency.

And let me say a word about the group of women we are honoring today. Twenty-seven candidates were selected for nomination by their organizations. They were mathematicians, computer programmers, electronic engineers; some had doctorates; one taught mathematics at a school no less prestigious than the University of Texas. Others taught themselves what they needed to know right out of high school.

Yet despite this highly impressive array of educational credentials and superb work performance, all are—at least from my perspective —young , between the ages of 23 and 37, to be precise.

Their relative youth brings me to my second point: Keep it up. We need you now doing the creative work of the technical expert. We will need you tomorrow as part of the Agency's Senior Management team. And here I'll offer another personal perspective: Too many times I find myself one of a handful of women, if not the only woman, sitting around the table in the Director's Conference Room. I've formed a habit of taking a head count. 75 men to 1 woman is not an acceptable ratio for these meetings. It must and it will change, but only if you make it happen.

Let me not be guilty of suggesting that you will have an easy time as you grow and progress in your chosen field. I can guarantee that you won't. As pioneers, each of you will face special challenges: many of you will lack adequate role models and as a result you'll be required to spend the extra energy demanded of all those who serve as our trailblazers. I can tell you this with confidence because I've been there. Just remember that your task is in many respects to change an entire culture that fails to encourage and may actively discourage the woman scientist or mathematician. And many will have the added challenge of working two equally hard jobs: raising young families right at the point when your career is at its most competitive. For this you will require hard work, military discipline and, most importantly, a sense of humor. But keep at it; you can do it. Perhaps a phrase from my high school Latin course will help: POSSUNT PROPTER COGEUNT — They can because they think they can.

Before I leave this point, let me mention something you won't have to fight: NSA Management. I believe you have as supportive an environment as you are likely to find anywhere today. Does that mean I guarantee you won't ever face discrimination? No. But I believe you will have the management support you need to conquer the problems you confront. Keep at it: you won't regret it if you do, but we will if you don't.

Lastly, spread the word. We need more technically versed women here at NSA and nationally.

I hope you'll encourage other women in their pursuit of technical and scientific careers, both here at NSA and outside as well. And I hope you'll get involved, when you can, in serving as role models and teachers of girls and young women still in school. And the earlier they see your example, the better. They need you to see what a kind of career the technically trained women can have.

Some will need your help to tackle the persistent problem of math anxiety, and other difficulties particularly faced by girls and young women who pursue technical careers. I believe much remains to be done in learning why it is that our young women find themselves, considered as a group, less comfortable than their brothers in learning math and science. Do they learn the principles of math and the sciences differently than boys? Possibly. Are they often intimidated and discouraged when they confront technical subjects? Probably. Can you make a difference in attacking this problem? I believe you can. Your example inspires; it encourages and creates the self-confidence essential to any fine performance.

Seeing your pictures as I came in today reminded me of just how powerful an example you can set. Some months ago, I visited the college my daughter has selected. Some of you may know Wellesley and its long tradition of encouraging academic excellence in women. For me it was a new experience, but one I will long remember. As I walked into the library's main reading room, I looked up to the clerestory windows bathed in lovely fall sunlight. There, hung all around the room, were the elegant oil portraits of all the college's past presidents, and each and every one had the face of a woman. I realized then with a thrill of surprise that in all my time in various fine educational institutions, I had never before been a room where those so honored were exclusively women. The message was clear in those confident gazes. They could do it.

About Classification
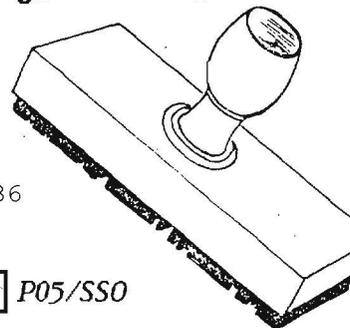
P.L. 86-36

P05/SSO

It is widely recognized that the volume of classified material NSA handles is enormous. It is therefore incumbent upon us also to recognize the importance of proper classification practices and to ensure that our actions reflect the sound judgment required.

Individuals responsible for making classification decisions should remember that overclassification increases our vulnerability in two ways:
◆ It increases the volume of documents to be protected.
◆ It degrades our awareness of documents which truly merit TOP SECRET classification.

The classification decision process includes four steps:
◆ An initial decision to classify information based on whether it concerns military plans, weapons, activities, or intelligence operations that are classifiable;
◆ A decision concerning whether unauthorized disclosures of the information can cause damage to the national security;
◆ A decision as to whether the information should be classified

TOP SECRET - information the disclosure of which could cause exceptionally grave damage to the national security;

SECRET - information the disclosure of which could cause serious damage to the national security, or

CONFIDENTIAL - information the disclosure of which could cause damage to the national security.

◆ A provision for an indefinite period of classification, unless an original classification authority establishes a date or event upon which automatic declassification should occur.

Note that classification decisions cannot be made merely to conceal violations of law, inefficiency, or administrative error; prevent embarrassment to a person, organization or agency; or to restrain competition.

COMPUTER SECURITY POLICY
FOR THE
INTELLIGENCE COMMUNITY
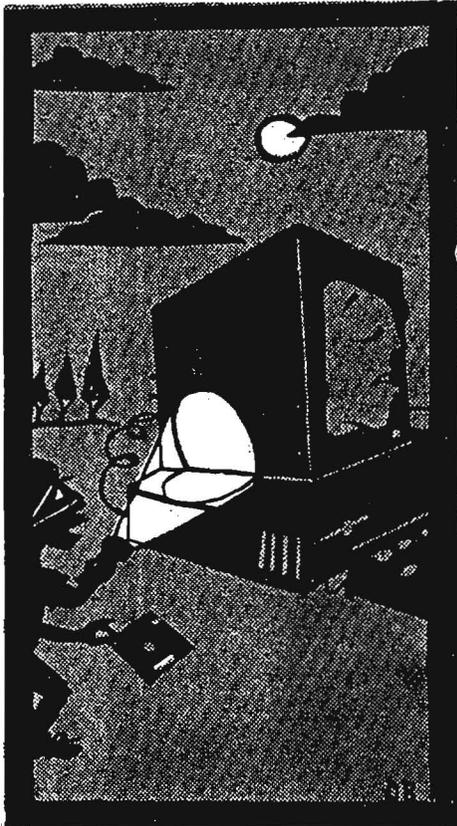
P1

(U) This article summarizes the current national level policy on protecting intelligence processed in computer systems or networks. It also discusses a problem that has plagued system security officers, accreditors, and users for years; how to determine the mode of operation of a computer system in the NSA/CSS environment.

BACKGROUND

(U) The Director of Central Intelligence (DCI) publishes security policy in the form of DCI Directives called DCIDs (pronounced D-skids). These cover all aspects of security, including personnel security, physical security, operations security, and computer security. DCID 1/16 "Security Policy for Uniform Protection of Intelligence Processed in Automated Information Systems and

Networks" covers computer security. Emphasis is placed on the *uniformity* of protection mechanisms, which becomes extremely important when we talk about interconnecting Automated Information Systems (AISs). Emphasis is also placed on the protection of *intelligence* processed in AISs and networks. Thus, this DCID applies not just to all classified information, not just to all sensitive information, but to all classified information that involves, or is derived from, intelligence sources or methods. The first DCID treating computer security as a separate topic was published in 1972. The DCID in effect before 19 July 1988 was dated 4 January 1983, but was substantially the same as the DCID published in 1972. Obviously, there have been a lot of changes in the computer security field since 1972, and the current version of the DCID reflects those changes.

GENERAL POLICY GUIDANCE

(U) The purpose of the directive is to establish long-term (year 2000) goals and near-term requirements (year 1992) intended to improve the security of U.S. intelligence processed in AISs and networks. The goal is for all Intelligence Community AISs to become trusted systems incorporating trusted products by the year 2000. In the period leading to the achievement of the foregoing goal, interim measures to improve security will be taken, including accrediting or reaccrediting systems, implementing minimum near-term security requirements, and improving access control and monitoring.

ACCREDITATION

(FOUO) Each AIS and network covered by this policy must be accredited to operate in one of the four modes of operation: dedicated, system-high, compartmented, or multilevel. Accreditation requirements vary with the mode of operation. The minimum levels of trust (as specified by the DoD Trusted Computer System Evaluation Criteria, or "Orange Book") to be achieved as a goal by the year 2000 are: Dedicated Mode - C1; System-High Mode - C2; Compartmented Mode - B1; and Multilevel Mode - B2. Of course, a higher level of trust may be mandated for a particular

SECRET

MODES OF OPERATION

USER ATTRIBUTES	Dedicated	Compartmented	System-high	Multilevel
Need-to-know	All	Some	Some	Some
Access	All	All	Some	Some
Clearance	All	All	All	Some

AIS by its sponsor or responsible accrediting authority. The responsible accrediting authority for NSA/CSS is the NSA/CSS Senior Computer Security Coordinator (T03).

MODES OF OPERATION

(U) In DCID 1/16, the fundamental basis for deciding which protective mechanisms are appropriate for a given situation is a concept called mode of operation. There are four modes of operation used for AISs and networks processing intelligence: dedicated, system-high, compartmented, and multilevel. There are exactly two sets of information needed to determine the mode of operation of a system: (1) the classifications and categories of information on the system, and (2) the clearance level, formal access approvals, and need-to-know of all users of the system. These two sets of facts are the sole determinant of the mode of operation of a system.

(FOUO) For purposes of this policy statement, categories of intelligence information include Sensitive Compartmented Information (SCI) and Special Access Programs for Intelligence (SAPIs). SCI includes either COMINT (sometimes referred

to as Special Intelligence or SI), TALENT KEY-HOLE, or BYEMAN information, or any combination thereof. Those programs that require, as a condition of access, the signing of a nondisclosure statement are considered to be SAPIs. Therefore, in NSA/CSS, programs in the Very Restricted Knowledge system (VRKs) and the GAMMA subcompartment are considered to be SAPIs and are given the status of SCI compartments and subcompartments in terms of the security levels required for their protection.

(U) For a system processing SCI, clearance is based on criteria contained in DCID 1/14, "Minimum Personnel Security Standards and Procedures Governing Eligibility for Access to Sensitive Compartmented Information." In NSA/CSS that means a user must have either a green, gold, or black badge.

(U) The attached chart shows how user attributes (need-to know, formal access approval, and clearance level) for the information on the system determine the mode of operation of the system. These modes of operation are described in further detail below.

SECRET

DEDICATED MODE

(U) When *all* users are cleared for *all* data on the system, and *all* users have formal access approvals for *all* data on the system, and *all* users have the need-to-know *all* data on the system, the system is operating in the *dedicated* mode of operation. The system is not required to provide *any* technical security at all. All the security for the system is based on traditional security disciplines, such as physical security, personnel security, communications security, and so forth.

SYSTEM-HIGH MODE

(U) When *all* users are cleared for *all* data on the system, and *all* users have formal access approvals for *all* data on the system, and *at least one user* has the need-to-know *only some* of the data on the system, the system is operating in the *system-high* mode of operation. The system is not required to provide much technical security, only a reasonable amount of user separation. Systems operating in the system-high mode are usually also trusted to provide some level of access control and accountability (i.e., logins and passwords) and, perhaps, some level of audit trails. Almost all the security for the system is based on traditional security disciplines.

COMPARTMENTED MODE

(U) When *all* users are cleared for *all* data on the system, and *at least one user* has formal access approvals for *only some* of the data on the system, the system is operating in the *compartmented* mode of operation. The system is required to provide some technical security over that needed for the system-high mode. Systems operating in the compartmented mode must provide a mandatory separation of users and data based on some sort of formal compartmentation and labeling. They must also provide a high level of access control and accountability (e.g., logins and passwords, biometrics, etc.), and a detailed and reliable level of audit trails. Even with this improved technical security, much of the security for the system is based on traditional security disciplines.

MULTILEVEL MODE

(U) When *at least one user* is cleared for *only some* of the data on the system, the system is operating in the *multilevel* mode of operation. The system is required to provide a high level of technical security; in addition to the technical security needed for the compartmented mode, systems operating in the multilevel mode must provide a mandatory separation of users and data based on classification and clearance. They must also provide an extremely high level of access control and accountability (e.g., logins and passwords, biometrics, etc.), and a detailed and reliable level of audit trails. Even with all this technical security, some of the security for the system is based on traditional security disciplines.

P.L. 86-36

ACCESS BY FOREIGN NATIONALS
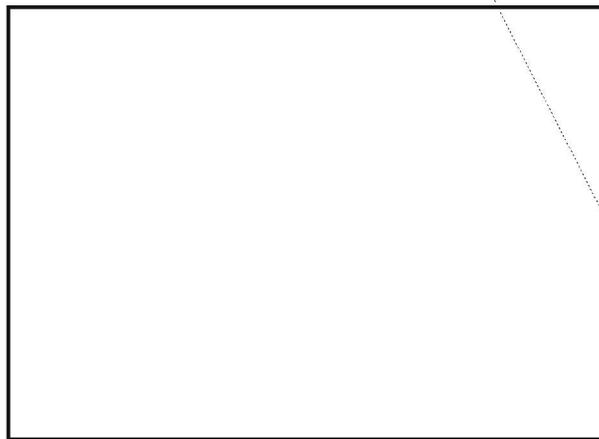
CONCLUSION

(U) If you have had experience working with previous policy statements on computer security, you will note that we have come a long way in defining the parameters and technical measures for protecting intelligence information. If this article has sparked your interest, I recommend that you obtain a copy of DCID 1/16 and its supplement, "Security Manual for Uniform Protection of Intelligence Processed in Automated Information Systems and Networks", dated 19 July 1988.

ACROSS THE POND

S331

P.L. 86-36
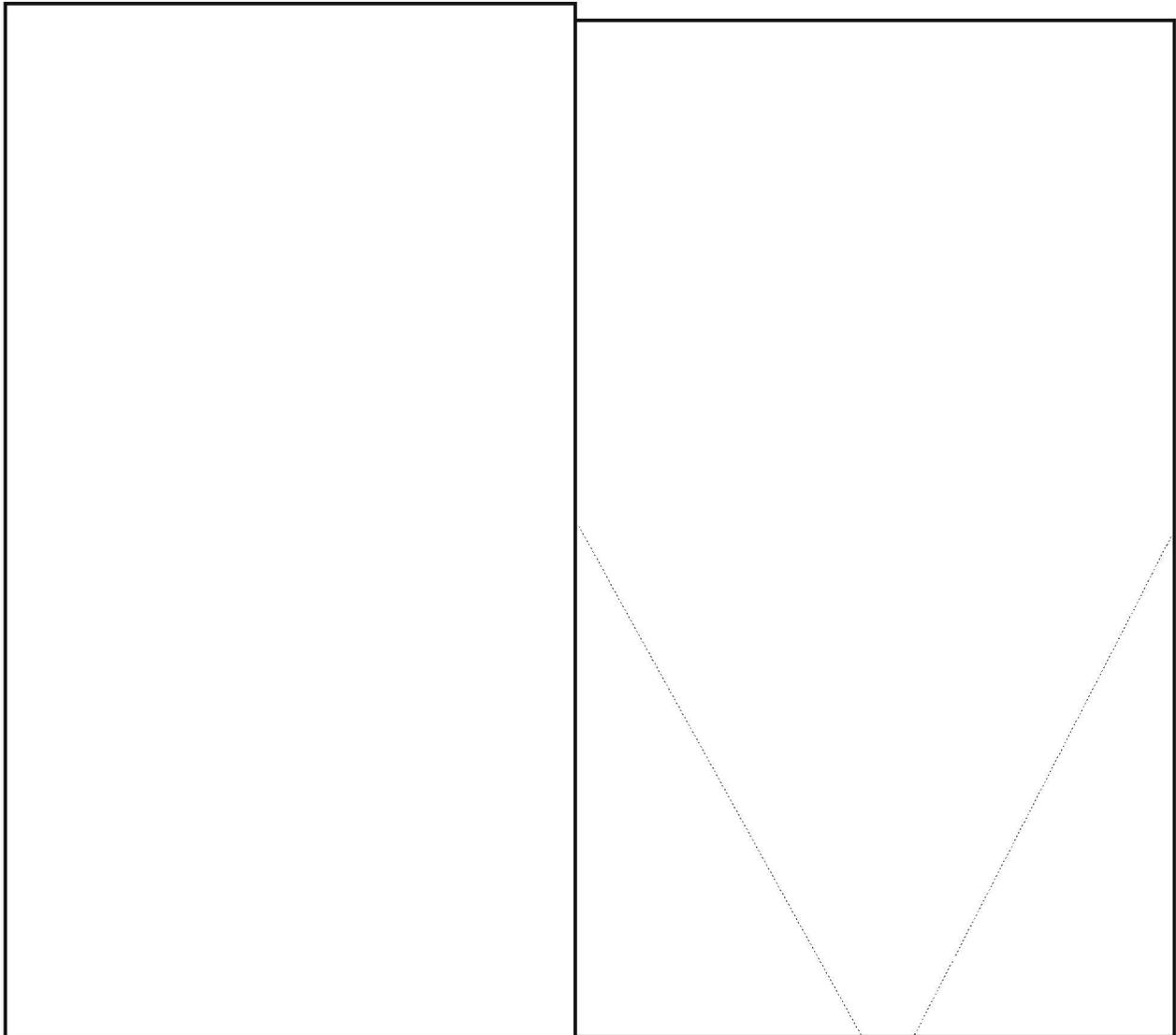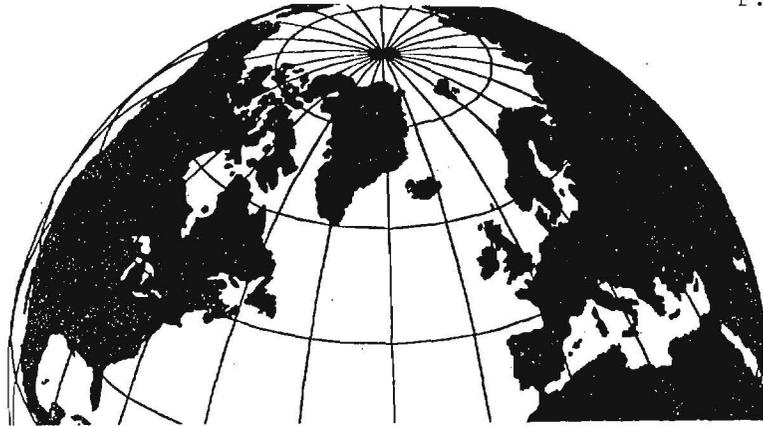
P.L. 86-36
EO 1.4.(c)
EO 1.4.(d)
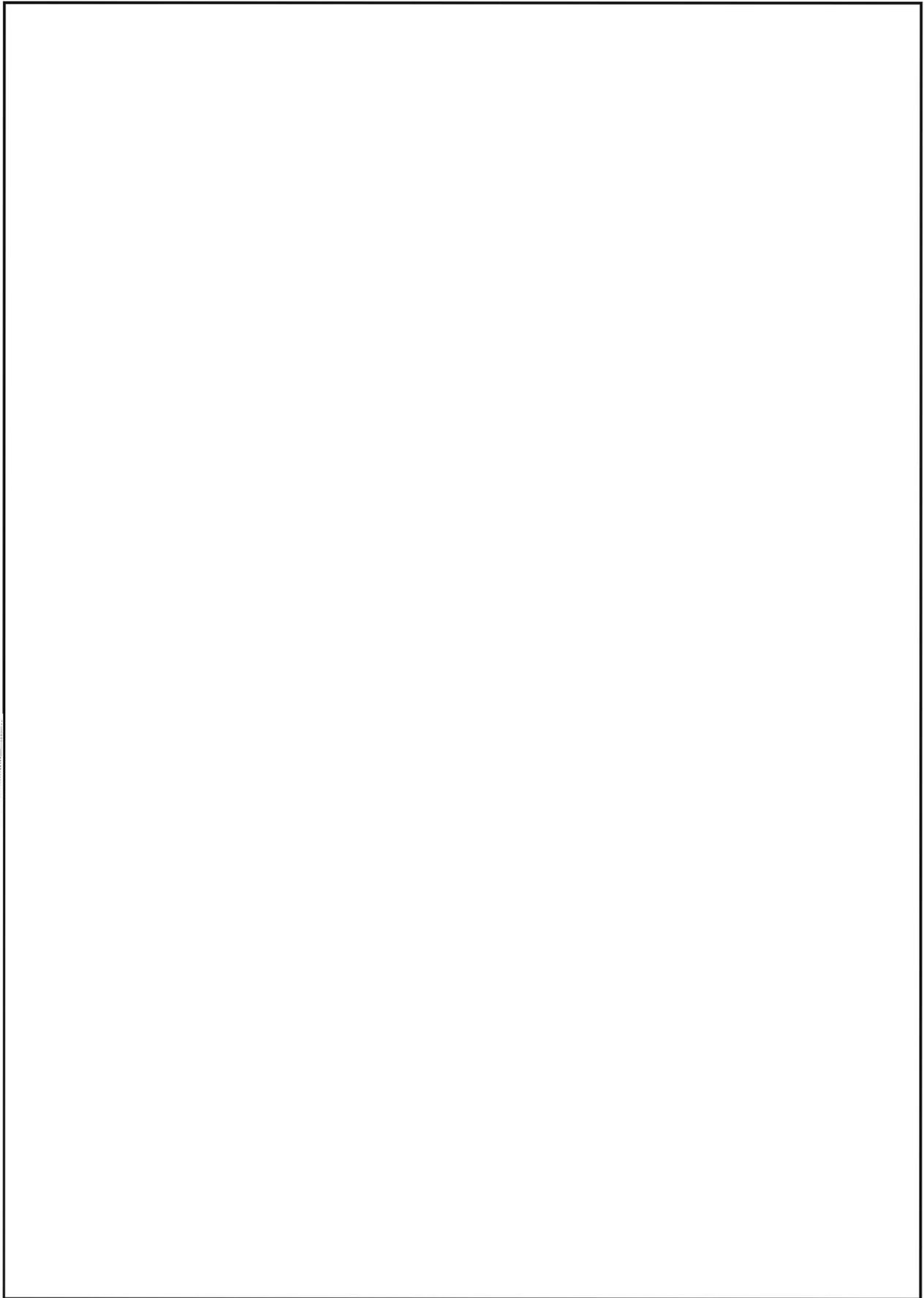
SECRET
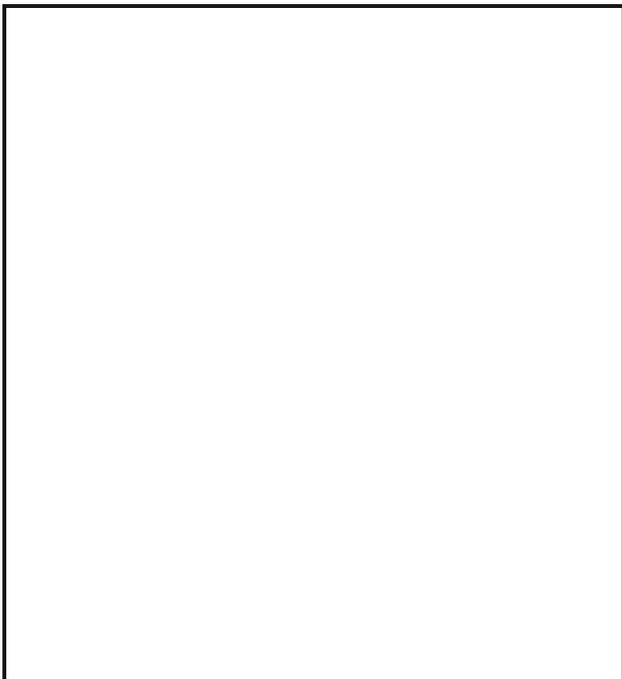
distance. Most shops close at 5:30 p.m. and most restaurants and theaters open at 7:00 p.m. Surrounding Cheltenham are the beautiful Cotswold hills, which provide an immediate retreat from the bustling, rather crowded business and residential areas.

(U) I traveled extensively throughout the country on coaches and in my Mini, in which I could drive on the narrowest of country roads and park in the smallest spaces in the cities. Armed with my AA (Automobile Association) membership, atlas, and tour books, and usually in the company of British friends, I thoroughly explored the Cotswold hills and Gloucestershire, home of the famous Gloucester Cathedral approximately 10 miles southwest of Cheltenham; visited London, Bath, Oxford, and Cambridge; and traveled to the Lake district in the northwest; Yorkshire in the northeast (where I stayed at Menwith Hill near Harrogate); Devon (including Torquay in "the English Riviera") and Cornwall in the southwest; Canterbury Cathedral, Windsor Castle, and Hastings in the southeast, and the Isle of Wight in the south.

(U) All over Great Britain, including Scotland and Wales, and also Southern Ireland and the Channel Islands, I saw everything from palaces and manor houses to tiny village pubs, rich in historical significance, nestled in magnificantly scenic countryside, a photographer's paradise. Although a few places, such as Stonehenge, now have fencing to protect them against vandalism, the countryside is kept very clean and neat by residents and travelers, as well as by organizations like National Trust and English Heritage. Also, much to my surprise and delight, travelers even encounter a fair amount of sunshine from time to time.

(U) A geographic and social feature of England of which I had not been previously aware was the contrast between the North and the South. The north of England is rugged and hilly with plenty of good low-priced housing, but a shortage of jobs, mostly in factories, which are closing down in large numbers. By contrast, the flatter, more urbanized south(east) has ever-increasing high-paying jobs, largely in new international computer firms, but only very high-priced houses and hardly anything like our apartments for people who would be just starting their careers or would be willing to move there from the north.

(FOUO) In spite of some of those morale dampers, the workers I met were generally cheerful and dedicated to using their highly developed abilities to the fullest extent. Personally, I was treated very well by my British supervisors and co-workers and made many friends. I rarely met anyone who "tooted his own horn," and I discovered that a major attribute of "English reserve" is patience which does not allow hasty judgments but gives people a chance to prove themselves.

(U) Outside of work, I met many really friendly English people in Cheltenham, a heterogeneous town about 3 miles in diameter located almost 100 miles west of London in the county of Gloucestershire, in the South Midlands. I spent considerable time socializing with friends from work, church, the Royal Scottish Dance Society, and other organizations.

(U) I rented a 7-room house in Cheltenham with all the modern conveniences one needs, and which was located within walking distance of work and the town center. In town were an abundance of lovely shops with the latest fashions (especially reasonably priced woolens), parks and gardens (one with a boating lake), schools, churches, doctors and dentists, a choice of grocery stores and produce markets, bookstores, camera shops, sport shops, British Gas and the Midlands Electricity Board, banks and administrative offices, appliance centers, theaters, restaurants, antique shops, and anything a person needs, all within walking

SECRET

HANDLE VIA COMINT CHANNELS ONLY

(U) Concerning prevalent English views of America, I noted that many people do indeed consider our country as being somewhat similar to theirs, but just more technologically advanced and less polished and less rich in cultural heritage. I noted also, however, that a surprising number, even the large number who like the Americans whom they have met, also strongly identify our country with the crime, aggressiveness, extreme wealth and wastefulness, neurosis, and superficiality that is seen in the very popular soaps like *Dallas* and *Dynasty,* police films, etc., which they receive from us. Although their news coverage is quite objective, some commentaries and documentaries, when seen in light of these one-sided films, are often misconstrued by the populace to the detriment of our country.

(U) Many people attribute several of England's recent social problems to U.S. influence. Some also act cool, at least initially, towards Americans, especially in the vicinity of Royal Air Force/U.S. military bases like Fairford, apparently because of their sometimes justified stereotype of Americans abroad. Nevertheless, I met even more people who were fascinated with everything American and wanted to visit America, or had already visited America and had a marvelous time. In my experience, the majority expressed admiration for the openness and friendliness of the Americans they had met. Others even changed their minds and decided to include a visit to America in future holiday plans.

(U) My tour in Cheltenham was, on the whole, very beneficial and enjoyable. Although it seemed to take a long time to adjust to the new surroundings and cultivate friendships, the results were well worth my efforts. I learned a great deal from the English people, not only in office matters, but also in human relations, and I discovered *tremendous opportunities for travel.* I highly recommend this tour of duty. □

NEW MAILING ADDRESS

CRYPTOLOG
P1
NORTH

A NOTE TO CONTRIBUTORS

Though CRYPTOLOG is an informal publication, it closely observes the etiquette of publishing. One of the provisions concerns anonymity. According to that rule, the Editor must know the source of the information but may not reveal the identity without permission; and without identification any submission is discarded, whatever its apparent worthiness. CRYPTOLOG has followed this rule many times over the years.

So, dear author, if you've sent a contribution without identifying yourself, you won't see it in print. Don't bother to scan issue after issue for your submission. You must make yourself known to the Editor. No exceptions. Submissions received without identification are tossed out.

If you have already submitted something anonymously, resubmit it with your name, organization, building, and secure phone number—that is, if you want to see it in print.

And please include, along with the hard copy, a floppy with specifics about the computer used, the operating system, and the word processing software.

Unclassified

THE EXCITEMENT OF INFOSEC

D2

P.L. 86-36

Some time ago, while I was having lunch with the Director of Security of one of our NATO allies and we were discussing the rash of books on intelligence agencies such as the CIA and Britain's MI-5 and MI-6 that were flooding bookstores, he asked, "Why aren't there more best selling books on INFOSEC?" I replied, "It's because the best days we have in INFOSEC are when nothing exciting happens in the outside world. When we are successful, which we are most of the time, the result is a non-event."

During the Walker spy trial, Earl Clark, an NSA INFOSEC expert, said, "Give me access to your codes, give me access to your ciphers, and you won't have any secrets." INFOSEC has all the secrets of US national security as well as the secrets of NATO and those of our allies around the world to protect. The responsibilities are awesome. On a good day for INFOSEC, the externals are placid, but make no mistake, the internals are boiling. That's the excitement of INFOSEC.

The internal story is unknown, and it must necessarily remain so to the outside world. It is possible, however, to give some appreciation of the scope of the INFOSEC task with respect to the various elements, each fascinating in its own right, which collectively must be integrated into the total security pattern which constitutes INFOSEC.

Consider the challenge to the cryptomathematician: Design a cryptoalgorithm to encrypt our most sensitive secrets, and having encrypted

them, we will give the resulting text to our most mathematically and technically sophisticated opponents and let them subject it to their most high-tech attacks. It must protect the information for decades against such continuous attack. That's not all. It must do this under the assumption that the opponent has the algorithm but not the key.

To cryptoequipment engineers we say, "Embody the algorithm in an equipment that is fail-safe," and to the evaluator we say, "Analyze the algorithm and the cryptoequipment that contains it and give it a seal of approval." Impossible as it seems, it is a task that must be coped with successfully if we are to have the ability to securely command and control our forces and to protect our strategic interests.

There are many situations, particularly in tactical operations, where valuable information can be derived, not by breaking the encrypted transmissions but by analyzing the stereotypic formats, the quantitative message data, and other externals. The task of protecting against such exploitation is the domain of transmission security. This is an entirely different type of challenge, the searching for seeming minutiae that could actually be a bonanza to hostile intelligence services.

One aspect of this, or for some an INFOSEC category of its own, is providing secure sequences for ECCM transmissions that are secure against enemy analytical reconstruction.

Hostile intelligence operations can concentrate on the attack mechanisms of their choosing. The job of INFOSEC is to protect against practical attacks. A technically pervasive phenomenon, a known physical fact of life, is that electronic and electro-mechanical equipment when processing information necessarily create emanations which can be detected if not protected against. TEMPEST is the field of INFOSEC devoted to the protection against unwanted, unintentional, compromising emanations. The technical challenge to determine how best to detect such emanations, to identify those that may be compromising, and then to devise corrective measures is complex. However, the real challenge is how to determine the cost-effective compromise. At what point have we made such an attack unprofitable? INFOSEC is always involved in optimization trade-offs, but it is a two-party game of exceedingly high stakes.

TEAPOT is a recently coined terms for another aspect of the compromising emanations problem. The difference between it and TEMPEST is that the emanations are hostilely induced by "bugs" planted in the equipment. In the TEA-POT category is the widely publicized GUN-MAN operation of the recent past, a rare case of our sharing the excitement with the outside world. In the GUNMAN operation we removed tons of equipment from our Moscow embassy and replaced it with clean equipment in one rapid move before the Soviets could react.

Physical security in INFOSEC includes the protection of the cryptomaterials: the codes, ciphers, cryptologics, keys, cryptoequipment. When you consider the high value the Soviets place on the acquisition of our cryptomaterial, coupled with the vast amount of codes, ciphers and keys in hard copy form around the world, you can appreciate the enormous size of the this task. If protection breaks down, security breaks down. That is why INFOSEC is a top priority target of the Soviet espionage apparatus.

Personnel security goes hand in glove with physical security since it is this route by which physical security is often attacked. There are no stricter security constraints and checks on any personnel in the U.S. Government than on those working at the heart of INFOSEC.

The rapid expansion of computers and the field of information processing has enormously complicated the qualitative and quantitative problem of protecting classified and sensitive information, and at an exponentially increasing rate. The previously discussed aspects of INFOSEC, as complex and challenging as they may be, have trend lines and data bases helpful in planning. If Communications Security, (COMSEC), is in a state of combustion, COMPUSEC is in a state of explosion. Harnessing an explosion can be almost too exciting. Again, it is a game, a deadly two party game with extraordinarily high stakes. It is vital to know the enemy capabilities if we are to be successful in countering them effectively. That is the field of threat analysis. Doctrine provides the procedural and regulatory sinew binding the INFOSEC capabilities into a coherent body. The production of literally mountains of codes, ciphers, keys, and other crypto-material with the utmost of security and accuracy is fundamentally important to an effective INFOSEC program. Each of these areas of INFOSEC is a story in itself.

INFOSEC is not, of course, an end in itself. It is only useful when applied in communications and electronics systems. This opens another whole dimension to the scope of INFOSEC. It is absolutely essential for both systems security and for the effectiveness and efficiency of the systems into which it is integrated that the INFOSEC professionals not only fully understand the technology of those systems, but also the operations those systems are supporting. Thus, INFOSEC professionals are spread throughout the world in a wide variety of roles. Take one example: Imagine the situation where a satellite launch is on hold for some unidentified technical problem and your equipment is the only Government Furnished Equipment in the whole system. Now that's real excitement.

And now the final INFOSEC role, systems security evaluation. In accomplishing this task, all the above discussed areas and their complex interactions must considered. Coupled with this must be the consideration technology, the varied environments and the wide range of applications, and the ever-present hostile threat. This must be integrated, assessed, and a determination made to give the seal of approval to a system, "OK to pass our nation's most vital secrets in this system." The pressures on INFOSEC are great, the task seemingly impossible, and the external recognition and rewards necessarily almost nil. But balancing all that out is the EXCITEMENT OF INFOSEC.

How to build a USER-SEDUCTIVE Architecture

W45

P.L. 86-36
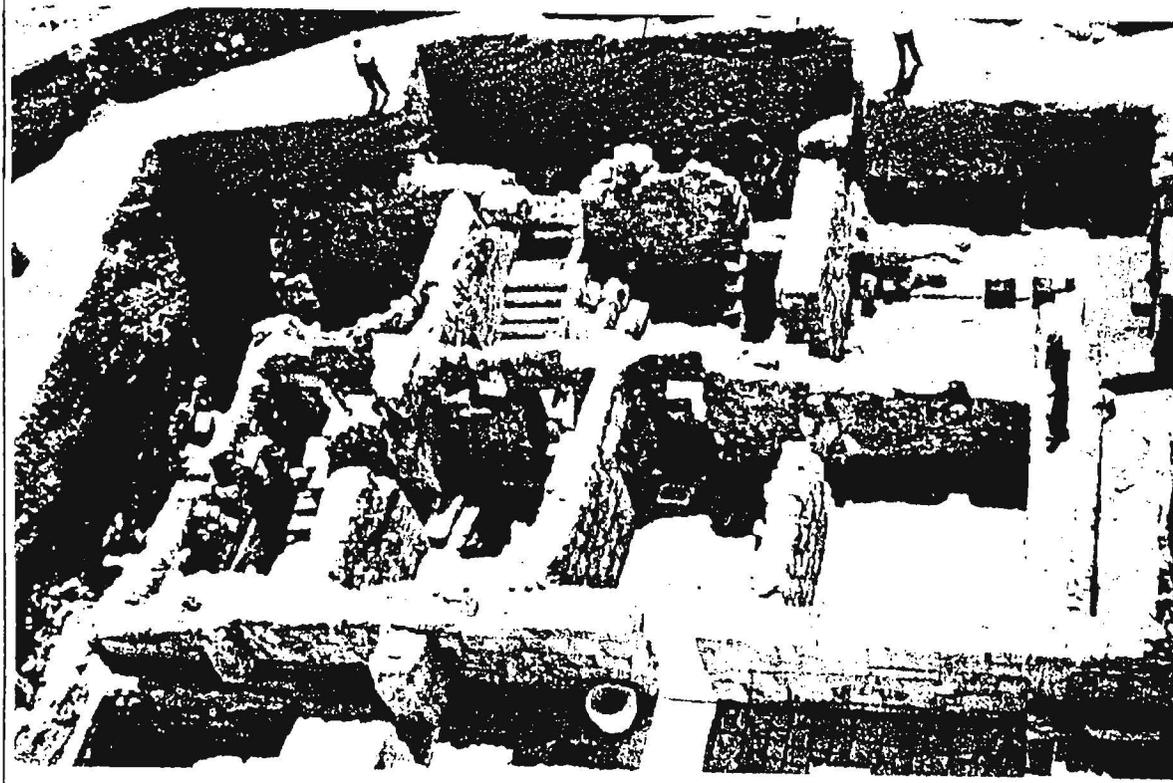
(U) By integrating off-the-shelf software, some standard Agency hardware, and programs already developed by the R Directorate, we developed the basis for the OPELINT Architecture that can be used at any level within the ELINT community. It has new ideas, capabilities, and opportunities in which analysts and programmers can be imaginative, creative, and explore new ideas and techniques. We incorporated into the Architecture some computer models that give program managers, decision makers, and analysts the capability to develop different collection and processing strategies, along with high quality computer simulations that give visual results of their decisions. The system is based on high quality graphics designed for easy use, with concealed text to explain only those areas where users might need more detailed information.

(U) We have been able to do this without a multi-million dollar budget and without having to purchase large mainframe computers or write massive programs or arrange large scale contracts.

(U) The process seemed simple at the outset. All we had to do was gather information on all the collection and processing systems that provide operational ELINT, put it together, and print out a nice document (that would probably end up being cumbersome, out of date before the end of the week, and whose only useful function would be as a doorstop). As we were collecting information two things quickly became evident: (1) most information was hopelessly out of date (some as much as 6 years); and (2) the scope of the Architecture meant that we were inundated with information.

P.L. 86-36

We plan to develop a system whereby the user can provide information directly to the data base and, after review and formatting, updating will be automatic. This will give the users the opportunity to be involved in the development and maintenance of the overall system.

(U) Then W45 was faced with a conundrum: what Architecture do you use to create an Architecture?
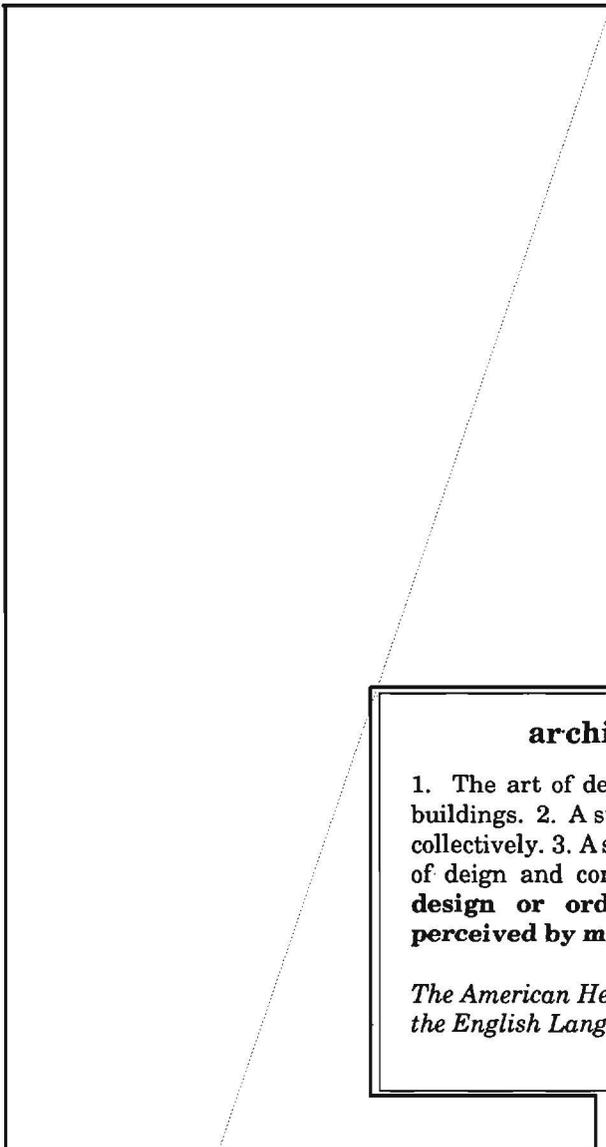
(U) We can anticipate shortfalls by comparing current requirements with new target trends. We then help planners and programmers make smart decisions about eliminating the shortfall. Their decisions then become a new plan that is factored back into the original equation either as a new capability or as a future plan.

(C) The primary purpose of an operational ELINT Architecture is to provide a baseline from which decision makers can make those smart choices.

(U) Once we determined what we would include in the Architecture, we had to decide on what form we would use to present it. This became a key issue because the Architecture will be used worldwide—at different levels of command and by people making decisions on various types of ELINT systems. The mass of information collected to date indicated that the finished product would be several large volumes of printed material. Since the ELINT community is a vast network of commands and organizations, writing, printing, and distributing hardcopy documentation would be a major effort, similar to publishing a full-length book. Instead, we decided to place the Architecture online and provide access through various means; printed versions of the documentation would be limited to a master copy and several hardcopies for the archives.

---

architecture

1. The art of designing and erecting buildings. 2. A structure of structures collectively. 3. A style style and method of deign and construction. **4. Any design or orderly arrangement perceived by man.**

The American Heritage Dictionary of the English Language

(U) A useful by-product of the baseline Architectures has been the creation of an ELINT Glossary. As we culled information for the baselines, we published ELINT terminology and acronyms in a hardcopy document. We are now looking for a way to make an interactive glossary for our online version of the Architecture.

(U) Once the objective A.rchitectures are developed, we must update them whenever there is a change.

CONFIDENTIAL

(C) To provide guidance in the areas of computer modeling, interactive glossary development, and database establishmen,t W45 purchased the services of a special engineering technical assistance contractor. The contractor may be able to use or adapt some computer models already on hand or under development by R08 and R3, thus reducing costs, saving time, and assuring compatability with other programs.

(C) The final area in which we want to concentrate is presentation. Once the ELINT baseline is finished and modeling is under development, we need to organize the information to prepare briefings, write new plans, and help document new programs. For this purpose we are purchasing software that will provide high quality graphic simulations of ELINT systems and run on an Agency standard high-performance workstation. The system will output the graphics as a composite video signal. By creating and editing video tapes, we hope to develop a library of simulations that can be combined with still graphics for high quality presentations.

(U) The process seemed simple enough to us, and we planned also to develop a floppy-diskette version that for tactical units in the field.

(U) Our real problems began when it was pointed out that while our idea was good, there were problems with printing in hard copy the diagrams and charts essential to the understanding of ELINT: flow charts, organizational diagrams, mission tracks, and even hardware configurations. Once again the hunt began for off-the-shelf software. Instead of incorporating diagrams with the Architecture, we are now developing a graphics-based program that allows the users to call up concealed or suppressed textual explanations of each graphic. In addition, we are developing a program that will allow the users to capture a series of graphics and assemble them in any order they wish for briefings and presentations.

(U) All of this software currently operates on an IBM or IBM-compatible AT with an enhanced graphics capability. In the next phase, when we have found an imaginative programmer to do it, we will convert the Architecture to run on the Agency's high-performance workstation.

(C) Our development of an OPELINT Architecture does not end with the establishment of a baseline. We now must find ways to make it compatible with existing Architectures, develop a program to add an interactive glossary, develop computer models, and look at ways to use the Architecture to help planners and decision-makers better understand ELINT.

(U) In summary, the Architecture is based on three ideas: (1) the development of an operational ELINT baseline and objective Architectures; (2) the creation of computer models operating on a community-compatible data base; and, (3) the development of computer simulations that can be combined with other visual material for demonstrating new concepts to decision makers and fiscal authorities.

(C) Although far from finished, the OPELINT Architecture has already begun to pay off for W4. We have been able to use the information developed to create a comprehensive Operational ELINT Program Plan, and we have opened some new avenues in the areas of graphic presentations for operational uses. There are still some areas we need to explore and that offer opportunities and challenges for imaginative analysts, programmers and planners.

(U) It is my hope this article not only aids others in their efforts to develop an Architecture, but also elicits responses from those of you who are working on existing programs that may aid us in the execution of ours.

```
New Address for CRYPTOLOG

P1
NORTH
```

GOLDEN OLDIE

REFLECTIONS ON SELECTED NILOTIC LANGUAGES

Nuer, which is spoken along the upper White Nile,

Has more complicated morphophonemic alternations than there are theories about Mona Lisa's smile.

And it may well be that they can be analyzed in terms of a pattern,

But, if so, the system is about as complex as the control mechanism of a space vehicle designed to travel to Venus by way of Saturn.

It is also possible that Shilluk is the language I have been looking for, which is neither a contour nor a discrete level nor a terraced level but a split level tone language,

But in any case I am very glad that Shilluk is not my own language,

Because if it were I would be wandering around naked with a spear in one hand and a nail through my lower lip for decoration,

A complex of cultural traits of which my wife would take a dim view because she is not fond of lethal weapons and is sure that a nail through the lower lip would seriously interfere with expectoration.

And although I have never heard Dho-Luo, I am extremely dubious about the feasibility of developing competence in it with the lesson manual entitled "Dho-Luo Without Tears,"

Because I have never heard any Nilotic language that isn't at least as hard as saying "Peter Piper picked a peck of pickled peppers" after seven beers.

In fact, my sentiments on another can be summed up by the simple observation that Dinka

Is a stinka.

In short, although these and other languages of the Sudan and Uganda and Kenya are indisputably neglected,

In the light of all the evidence I have collected,

I am not just before proposing a government-supported Language and Area Center specializing in Nilotic,

Because even I have a feeling, possibly brought on by advancing years, that there is such as thing as being too ruddy exotic.

March 14, 1961

P.L. 86-36

BELATED THANKS

U 25, 26.
(Oceangoing type.)

U 37—44.
(Oceangoing type.)

U 77, 78
(Oceangoing type.)

P13D

For most of us us Sigint-oriented types at NSA, appreciation for the many-faceted art of cryptanalysis is something that just seems to go with the turf. We *have* to give it lip-service, at least, since after all we do work here. Well, the other day, while cleaning out my desk (preparatory to getting ready to retire) something happened to drastically change my perspective. I came across an old copy of *Cryptologic Spectrum* (Winter, 1978) which featured a series of fascinating articles on "Ultra and the Battle of the Atlantic." *More* than fascinating to me, because one of them, called "The German View" by Prof. Jurgen Rohwer, made me realize that C/A might very well have been the instrument for saving my own life—and those of a number of other nervous GIs aboard our troop ship—way back in World War II days, in February 1943. Suddenly, in flashback, I experienced a visceral appreciation of what the art of cryptanalysis could really do, worked out on a scale of life and death.

Looking back on the Battle of the Atlantic from my own perspective aboard a tiny Liberty ship, I can imagine few situations better calculated to instill a feeling of Man Against the Elements. Add to that the gnawing fear of *das Boot*--the dread of a torpedo coming from somewhere out there. We were heading northeast in early February on a convoy out of Boston, both seasick and scared, en route for Iceland. We heard that a previous convoy had been hit—the one in which the *Dorchester* (with the three heroic chaplains aboard) had gone down. Reading Mr. Rohwer's article, I discovered something none of us knew at the time: Bletchley Park's OIC (Operational Intelligence Centre)—which we of course had never heard of!—had begun to make inroads on the German Naval Triton cipher, after a bad six months during the second half of '42 when our people had a decrypting blackout and the U-boats were intercepting about one-third of the running convoys. At last, as he says, "Bletchley Part succeeded...in solving the main problems of the Triton cipher circuit... (and) from the 5th to 28th February the traffic was decrypted with seldom more than 24 hours time lag." A comforting thought—just in time to protect *our* convoy...

These C/A successes meant that the British Submarine Tracking Room at the OIC could get warnings out to reroute the convoys and keep us away from the U-boat patrol lines. If we had known this then, a lot of us aboard the tiny USS *Chateau Thierry* (barely 10,000 miserable tons in size) would certainly have slept better. But there was still a problem with the wolfpacks, according to Prof. Rohwer: at this time, the German *B-Dienst* (the counterpart of the OIC at Bletchley Park) was also having its most successful period: "it could decrypt many of the routing and re-routing signals and many of the U-boat situation reports sent daily by the Admiralty or COM-INCH (Commander-in-Chief, Tenth Fleet)..." He goes on to say that even if only 6 to 10% of the German decrypts were done in time to be of operational value, they could have "important

consequences." He then cites February 1943 as a good example of this—the fateful time when the Battle of the Atlantic was beginning to come to a climax, and when our own little troop ship was riding into it.

Prof. Rohwer cites a certain convoy HX.228, heading northeast, about which the *B-Dienst* had decrypted a position report. (Might this have been us? It was the right time and direction.) At this point Admiral Dönitz's people ordered the wolfpack *Ritter* to move north to find them (or us!). Fortunately, our heroes at Bletchley Park decrypt the order—God bless 'em—and the convoy gets an order to move north of the *Ritter* patrol line. Another piece of luck for HX.228— the U-boat command about this same time intercepted and decrypted position reports on two other convoys.ON.166 and ON.167--heading west, away from our convoy. Group *Ritter* was thereupon ordered to change course and set up a new patrol line along longitude 30° West to cover the expected passsing routes of those two other convoys—as I interpreted Prof. Rohwer's account.

What followed must have been one of the most dramatic scenarios that fate could ever devise—more gripping than anything Hollywood could have dreamt up—move and countermove —as Bletchley Part and the U-boat command decrypted each other's messages, the hunted and the hunters maneuvering for position in a life-and-death chase through leaden seas.   Finally it came to a grim conclusion. Prof. Rohwer reports: "In a fierce convoy battle which lasted for six days and covered 1100 miles the U-boats sank 15 ships and lost two of their number to the counter-attacks of the escorts."

Many died, many were spared. Looking back, I wish I knew for sure whether HX.228 was really *my* convoy. I do know I want to look for more clues in Patrick Beesly's much-praised book, *Very Special Intelligence. The Story ofthe Admiralty's Operational Intelligence Centre 1939-1945*, published in 1977. (Mr. Beesly was deputy chief of the Submarine Tracking Room of the OIC at that time.) Anyway, regardless of the exact role that C/A played in getting our own particular convoy to Reykjavik safe, sound (and hungry), I'm now totally convinced that innumerable Allied lives were saved because of the hard work of the people at Bletchley Park, OP-20-G (the US counterpart of the OIC), and

in other places where the cryppy's trade was relentlessly pursued during those trying days. So, 46 years after the fact, I send belated kudos to all of them from a grateful, non-cryppy survivor. In normal times, we can only dimly surmise how our lives are daily affected by unknown, unseen people and events. But in times of war, how much more dramatic and powerful, if the curtain is lifted and we catch a glimpse of it.   (And, in the case of my own particular convoy, who knows but that some strange quirk of fate might have decreed that a decryption by our adversary, the *B-Dienst*, could become a strangely crucial factor in getting the wolfpack off *our* trail and on to someone else's? If so, perhaps I could even say, somewhat selfishly: *Vielen Dank* to you too, our one-time enemy.)   On the basis of my own experience, relived retrospectively, I can now agree with Prof. Rohwer on the very deepest level, when he says: "... without the work of many unknown experts at Bletchley Park, the turning point of the Battle of the Atlantic would not have come as it did...but months, perhaps many months, later."

Solution to:

NSA-Crostic #68
1st Issue 1989

by [   ]

[Finally, and perhaps most important,] translators must learn to unleash their minds. Rote translation works for some texts all of the time and all texts some of the time. It is at this juncture that creativity [- - the choice of right phrase or word in English to match the thought and flavor of the original --] becomes crucial.

[   ] "The Creative Translator," *Collected Articles on Translation*, 1973

SECRET SPOKE

CAMBODIA IN PEACE AND WAR

P.L. 86-36

КАМБОДЖА
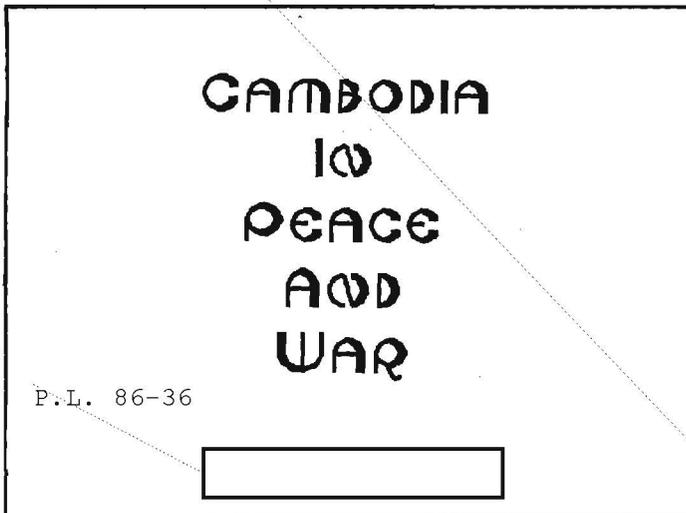
(S-CCO) After spending over 12 years on the Vietnamese Communist (VC) target, I had the opportunity in July 1968 to assume the duties of Chief of the Cambodian Branch in what was then B12. I took on this assignment because I saw it as both a career-enhancing job and an opportunity to work in a new environment. But I can remember my former Division Chief asking shortly after I took on this assignment, "How are things in Sleepy Hollow?"

(S-CCO) I must say that in July 1968 the Cambodian problem could not be compared in any way with the Vietnamese problem. With the war going on in Vietnam and the US involvement steadily increasing, Cambodia (also known as Khmer) was not a target of importance. [   ] brought about a great change in Cambodia in the way the target was handled here at NSA and in the Intelligence Community as a whole. The following is a description of this change and the developments we had to contend with to support our customers on a real-time basis.

(S-CCO) My main intent in writing this article is to provide information that might be of value to production organizations which may have to face a similar target change in the future. I dedicate it to all the men and women who worked on the Cambodian problem, both at NSA and overseas, and willingly put in many hours of their own time to ensure that the job was accomplished in a quick and thorough manner. I will not mention any of them by name simply to avoid leaving anyone out.

SUMMARY OF THE EFFORT

(S-CCO) In July 1968, the Cambodian Branch consisted of three teams [   ] most of whom were civilians. This included Traffic Analytic, Linguistic/Reporting and ADP Support Teams. Most of our effort was concentrated on the analysis, processing and reporting of Cambodian Armed Forces (FANK) communications.

(S-CCO) Cambodia at that time enjoyed a low priority, and in many cases hard copy reports satisfied the requirements. These were annual disposition of forces/order of battle reports and periodic reporting on political and military activities and changes in the military organization and equipment. Electrical reports were issued by field elements (USM-626, Saigon, South Vietnam; [   ] and U.S. Technical Research Ships) as well as by B12.

(SC) The only significant requirements related to the use of Cambodian territory by the VC forces, either as safe havens or for the infiltration of supplies, equipment, and personnel through Cambodia into South Vietnam.

SECRET SPOKE

P.L. 86-36
EO 1.4.(c)

any evidence to support this theory and reported electrically on any indications of Vietnamese forces operating in Cambodian territory.

(SC) Also, there was relatively high interest in Cambodian insurgent activities. Three insurgent groups, the Khmer Serie, Khmer Leou and Khmer Rouge (Cambodian Communists) were active at the time, primarily in the Cambodia-Thailand border areas. We often reported on FANK operations and activities of insurgent elements. None of these insurgent groups appeared to be of any great threat to the Cambodian government or its armed forces.

(S-CCO) In my opinion the FANK communications network was a relatively easy target to collect, process, and analyze. Most communications were in HF manual Morse
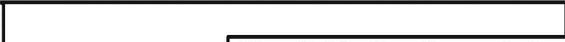
The field sites forwarded most traffic by courier except that which was encrypted or they felt was possibly important enough to meet product reporting requirements.

TRAFFIC ANALYSIS

(S-CCO) The traffic analysts working on FANK communications performed the normal tasks expected of them by maintaining continuity on all targets of interest, preparing net diagrams and message/address group logs, and providing technical support to field elements via technical messages and hard copy working aids and COMINT Technical Reports.

(SC) The FANK did not use signal plans as sophisticated as those used by the VC. Most HF communication groups were simplex stars using day and night frequencies, international procedure signals and French language operator chatter. They often compromised address groups used in message preambles and occasionally in operators chatter and gave advance warning of callsign/frequency changes. This coupled with the fact that we were eventually able to obtain, through the Defense Attaches Office (DAO) in Phnom Penh, copies of the FANK overall signal plan, made the traffic analysts task relatively easy. Our only problem was dealing with the large volume of material.

(S-CCO) Examples of the FANK message preambles and HF communications groups are shown in Figure 1.

CRYPTANALYSIS AND LANGUAGE EXPLOITATION

(S-CCO) Approximately ninety-five percent of the correspondence passed over FANK communications was French plaintext messages. A small amount of Cambodian language material was noted as well as some encrypted traffic. We were able to exploit all traffic with except that which was badly garbled.

(S-CCO) The Linguistic/Reporting Team consisted of about ten French linguists and a couple of Cambodian linguists. They were able to scan and translate messages of significance and flag them for inclusion in a machine data base for subsequent reporting in hard-copy reports. The messages meeting timely reporting requirements appeared in electrical summary reports. The messages put into the machine data base were flagged by subject so that a pull could be made on them when a hard copy report on a specific subject was in preparation.

(S-CCO) The FANK messages were normally less than a hundred groups in length and generally limited to one subject. This made it fairly easy to categorize the messages and identify the originators and addressees. Often specific addresssees were spelled out in the preambles and/or the beginning of the message texts. So, just as with the traffic analysts' job, the linguists/reporters had a relatively easy but heavy work load.

THE KHMER COMMUNIST PROBLEM

(S-CCO) The development of the Khmer Communist (KC) problem was a very gradual one that took place over about a two-year period. By early 1969, we began seeing reference in FANK communications to increased KC activities in scattered areas of Cambodia. This, supported by collateral information, indicated that the KC were growing much stronger and were probably being assisted by the VC. Extensive collection efforts were undertaken to intercept KC communications which we suspected would be HF manual Morse and HF/LVHF radiotelephone.
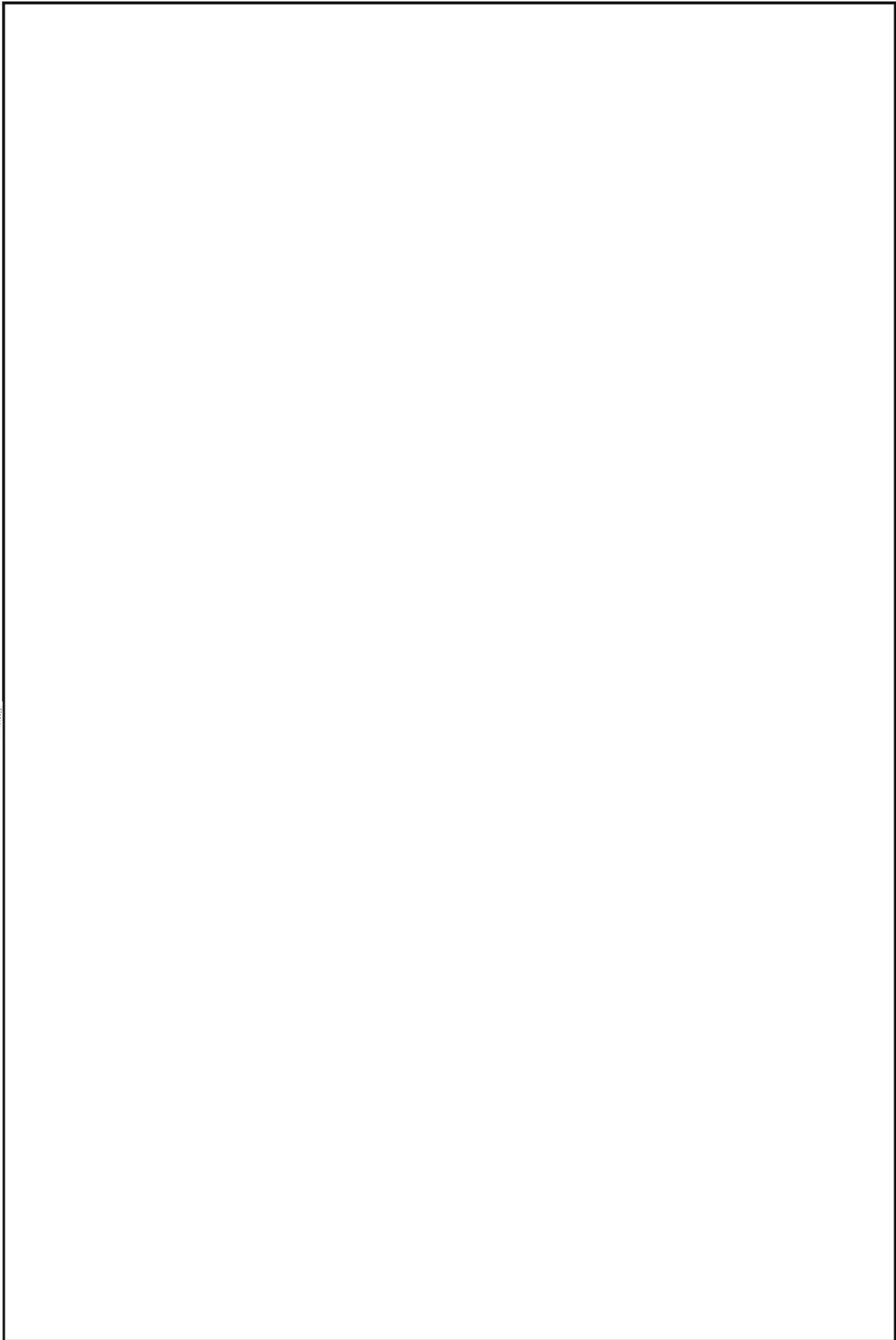
SECRET SPOKE

(S-CCO) By early 1970 we had not been able to identify any KC communications. We suspected that if the KC were being assisted by the VC, the Vietnamese would provide them with necessary radio equipment and probably signal plan and cryptographic training and materials, and so the KC communications characteristics and procedures would be similar to that used by the VC. Eventually this proved to be the case.

(SC) In late 1969, while conducting general search efforts at Udorn, Thailand, USM-7 intercepted an HF manual Morse link initially identified as VC.

(S-CCO) Because of the signal plan used on this link [   ] we tried to have these communications worked in the Vietnamese Office. Some even thought that the link was actually down-working of a VC communications group which had most of its stations located in Cambodian territory. Further, some believed that the VC were running the show in Cambodia and that this communications group served their command and control apparatus in Cambodia. Later we concluded that the VC group probably served the VC cadre who were helping the KC establish their infrastructure and were providing communications and technical assistance. Nevertheless, this initially proved to be a stumbling block in our attempts to move on with the development of the KC communications network.

(S-CCO) By this time, interest in Cambodia had increased significantly as the KC military and political arms started putting more and more pressure on the Cambodian Government and its armed forces. Intelligence requirements changed drastically, with high emphasis being placed on obtaining information on KC activities. Meanwhile, the intelligence requirements on FANK communications lessened as the US increased aid to Cambodia and a larger contingent of US personnel was assigned to Cambodia. They were to help train FANK personnel and develop a more effective intelligence gathering operation and expedite the flow of intelligence information to FANK units involved in day-to-day combat operations.

(S-CCO) During the next two years, we found ourselves having to deal with a number of problems besides the isolation and development of the KC communications network. The traffic analytic work force had to be trained to do more detailed analysis on the more complex and sophisticated KC communications; the language pool had to be changed from primarily French to a Cambodian cadre; and those involved in preparing product reports (other than translations) had to shift gears to produce more timely reports and, as with the traffic analysts, learn to deal with a target that was not going to reveal as much in message externals and text as the FANK did.

(S-CCO) During 1970-1972 one of our primary goals was resolving our language problem. This, as you can imagine, was not an easy task since good Cambodian linguists couldn't be hired off the street. The Service Cryptologic Elements were our initial sources of linguists and both INSCOM and NSG were able to provide assistance over a period of time. They adjusted their language training program to train more Cambodian and fewer French linguists and the Language Intern Panel was able, eventually, to provide us with one linguist.

(S-CCO) By late 1972, our Cambodian language pool had grown [   ] which included a number of military personnel. We were able to hire some of these military personnel after they had satisfied their military commitment. Fortunately, this was back in the days when a military person could clear the building on a Friday and come back as a civilian the next Monday.

(S-CCO) During the next couple of years, the Cambodian Branch was able to meet all time-sensitive reporting requirements and at the same time develop a mean and lean work force. By early 1974, the branch consisted of three teams [   ] Traffic Analysis,

SECRET SPOKE

Language Exploitation, and Reporting. The three teams had to work closely together to ensure that all communications were collected, processed, analyzed, and reported on as quickly as possible.

(S-CCO) Meanwhile, an extensive collection and analytic effort had also been established at USM-7, Udorn, Thailand, where the bulk of the KC communications were being collected. Every effort was being made to give the field the capability to perform first instance reporting on the target. However, because of language difficulties, only a limited reporting effort could be maintained. Thus, most of the reporting effort fell on the Cambodian Branch at NSA. The efforts involved in this process are described in the following paragraphs.

TRAFFIC ANALYSIS

(SC) As I indicated earlier, the traffic analysts working on the Cambodian problem had to significantly alter their method of analysis in order to contend with the KC communications. The KC communications groups consisted of two or more complex links which changed callsigns and frequencies on an ad hoc basis. Fortunately not all links changed at the same time, so it was possible to concentrate on certain areas following signal plan changes.

(SC) The KC HF manual Morse network grew gradually as the KC established their infrastructure in different parts of the country. This allowed us to develop the communications while teaching the analysts to perform more detailed analysis. Since the KC used complex links, the intercept operators often intercepted only one end of a link; in many cases they didn't have the capability to copy both ends at once. So, it was a case of searching through the unidentified material we received in the Cambodian Branch, as well as material held in the unidentified files in the Vietnamese Office to find backlink activity. A time-consuming task, but often profitable. Through up- and down-time matches and message servicing activity, we were able to recover many backlinks and aid the intercept operators in collecting those stations passing message traffic.

(SC) Because all KC traffic was readily exploitable, the traffic analysts thought that their task would be relatively simple as it had been with FANK communications. However, this proved not to be the case. The KC used

cover names and designators in the message texts for originators and addressees of messages as well as for political and military organizations. The analysts had to be very careful when using the internal address information to co-locate terminals because it was often difficult to determine if a message was being sent by one control station to two or more subordinates or if the message was originated by a senior authority and being relayed to a second- or third-level subordinate. The KC did not use broadcast communications to send a message to two or more subordinates; the same message was sent separately to each.

(S-CCO) Another interesting traffic analytic problem developed a few months after the first KC link was isolated. Also, different message preambles containing what appeared to be routing indicators started to be used on messages passed over the link. The traffic analysts as well as the linguists were miffed by this development and it wasn't until we obtained some HF direction finding (DF) results that we started to get a hint at what we were dealing with.

SECRET SPOKE

(S-CCO) By the end of the war, the traffic analysts had gained a lot of experience by working on a more difficult problem and learned how to use a variety of techniques to assist in their analysis. This included not only HFDF results, but airborne DF (ARDF) as well. ARDF missions flown by US Army platforms over South Vietnam were able to provide accurate fixes on many of the KC targets while they flew missions along the Cambodian border. In addition, they were able to gain some experience in working on KC VHF/LVHF voice communications which were occasionally intercepted by airborne platforms. Unfortunately, these communications were very difficult to intercept because of the low power and the type of equipment used by the KC, and never really provided any meaningful intelligence information.

(S-CCO) Figure 2 shows an example of a KC message preamble and HF manual Morse communications group; note the difference between the KC and the FANK practises.

LANGUAGE EXPLOITATION

(SC) As we became more familiar with chart usage and patterns noted in the chart internals, it was possible to develop   process which not only did not help, but it hindered our language exploitation efforts.

(SC)   the linguists had developed romanized equivalents for the Cambodian script characters, patterning them after the romanized Cambodian we occasionally noted in FANK Cambodian plaintext messages.

the new linguists had great difficulty.

(S-CCO) All of the new linguists had been trained in Cambodian using materials such as newspapers and books in Cambodian script and were taught to use Cambodia dictionaries written in that scrip, for Romanized Cambodian was not used in non-SIGINT material. Thus, to   one of the first things they had to do was to learn the romanized equivalents. This was time-consuming, the more so because before these new linguists could translate the message, they converted the romanized text back to Cambodian script by hand. Another key problem was garbles that often appeared in intercepted material. Even if there were only a few in a particular message, the new linguists were unable to render the text readable.

(S-CCO) Another thing the linguists had to deal with was the different vocabulary used by the KC. Besides using words not normally seen in FANK communications, they also used a large number of abbreviations, cover terms and cover names for persons, places and things. The linguists had to become familiar with all these before they could produce meaningful translations.

(SC) To further complicate the linguists' task there was the problem of part messages. KC messages were long, sometimes more than 1000 groups.   Moreover, unlike the FANK, the KC might cover a number of subjects in one message. Thus, the linguists would have to determine the right order of the message parts before they could translate the message in the right context.

(SC) The linguists first scanned the cleaned up message for any information that met timely reporting requirements. In most cases this was tactical information which provided data on KC attack plans or unit locations. Normally, they prepared a short narrative SPOT Report or CRITIC providing the information to the DAO in Phnom Penh and our other customers, then followed it up as soon as possible with a full text translation. The Defense Attaché Office in Phnom Penh had a standing requirement for full-text translations of all messages containing
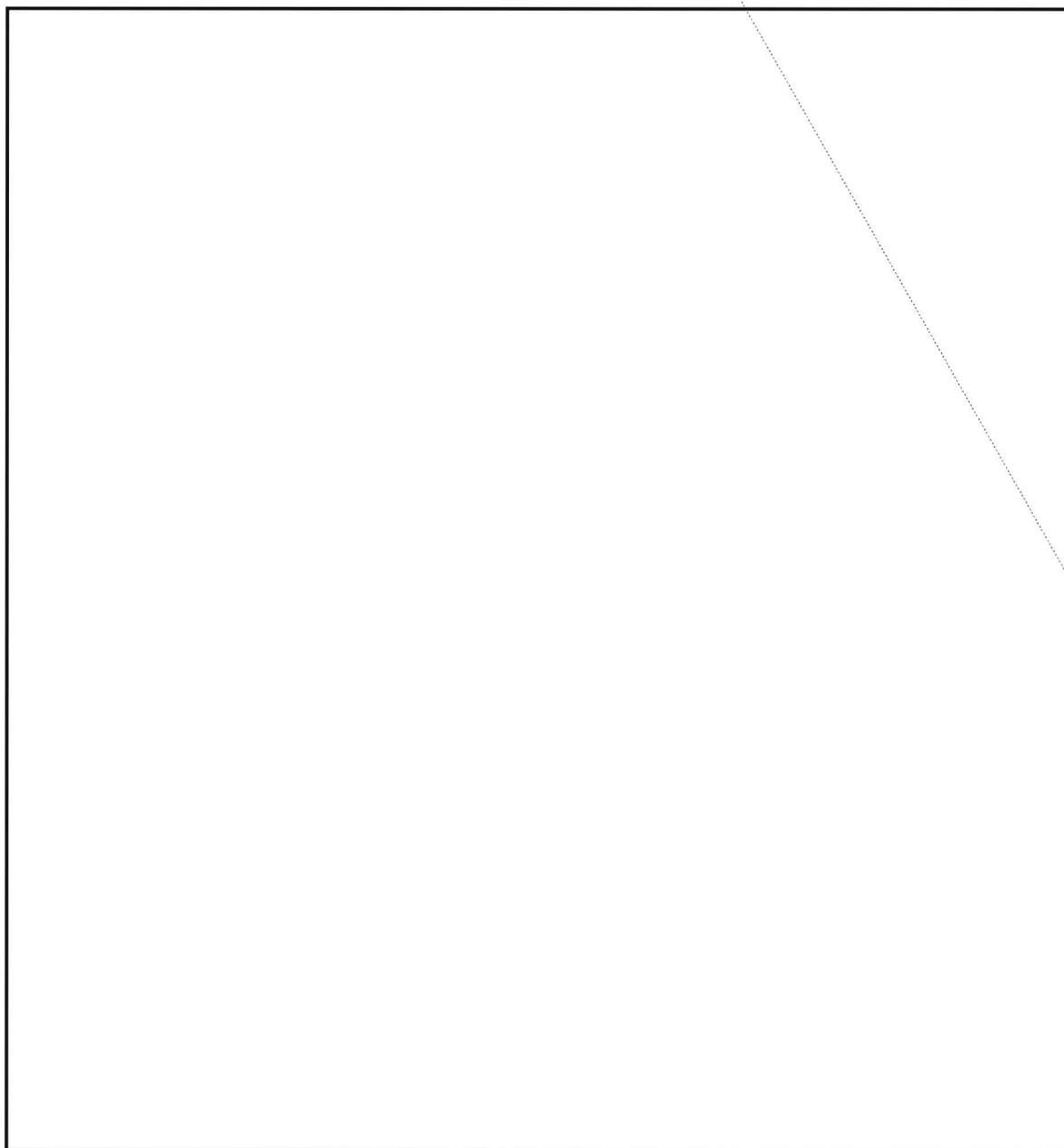
any intelligence information of value to them. The DAO then sanitized the information and forwarded it to the appropriate FANK authorities for the necessary military action.
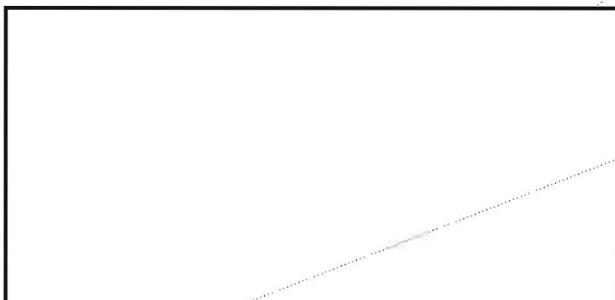
(S-CCO) Besides having to contend with all of the above, the Language Team also had to provide on-the-job training for Cambodian linguists assigned to USM-7. Through early 1975, several groups of Army linguists spent three to

SECRET SPOKE

SECRET SPOKE 23 JUN '74

T13 5868

T13 5868

T13 5868

T13 5868

T13 5868

T13 5868

T13 5868

T13 5868

T1

5868

DDIR
ESS

ZCZCT13 5868AXA841   LOG LN NO. 066
P.L. 86-36

RR

DE   #0002 1800312

ZNY GWASH

ZKZK RR RNI DE

R 2803ICZ JUN 74

FM SSO PHNOM PENH

TO DIRNSA/B3 (CAMBODIA)

INFO:

ZEN

S E C R E T SPOKE CAM C88S INTEL

FROM AMBASSADOR JOHN GUNTHER DEAN, US EMBASSY, PHNOM PENH, CAMBODIA

ACT: B3
INFO: B
DDO
DIR

EO 1.4.(c)
P.L. 86-36

SECRET SPOKE

four months each in the Cambodian Branch prior to going overseas. Much time had to be spent with these linguists to get them to the point where they could adequately handle the KC traffic sent back to USM-7 Although they did their best in the

EO 1.4.(c)
P.L. 86-36

SUBJECT: MESSAGE OF APPRECIATION

IT IS ALL TOO SELDOM THAT PERSONNEL PERFORMING AN INDISPENSABLE TASK RECEIVE THE RECOGNITION RIGHTFULLY DESERVED. I WOULD LIKE TO TAKE THIS OPPORTUNITY TO EXPRESS MY PERSONAL APPRECIATION FOR THE OUTSTANDING INTELLIGENCE PRODUCT

PRODUCED BY THE CAMBODIAN SECTION OF B3. I CAN HONESTLY SAY THAT WITHOUT YOUR CONTRIBUTION, THIS MISSION-S EFFECTIVENESS WOULD BE GREATLY REDUCED. TO ILLUSTRATE THIS POINT, ONE MERELY HAS TO RECALL YOUR RECENT CRITIC SERIES MESSAGES OF 8 JUNE IN WHICH YOUR TIMELINESS MAY HAVE AVERTED A CATASTROPHE. IN ADDITION TO MY OWN PRESENCE, THE FANK GENERAL STAFF, THE REPUBLIC CABINET, AND THE ENTIRE DIPLOMATIC CORPS HAD BEEN SCHEDULED TO ATTEND CREMATION RITES FOR THE ASSASSINATED MINISTERS ON 9 JUN. AS A RESULT OF THE CRITIC, RITES WERE RESCHEDULED FOR 10 JUN UNDER MORE STRINGENT SECURITY CONDITIONS.

IN FULFILLING OUR GOVERNMENT'S OBJECTIVES REGARDING CAMBODIA-- A JUST AND LASTING PEACE-- YOU MAY TAKE SPECIAL PRIDE. YOUR ENDEAVORS HAVE DAILY DIVIDENDS IN THE FIELD AS WELL AS IN PHNOM PENH ITSELF.

THANK YOU FOR A TRULY PROFESSIONAL JOB.

XGDS-2

#0002

Page 2 of 2

field, they never reached a level of proficiency sufficient to handle first-instance reporting on the KC problem.

REPORTING

(S CCO) In addition to the SPOT Reports, CRITICS, and translations published by the linguists, it was also necessary to produce summary reports on KC activities. These reports dealt with message content as well as developments noted in communications. A separate team was established to report these activities in addition to the Traffic Analytic Team. One of the major problems we had early on was identifying the users of the KC radio network; reporters familiar with FANK practises provided specific information which made it relatively easy to identify radio station users. This was not the case with the KC.

(S CCO) I can remember many instances where we had long, tough discussions on whether we had enough information to at least tentatively identify a user on the basis of the structure of a communications group, the area it was located in, and the contents of message text. Because the KC seldom identified themselves as the FANK did, it was hard to get reporters to go along with less than "A" validity. But, after a while this became normal practice and to my knowledge the identifications we came up with were pretty much on the mark.

(S CCO) The summary reports prepared by the Reporting Team were done to satisfy customer requirements not requiring real-time reporting. This included periodic disposition of forces reports, special topic summaries and inputs for the Southeast Asia SIGINT Summary and the NSA SIGINT Summary. Often we were able to use collateral information to supplement what we obtained from SIGINT.

CONCLUSION

(S CCO) I hope after reading this article you will have a better appreciation of what the Cambodian Branch had to contend with during the war in Cambodia. I wish I could conclude with a statement that would give a happy ending to this story. But, unfortunately, the Lon Nol Government of Cambodia fell to the KC on 15 April 1975. We had given it our best effort, as I'm sure the US personnel in Cambodia had done.

(S CCO) We published thousands of pieces of product with information which should have been of value to the FANK authorities in countering KC activities. Unfortunately, there were often indications that when FANK forces were tipped off to an impending KC attack, instead of reinforcing their positions or conducting a counter attack, they would simply withdraw from the area to avoid a confrontation. Because of this and the constant heavy workload, we often felt frustrated with the job at hand.

(SC) However, there were times when our efforts were duly rewarded. One prime example was an event that occurred in early June 1974 shortly after a number of Cambodian Government officials were assassinated by the KC. A funeral ceremony was going to be held in Phnom Penh which the U.S. Ambassador, the FANK General Staff, the Cambodian Cabinet and the entire diplomatic corps were scheduled to attend. A KC message was intercepted indicating that they were going to conduct a rocket attack on Phnom Penh during the ceremony. We issued a CRITIC immediately upon recognition of the KC intent. As a result of this CRITIC, the ceremony was rescheduled and according to the words of the US Ambassador, "your timeliness may have averted a catastrophe." The Ambassador's message to DIRNSA on this subject is shown as Figure 3.

(S CCO) Needless to say, the job of working on the Cambodian problem during the war was very rewarding from many aspects, most important, the privilege to work with a fine group of people. During the period covered by this article, B Group had a few reorganizations which fortunately didn't affect the Cambodian Branch.. The division we were part of was moved between a number of offices and changed designators each time--B12, B65, and B33.

BULLETIN BOARD

UPDATE ON CAPULET

(FOUO) Individuals who have CAPULET software may obtain updated versions from [   ] B824, HQ 1A205, 963-5184. CAPULET is a suite of CA programs for the IBM XT or AT

P.L. 86-36
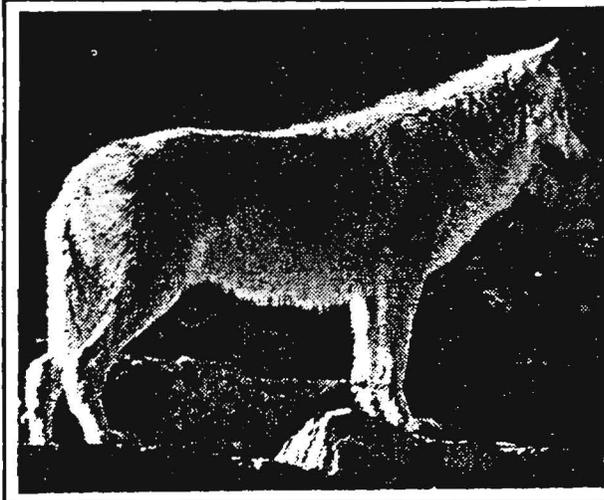
HARDWARE REVIEW

OCR DEVICES

P.L. 86-36

Reviewed by: [   ] *R8311*

(S-CCO) [   ] R831 has been testing Optical Character Recognition (OCR) devices to determine the feasibility of using such machines for producing digitized working aids [   ] [   ] Presented here are some principles of OCR, our test methodology, the results of our tests on three devices and of an accuracy test for several languages that was run on the best device.

ON OPTICAL CHARACTER RECOGNITION

(U) The technology which allows printed matter to be transformed into computer (digitized) format is called Optical Character Recognition (OCR). There are two kinds of OCR software. The first is dot matrix recognition which works by analyzing the location of dark and light spots on an image and matching them as closely as possible to a stored set of dots. A more sophisticated technology is called pattern matching, sometimes termed Intelligent Character Recognition (ICR), and relies on analyzing an image for its component parts, much as a human does when reading. Simply put, a scanning device looks at the pattern for each character of text, determines what this pattern represents and assigns a value to that pattern, generally using ASCII text codes. For example, when the scanner sees a vertical bar crossed at the top with a horizontal bar it interprets the pattern to be that for the upper case letter "T" and assigns the appropriate ASCII value to the image.

(U) The accuracy of such interpretations is further enhanced, to a limited degree, by lexicons and artificial intelligence rules which help determine if a particular sequence of characters is in fact a plausible combination. Artificial intelligence, however, is much less reliable than the human eye in viewing text as meaningful words. A computer, as well as a human, can tell just as confidently that in English the character sequence "happ." should end with the letter "y." But the human can go one step further and also accurately fill in the missing letters for the phrase "Happ. N.. Year!" while the computer, unless programmed with context-specific information, probably cannot. To a computer, words such as "Now", "Net", etc. are just as likely as the word "New."

(U) R831 purchased for testing three optical character readers of the many on the market. Although there are a number of different types of readers available, including hand-held models, desktop PC-based models and large stand-alone models, the selection criteria were geared primarily to the need to digitize multilingual publications, especially those having Cyrillic text. Given these considerations, the models selected were: Intelligent Optics Corporation's (IOC) SPEEDREADER, Kurzweil Computer Products' DISCOVER 7320 and Kurzweil's 4000 Intelligent Scanning System (ISS).

DESCRIPTIONS OF THE DEVICES

IOC SPEEDREADER

(U) The IOC SPEEDREADER is a desktop, pattern-matching page reader which can input text or graphics from a printed or typed page to a personal computer or word processor. Any

P.L. 86-36

alphabet of a typewritten or typeset format can be recognized in a variety of fonts, although point size is limited to the range 8 to 12. Scanning speed for high-quality text is approximately 30 characters per second for typeset and 40 characters per second for typewritten material. In the graphics scanning mode the SPEEDREADER scans either line art or continuous tone images. It has an automatic sheet-fed scanner which can be loaded with up to thirty pages of material. Files of longer than thirty pages can be created using an append facility. Minimum page size is 5.5 by 5.5 inches and maximum is 8.5 by 14 inches.

(U) Character recognition is achieved when the scanner compares the patterns appearing in the text to a font selected from the document processing software and then provides the appropriate ASCII text code. This font of patterns can be created by the user or chosen from a library of starter fonts contained in the software. Creating a user font, called "training," is a process which requires assigning an ASCII text code to each of the patterns appearing in the text. This procedure might only require 30 minutes or so for very simple text but can be considerably longer for text having multiple typefaces. The assigning of text codes to patterns is left to the discretion of the operator but is limited to two codes per pattern with a storage capacity of 256 patterns. The operator might find this to be insufficient if, as in the case of a document with several typefaces and/or alphabets, it is necessary to maintain typeface integrity in the output.

(U) In training SPEEDREADER, the user must take time to account for all the characters and symbols appearing in the text. This is important because if the scanner detects a pattern on which it has not been trained, the output text displays either an incorrect character or a special symbol indicating an unknown character. Since the SPEEDREADER has no capability to add to a trained font while scanning is in progress, in order to correct this problem the operator has to stop scanning, add the new character to the trained font and re-scan the text.

(U) SPEEDREADER's output is a text file including the control characters inserted by the word processing program. These control codes include hard and soft returns to preserve paragraph integrity but allow wordwrap functions, tabs, indents, centering, underlining, space compression, subscripts and superscripts.

SPEEDREADER supports most of the more popular word processing software packages on the market.

(U) All things considered, the SPEEDREADER, although the slowest of the three scanners tested, performs satisfactorily as long as the input text is of good quality and contrast, a caveat one could apply to all scanners. Despite the limited capacity for trained characters, and the lack of a concurrent editing function, the SPEEDREADER is a flexible device which does an adequate job under most circumstances.

DISCOVER

(U) DISCOVER 7320 is a PC-based, menu-driven, pattern-matching scanner. It reads both normal and landscape-oriented text in a variety of fonts ranging in size from 8 to 24 points. A maximum speed of 50 to 60 characters per second can be achieved, depending on the complexity of the character fonts and the quality of the scanned document. Software, containing system lexicons of approximately 50,000 words, is available for English, Spanish, Dutch, French, German, Italian or Swedish. In addition, user-defined lexicons of up to 10,000 words each can be created to supplement the system lexicon.

(U) DISCOVER is also an image scanner which can scan drawings and photographs with 15 levels of contrast control. There is a built-in sheet feed which will handle up to ten pages of material. DISCOVER can operate in a background mode under MS DOS thereby allowing document scanning and access to other software packages and PC functions simultaneously.

(U) Other features include an assurance threshold, which is a system confidence level telling the software how sure it must be of an identification before considering a text character as being recognized; a window option, which allows the operator to specify a zone whose contents should be captured when the page is scanned; and a column recognition function, which allows the system to analyze the pages for the specified number of columns and separate them into discrete zones.

(U) DISCOVER 7320 is probably the fastest scanner and the easiest to configure, but it is the least flexible. It is not trainable and can be used realistically only for monolingual

documents containing text in one of the preprogrammed languages. Also, even though the DISCOVER 7320 reads different typefaces, it gives no beginning- nor end-of-font markers in the output text to aid parsing. Nevertheless, if scanning requirements are such that these perceived drawbacks are not important, then DISCOVER 7320 can provide a reliable product

KURZWEIL 4000 ISS

(U) Kurzweil 4000 is a stand-alone pattern-matching scanner which can accommodate multiple fonts within a single document and read virtually any type font in sizes ranging from 6 to 24 points. Standard equipment features a display terminal, hard disk drive and a floppy disk drive used for file backup. Optional equipment includes an electronic tablet for document mark-up and a document feeder capable of holding up to 25 pages. Input can be either from bound books or individual pages measuring up to 11 by 14 inches. Separated pages can be placed individually on a glass surface directly over the optical scanning head or fed into the scanner by means of the document feeder. In the case of bound books, the pages must be placed on the scanning window. Depending on print quality and format of the original material, scanning speed is between 20 and 50 characters per second.

(U) The standard language configuration is for English but additional optional language packages are available for German, Italian, French, Dutch, Danish and Spanish. Language packages come with lexicons containing approximately 33,000 words. Since it is a trainable device it is also possible to program the 4000 to read other languages as long as the text is composed of discrete characters.

(U) Output is a file which can be transmitted to magnetic tape or hard disk, displayed on the terminal screen or transmitted to a peripheral device using appropriate communications software. The 4000 will also receive documents from asynchronous ports and magnetic tape machines. A more detailed examination of how the Kurzweil 4000 functions follows below.

TESTING AND EVALUATION

(U) The process of converting printed material into computer format requires more than simply choosing a scanning device and feeding a document into it. At least four mutually dependent phases of the digitizing operation

need to be evaluated before deciding which scanner will produce the most accurate and reliable output. In addition to considering (1) the features and capabilities of the scanning device, and (2) the format and clarity of the printed text, the user must give equal consideration to (3) the amount and method of output editing to be done as well as (4) the format and intended use of the digitized product.

(U) It is also possible to have a single restriction as the deciding factor in the selection of a scanner. If, for instance, there is a need to capture a publication having significant intrinsic or monetary value, removing the pages of the document from the binding is not practical. Digitizing in this case is then limited to a device capable of scanning bound book material. Only after all of the variables are evaluated can a user make a reliable determination as to the best scanner for a job.

(C-CCO) Performance evaluation of the three scanners was conducted on a wide variety of documents such as office memorandums, monolingual and multilingual publications, printer text and magazine articles. In addition to English, languages included Amharic, Arabic, French, German, Italian, Polish, Portuguese, Spanish and Russian. The quality of the material ranged from very good for clear, high contrast print of a newer publication, to very poor for older, worn documents whose print has faded through use and age. One goal of the evaluation was to judge the relative capabilities of the three scanners, but the primary emphasis was on determining which scanner produced bilingual digitized output in a form most suitable for incorporation into a relational database.

CONCLUSIONS

(U) Given the rigid constraints of producing bilingual digitized text for a relational database, the flexibility of the Kurzweil 4000 appears to make it the most effective device for the job. This should not be interpreted as a blanket endorsement of the Kurzweil 4000 over the other scanners but merely a best-bet choice for this particular scenario. Less versatile, but nevertheless still a viable alternative, is the IOC SPEEDREADER. The remaining scanning device, the Kurzweil DISCOVER 7320, is least suitable for the purposes described.

	DISCOVER 7320	KURZWEIL 4000	IOC SPEEDREADER
Model Type	Desktop	Stand-alone	Desktop
Cost	$11,950	$29,107 **	$5,495
Speed	50-60 cps	20-50 cps	30-40 cps
Font Size (points)	8-24	6-24	8-12
Graphics Capability	Y	N	Y
Scan Bound Book	N	Y	N
Language Lexicons	50,000 WD	33,000 WD	N
Trainable	N	Y	Y
Concurrent Editing	N	Y	N
Column Recognition	Y	Y	Y
Multifont Recognition	Y	Y	Y
Parsable Output	N	Y	Y
Scan Bilingual Text	N	Y	Y

** The cost for the Kurzweil 4000 includes the optional electronic tablet ($3097), automatic document feeder ($2655) and four language lexicons ($1425 each)

Table 1. COMPARISON OF OCR DEVICES (U)

ESTIMATES OF ACCURACY

(U) Table 1 provides an estimation of the accuracy one could expect from each of the optical character readers. Percentage of accurately-read characters is based on results of test material or represents computations derived from a comparison of characters within scripts. It assumes that all text is monolingual and of high quality.

PRODUCTION OF A DIGITIZED DICTIONARY

(S-CCO) After preliminary testing of the three scanners was completed, R831 decided to begin production of a digitized version of a text for inclusion in a prototype of the [   ] database system. Callaham's Russian-English Chemical and Polytechnical Dictionary was chosen as the candidate for this effort because of its high priority status in the queue of working aids for incorporation into the [   ] database system and because of its high-contrast, good quality text. Being a bilingual dictionary, it also gives a true representation of the challenges R831 faces in converting hardcopy publications to relational database form.

(U) The Callaham is 852 pages long with two columns of text per page. Most columns contain 62 lines of text and an average of approximately 32 characters (including spaces) per line. This is roughly 3,400,000 characters of input, which produces output of approximately 4,000,000 characters. Included in the output figure are begin-end font indicators and multiple codes for certain individual letters. The text is consists of standard (Roman) English characters (approximately 58 percent), English italics (7 percent) and bold Cyrillic (35 percent). Special characters include diacritical marks (acute accent, grave accent, circumflex, tilde, hacek, diaeresis), degree symbol, percent sign, virgule (/), ellipses (...), em dash, superscript, subscript and Greek letters.

(U) The first step in the procedure for scanning a new document is to define the initial set up parameters. This process involves stepping through a series of menus in order to set the framework for training. The user is prompted to define a document name, training set name, training mode, lexicon, page incrementation option and use of super/sub scripts and underlines. Some textual ambiguity information is also required in order to help the scanner distinguish between 0/O, l/1 and I/l.

(U) After the initial parameters have been defined, a training set for the document is formulated. Training is the process by which the Kurzweil system learns to recognize the set of characters which constitute the document text. The machine is configured to the manual

~~SECRET~~

training mode in which the operator's role is to review text character'by'character or line-by-line and verify that the images displayed on the screen are clear and well-formed. Training should not be done on partial or ill-formed characters. Those characters which do not meet the confidence level of the system recognition logic are displayed in inverse video and are called interventions. All interventions should be verified as correct by the operator or assigned the appropriate identification. All images, not just interventions, should be reviewed by the operator to ensure the training set is not corrupted with incorrect values.

(U) When the operator is satisfied that most of the text characters have been incorporated into the training set it can be copied to system memory for use in production. Training sets are stored in memory and can be used repeatedly as long as the characteristics of the text remain constant.

(U) There is also an automatic training mode which is used to analyze and display characters to the operator once a base level of recognition has been established, that is, after a training set has been created and production is in progress. An existing training set can be supplemented during production through the use of a special function key as long as the storage capacity of 432 images is not exceeded. Exceeding that limit results in a corrupted training set.

(U) Classification of fonts is also done during training. Fonts are numbered 0 through 9 and are flagged on the screen and in the output file as "<n" at the beginning of the font sequence and ">n" at the end. The predominant font in the text, which, in the case of the Callaham Dictionary, is standard English, is assigned font zero. Font zero is maintained only internally and appears in the output file without a font flag. Superscript, subscript and underlines are flagged "<s...>s", "<i...>i" and "<u...>u" respectively.

(U) With the establishment of a satisfactory training set, the operator can now begin the production process by stepping through the menus and setting the scanning parameters. Part of this procedure requires the operator to define page and column limitations by using the electronic tablet. This entails marking the upper left and lower right corners of each

	DISCOVER 7320	KURZWEIL 4000	IOC SPEEDREADER	
Amharic	0	85	75	
Arabic	0	5	3	
Chinese	0	5	3	
Czech	60	97	95	
Danish	85	97	95	
Dutch	95	97	95	
English	95	97	95	
French	95	97	95	
German	95	97	95	
Greek	50	97	95	
Hebrew	0	97	95	
Hungarian	70	97	95	
Italian	95	97	95	
Japanese	0	5	3	
Korean	0	18	11	
Polish	65	97	95	
Portuguese	55	97	95	
Russian (Cyrillic)	5	95	90	
Spanish	95	97	95	
Swedish	85	97	95	

Table 2. Estimates of Accuracy

~~SECRET~~
~~HANDLE VIA COMINT CHANNELS ONLY~~

column of each page of text. When all the page/column boundaries have been defined, the material is ready to be scanned. The operator loads the pages of text into the feeder, or places individual pages or bound documents directly on the scanning window, and sets the optical scanning head into motion. For the Callaham, the pages were removed from the binding and scanned individually. In this case at least, placing the individual pages directly on the scanning window seemed to produce slightly better results than using the document feeder.

(U) Depending on the degree of accuracy required, the operator has a number of options for scrolling through the text as it appears on the screen. The "SCROLL UNLIM" option allows uninterrupted scrolling of text in the editor. This is the quickest way to scan a document but produces the least accurate output since corrections, changes or additions to the training set or output text can e made only bby accessing the file via a word editor or word processor.

(U) A second option is the "SCROLL GOOD" mode which allows scrolling to proceed uninterrupted as long as the rate of interventions (questionable images) is low. If the rate of interventions increases, operator verification is required.

(U) A third mode is the "SHOW (right arrow)" option which allows the operator to skip from intervention to intervention and make appropriate corrections or additions. Unfortunately, using this method alone does not allow for correction of those errors not appearing as interventions, a common occurrence. Other arrow keys allow the cursor to be moved one position at a time and help expedite the editing process.

(U) The final option is the "ACCEPT" mode which scrolls text one line at a time. The "ACCEPT" mode, especially when used in combination with the "SHOW (right arrow)" option and arrow cursor keys, produces the highest accuracy output but requires the most time since the operator is verifying both interventions and non"interventions.

POST-PRODUCTION EDITING

(U) It is estimated that unedited optical scanning of Callaham's Dictionary would result in an accuracy rate (per character) of 85 to 90 percent. Choosing the "ACCEPT/SHOW (right

arrow)" scrolling options, thereby editing the document on a line-by-line basis during scanning, increases the accuracy rate to approximately 96-98 percent. But even such a modest error rate results in approximately 80,000 to 160,000 incorrect characters in the output. Three different methods were used to detect and correct these remaining errors.

(U) The most efficient method is a computer program which is written to reflect inconsistencies in the presentation of the various elements of information contained in the dictionary. For instance, the Callaham has an identifiable, finite number of abbreviations, always appearing within parentheses, which provide contextual information. It is a simple matter then to write a computer program to verify that all abbreviations within parentheses match a list of context abbreviations. Another way we used a computer program in editing was to develop a partial spell checker to help verify the orthographic sequence of dictionary headwords. Approximately five percent of the errors produced through machine scanning were identifiable programmatically.

(U) Another technique in editing errors is employing the Search and Replace functions of a word editor. The nature of optical character recognition, especially when applied to a large volume of text, often produces errors which appear in a consistent, repeated fashion. It is those errors which fall into this category that can best be edited using the Search and Replace functions. For example, in the Callaham the scanner often was unable to distinguish between the character sequences "f.", "fi" and "f" when appearing in italic script. As a result, the part of speech "prefix" was frequently rendered as "prefx". It is a simple matter, then, to globally replace all occurrences of the incorrect sequence using the Search/Replace function. The operator, however, must take care to ensure that the sequence being replaced is not a valid sequence elsewhere in the text. If global replacement is not possible, then replacing the sequence in question must be treated on a case-by-case basis.

~~(C-CCO)~~ Another instance of how the Search/Replace function can be used is when text, although technically correct in the output, needs to be modified to accommodate a database design. For example, since the Callahan has two narrow columns of text per page, words are often hyphenated. When this

output is incorporated into a database such as for [   ] these hyphens must be removed so that an interrogation of the data yields comprehensive retrieval information. This problem is compounded, however, by the fact that many words need to maintain hyphenation. Further complication is encountered because the operator often cannot tell the difference between a word which should have the hyphen removed and one which maintains hyphenation. To illustrate, the average OCR operator may not know if a chemical term such as "hexamethyl-enetetramine" contains a hyphen because it is a compound word or if it is because of columnar limitation in the source document.

(U) Many different types of errors are encountered during digitizing, but generally they can be broken down into two broad categories: those caused by scanner programming and those resulting from problems with the input text. The predominant error by far is incorrect character recognition, that is, the optical scanner detects an image but assigns an incorrect value to that image. This occurs most often when the scanner is forced to perform a task it was not designed to do. The Kurzweil 4000 is intended to scan monolingual text in English, French, German, Dutch, Italian, Spanish, Swedish or Danish.

(U) By trying to scan bilingual text the machine is being forced to recognize characters in a language for which it has not been programmed. Nevertheless, scanning even bilingual text yields a reasonably high accuracy rate as long as the character sets for both languages are based on the Latin alphabet. For example, scanning an Italian-English dictionary is not a major problem for the Kurzweil 4000 because both Italian and English use Latin script. Italian has five characters which are not used in English and English has five characters which are normally used only in borrowed words in Italian publications. Scanning could be accomplished by determining the predominant language in the dictionary, initiating the lexicon for that language and teaching the optical reader to recognize the extra letters from the secondary language.

(U) Accuracy begins to suffer, though, when it becomes necessary to scan bilingual text having significantly different alphabets such as in a Russian-English dictionary. The only method of scanning text presented in non-Latin script such as Cyrillic is to utilize a transliteration

scheme of Latin characters. This can create recognition problems for the scanner, however, since a single image might appear in both character sets and, therefore, have two values. An illustration of this is the image "C" which appears in both the English and Cyrillic alphabets. The scanner needs to assign one value when this image appears in Cyrillic and a different value when it appears in the English character set.

(U) Incorrect character recognition occurs also when images are very similar to each other. The problem appears to be compounded in bilingual text when multiple language fonts are used and one of the fonts represents a non-Latin alphabet. The Cyrillic letters И, Н, П, have similar images and are often indistinguishable to the scanner. The result is an error rate of approximately 10 to 15 percent which is correctable mainly through manual scanning of the output text. But recognition of characters with very similar images in the preprogrammed lexicon results in highly accurate output (at least 97 percent).

(U) Another fairly common error encountered in processing the Callaham is misplacement of the language font indicators. The Kurzweil 4000 often had problems determining when one font ended and another began, especially with a Cyrillic/English italics sequence. This is a very difficult type of error to detect in editing because the flaw does not always happen in a discernibly consistent fashion.

(U) Other errors caused by the scanner itself include incomplete character values (i.e., only two characters of a three-position code are returned in output), omission of certain characters, spurious insertion of characters, and lines of text skipped (rare).

(U) The digitized output of the Callaham also showed a significant amount of errors which were more a result of publishing errors as opposed to scanner interpretation of images. These problems included skewed text, misprints and misspellings in the publication, as well as downright erroneous information.

(U) Although the acceptable error rate varies from document to document, the operator usually will want to identify and correct as many of the errors as is practical. As the type of error becomes more obscure the likelihood that the error will be found decreases. Similarly, the amount of time it takes to find

DOCID: 4010026

SECRET

the error increases. In a large publication such as the Callaham Dictionary, for which a high degree of accuracy was necessary, processing time (scan and edit) averaged approximately one hour per page of input. This produced an extremely low error rate, conservatively estimated to be less than one error per 1000

characters, which facilitated easy assimilation into a relational database.

ERROR COMPARISON

(U) Table 3 compares the type and approximate percentage of errors encountered while scanning the Callaham Dictionary as well as the methods used to correct them.

		METHOD	PERCENT OF CORRECTION	
Error Category	Percentage of Error	On-line During Production	Word Editor	Computer Program
Incorrect Character Standard English	1	90	5	5
Incorrect Character English Italics	15	70	15	15
Incorrect Character Bold Cyrillic	75	80	15	5
Incorrect Font ID	2	50	25	25
Incorrect Font Placement	2	50	25	25
Character omitted	2	75	20	5
Miscellaneous	3	75	20	5
Partial value				
Spurious Characters				
Text omitted				
Publishing errors				
Illegible text				

Table 3. Error Comparison

CRYPTOLOG has moved!

Mailing address:
CRYPTOLOG
P1 NORTH

2nd Issue 1989 * CRYPTOLOG * page 36
SECRET
HANDLE VIA COMINT CHANNELS ONLY

From the Past

HARVEST Machine Support in the Publication of Working Aids to Consumers (U) --
Many consumers require listings and references as working aids which must be
published and updated periodically. If a listing contains 9,000 records, for
example, 300 additions and changes resulting from current analysis may have to
be made to the basic listing every month.

Ordinarily, the original listing is published and distributed to consumers
once a year. When additions and changes are compiled monthly, they are machine
processed against the basic file through some file maintenance program. Work
listing are then sent to the sponsor for the next month's updating. The sponsor
usually rewrites the current changes and additions in a format suitable for
presentation to the consumer and proceeds to type his own multilith masters. The
consumer then has the burden of making pen-and-ink changes to his copy of the
basic document.

Under a new HARVEST process, the changes and additions, after they have
been used to update the basic file, are utilized to prepare a set of supplemen-
tary multilith masters. This is done by extracting from the newly updated basic
file all records which have been affected. These extracted records are then
arranged in exactly the same format and sequences as they appear in the basic
document. These by-product masters are then sent to the sponsor along with
his new listings of the basic file. He needs only to add his introduction and
release them for publication. All monthly change files are retained in the
machine system so that they can be accumulated to produce a cumulative master
file with only the latest changes included. This, in effect, produces for the
consumer, one current supplement containing all changes and additions to date.
Obviously, he can then destroy his preceding month's supplement. For him to
find the latest information, he need only look first in his current supplement.
If he finds the record, he knows he need not look any further; he has the latest
data. If he does not find it, he must look in his basic document knowing that
the basic entry is still valid. This technique does away with pen-and-ink
changes.

Multilith masters, as a rule, result in a rather bulky and awkward document.
Since pen-and-ink changes have been eliminated, a use of multilith masters can
also be eliminated by replacing them with #16 paper. As many as 108 records can
be listed on the #16 paper, whereas the usual number of records on multilith is
a maximum of 72. L24 photographically reduces the #16 paper image to a letter
sized sheet - an 8 x 10½ inch page. The resulting publication is much more
compact and usable. For example, in one publication, the multilith method
produced 534 pages. The same product using the #16 paper resulted in a document
containing only 152 pages. Incidently, in the multilith method, L24 estimated
that 190 man-hours were required and the cost of printing and binding was $1,757.33
to produce 533 copies. Under the latter method, the man-hours required were 52,
at a cost of $652.72 to produce the same number of copies. One hundred and thirty
seven man-hours were saved and $1,104.60 were cut from the production cost.
According to L24, the image area and number of copies should determine when #16
paper should be used. When the number of copies is small, and the format of the
report (image) can be accommodated on standard multilith masters, the masters
should be used.

Some working aids have elaborate introductory narratives. In some cases
these have been put on magnetic tape as a separate file from which multilith
masters or #16 paper listings are made, thereby saving the sponsor from typing
his introduction.

The technique described in this article could be applicable to any machine
process which has a file maintenance program.

getting started

On writing an article

1. My object is to:

☐ Make a report on something that has been accomplished

☐ Call attention to something that is wrong

☐ Suggest a better method or idea

☐ Share a personal experience

☐ Enlist support

☐ Explain a process

☐ Ask a question

☐ Defend a principle

☐ Report news or announce a coming event

☐ Recognize an achievement

☐ Amuse and entertain

☐ React to something someone else has written

☐ None of the above, but something else, namely: _____

2. My working title is:

☐ A better way to _____

☐ Hurray for _____

☐ The fallacy of _____

☐ It's time to _____

☐ The scandal of _____

☐ Are we paying too much for _____?

☐ After _____, What?

☐ A proposal for _____

☐ What happened at_____

☐ Why I agree/disagree with _____who wrote _____

☐ A funny thing happened on the way to _____

☐ None of the above, but _____

3. I will consider this, in my own mind, an open letter to _____ and will address to primarily to her/him/them.

4. ☐ This would be a particularly good time for such an article to appear because _____ *(or)*

   ☐ This subject is timeless

5. ☐ I am especially well qualified to write on this because _____
        (or, on the other hand)

   ☐ I want to speak up on this even though I am no expert on it, because_____

6. Some tentative suggestions for a final title are: _____
   _____
   _____

7. I think it would be most appropriate for publication in:
   ☐ CRYPTOLOG   ☐ The NSA Newsletter
   ☐ The Cryptologic Quarterly   ☐ PQS Update
   ☐ CLArion   ☐ Vox Topics

To the Editor:

(U) A version of my trip report on the American Mathematical (AMS) meeting in Providence last August appeared recently in CRYPTOLOG, and, *to my surprise, cause a controversy.* Buried on p. 23, in a description of a talk critical of AMS by former presidential science advisor Edward David, was the following section, intended as a sarcastic criticism of the mathematical establishment that runs the AMS:

> Mathematicians must find better ways to attract young people to our field - NOT as currently, take it or leave it. The young have been leaving it in droves for adjacent fields . . . Most students think we aim at winnowing out the weak and grinding down the ungifted. [They are right. Most outstanding mathematicians view the world differently from David, not a mathematician. There is only a comparative handful of people in any generation capable of making important breakthroughs in theoretical pure mathematics. They are of interest to us. The rest are a waste of our time.] Academic education should be a pump, not a filter.

(U) My own views (and David's) are clearly represented by the last sentence. The "we," "our," and "us" are meant to be the sort of "outstanding mathematicians" that have power in AMS and that constituted David's audience, the people who are driving young people from the field. Surely it was never intended that "us" would be taken to mean NSA.

(U) Neither I nor anyone who knows me will take it that I think of myself as in the category of "outstanding mathematicians." I saw enough of them in graduate school to doubt I would want to be one of them if I could. While my brother-in-law in academia would not agree, I personally see little reason for society to tolerate people who undermine the general quality of university education because they care only about obtaining "breakthroughs in theoretical pure mathematics." While exposure to people doing research is part of training new mathematicians, it is only part of the process.

(U) I hope that this straightens out any misunderstandings. I do have to admit that when

LETTERS

I attend a math conference, I always feel a bit of conflict between my role as member (of rather mediocre abilities as any star of the AMS will freely tell me) of the mathematics society and my role as government bureaucrat seeing the considerable failings of the professional quality mathematicians from society's point of view. But I have long since stopped cult worship of the "great men" of mathematics, and taken a broader view that mathematicians, like everybody else, must justify their salaries by arguments that involve more than their talent at "glass bead games." In this I believe I echo Admiral Studeman's speech to the AMS in Phoenix.

David Harris, R512

(U) We regret to report the death of David Harris, a valued contributor to the technical health of NSA. An appreciation will be published in a future issue.

To the Editor:

(U) The "anonymous" *Cryptopoem* published in the 1st Issue 1989 was in fact written by [redacted] sometime prior to 1971.

P.L. 86-36

[redacted] A532

EO 1.4.(d)

To the Editor:

(FOUO) [redacted] mentioned in 1st Issue 1989, page 13, there is a copy available in the T53 Technical Library (FANX II, 968-8611), Accession Number S-170,832. There are also a few spare copies in the R51 Mathematics Library (FANX III, 968-8580).

(U) If there is sufficient demand for this document, we will look into another production run. In the interim, we see no reason why this document could not be reproduced as needed. Please address your inquiries to the undersigned.

[redacted] R51, FANX III

To the Editor:

(U) My initial reaction to [redacted] article "Where are Our Textbooks?" (CRYPTOLOG, 1st Issue 1989) was that Mr. Gaddy was out of touch with the operational world because there are plenty of good technical documents being written at NSA. I mentally started ticking off textbooks, thought pieces, working aids, and other types of technical documents whose absence Mr. Gaddy was bemoaning. Then I realized most of these examples had never been formally published.

(U) Why aren't a number of our best technical documents being published? Quite simply, the process is so lengthy and frustrating that many good analysts will not put up with the hassle. These people value their time and talents and prefer to expend their efforts on technical problems rather than on pushing a paper through the publication pipeline. Consequently, we have the situation where many analysts are not submitting their papers for publication but are just making copies for their colleagues. Such documents form what is a growing "underground" technical library.

(U) If we want our technical documents to be formally published and available to the general NSA technical population, we are going to have to make the publication process less painful. As long as we force analysts into a "publish or produce" dilemma, we are going to have a shortage of publications.

[redacted] P15

P.L. 86-36

Dear Charlie,

(U) Now that you're in P1 you've landed in clover, documentation-wise. P1 has a technical series and encourages its people to write for it. There's a minimum of fuss, and, moreover, each publication can have its own distribution.

(U) Other organizations have technical series that are similar.

[redacted] *Chief, P1*

P.L. 86-36

CRYPTOLOG

welcomes your letters and comments.

Send them to The Editor, P1, NORTH

NOTE THE NEW MAILING ADDRESS

On the Lighter Side

Introducing Exciting Friedman **BINGO!**

Another boring briefing in Friedman? Never again, with the POLEMIC's new FRIEDMAN BINGO game, guaranteed to keep even those with the heaviest eyelids awake and on the edge of their seats during any talk.

Just cover up a square any time you see one of the events on the FRIEDMAN BINGO game. Get five in a row (horizontally, vertically, or diagonally) and you're a winner!

Take FRIEDMAN BINGO to your next briefing. Never a dull moment again at CMI, HINT, KRYPTOS, CISI, CLA, or any other presentation!

B	I	N	G	O
Lights Dim	Phone Rings in Projection Booth	4 or More Interns in Your Row	Chair Squeaks During Talk	No One Asks Questions
Man Snoring in Back	Sign Language Interpreter Present	Slides Change without Speaker Saying "Next"	A Former Supervisor is in Attendance	Microphone Falls Off Speaker
Speaker Says "uh ..." Twice Consecutively	Slide Upside Down	FREE SPACE	Your Supervisor Walks in Very Late	Book Awarded is NOT Math-related
"Wiring Diagram" on the Screen	Speaker Adjusts Height of Lectern	You see Sign "This Entire Row Reserved"	2 People Asleep in Your Row	Speaker Pours Water From Pitcher
Number of People in Your Row Divisible by 3	Introducer of Speaker is Introduced	Speaker Begins With Joke	Someone Leaves Talk Early	Director is in Attendance

Courtesy of POLEMICS

SECRET