Kicking the Hardcopy Habit (U)

(U) Most of the analysts I have known insisted on hardcopy to perform their daily tasks. Regardless of the amount of data they could be reasonably expected to analyze, they usually printed volumes more, often in multiple copies. These listings were anxiously awaited and generally put to use upon arrival. In many cases, due to the pressures of time or presence of other duties, only part of this material could be fully digested, and the remainder was virtually untouched. Although some printouts were disposed of when tomorrow's data arrived, many were retained for possible future reference. Eventually the shelf life of these lists outlived their utility and they were discarded.

(U) Although this may sound somewhat wasteful, it should be clearly stated that the intentions of these analysts were honorable. They honestly believed that they had pared their retrieval requirements to the bone and that the data they had asked for represented only what they actually required to perform their analysis. And, it should also be noted, they were often surprised at the information overload which their "minimum daily requirements" represented.

(U) There was a Catch-22 associated with attempts to reduce the volume of this hardcopy, however. Reducing the print volumes ran the risk of depriving the analysts of potentially significant data unless the traffic could be made available to them in some other medium. Employing anything other than hardcopy for data distribution encountered head on the problem of analytic acculturation. These analysts, for the most part preferred hardcopy, period.

(U) Asking a confirmed analyst of this ilk to give up his daily ration of hardcopy is almost like asking a smoker to give up his habit. Telling either one how much better off he will be falls on deaf ears, and the more you marshall facts and reports corroborating the wisdom of your suggestion, the more entrenched he becomes. Hardcopy is so integral to analysis that most analysts I know would happily risk drowning in a sea of paper rather than give serious consideration to alternatives.

~~(C-CCO)~~ Such is not the case in G7, however, for the analysts working the [redacted] problem are enthusiastically converting to a softcopy distribution system for their daily traffic. This system, called OVENWARE, consists of a minicomputer which organizes the data files for each analyst, a CRT at the analyst's desk for scanning these files and inexpensive printers in the analytic area for those few messages which truly must be converted to hardcopy form.

~~(C-CCO)~~ Read on, my analytic friends, for to understand this apparent heresy, you must first have an appreciation of the hardcopy problem your counterparts in G7 faced. The [redacted] cradle-to-grave process is relatively straightforward. [redacted] traffic is intercepted by a number of collectors and forwarded electrically to NSA. The data is received by the HOLDER communications processing system where it is temporarily staged and periodically transmitted to the STARSHELL system for processing.

~~(C-CCO)~~ [redacted] traffic can contain almost anything, originate virtually anywhere and refer to any subject imaginable. To bring some order to this chaos, the traffic is processed by the CHESWICK software system on STARSHELL. CHESWICK performs message separation and then does a full-text scan on each message, matching words, terms and phrases against resident dictionaries. Every "hit" equates to a category," and every "category" equates to a specific analyst in a particular office. An individual message, therefore, might "hit" on a series of categories and, as a consequence, must be distributed to several analysts.

(U) It sounds simple enough, but still doesn't explain giving up hardcopy. The problem lies in the ambiguity inherent in [redacted] intercept itself. Textual analysis by computer is not an exact science, and CHESWICK, despite all its sophistication, cannot

yet distinguish between those messages of intelligence significance which "hit" and those utterly without redeeming value which also "hit" on the same category. As such, all hits must be passed on to the analyst for this determination to be made, and that is where softcopy comes into play.

~~(C-CCO)~~ While it was my intent to minimize the percentages, statistics, and volumes associated with [redacted] distribution, some facts and figures are necessary to convey the full flavor of the problem. G7 is the ultimate recipient of over 85 percent of our daily [redacted] traffic. The STARSHELL operations people process CHESWICK as often as twenty times a day, but all of the printing is done overnight. [redacted] printing averages about 85 to 90 boxes of paper per night, with peaks exceeding 100 boxes not uncommon.

~~(FOUO)~~ The cost of all this printing is phenomenal. The paper costs exceed $13 per box, equating to $1300 per day in paper alone. It takes 30 minutes to print each box — roughly 45 to 50 machine-hours of steady printing each night. In addition to the cost of the printers involved, there is operator time — 6 minutes per box to set up, stack, unstack, etc., which adds up to 9 hours or more of operator time just to print all this traffic for morning pickup.

~~(C-CCO)~~ The daily fetch-and-carry ritual has an impact on G Group manpower. Simply transporting 70 or so side feet of computer printouts from operations to a central pickup point in G takes about 3 hours per day. There the traffic is separated for each G7 division; subsequent pickup and dispersal to the category analyst involves another 6 man-hours per day.

~~(C-CCO)~~ The daily printing and distribution costs alone for [redacted] traffic can be summarized as follows:

Paper Costs	$1300/day	
Printer Costs	300/day	
Operator Time		9 hours/day
Pickup Time		3 hours/day
Dispersal Time		6 hours/day
TOTALS	$1600/day	18 hours/day

~~(C-CCO)~~ Now that we have expended $1600 and 18 precious man-hours to get this intercept to an analyst, what becomes of it? Out of all this traffic, approximately 90 percent is discarded immediately as having no intelligence value; something less than 5 percent is ever used in product.

~~(C-CCO)~~ G7 produces about 20,000 burnbags per year and each analyst devotes at least 3 minutes a day to burnbag functions. With over [redacted] disposing of unneeded intercept alone amounts to over 6 man-hours per day lost to analysis. The figures don't stop there either, as 20,000 burnbags amount to about 100 tons annually. L5 personnel have figured the total disposal cost of our classified paper at $254 per ton, equating to another $25,000 a year just to dispose of the [redacted] traffic having no Sigint value.

~~(C-CCO)~~ Faced with volumes of this magnitude, the [redacted] problem cried out for support. The OVENWARE systems, of which three will be operational this year, are the first steps in this direction. OVENWARE provides [redacted] analysts with one essential function — softcopy distribution of their traffic. The system has a communications link to STARSHELL, and whenever the CHESWICK process is executed, [redacted] comes roaring down the line to OVENWARE.

~~(FOUO)~~ Users log on their terminals and are informed which of their files have been received and are available for scanning. Although there are provisions for selecting traffic for display (e.g., by category, by keyword, by keyword within category), most analysts peruse their files sequentially, generally starting with the most recent transactions available.

~~(FOUO)~~ While scanning, the analyst has four disposal options available. He can take no action on a particular message, or he can delete, print, or save it. Since the local printers are shared by the users, printing is not immediate. Messages marked for printing are queued until the analyst has completed his scan and either starts to read another file or logs off his terminal. When printing takes place, each analyst's traffic is kept separate from messages requested by other analysts.

~~(FOUO)~~ Taking the "no-action" course will leave the message in the transaction file for rescanning at the discretion of the user. Deleting messages entails marking them for removal but not actually purging them. Subsequent browsing through this file will "skip over" those messages marked for deletion, but they can still be recovered for display should the situation warrant. Physical purging takes place overnight.

~~(C-CCO)~~ The fourth disposal option — save the message — has proved to be a real boon to term analysis. Users have the capability to create an unlimited number of save files, affording each the opportunity to pigeonhole his messages for subsequent use. The analysts topically aggregate subfiles of homogeneous traffic during the scanning cycle and later read the save files for their analysis and reporting functions. In that [redacted] traffic can be in various languages, innovative G7 users have created translation pools into which they insert foreign language messages from their categories of analytic specialties

(b) (1)
(b) (3)-50 USC 403
(b) (3)-18 USC 798
(b) (3)-P.L. 86-36

for review and translation by qualified office linguists. Similarly, analysts create save files to pass messages of potential interest to one another. A message in one user's category may be of no value to him, but contextually could be significant to another analyst who did not receive it because it did not "hit" on any of his categories. This message is then passed to him through a save file for his review and analysis.

~~(C-CCO)~~ Since many categories of [redacted] traffic often concern proposals and subsequent responses, the save files provide an excellent means of organizing the traffic for analysis. A proposal received today may not be responded to for a month or more. With OVENWARE, G7 users have a flexible mechanism to connect related pieces of information received weeks or months apart. They simply create a topical save file and store all pertinent messages there, available for easy recall when more traffic on the same or a related subject is noted.

~~(FOUO)~~ OVENWARE has a number of other features, but they are not germane to the point of this article. In summary, OVENWARE has provided a number of significant benefits to the Agency at large. With data transmitted upon completion of the dictionary match process, transactions are available for scanning up to 18 hours sooner than before. The system similarly provides scanning efficiencies over the labor-intensive manual scan and provides improved product reporting facilities to G7 analysts.

~~(C-CCO)~~ By obviating [redacted] printing, tremendous benefits accrue in cost avoidance, reduced waste, and manpower savings. Savings in paper alone exceed $100,000 per system per year, and in less than three years OVENWARE systems have recouped their purchase price in cost avoidance items alone. Additionally, these systems will result in an annual net savings of three man-years in such critical skill areas as experienced [redacted] analysts and short-supply computer operations personnel.

~~(C)~~ In the area of reduced waste, 100 tons a year sounds like a prodigious amount. I'd like to report that eliminating most of that volume will have a substantial impact on our classified waste system. Unfortunately, that is not the case. L51 picks up about 25 tons of classified paper trash each workday, so even if OVENWARE totally eliminated G7's paper load, it would amount to less than 4 days worth of L51's annual collection.

(U) Twenty-five tons per day of classified paper waste is a staggering figure which clearly indicates our hardcopy production is excessive. The manpower savings and cost avoidance benefits accruing from softcopy distribution are enormous, and are undoubtedly applicable across a wide spectrum of Agency production and support elements. There are numerous projects currently under way designed to further automate our efforts and better support each of our missions. If softcopy distribution isn't integral to your planned or developing system, take the opportunity now to kick your hardcopy habit. From a cost perspective alone, volume hardcopy is seldom worth the paper it's printed on, and with austerity the watchword, softcopy distribution can provide you significant savings. If you are reluctant to rely completely on automation, consider partial softcopy distribution. Even incremental implementation can provide a substantial cumulative impact. When you consider all the heretofore hidden costs associated with the standard way of doing business, you'll have to agree that softcopy distribution is more cost effective. If you can switch to softcopy, as G7 has done, you won't regret it.

(U) [redacted] rofessionalized as a data systems analyst, currently serves as Chief, T424 (Project Management Division). Since joining the Agency in 1961, he has had a variety of data processing assignments, primarily in A, C and T elements and including a tour in Germany. [redacted]

(b)(3)-P.L. 86-36

(b)(1)
(b)(3)-50 USC 403
(b)(3)-18 USC 798
(b)(3)-P.L. 86-36