

DIRECTOR OF  
CENTRAL  
INTELLIGENCE

# Preparing US Intelligence for the Information Age

*Coping With the Information Overload*

*Scientific and Technical  
Intelligence Committee*

*STIC 93-001  
January 1993*

## Note to Readers

The mission (DCI Directive 3/10) of the Scientific and Technical Intelligence Committee (STIC) is to advise and assist the Director of Central Intelligence (DCI) in the discharge of his duties and responsibilities with respect to the production, coordination, and evaluation of intelligence on foreign scientific and technical developments and to promote the effective use of Intelligence Community resources for this purpose.

### Members

|   |  |
|---|--|
| Dr. James T. Leong<br><i>(Chairman)</i> | Central Intelligence Agency                            |
| Dr. Robert L. Bingham                   | Department of Energy                                   |
| Dr. John G. Dardis                      | Department of State                                    |
| Dr. Joe E. Dollar                       | Department of the Air Force                            |
| Mr. Raymond B. Firehock                 | Arms Control and Disarmament Agency                    |
| Dr. John M. MacCallum, Jr.              | Office of the Under Secretary of Defense (Acquisition) |
| Mr. Melvin Potosky                      | National Security Agency                               |
| Dr. William Reese                       | Department of the Navy                                 |
| Dr. George A. Sinnott                   | Department of Commerce                                 |
| Dr. Bertram B. Smith, Jr.               | Department of the Army                                 |
| Mr. George Weber                        | Defense Intelligence Agency                            |

### Alternate and Special Members

|   |                             |
|---|-----------------------------|
| Mr. Robert P. Gallagher                               | Department of Commerce      |
| Mr. Robert L. Gerrity                                 | Department of the Navy      |
| Ms. Glenda K. Griffin                                 | Department of the Army      |
| Mr. David S. Jackson                                  | Central Intelligence Agency |
| Mr. Joseph Masonis                                    | US Secret Service           |
| Ms. Mary B. Scott                                     | Department of the Army      |
| Capt. Lauren K. Wagner                                | Department of the Air Force |
| Mr. Daniel R. Cantley<br><i>(Executive Secretary)</i> | Central Intelligence Agency |

## Preface

The STIC Open-source Subcommittee is comprised of members from throughout the Scientific and Technical Intelligence Community whose objective is to lead efforts advocating more effective use of open-source information in all-source analysis. Over the past 14 months, the subcommittee has identified some fundamental shortcomings in the capabilities of the Intelligence Community to deal with the volume of open-source information that is currently available -- to say nothing of the massive increases expected in the not-too-distant future.

Parts of the Community have surmised that our problem is not collecting more open-source information but rather just effectively using what we already collect. We certainly collect more information than the all-source analysts can read, but that does not mean we collect all the open-source information that is needed. Our problem is collecting all that we need and then effectively sifting it down to a manageable set of information that supports our analyses.

The situation creates an interesting paradox of information overload. Many analysts already decry their inundation with more information than they can possibly analyze. However, the recent shift in interest to "rest-of-world" analysis dictates the need for not less, but dramatically more information in the future. The problem then is twofold. First, we must increase the user's access to open-source information. Second, we must provide the automated tools necessary to allow effective exploitation and analysis of that information.

The need to address the issue of automated tools development cannot be overstated. The raw information is there and the technology to acquire it is available. However, effective methods for its timely use in intelligence analysis are limited. Once acquired, information must be prepared or preprocessed for analysis. That is, the small amount of "good"

information must be sifted out from the vast quantities of raw information. This preprocessing is one key to establishing open-source information as the "first source" of information for the all-source analyst. The other key is the analysts' need for automated tools to assist in analyzing it.

All disciplines require specialized tools for analysis, be they modeling and simulation packages or imagery mensuration capabilities. Open-source information is no different, and many of the tools that apply to open-source information will also be applicable to a broad array of intelligence information. Thus, a concerted effort must be started to develop tools for more effective, systematic and timely analysis of open-source information. Furthermore, the development of such tools is only the first step. They must also be made available throughout the Intelligence Community and analysts must be trained to use them effectively.

The primary purpose of publishing this paper is to promote awareness of the problems this subcommittee has identified. Future study will focus on identifying deficiencies in particular mission areas and on defining specific issues that must be addressed.

This is a republication of an article published by a former member of the subcommittee, Mr. Andrew Shepard, CIA, that was first distributed on a limited basis in April, 1992.

## Key Judgments

- The collection of open-source information is inadequate to support the analyses expected of the Intelligence Community. Moreover, the gross inadequacy of automated exploitation and analysis tools leads to the false premise that too much information is being collected.
- A special emphasis to develop breakthroughs or paradigm shifts in approaches is necessary to cope with the combined need for more information and the information overload.
- Automated tools will be needed for open-source information exploitation, library and other intermediary activities, and analysis.
- Underlying all of this is the a special need to provide a significant investment in prototyping more efficient ways of exploiting information. This will permit informed decisions to be made that range from technology development to analyst training.

## Introduction

The STIC Open-Source Subcommittee believes there is an urgent need to develop automated tools for coping with information overload. This report, the first in a three-part series undertaken by the subcommittee, will give an awareness of the extent of the problem.

The next report will focus on a mission needs analysis for tools. The final report will detail some specifications of the capabilities required of the tools.

## Background

*"There is increasing evidence that we are being bogged down today as specialization extends. The*

*investigator is staggered by the findings and conclusions of thousands of other workers--conclusions which he cannot find time to grasp, much less remember, as they appear."*

*Vannevar Bush, 1945*<sup>1</sup>

*"To handle all these data, from both human and technical sources, a dizzying bureaucracy has grown up which applies the factory principle of the division of labor, breaking production into a series of steps. Many corporations today are learning that this form of sequential production is inadequate. Finding the right piece of information, analyzing it correctly, and getting it to the right consumer in time are turning out to be bigger problems than collecting it in the first place.*

*Alvin Toffler, 1990*<sup>2</sup>

The global production of information has grown rapidly in the past several decades and promises to grow even faster in the decade ahead, as information technology continues to make production and storage easier and cheaper. At the same time, changes in global politics have made open-source information more important to the intelligence consumer. The Intelligence Community must make more effective use of open information to keep US decisionmakers abreast of threats and opportunities. But a major problem blocks our path: we do not have sufficient time to read all the openly available information that now is on hand.

To exploit available information, we must make a major breakthrough in our ability to absorb US and foreign documents. We need to figure out how analysts can benefit from information in documents they have never read. We need a process that accommodates ten or a hundred times more information than the analyst has time to monitor as it comes in, and yet one that gives every unseen document a chance to influence the judgments we pass along to consumers.

To call this a challenge understates the enormity of the problem; to call it impossible concedes defeat and recklessly invites others to steal a march on us. Decades-old procedures for handling information cannot be tweaked to solve the problem; we must begin redesigning how we work.

## Reviewing the Problem

Coping with information overload is seen in different ways by different people, notwithstanding innumerable studies over the years that have tried to determine how to use information more effectively. Some people see no problem -- we simply read as much information as time permits, and then draw our conclusions; the product is therefore as good as we can afford to make it. For others, information overload is seriously undermining the quality of intelligence. This paper provides a user's view of the data-overload problem, with the aim of influencing what we might choose to do about it.

The list that follows represents part of an analyst's inbox. (The analyst is not the only worker affected by information overload, but provides a good example.) It shows part of the newly arrived documents on topics the analyst was tasked to monitor. And it is prioritized according to the probable relevance and importance of each document competing for the analyst's time and attention. During the day, the analyst works through the list of documents, taking appropriate actions that include filing, sharing, or ignoring some documents; following up on others with specific collection requirements; correlating some with past reporting to provide better context or to check some hunches; and rendering one or more into a draft of finished intelligence that would be useful to a consumer. By the end of the typical day, the analyst has only reached number 79.

Friday, June 13: You have 312 new documents in your inbox.

72. Foreign media report: New proposals on land reform
73. Imagery report: Continued activity at the new install
74. NY Times: French authorities suspect terrorist involve
75. State cable: Deliberations stall progress on economic
76. Clandestine report: Dictator likely to fire chief of sta
77. Wall Street Journal: Multinationals increasingly reac
78. Science: Desalination plants in Middle East will not b
79. Clandestine report: Plans for land reform will be d
80. Newsweek: Capital investment in irrigation dependen
81. Foreign media report: Qui trouve les problems n'aura
82. State cable: Chances for collaboration diminishing a
83. LA Times: Foreign interest shifting from US to Sout
84. Hudson Institute: A Comparative Analysis of Gulf S
85. Foreign media report: Ethnic divisions seen as fiction
86. Imagery report: Construction at a standstill at termin
87. Foreign media report: Professor receives special reco
88. Wall Street Journal: Northern Africa seen as next lik
89. NY Times: Clash over patent rights imperils coopera
90. Clandestine report: Agreement on oil exploration link
91. State cable: Information sharing tied to export appro
92. CNN: Foreign Minister blames bureaucratic snafu for

A key question raised here is, "What becomes of the information in the unseen documents, (80-312)?" The chances are that they will never be seen by this analyst, who will get a newly prioritized list of documents to review the next day. There are several reasons to be concerned.

All the documents delivered to the analyst are potentially important, having already met some criteria that defined them as relevant to issues needing to be monitored. Because the prioritization process cannot be perfect, it is highly probable that some of the information in the unseen documents is more important--more worthy of action--than information in the top 79 documents. Although some documents should be ignored, most are likely to be worth filing for future use.

The prioritization process is necessary but inherently dangerous. It is based on analyst expectations about the data that would be relevant to particular issues--and therefore filters incoming data according to its consistency with analyst expectations. The documents that are most likely to point the analyst to errors in his expectations may end up near the

bottom of the list. In a situation of increasing information overload, where the analyst reads a diminishing fraction of all the pertinent documents each day, reliance on the prioritization is likely to make expectations increasingly self-fulfilling.

When an analyst knows that there are 312 documents purporting to address consumer requirements and that unseen documents cannot figure into his/her analysis, there is chronic pressure to review as many of the documents as possible. One consequence is that the analyst spends less time with each document, in order to review as many as possible. In this rush, the content of even highly ranked documents is less likely to be understood in depth before decisions are made on their disposition. Clues may be missed.

The problem worsens in periods of heightened demand, when the analyst is likely to see incoming traffic increase several fold. Surge capability is impeded.

The need of analysts to see as many documents as possible before giving up also has undesirable effects on other areas of analyst responsibilities. As long as there is unread mail in the analyst's inbox, it is a compelling distraction that reduces the time spent on collection tasking, coordination with other analysts, research on consumer needs, monitoring of contract work, maintenance of professional skills, and on analysis itself.

Information overload tends to shift the balance of research away from in-depth analysis and toward current intelligence. Analysts can be successful either by reading documents in depth and correlating them with past reporting to check for subtle changes--impressing managers with their analytic insight--or by scanning large volumes of data to spot items that stand out as obvious candidates for repackaging as current intelligence, which boosts production. Analysts strike a balance that responds to the local reward structure, but the pressure to see more of

the available documents probably is a factor that favors production of single-source intelligence pieces.

Information overload has been identified as a factor in analytic errors of judgment, apart from its effect on the amount of time available for in-depth analysis. Our human tendency is to interpret evidence as support for preconceived notions, even if it supports many possible notions--and we tend to completely overlook evidence that does not fit with our ideas. As we race through increasing amounts of daily traffic, the potential for "cognitive dissonance" is high.

The list of 312 documents does not include many of the other sources of information that could improve analytic judgment. Documents such as books, video documentaries, certain magazine articles, trip reports, encyclopedic reference works, phone directories, and maps are not included if they do not fit the category of documents that an analyst wants to see on his screen as he is trying to keep up with foreign events. Without a mechanism for such information to enter the analytic train of thought, the analyst becomes a product of the inbox, acquiring a perspective that may be too narrow.

Some important information does not exist in any single document, and is discernible only by comparing many documents. The search for patterns of foreign activity that show up only when a large number of documents are compared is time consuming and, therefore, is probably one of the most frequent casualties of information overload.

This problem will worsen. Global information services and communication systems are making news travel fast, and they will soon make news travel selectively, directly to people who care most about particular news items.<sup>3-8</sup> At the same time, we see increasing globalization of companies, industries, and economies - which is complicating the distinction between government and

corporate interests in foreign affairs.<sup>9,10</sup> The effect of these two trends is that ordinary people in democratic countries are gaining power to influence foreign affairs policies both through their governments and through pressure on corporations (through such devices as pension fund investments). More to the point, people who have a policy agenda will be able to mobilize popular sentiment more effectively. Global developments that can be cast in terms that matter to ordinary citizens will more often require a response from their governments--on issues that range from students facing down tanks to foreign factories spewing pollution. Five to 10 years from now, therefore, our leaders probably will have to devote more attention to more kinds of openly reported events. As the real-time nature of commercial news forces policymakers to respond more rapidly to foreign actions and statements, intelligence will be pressed to add its value more rapidly as well. This pressure translates into a need for analysts to review incoming information and correlate it with previous reporting more quickly. At present, we leave it up to the analyst to make the critical decisions about how much time to spend reviewing incoming documents, and how best to use that time. We train, guide, and reward the analyst in ways geared to bring about incremental improvements in productivity, creativity, openmindedness, and analytic production--but we have yet to give analysts tools that would make a profound difference. And we have yet to establish a coherent search for new tools and operating concepts.

### **In Search of a Breakthrough**

Because present operations appear inadequate for the future, we need to reexamine the process that US Intelligence uses for monitoring foreign developments. Today's analysts are doing their heroic best, but modern information technology allows a far greater capability. It will not be plausible in the future to claim that we exploit as much of the available information as possible, if we have failed

to apply the technology that now goes with our trade.

What, then, are some of the options available for making a breakthrough in the overload problem? Ideas abound, some dating back to the 1940s. It is beyond the intended scope of this paper to describe them. Instead, a few guidelines are suggested to aid in the development of options. The most important decision at this point is simply to set high enough goals. For example:

- We need to be working toward a capability to handle orders of magnitude more data; we cannot afford to spend time and money on proposals designed for only modest gains.
- We need automated correlation of new information; we do not have enough employees for procedures that require a document to be read as it enters the building. We also need to bring to the analyst's attention information distributed across multiple documents.
- When one person takes action on a document, our information systems need to make it possible for others to benefit without necessarily having to read the document.
- We need efficient ways to check hunches and develop alternative scenarios--which means we need information retrieval and display systems that allow users to pursue complex trains of thought without disruptions caused by system complexities or the inaccessibility of data.
- We need to deliver news to analysts on topics that worry policymakers--and this increasingly will include news communicated by television and electronic news services. We

need an intelligence process that handles all kinds of data in its prioritization, delivery, retrieval, and production of information.

No one can know exactly how to achieve such ambitious goals without trying some pilot projects first. By comparison, it is possible to begin any number of more modest initiatives with high confidence of their effect--and therein lies a dilemma. The initiatives we can be confident of accomplishing at a specific time for a specific cost are modest in their effect; and the initiatives that would produce the gains we need are not ready for management commitment. If we are to overcome the problem of data overload, we cannot tell top management that we know what to build and how to reorganize; we can only propose a way to figure it out, over several years. The most we can do at the outset is to set a strategic direction and pursue a program of development and discovery so as to create reliable data on costs, risks, and benefits that managers need to see before committing to large-scale applications. We need to begin today.

It may be helpful to renew the relevant experience of external organizations. Numerous organizations have found that today's technology allows performance breakthroughs and that the necessary changes to business practices are manageable, even when the solution is not obvious at the outset. Below, for example, are some of the conclusions drawn by external researchers who have studied past efforts to solve problems such as ours:

*"The time is coming when the investment conservatives will gain confidence that even this strange new electronic world is manageable. This discovery will evolve from painful experiences which prove that the fundamentals of good management apply as much to the economics of information work as to the economics of manufactured products."*

*(Strassman, 1985)<sup>11</sup>*

*"So far most computer users still use the new technology only to do faster what they have always done before. But as soon as a company takes the first tentative steps from data to information, its decision processes, management structures, and even the ways its work gets done begin to be transformed. In fact, this is already happening, quite fast, in a number of companies throughout the world."*

*(Drucker, 1988)<sup>12</sup>*

*"Whether in a Japanese insurance company or an American automaker, electronic data interchange forces major changes. When a company goes electronic, jobs change; people move around; some departments gain clout, others lose. The entire relationship of the firm to its suppliers and customers is shaken up."*

*(Toffler, 1990)<sup>2</sup>*

*"The usual methods for boosting performance --process rationalization and automation --haven't yielded the dramatic improvements companies need. In particular, heavy investments in information technology have delivered disappointing results -- largely because companies tend to use technology to mechanize old ways of doing business. They leave the existing processes intact and use computers simply to speed them up. It is time to stop paving the cow paths. We should "reengineer" our businesses: use the power of modern information technology to radically redesign our business processes in order to achieve dramatic improvements in their performance."*

*(Hammer, 1990)<sup>13</sup>*



## **Conclusion**

The data overload problem has serious adverse effects on the ability of US Intelligence to accomplish its mission. These effects, which are likely to worsen, may not be fixable through incremental adjustments to existing procedures. To begin making real progress in coping with data overload, we need to begin with a management commitment--not to a specific system design, but to a program of developing and testing some new concepts of information handling.

The development of tools must span the needs of the all-source analysts, the information specialist intermediaries, and the many nonanalyst exploiters of open-source information.

## References

1. "As We May Think," Vannevar Bush, *Atlantic Monthly* 176(1), pp 101-108, June 1945. The article offered the following vision for efficient exploitation of voluminous information files: Wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them... The lawyer has at his touch the associates' opinions and decisions of his whole experience and of the experience of friends and authorities. The patent attorney has on call the millions of issued patents, with familiar trails to every point of his client's interest. The physician, puzzled by a patient's reactions, strikes the trail established in studying an earlier similar case, and runs rapidly through analogous case histories, with side references to the classics for the pertinent anatomy and histology. The chemist, struggling with the synthesis of an organic compound, has all the chemical literature before him in his laboratory, with trails following the analogies of compounds, the side trails to their physical and chemical behavior.
2. *Powershift: Knowledge, Wealth and Violence at the Edge of the 21st Century*, Alvin Toffler, Bantam Books, 1990.
3. "The Next Revolution," by John Schwartz, *Newsweek*, 6 April, 1992. pp. 42-48.
4. "Newspaper Computer News Service Inaugurated," *The Daily Yomiuri*, Tokyo, 23 October 1991, p. 8. OUO
5. "Competition Fuels Technological Explosion in Telecommunications," *Washington Post*, January 10, 1992.
6. "The French Videotex Network: How To Reach 6 Million MININTEL Users," Advertisement for seminar by B.Rious and F. Libmann, FLA Consultants; materials prepared for the May 1992 National On-Line Meeting, International Organization of Library Science.
7. "Phones Into Orbit," *The Economist*, 28 March, 1992.
8. "Viewpoint" by Senator Albert Gore, Jr.; *Communications of the ACM*, November, 1991.
9. "The Stateless Corporation," *Business Week*, 14 May, 1990
10. "W.R. Grace Shifting Its Scientific Growth And Capital Overseas," *New Technology Week*, April 13, 1992.
11. *Information Payoff--The Transformation of Work in the Electronic Age*, Paul A. Strassman, Free Press, 1985.
12. "The Coming of the New Organization," by Peter F. Drucker, *Harvard Business Review*, January-February, 1988.
13. "Reengineering Work: Don't Automate, Obliterate," by Michael Hammer, *Harvard Business Review*, July-August, 1990.